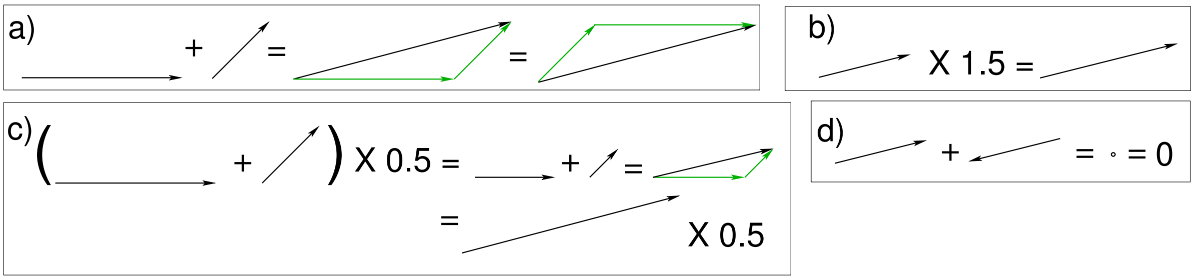
PHYS30201 Mathematical Fundamentals of Quantum Mechanics 2016-17: Solutions 1

1. (i) Let $|v\rangle = (v_1, v_2, v_3)$ etc. $\alpha|v\rangle = (\alpha v_1, \alpha v_2, \alpha v_3)$ which is another complex number triplet for any α; $|v\rangle + |w\rangle = (v_1 + w_1, v_2 + w_2, v_3 + w_3)$ is another. The commutivity and associativity of complex arithmatic ensure the correponding properties of the number triplets: $|w\rangle + |v\rangle = |v\rangle + |w\rangle$; $(|u\rangle + |v\rangle) + |w\rangle = |u\rangle + (|v\rangle + |w\rangle)$; $(\alpha\beta)|v\rangle = \alpha(\beta|v\rangle)$; $1|v\rangle = |v\rangle$; $(\alpha + \beta)|v\rangle = \alpha|v\rangle + \beta|v\rangle$; $\alpha(|u\rangle + |v\rangle) = \alpha|u\rangle + \alpha|v\rangle$. The null vector is $|0\rangle = (0, 0, 0)$, so $|0\rangle + |v\rangle = |v\rangle + |0\rangle = |v\rangle$. The inverse is $| - v\rangle = (-v_1, -v_2, -v_3)$ and $|-v\rangle + |v\rangle = |v\rangle + |-v\rangle = |0\rangle$.

   (ii)



   In the picture above, a) and b) illustrate the rules of addition and scalar multiplication, both of which genererate another vector; a) also shows that addition is commutative; that it is also associative (the order of addition of three or more vectors doesn't matter) is obvious but not shown. Multiplying by unity leaves the vector unchanged (not illustrated). c) illustrates the 1st distributive rule for scalar multiplication, the second is not illustrated but involves showing that, say, multiplying a vector by 2 and by 3 and adding the results is like multiplying it by 5 to start with. The zero vector is an arrow with zero length—a point—and adding it to any vector leaves it unchanged (not illustrated). d) shows that the inverse of a vector is another with the same length but reversed direction, with the sum of the two being the zero vector.
   In both (i) and (ii), results such as the uniqueness of the null and inverse vectors are obvious.

2. (i) Yes. The sum of two functions or the product of a function and a number is another function, and it will satisfy the boundary conditions. Commutivity and associativity for addition, the distributive laws for scalar multiplication and multiplication by unity as the identity all follow from the rules of arithmetic. The zero vector is $f(x) = 0$ (that is, the function equals the number 0 for all x). The inverse vector is $g(x) = -f(x)$ (not the inverse function $f^{-1}(x)$); both the zero vector and the inverse satisfy the boundary conditions.
   (ii) Yes. Focusing only on the differences from the first case, we note that the sum of two periodic functions and a scalar times a periodic function (including -1 for the inverse) all give periodic functions; also $f(x) = 0$ is periodic.
   (iii) No, because the sum $g(x)$ of two functions each with $f(0) = 4$ has $g(0) = 8$, so g is not a vector in the space. (Similarly for scalar multiplication; there is also no zero vector or inverse.)

3. A generic 3rd order polynomial (TOP) has the form $|c\rangle = c_0 + c_1 x + c_2 x^2 + c_3 x^3$, with complex coefficients c_i. The sum of two TOPs $|c\rangle$ and $|d\rangle$ is another $|f\rangle$ with $f_i = c_i + d_i$; similarly $\alpha|c\rangle$ is a TOP with coefficients αc_i. If $\alpha = 1$ the TOP is unchanged. Commutivity, associativity and the distributive laws follow from the rules of arithmetic. The zero vector is the TOP with all coefficients equal to zero. The inverse vector has coefficients $-c_i$.
   The inverse vector corresponding to $3 + ix + (1 - 2i)x^3$ is $-3 - ix - (1 - 2i)x^3$.
   This is a 4-dimensional space, with the obvious basis being 1, x, x^2 and x^3. (Recall that these are all linearly independent: it is not possible to find numbers c_i such that $c_0 + c_1 x + c_2 x^2 + c_3 x^3 = 0$ for all x in some finite range, as opposed to at up to three individual values of x.) Another

obvious basis is Legendre polynomials: $1, x, \frac{3}{2}x^2 - \frac{1}{2}$ and $\frac{5}{2}x^3 - \frac{3}{2}x$. (This is an orthogonal basis with the usual definition of the inner product for functions.)

If only odd functions are considered, we still have a vector space (check the rules) but a 2-dimensional one this time. Both the bases we wrote down previously consisted of functions with definite symmetry, so we just select the odd ones. (If you think of 0 as 0×1, ie $c_0 = 0$, you might worry that the zero vector is not in the space, but $0 = 0 \times x + 0 \times x^3$ here.)

4. i) Assume that there is a vector $|b\rangle$, $|b\rangle \neq |0\rangle$, such that $|v\rangle + |b\rangle = |v\rangle$. Now add $|-v\rangle$ to each side and use $|-v\rangle + |v\rangle = |0\rangle$ which is the definition of $|-v\rangle$. Then we get $|0\rangle + |b\rangle = |0\rangle$. But by the definition of $|0\rangle$ the LHS is $|b\rangle$. Hence we have $|b\rangle = |0\rangle$, which contradicts our initial assumption. Hence no such vector $|b\rangle$ can exist and the zero vector is unique.

   ii) From the rules, $1|v\rangle = |v\rangle$. But $1 = 1 + 0$, so $|v\rangle + 0|v\rangle = |v\rangle$. But this means that $0|v\rangle$ satisfies the definition of the zero vector, and since the latter is unique, it IS the zero vector.

5. a) Dependent since $2(1, 1, 0) + (1, 0, 1) - (3, 2, 1) = 0$.
   b) Independent: the only solution to $a(1, 1, 0) + b(1, 0, 1) + c(0, 1, 1) = 0$ is $a = b = c = 0$.
   $(2, 4, 6) = a(1, 1, 0)+b(1, 0, 1)+c(0, 1, 1) \Rightarrow a+b = 2, a+c = 4, b+c = 6 \Rightarrow a = 0, b = 2, c = 4$.

6. Let $\{|i\rangle\}$ be a set of N linearly independent vectors which is not a basis. If it is not a basis, there is at least one vector $|v\rangle$ in the space which cannot be written as $\sum_i^N v_i|i\rangle$. But that means that the set $\{|1\rangle...|N\rangle, |v\rangle\}$ is a set of $N + 1$ linearly independent vectors. But an N-dimensional space cannot by definition have such a set, so we have reached a contradiction. The initial set must have been a basis, and the addition of any other vector gives a set which is not linearly independent.

7. i) (a) The positive x axis is NOT a subspace, because no vector has an inverse—if $a\mathbf{i}$ is in the space, a must be positive, but then $-a\mathbf{i}$ is not in the space. (Multiplication by (negative) scalars will also take one outside the space.)
   (b) the plane $z = 1$ (vectors of the form $x\mathbf{i} + y\mathbf{j} + \mathbf{k}$) is NOT a subspace. Almost none of the rules of a vector space are satisfied; the sum of two such vectors will be in the plane $z = 2$, for instance. Most obviously the zero vector is not in the space.
   (c) the plane $x+y+z = 0$ (vectors of the form $x\mathbf{i}+y\mathbf{j}-(x+y)\mathbf{k}$) IS a subspace. This consists of all vectors in the plane perpendicular to $\mathbf{i} + \mathbf{j} + \mathbf{k}$ which passes through the origin. The sum of any two vectors in this plane is another in the plane, as is a real number (including -1) times such a vector; finally, the zero vector is in the plane.

   ii) First, note that the set of all vectors orthogonal to a given vector $|v\rangle$ includes the zero vector, since $\langle 0|v\rangle = 0$. So this is like the third example above, not the second. If we started with a basis of the original space which included $|v\rangle$, we could by Gram-Schmidt orthogonalisation construct an orthogonal basis starting with $|v\rangle$, which would consist of $|v\rangle$ and $N - 1$ other vectors orthogonal to $|v\rangle$. These form a basis for an $N - 1$ dimensional subspace, as can easily be seen—any linear combination of these $N - 1$ vectors is in the subspace, as is $|0\rangle$ and the inverse of any vector in the subspace.

8. (i) In any of $\mathbb{V}^N$, $\mathbb{V}^M$ and $\mathbb{V}^N \otimes \mathbb{V}^M$ we can write $|0\rangle = 0|p\rangle$ where $|p\rangle$ is an arbitrary vector in the space. So $|v\rangle \otimes |0\rangle = |v\rangle \otimes (0|w\rangle) = 0(|v\rangle \otimes |w\rangle) = 0$ and similarly starting with $|0\rangle \otimes |w\rangle$. The inverse of $|v\rangle \otimes |w\rangle$ can be written $|-v\rangle \otimes |w\rangle$, $|v\rangle \otimes |-w\rangle$, $-(|v\rangle \otimes |w\rangle)$, or $-(|-v\rangle \otimes |-w\rangle)$. They are all the same vector.
   (ii) The individual spaces are 2-dimensional and the product space is 4-dimensional. xy, $(x + x^3)(y + 2y^3)$, $xy - xy^3 + x^3y - x^3y^3$ and $xy + 2x^3y^3$ are all vectors in the product space. All are seperable except the last (the third can be written $(x + x^3)(y - y^3)$).

9. $|a\rangle \to \begin{pmatrix} -2 \\ 2i \end{pmatrix}$, $|b\rangle \to \begin{pmatrix} 2+3i \\ 2i \end{pmatrix}$, $|c\rangle \to \begin{pmatrix} 3i \\ 4i \end{pmatrix}$

$$\langle a|b\rangle \to (-2, -2i)\begin{pmatrix} 2+3i \\ 2i \end{pmatrix} = -6i \qquad \langle b|a\rangle \to (2-3i, -2i)\begin{pmatrix} -2 \\ 2i \end{pmatrix} = 6i$$

$$\langle c|c\rangle \to (-3i, -4i)\begin{pmatrix} 3i \\ 4i \end{pmatrix} = 25 \Rightarrow |c| = 5.$$

We wouldn't normally calculate both $\langle a|b\rangle$ and $\langle b|a\rangle$! While we are at it, note $|a| = 2\sqrt{2}$ and $|b| = \sqrt{17}$ so we can check that the triangle and Schwartz inequalities are satisfied.

10. $|\psi\rangle = C\big(|1\rangle + 2i|2\rangle + (1+i)|3\rangle\big)$ so

$$\begin{aligned}\langle\psi|\psi\rangle &= |C|^2\big(\langle 1| - 2i\langle 2| + (1-i)\langle 3|\big)\big(|1\rangle + 2i|2\rangle + (1+i)|3\rangle\big) \\ &= |C|^2\big(\langle 1|1\rangle + 2i\langle 1|2\rangle + (1+i)\langle 1|3\rangle - 2i\langle 2|1\rangle + 4\langle 2|2\rangle + \ldots + 2\langle 3|3\rangle\big) \\ &= |C|^2(1+0+0+0+4+0+0+0+2) = 7|C|^2.\end{aligned}$$

We used the fact that for an orthonormal basis $\langle 1|1\rangle = 1$, $\langle 1|2\rangle = 0$ etc., usually we would not write all these terms out explicitly but jump straight to $\langle\psi|\psi\rangle = |C|^2(1+4+2)$. Anyway we need $|C|^{-1} = \sqrt{7}$, but $\sqrt{7}$, $-\sqrt{7}$, $i\sqrt{7}$, $\sqrt{7}e^{i\pi/5}\ldots$ etc are all possibilities.

11. The original set is $\{|a\rangle = (1,1,0),\ |b\rangle = (1,0,1),\ |c\rangle = (0,1,1)\}$. Let the new orthonormal basis be $\{|1\rangle, |2\rangle, |3\rangle\}$. $|1\rangle$ is just normalised $|a\rangle$: $|1\rangle = \sqrt{\tfrac{1}{2}}(1,1,0)$.

For $|2\rangle$ we start with $|b\rangle$ and subtract the bit perpendicular to $|1\rangle$. For that we need $\langle 1|b\rangle = \sqrt{\tfrac{1}{2}}$, so

$$|2\rangle = C\big(|b\rangle - \langle 1|b\rangle|1\rangle\big) = C\big((1,0,1) - \tfrac{1}{2}(1,1,0)\big) = \tfrac{1}{2}C(1,-1,2) = \sqrt{\tfrac{1}{6}}(1,-1,2)$$

(to start with we left the normalisation C undetermined then fixed it at the end). Finally we need $\langle 1|c\rangle = 1/\sqrt{2}$, $\langle 2|c\rangle = 1/\sqrt{6}$ and

$$\begin{aligned}|3\rangle &= D\big(|c\rangle - \langle 1|c\rangle|1\rangle - \langle 2|c\rangle|2\rangle\big) \\ &= D\big((0,1,1) - \tfrac{1}{2}(1,1,0) - \tfrac{1}{6}(1,-1,2)\big) = \tfrac{D}{6}(-4,4,4) = \sqrt{\tfrac{1}{3}}(1,-1,-1)\end{aligned}$$

where we chose D negative for tidiness. At the end of this rather lengthy procedure we should check all the inner products do vanish! If we'd started with the original vectors in a different order, we would have obtained a different result.

12. i) If $|\phi\rangle = |a\rangle - \big(\langle b|a\rangle/|b|^2\big)\,|b\rangle$, then

$$\begin{aligned}\langle\phi|\phi\rangle &= \left(\langle a| - \frac{\langle b|a\rangle^*}{|b|^2}\langle b|\right)\left(|a\rangle - |b\rangle\frac{\langle b|a\rangle}{|b|^2}\right) \\ &= \langle a|a\rangle - 2\frac{|\langle b|a\rangle|^2}{|b|^2} + \frac{|\langle b|a\rangle|^2}{|b|^4}\langle b|b\rangle = |a|^2 - \frac{|\langle b|a\rangle|^2}{|b|^2}\end{aligned}$$

But $\langle\phi|\phi\rangle \geq 0$, so the result $|\langle b|a\rangle| \leq |a|\,|b|$ follows. (In two and three dimensions, this translates to $|\cos\theta| \leq 1$, where θ is the angle between the vectors.)

ii)

$$\begin{aligned}|a+b|^2 &= \langle a+b|a+b\rangle = \langle a|a\rangle + \langle b|b\rangle + \langle a|b\rangle + \langle b|a\rangle = |a|^2 + |b|^2 + 2\mathrm{Re}\,\langle a|b\rangle \\ &\leq |a|^2 + |b|^2 + 2\,|\langle a|b\rangle| \\ &\leq |a|^2 + |b|^2 + 2\,|a|\,|b| = (|a|+|b|)^2 \\ \Rightarrow |a+b| &\leq |a| + |b|\end{aligned}$$

where we used the fact that the real part of a complex number is not greater than its magnitude, and then the Schwartz inequality, and finally the fact that both quantities are non-negative so we can take the square root and still preserve the inequality.

13. i) Clearly the sum of any two 2×2 matrices is just another matrix, and the usual rules of arithmetic mean that the commutative, associative and distributive rules of addition and scalare multiplication, and multiplication by unity as the identity, all hold. The zero vector is $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and the inverse vector of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$.

ii) (I) Skew-symmetry is proved as follows: $\langle M|N\rangle = \frac{1}{2}\mathrm{Tr}(\mathbf{M}^\dagger\mathbf{N}) = \frac{1}{2}\mathrm{Tr}((\mathbf{M}^\dagger\mathbf{N})^\dagger)^* = \frac{1}{2}\mathrm{Tr}(\mathbf{N}^\dagger\mathbf{M})^* = \langle N|M\rangle^*$ as required for the inner product in a complex vector space.
(II) Positive definiteness is demonstrated by $\frac{1}{2}\mathrm{Tr}(\mathbf{M}^\dagger\mathbf{M}) = \frac{1}{2}(|a|^2 + |b|^2 + |c|^2 + |d|^2) \geq 0$.
(III) Linearity on the ket side follows from linearity of the trace.

iii) Sample calculations to show $\langle i|j\rangle = \delta_{ij}$ are shown below (watch out: $\langle 3|$ and $|4\rangle$ are represented by the same matrix.)

$$\langle 2|4\rangle = \tfrac{1}{4}\mathrm{Tr}\left[\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}\right] = \tfrac{1}{4}\mathrm{Tr}\begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix} = 0$$
$$\langle 3|3\rangle = \tfrac{1}{4}\mathrm{Tr}\left[\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\right] = \tfrac{1}{4}\mathrm{Tr}\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} = 1$$
$$\langle 4|3\rangle = \tfrac{1}{4}\mathrm{Tr}\left[\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\right] = \tfrac{1}{4}\mathrm{Tr}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

There is no need to check $\langle j|i\rangle$ as well as $\langle i|j\rangle$ as they are the same from rule (I).
The 4 orthonormal vectors are linearly independent and hence form a basis in this 4D space.

$$m_1 = \langle 1|M\rangle = \sqrt{\tfrac{1}{8}}\mathrm{Tr}\left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = \sqrt{\tfrac{1}{8}}(a+b+c+d)$$
$$m_2 = \langle 2|M\rangle = \sqrt{\tfrac{1}{8}}\mathrm{Tr}\left[\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = \sqrt{\tfrac{1}{8}}(a-b-c+d)$$
$$m_3 = \langle 3|M\rangle = \sqrt{\tfrac{1}{8}}\mathrm{Tr}\left[\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = \sqrt{\tfrac{1}{8}}(a-b+c-d)$$
$$m_4 = \langle 4|M\rangle = \sqrt{\tfrac{1}{8}}\mathrm{Tr}\left[\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = \sqrt{\tfrac{1}{8}}(a+b-c-d)$$
$$\begin{aligned}\langle M|M\rangle &= \tfrac{1}{8}\left(|a+b+c+d|^2 + |a-b-c+d|^2 + |a-b+c-d|^2 + |a+b-c-d|^2\right) \\ &= \tfrac{1}{2}(|a|^2 + |b|^2 + |c|^2 + |d|^2)\end{aligned}$$

14. In all the following we use linearity of the inner product extensively.

i) $\langle i|a\rangle = \langle i|\left(\sum_j a_j|j\rangle\right) = \sum_j a_j\langle i|j\rangle = \sum_j a_j\delta_{ij} = a_i$
$\langle b|i\rangle = \left(\sum_j b_j^*\langle j|\right)|i\rangle = \sum_j b_j^*\langle j|i\rangle = \sum_j b_j^*\delta_{ji} = b_i^*$.

ii) $\langle b|a\rangle = \left(\sum_i b_i^*\langle i|\right)\left(\sum_j a_j|j\rangle\right) = \sum_{ij} b_i^* a_j\langle i|j\rangle = \sum_{ij} b_i^* a_j\delta_{ij} = \sum_i b_i^* a_i$.

iii) $\left(\sum_i |i\rangle\langle i|\right)|a\rangle = \sum_i |i\rangle\langle i|a\rangle = \sum_i |i\rangle a_i = |a\rangle$, so since $|a\rangle$ is arbitrary, $\sum_i |i\rangle\langle i| = \hat{I}$.

iv) $\langle b|\hat{A}|a\rangle = \left(\sum_i b_i^*\langle i|\right)\hat{A}\left(\sum_j a_j|j\rangle\right) = \sum_{ij} b_i^*\langle i|\hat{A}|j\rangle a_j = \sum_{ij} b_i^* A_{ij} a_j$

v) $\left(\sum_{ij} A_{ij}|i\rangle\langle j|\right)|a\rangle = \sum_{ij} A_{ij}|i\rangle\langle j|a\rangle = \sum_{ij} A_{ij}a_j|i\rangle$.
Also, $\hat{A}|a\rangle = \sum_i c_i|i\rangle$ where $c_i = \langle i|\hat{A}|a\rangle = \langle i|\hat{A}\big(\sum_j |j\rangle\langle j|\big)|a\rangle = \sum_j \langle i|\hat{A}|j\rangle\langle j|a\rangle = \sum_j A_{ij}a_j$.
So the two are equivalent and since $|a\rangle$ is arbitrary, $\sum_{ij} A_{ij}|i\rangle\langle j| = \hat{A}$.

vi) $\langle i|\hat{B}\hat{A}|k\rangle = \langle i|\hat{B}\big(\sum_j |j\rangle\langle j|\big)\hat{A}|k\rangle = \sum_j B_{ij}A_{jk}$.

vii) $\langle a|\hat{C}|b\rangle = \sum_{ij} a_i^* C_{ij} b_j$. Also, $\langle b|\hat{A}|a\rangle^* = \left(\sum_{ij} b_j^* A_{ji} a_i\right)^* = \sum_{ij} a_i^* A_{ji}^* b_j$. Comparing the two, since $|a\rangle$ and $|b\rangle$ are arbitrary, $C_{ij} = A_{ji}^*$. (Note we chose our dummy indices to aid direct comparison of the final expressions.)

15. $\langle 1|\hat{G}|1\rangle = \langle 2|\hat{G}|1\rangle = \langle 2|\hat{G}|2\rangle = 1$;
$\langle 1|\hat{G}|2\rangle = -1$; $\langle 3|\hat{G}|1\rangle = \langle 3|\hat{G}|2\rangle = \langle 1|\hat{G}|3\rangle = \langle 2|\hat{G}|3\rangle = \langle 3|\hat{G}|3\rangle = 0$. So

$$\hat{G} \to \begin{pmatrix} \langle 1|\hat{G}|1\rangle & \langle 1|\hat{G}|2\rangle & \langle 1|\hat{G}|3\rangle \\ \langle 2|\hat{G}|1\rangle & \langle 2|\hat{G}|2\rangle & \langle 2|\hat{G}|3\rangle \\ \langle 3|\hat{G}|1\rangle & \langle 3|\hat{G}|2\rangle & \langle 3|\hat{G}|3\rangle \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Note $\hat{G}$ is not Hermitian. $\langle 1|\psi\rangle = C, \langle 2|\psi\rangle = 2iC, \langle 3|\psi\rangle = (1+i)C$ with $C = 1/\sqrt{7}$:

$$|\psi\rangle \to \frac{1}{\sqrt{7}}\begin{pmatrix} 1 \\ 2i \\ 1+i \end{pmatrix} \qquad \hat{G}|\psi\rangle \to \frac{1}{\sqrt{7}}\begin{pmatrix} 1-2i \\ 1+2i \\ 0 \end{pmatrix}$$

16. (i) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is Hermitian and unitary.
(ii) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is neither.
(iii) $\begin{pmatrix} 1 & -i \\ i & 2 \end{pmatrix}$ is Hermitian (not unitary, even though the determinant is 1).
(iv) $\begin{pmatrix} 0 & i \\ 1 & 0 \end{pmatrix}$ is unitary but not Hermitian.

17. $\hat{A}^\dagger = \hat{A}$ and $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$

i) Let $\hat{A}|v\rangle = |w\rangle$. Then $\langle w| = \langle v|\hat{A}^\dagger = \langle v|\hat{A}$. So $\langle v|\hat{A}\hat{A}|v\rangle = \langle w|w\rangle \geq 0$. (For equality, $|v\rangle$ is either the zero vector or an eigenvector of $\hat{A}$ with zero eigenvalue.)

ii) $\left(\hat{U}\hat{A}\hat{U}^\dagger\right)^\dagger = (\hat{U}^\dagger)^\dagger\hat{A}^\dagger\hat{U}^\dagger = \hat{U}\hat{A}\hat{U}^\dagger$; $\qquad \left(\hat{U}^\dagger\hat{A}\hat{U}\right)^\dagger = \hat{U}^\dagger\hat{A}^\dagger(\hat{U}^\dagger)^\dagger = \hat{U}^\dagger\hat{A}\hat{U}$
so the transformed matrices are also Hermitian.

iii) $\langle v_i|v_j\rangle = \langle u_i|\hat{U}\hat{U}^\dagger|u_j\rangle = \langle u_i|u_j\rangle = \delta_{ij}$.

iv) $\det(\hat{U}\hat{A}\hat{U}^\dagger) = \det(\hat{U})\det(\hat{A})\det(\hat{U}^\dagger) = \det(\hat{A})\det(\hat{U}\hat{U}^\dagger)) = \det(\hat{A})$ (or use $\det(\hat{U})$ has the form $e^{i\phi}$ and $\det(\hat{U}^\dagger) = \det(\hat{U})^*$).
$\mathrm{Tr}(\mathbf{ABC}) = \mathrm{Tr}(\mathbf{BCA}) = \mathrm{Tr}(\mathbf{CAB})$ so $\mathrm{Tr}(\hat{U}\hat{A}\hat{U}^\dagger) = \mathrm{Tr}(\hat{A}\hat{U}^\dagger\hat{U}) = \mathrm{Tr}(\hat{A})$

18. i) By definition, and using $\hat{\Omega}^\dagger = \hat{\Omega}$,

$$\begin{aligned} \hat{\Omega}|\omega_i\rangle &= \omega_i|\omega_i\rangle && \text{and} && \langle\omega_j|\hat{\Omega} = \omega_j^*\langle\omega_j| \\ \Rightarrow \langle\omega_j|\hat{\Omega}|\omega_i\rangle &= \omega_i\langle\omega_j|\omega_i\rangle && \text{and} && \langle\omega_j|\hat{\Omega}|\omega_i\rangle = \omega_j^*\langle\omega_j|\omega_i\rangle \\ \Rightarrow 0 &= (\omega_i - \omega_j^*)\langle\omega_j|\omega_i\rangle \end{aligned}$$

so either $\langle\omega_j|\omega_i\rangle = 0$ or $(\omega_i - \omega_j^*) = 0$. If $i = j$ the first is not possible ($|0\rangle$ is not an eigenvector) so the second must be true: $(\omega_i - \omega_i^*) = 0$ so ω_i is real. Hence we have proved the first point. Now if $i \neq j$ the second cannot be true because (in this case) the eigenvalues are non-degenerate, so the first must be true: $\langle\omega_j|\omega_i\rangle = 0$ and the eigenvectors are orthogonal. (If there is degeneracy, this argument can only be used to show that the eigenvalues are real and the eigenvectors corresponding to distinct eigenvalues are orthogonal, but in fact a complete orthogonal set can be found.)

ii) Any vector $|a\rangle$ can be written in this basis as $\sum_i a_i|\omega_i\rangle$. Then $(\sum_i \omega_i|\omega_i\rangle\langle\omega_i|)\,|a\rangle = \sum_i \omega_i|\omega_i\rangle a_i$. Also, $\hat\Omega|a\rangle = \hat\Omega\,(\sum_i a_i|\omega_i\rangle) = \sum_i a_i\omega_i|\omega_i\rangle$. As these are the same, the operators are equivalent.

19. i) We start with $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$. When we multiply two matrices $\mathbf{C} = \mathbf{A}^\dagger\mathbf{B}$, the element in the ith row and jth column of $\mathbf{C}$ comes from the inner product of the ith column of $\mathbf{A}$ with the jth column of $\mathbf{B}$: $C_{ij} = (A^\dagger)_{ik}B_{kj} = (A_{ki})^*B_{kj}$. If $\mathbf{C}$ is the identity matrix, all these inner products vanish except the diagonal ones $i = j$, which are unity. In this case $\mathbf{A} = \mathbf{B} = \mathbf{U}$, so the columns of $\mathbf{U}$ are normalised and mutually orthogonal. If instead we consider $\mathbf{D} = \mathbf{A}\mathbf{B}^\dagger$, $D_{ij} = A_{ik}(B^\dagger)_{kj} = A_{ik}(B_{jk})^*$: i.e. $\mathbf{D}$ is formed from the inner product of the rows of $\mathbf{A}$ and $\mathbf{B}$, so the same argument goes through for the rows of $\mathbf{U}$.

ii)

$$\begin{aligned} \hat U|\omega_i\rangle = \omega_i|\omega_i\rangle \qquad &\text{and} \qquad \langle\omega_j|\hat U^\dagger = \omega_j^*\langle\omega_j| \\ \Rightarrow \langle\omega_j|\hat U^\dagger U|\omega_i\rangle &= \omega_j^*\omega_i\langle\omega_j|\omega_i\rangle \\ \Rightarrow \langle\omega_j|\omega_i\rangle &= \omega_j^*\omega_i\langle\omega_j|\omega_i\rangle \Rightarrow \langle\omega_j|\omega_i\rangle(1-\omega_j^*\omega_i) = 0 \end{aligned}$$

So either $\langle\omega_j|\omega_i\rangle = 0$ or $\omega_j^*\omega_i = 1$. If $i = j$ the first cannot be true so $\omega_i^*\omega_i = 1$, i.e. ω_i is a complex number of unit modulus. If $i \neq j$ the second cannot be true: if we multiply both sides by ω_j and use the previous result we get $\omega_i = \omega_j$ contrary to the initial assuption that $\hat U$ is non degenerate. So $\langle\omega_j|\omega_i\rangle = 0$. (If there is degeneracy, this argument can only be used to show that the eigenvalues have unit norm and the eigenvectors corresponding to distinct eigenvalues are orthogonal, but in fact a complete orthogonal set can be found.)

iii) A general complex $N \times N$ matrix has N^2 complex entries and so $2N^2$ real parameters. For a Hermitian matrix, the N diagonal elements H_{ii} are real and so we have N constraints of the form $\mathrm{Im}\,[H_{ii}] = 0$. The $N(N-1)/2$ elements below the diagonal (H_{ij} for $j < i$) are not independent, being just the complex conjugates of the one above, so we have $N(N-1)/2$ constraints of the form $\mathrm{Re}\,[H_{ij}] = \mathrm{Re}\,[H_{ji}]$ and $N(N-1)/2$ of the form $\mathrm{Im}\,[H_{ij}] = -\mathrm{Im}\,[H_{ji}]$. The number of free parameters is $2N^2 - N - N(N-1) = N^2$.

For a unitary matrix, the constraints come from orthonormality of the columns. There are N normalised columns giving N constraints, $\sum_i [u_{ij}^* u_{ij}] = 1$. (This has to be real, so there is no constraint from the vanishing of the imaginary part.) And there are $N(N-1)/2$ pairs of orthogonal columns, each this time giving 2 constraints, $N(N-1)$ in all. So there are N^2 constraints and again N^2 free parameters. The most general 2×2 unitary matrix can be written $\mathrm{e}^{i\alpha}\begin{pmatrix} \cos\theta\mathrm{e}^{i\beta} & -\sin\theta\mathrm{e}^{i\gamma} \\ \sin\theta\mathrm{e}^{-i\gamma} & \cos\theta\mathrm{e}^{-i\beta} \end{pmatrix}$, where α, β, γ and θ are all real angles.

20. i) $\sqrt{\tfrac{1}{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$, eigenvalue 1 and $\sqrt{\tfrac{1}{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$, eigenvalue -1.

ii) $\sqrt{\tfrac{1}{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, eigenvalue $a+b$ and $\sqrt{\tfrac{1}{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, eigenvalue $a-b$.

iii) Here the ONLY eigenvector is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, eigenvalue 1.

iv) $\begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}$, eigenvalue 1 and $\begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2) \end{pmatrix}$, eigenvalue -1.

v) $\sqrt{\frac{1}{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$, eigenvalue $e^{i\theta}$ and $\sqrt{\frac{1}{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$, eigenvalue $e^{-i\theta}$.

In (i), (iv) and (v) the eigenvectors are orthogonal as we expect for Hermitian and for unitary matrices. If a and b are real, (ii) is also Hermitian (the eigenvectors are orthogonal either way) and the eigenvalues for (i) and (ii) are real. (ii) is a frequently met case and the result for the eigenvectors is worth remembering! (v) is the rotation matrix in the plane, so it is not surprising that it has no real eigenvectors. It is unitary, though, so as expected the eigenvalues have unit modulus. In (iii) the characteristic equation is $(\omega-1)^2 = 0$, so there is a repeated root, but when this is plugged into the equation for the components of the eigenvectors one constraint remains, namely that the lower component vanish. (Compare with the case of the identity matrix, where there is no constraint and hence any vector is an eigenvector, though there are only two linearly-independent ones.) This matrix is neither Hermitian nor unitary, so it is not guaranteed to have two eigenvectors.
In all cases the sum of the eigenvalues is equal to the trace and the product to the determinant.

21. The characteristic equation is $\omega(\omega^2-2) = 0$ so the eigenvalues are $-\sqrt{2}, 0, \sqrt{2}$. The corresponding eigenvectors and the matrix of eigenvectors are

$$\frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \quad \sqrt{\frac{1}{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}; \qquad \mathbf{S} = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix}$$

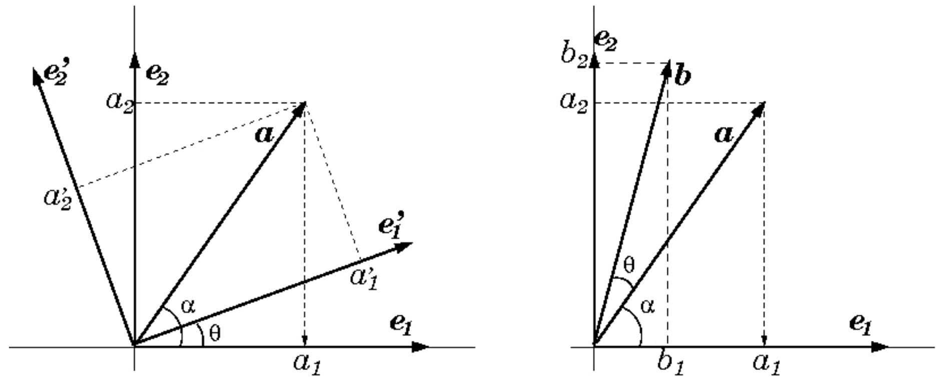
which is unitary.

$$\mathbf{S}^\dagger\mathbf{M}\mathbf{S} = \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix} = \begin{pmatrix} -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

22. There are no off-diagonal elements in the second row or column of $\mathbf{N}$, so $(0,1,0)^\top$ is an eigenvector with eigenvalue 6, and the other two (unnormalised) are $(1,0,1)^\top$ and $(1,0,-1)^\top$ from the result of qu. 20(ii) above, with eigenvalues 6 and 4. (This illustrates an important point—if you can guess the eigenvectors, it is trivial to verify your guess and pick up the corresponding eigenvalues at the same time. For a Hermitian matrix, if you can guess two you can get the third by orthogonality. If you can guess one, you can simplify finding the other two by requring them to be orthogonal to that one.)

$$\mathbf{MN} = \mathbf{NM} = \begin{pmatrix} 0 & 6 & 0 \\ 6 & 0 & 6 \\ 0 & 6 & 0 \end{pmatrix}$$

So $\mathbf{M}$ and $\mathbf{N}$ should have common eigenvectors, but looking at our answers only $(1,0,-1)^\top$ is in both lists. However any combination of $(0,1,0)^\top$ and $(1,0,1)^\top$ is also an eigenvector of $\mathbf{N}$, and this is only one possible orthogonal pair. $(1,\sqrt{2},1)^\top$ and $(1,-\sqrt{2},1)^\top$ are an equally good choice. Hence the eigenvectors of $\mathbf{M}$ are also those of $\mathbf{N}$, and the pair that are degenerate with respect to $\mathbf{N}$ are distinguished by their eigenvalue under $\mathbf{M}$.

23. 

Since $\mathbf{e}_1' = \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2$ and $\mathbf{e}_2' = -\sin\theta\,\mathbf{e}_1 + \cos\theta\,\mathbf{e}_2$,

$$\mathbf{S} = \begin{pmatrix} \langle e_1|e_1'\rangle & \langle e_1|e_2'\rangle \\ \langle e_2|e_1'\rangle & \langle e_2|e_2'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$(a_1', a_2') = \big(\cos(\alpha-\theta), \sin(\alpha-\theta)\big)$ follows from the left-hand diagram above, and agrees with the result

$$\mathbf{a}' = \mathbf{S}^\dagger\mathbf{a} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\alpha + \sin\theta\sin\alpha \\ -\sin\theta\cos\alpha + \cos\theta\sin\alpha \end{pmatrix}.$$

On the other hand $\mathbf{b}$ has coordinates $\big(\cos(\alpha+\theta), \sin(\alpha+\theta)\big)$ which can be obtained from $\mathbf{Sa}$.

24. Hermitian and unitary matrices can be diagonalised by a unitary transformation, but the trace and determinant are unchanged by the transformation. (See qu. 17(iv) above.) Once in diagonal form, the trace is the sum of the elements and the determinant is their product. Since these elements are the eigenvalues of the original matrix, the theorem is proved.

25. First note that $\mathbf{\Omega}^2 = \mathbf{I}$, so $\mathbf{\Omega}^3 = \mathbf{\Omega}$ etc. Then

$$\begin{aligned} \mathrm{e}^{ia\mathbf{\Omega}} &= \mathbf{I} + ia\mathbf{\Omega} - \tfrac{1}{2!}a^2\mathbf{\Omega}^2 - i\tfrac{1}{3!}a^3\mathbf{\Omega}^3 \ldots \\ &= \left(1 - \tfrac{1}{2!}a^2 + \tfrac{1}{4!}a^4 \ldots\right)\mathbf{I} + i\left(a - \tfrac{1}{3!}a^3 + \tfrac{1}{5!}a^5 \ldots\right)\mathbf{\Omega} = \cos a\,\mathbf{I} + i\sin a\,\mathbf{\Omega} \end{aligned}$$

26. An obvious (but not unique) choice for the basis in the product space is

$$\{|I\rangle = |1\rangle\otimes|1\rangle, \quad |II\rangle = |1\rangle\otimes|2\rangle, \quad |III\rangle = |2\rangle\otimes|1\rangle, \quad |IV\rangle = |2\rangle\otimes|2\rangle\}.$$

The labelling should make clear whether a ket is in the individual or produt spaces (and roman numerals are not related to identity operators!). A sample matrix element is
$\langle I|\hat\Omega\otimes\hat I|III\rangle = \langle 1|\hat\Omega|2\rangle\langle 1|\hat I|1\rangle = 1$.

In this basis the four basis kets are as usual represented as

$$|I\rangle \to \begin{pmatrix} \langle I|I\rangle \\ \langle II|I\rangle \\ \langle III|I\rangle \\ \langle IV|I\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad |II\rangle \to \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |III\rangle \to \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |IV\rangle \to \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

and

$$\hat\Omega\otimes\hat I \to \begin{pmatrix} \langle I|\hat\Omega\otimes\hat I|I\rangle & \langle I|\hat\Omega\otimes\hat I|II\rangle & \langle I|\hat\Omega\otimes\hat I|III\rangle & \langle I|\hat\Omega\otimes\hat I|IV\rangle \\ \langle II|\hat\Omega\otimes\hat I|I\rangle & \langle II|\hat\Omega\otimes\hat I|II\rangle & \langle II|\hat\Omega\otimes\hat I|III\rangle & \langle II|\hat\Omega\otimes\hat I|IV\rangle \\ \langle III|\hat\Omega\otimes\hat I|I\rangle & \langle III|\hat\Omega\otimes\hat I|II\rangle & \langle III|\hat\Omega\otimes\hat I|III\rangle & \langle III|\hat\Omega\otimes\hat I|IV\rangle \\ \langle IV|\hat\Omega\otimes\hat I|I\rangle & \langle IV|\hat\Omega\otimes\hat I|II\rangle & \langle IV|\hat\Omega\otimes\hat I|III\rangle & \langle IV|\hat\Omega\otimes\hat I|IV\rangle \end{pmatrix}$$

$$= \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right), \qquad \hat I\otimes\hat\Omega \to \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right), \qquad \hat\Omega\otimes\hat\Omega \to \left(\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right)$$

The third is the matrix product of the first two (in either order, since they commute). *Ignore the horizontal and vertical lines for now.*
Eigenvectors: $|1\rangle \pm |2\rangle$ are eigenvectors of $\hat{\Omega}$ with eigenvalues ± 1, so the 4 eigenvectors of $\hat{\Omega} \otimes \hat{\Omega}$ are $(|1\rangle \pm |2\rangle) \otimes (|1\rangle \pm |2\rangle)$ (where the two $\pm$ are independent), i.e.

$$|I\rangle + |II\rangle + |III\rangle + |IV\rangle \to \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad |I\rangle - |II\rangle - |III\rangle + |IV\rangle \to \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

both with eigenvalue $+1$, and

$$|I\rangle + |II\rangle - |III\rangle - |IV\rangle \to \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \qquad \text{and} \qquad |I\rangle - |II\rangle + |III\rangle - |IV\rangle \to \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

both with eigenvalue -1. Of course because there is degeneracy, these are not unique; for example the simpler vectors $|I\rangle \pm |IV\rangle$ and $|II\rangle \pm |III\rangle$ are also eigenvectors with eigenvalue ± 1.

It is easily checked that these are indeed eigenvectors of the corresponding matrix above. They are orthogonal as expected.

(We might have had the sense to choose the egenstates $(|1\rangle \pm |2\rangle) \otimes (|1\rangle \pm |2\rangle)$ as a basis in the first place, in which case the matrices above would simply be diagonal.)

For future reference, let's look at the structure of these matrices indicated by the horizontal and vertical lines, so that each 4×4 matrix is written as 2×2 blocks, each block being a 2×2 matrix. For instance in $\mathbf{I} \otimes \mathbf{\Omega}$ the block structure is diagonal and both diagonal blocks are identical, like $\mathbf{I}$, but the blocks themselves are just $\mathbf{\Omega}$. The pattern is reversed for $\mathbf{\Omega} \otimes \mathbf{I}$, and for $\mathbf{\Omega} \otimes \mathbf{\Omega}$ both the block structure and the blocks themselves are like $\mathbf{\Omega}$. This is a general pattern when using this basis:

$$\hat{A} \otimes \hat{B} \to \left(\begin{array}{c|c} A_{11}\mathbf{B} & A_{12}\mathbf{B} \\ \hline A_{21}\mathbf{B} & A_{22}\mathbf{B} \end{array}\right).$$