Linear Algebra I

Lecture 2

Physics 342
Quantum Mechanics I

Wednesday, January 27th, 2010

From separation of variables, we move to linear algebra. Roughly speaking, this is the study of vector spaces and operations on vector spaces. Our primary goal is to review the basic features of linear algebra and extend familiar ideas from, for example, three dimensional vectors to more abstract vector spaces.

2.1 A Vector

2.1.1 Definition

A vector is an object, a vector space, a collection of these objects, together with two basic operations and a set of scalars. The most familiar vector space, and one which we will refer to as an example, is $\mathbb{R}^N$ – this has operations $+$ and $\times$ (addition of vectors and multiplication of a vector by a number), and we draw scalars from $\mathbb{R}$ (more generally, from $\mathbb{C}$). We denote a vector $\mathbf{p}$ in $\mathbb{R}^N$ by $\mathbf{p} \in \mathbb{R}^N$. The "components" of the vector are real, and the real numbers act as the scalars underlying the defining operations: take $a \in \mathbb{R}$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{s} \in \mathbb{R}^N$, then

$$\mathbf{p} + \mathbf{q} \in \mathbb{R}^N \qquad \mathbf{p} + \mathbf{q} = \mathbf{q} + \mathbf{p} \qquad (\mathbf{p} + \mathbf{q}) + \mathbf{s} = \mathbf{p} + (\mathbf{q} + \mathbf{s}). \tag{2.1}$$

The first expression defines the operation "+" that takes two vectors in the space and generates a third vector in the space. The second expression defines $+$ to be commutative (order doesn't matter), and the third makes $+$ associative (grouping doesn't matter). The final property of the operator $+$ is the existence of a "zero", and hence inverses. We define $\mathbf{0}$ via: $\mathbf{p} + \mathbf{0} = \mathbf{p}$, and then we demand that for every $\mathbf{p}$, there exist $\mathbf{q}$ such that $\mathbf{p} + \mathbf{q} = \mathbf{0}$ – typically, then, we can write $\mathbf{q} = -\mathbf{p}$.

We can also take one of the scalars, $a \in \mathbb{R}$ and "multiply", those are still in $\mathbb{R}^N$: $a\,\mathbf{p} \in \mathbb{R}^N$. The multiplication is distributive (over vector addition) and associative (w.r.t multiplication by another $b \in \mathbb{R}$): $a\,(\mathbf{p}+\mathbf{q}) = a\,\mathbf{p} + a\,\mathbf{q}$, $a\,(b\,\mathbf{p}) = a\,b\,\mathbf{p}$. The zero vector has $a\,\mathbf{0} = \mathbf{0}$ (it must, after all, be a vector), and $0\,\mathbf{p} = \mathbf{0}$. Finally, the inverse of $\mathbf{p}$ can be written as $-1\,\mathbf{p}$ as suggested by the analogy with addition.

2.1.2 Inner Product, Basis, and Decomposition

There is an operation that takes two vectors and returns a scalar – in our defining example of $\mathbb{R}^N$, we say: $\cdot : (\mathbb{R}^N, \mathbb{R}^N) \longrightarrow \mathbb{R}$. The properties of an inner product (of which, you should be thinking in the back of your mind, the dot product is an example) are (for $\mathbf{p}$, $\mathbf{q}$, $\mathbf{r} \in \mathbb{R}^N$, a, $b \in \mathbb{R}$):

$$\begin{aligned} \mathbf{p}\cdot\mathbf{p} &\geq 0 \\ \mathbf{p}\cdot\mathbf{p} &= 0 \text{ iff } \mathbf{p} = \mathbf{0} \\ \mathbf{p}\cdot(a\,\mathbf{q} + b\,\mathbf{r}) &= a\,\mathbf{p}\cdot\mathbf{q} + b\,\mathbf{p}\cdot\mathbf{r}. \end{aligned} \tag{2.2}$$

Define a "linear combination" of vectors as a scalar-weighted sum of vectors: $\mathbf{p} = a\,\mathbf{q} + b\,\mathbf{r}$, for example. Then we say that two vectors $\mathbf{p}$ and $\mathbf{q}$ are *orthogonal* if: $\mathbf{p}\cdot\mathbf{q} = 0$. For $\mathbf{p} = a\,\mathbf{q} + b\,\mathbf{r}$ with $\mathbf{q}\cdot\mathbf{r} = 0$ (i.e. $\mathbf{q}$ and $\mathbf{r}$ are linearly independent), we have, assuming $\mathbf{q} \neq \mathbf{0}$:

$$\mathbf{p}\cdot\mathbf{q} = a\,\mathbf{q}\cdot\mathbf{q} + b\,\mathbf{q}\cdot\mathbf{r} = a\,\mathbf{q}\cdot\mathbf{q} \neq 0, \tag{2.3}$$

so $\mathbf{p}$ and $\mathbf{q}$ are not orthogonal. If we instead take $\mathbf{p} = b\,\mathbf{r}$, then $\mathbf{p}\cdot\mathbf{q} = 0$, so $\mathbf{p}$ and $\mathbf{q}$ are orthogonal (and hence linearly independent). The story we tell ourselves in the case of two orthogonal vectors is: "there is no $\mathbf{q}$ in $\mathbf{p}$."

Basis

Suppose we have a linearly independent, orthogonal, set of vectors $\{\mathbf{e}_i\}_{i=1}^N$ in $\mathbb{R}^N$. If $\forall\,\mathbf{p} \in \mathbb{R}^N$, we can write $\mathbf{p}$ as a linear combination of the set $\{\mathbf{e}_i\}_{i=1}^N$, then we say that this set forms a *basis* for $\mathbb{R}^N$. In other words, we have, for some set of coefficients $\{a_i\}_{i=1}^N$ with $a_i \in \mathbb{R}$:

$$\mathbf{p} = \sum_{i=1}^{N} a_i\,\mathbf{e}_i \qquad \forall \mathbf{p} \in \mathbb{R}^N. \tag{2.4}$$

If we have such a basis set, then the vector $\mathbf{p}$ can be represented by its "components", that is, by specifying the basis and the coefficients $\{a_i\}_{i=1}^N$. This gives us the standard representation in $\mathbb{R}^N$, we set

$$\mathbf{e}_1 \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad \mathbf{e}_2 \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad \ldots, \tag{2.5}$$

and then the vector $\mathbf{p}$ can be described by the column vector:

$$\mathbf{p} \doteq \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} + \ldots \tag{2.6}$$

Notice that the dot product is now clearly identified with a "projection" – if we want to know the "amount of $\mathbf{e}_4$ in $\mathbf{p}$", given (2.4), then $\mathbf{p} \cdot \mathbf{e}_4 = a_4$ is the answer. The decomposition of two vectors $\mathbf{p}$ and $\mathbf{q}$ gives us the more familiar (computationally useful) definition of the dot product – since $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$[1] (we can scale the basis vectors so that $\mathbf{e}_i \cdot \mathbf{e}_i = 1$), if we have

$$\begin{aligned} \mathbf{p} &= \sum_{i=1}^N a_i \, \mathbf{e}_i \\ \mathbf{q} &= \sum_{i=1}^N b_i \, \mathbf{e}_i \end{aligned} \tag{2.8}$$

then

$$\mathbf{p} \cdot \mathbf{q} = \sum_{i=1}^N a_i \, b_i. \tag{2.9}$$

Finally, given a vector $\mathbf{p}$ and a basis set $\{\mathbf{e}_i\}_{i=1}^N$, we can pick out the $\mathbf{e}_j$ component of $\mathbf{p}$ via $\mathbf{e}_j \cdot \mathbf{p}$, so we can write the vector as the sum:

$$\mathbf{p} = \sum_{i=1}^N (\mathbf{e}_i \cdot \mathbf{p}) \, \mathbf{e}_i, \tag{2.10}$$

[1] Remember the definition of the Kronecker delta:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} . \tag{2.7}$$

this is obvious if we already know the decomposition in terms of the basis $\{\mathbf{e}_i\}_{i=1}^N$, but a useful form if we do not.

2.1.3 Examples

Cartesian vectors ($\mathbb{R}^3$)

We can take the usual Cartesian vectors from, for instance, your studies of electricity and magnetism – a generic vector is given by its decomposition into $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ (these are the basis vectors, $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$) – take two such vectors

$$\begin{aligned} \mathbf{p} &= \hat{\mathbf{x}} + 2\,\hat{\mathbf{y}} - \hat{\mathbf{z}} \\ \mathbf{q} &= -\hat{\mathbf{y}} + 5\,\hat{\mathbf{z}}, \end{aligned} \tag{2.11}$$

in this case, $\mathbf{p}\cdot\mathbf{q} = -7$. We can find the "component of $\mathbf{p}$ in the $\hat{\mathbf{x}}$ direction" by taking $\mathbf{p}\cdot\hat{\mathbf{x}}$ – this is $= 1$. As a check of (2.10), consider the dot products of $\mathbf{q}$:

$$\hat{\mathbf{x}}\cdot\mathbf{q} = 0 \quad \hat{\mathbf{y}}\cdot\mathbf{q} = -1 \quad \hat{\mathbf{z}}\cdot\mathbf{q} = 5, \tag{2.12}$$

then it is trivially the case that:

$$\mathbf{q} = (\hat{\mathbf{x}}\cdot\mathbf{q})\,\hat{\mathbf{x}} + (\hat{\mathbf{y}}\cdot\mathbf{q})\,\hat{\mathbf{y}} + (\hat{\mathbf{z}}\cdot\mathbf{q})\,\hat{\mathbf{z}}. \tag{2.13}$$

A Vector Space of Functions

Consider the space of all functions decomposable into a cosine series (sufficiently smooth, with some finite domain $x \in [0,1]$). That is, functions $f(x)$ for $x = 0 \to 1$ that have:

$$f(x) = \sum_{n=1}^{\infty} \alpha_n \cos(n\,\pi\,x), \tag{2.14}$$

which we will denote $\mathbf{f}$. All of the requirements for a vector space are met: We can add two such functions to get a third, we can multiply by real or complex numbers, etc. Suppose we think of the functions $\cos(n\,\pi\,x)$, themselves in the space, as a basis (infinite dimensional). That is, define:

$$\mathbf{e}_j = \sqrt{2}\,\cos(j\,\pi\,x). \tag{2.15}$$

For our inner product, define:

$$\mathbf{f}\cdot\mathbf{g} = \int_0^1 f(x)\,g(x)\,dx. \tag{2.16}$$

Then the basis vectors satisfy:

$$\mathbf{e}_j \cdot \mathbf{e}_k = 2 \int_0^1 \cos(j\,\pi\,x)\,\cos(k\,\pi\,x)\,dx = \delta_{jk}. \tag{2.17}$$

Given an arbitrary function in the space, we know (by definition here) that it can be written as:

$$\mathbf{f} = \sum_{i=1}^{\infty} \alpha_i\,\mathbf{e}_i, \tag{2.18}$$

and, as always, $\mathbf{e}_j \cdot \mathbf{f}$ tells us "how much $\mathbf{e}_j$ there is in $\mathbf{f}$". Now we can see the utility of (2.10). For example, take

$$\mathbf{f} \doteq f(x) = \cos(\pi\,x)^3 - 3\,\cos(\pi\,x)\,\sin^2(\pi\,x). \tag{2.19}$$

As a vector in the space, it has a decomposition in terms of the basis, and we can pick out components by taking inner products. The first few are:

$$\begin{aligned}
\mathbf{e}_1 \cdot \mathbf{f} &= \sqrt{2}\int_0^1 \left(\cos(\pi\,x)^3 - 3\,\cos(\pi\,x)\,\sin^2(\pi\,x)\right)\,\cos(\pi\,x)\,dx = 0 \\
\mathbf{e}_2 \cdot \mathbf{f} &= \sqrt{2}\int_0^1 \left(\cos(\pi\,x)^3 - 3\,\cos(\pi\,x)\,\sin^2(\pi\,x)\right)\,\cos(2\,\pi\,x)\,dx = 0 \\
\mathbf{e}_3 \cdot \mathbf{f} &= \sqrt{2}\int_0^1 \left(\cos(\pi\,x)^3 - 3\,\cos(\pi\,x)\,\sin^2(\pi\,x)\right)\,\cos(3\,\pi\,x)\,dx = \frac{1}{\sqrt{2}} \\
\mathbf{e}_4 \cdot \mathbf{f} &= \sqrt{2}\int_0^1 \left(\cos(\pi\,x)^3 - 3\,\cos(\pi\,x)\,\sin^2(\pi\,x)\right)\,\cos(4\,\pi\,x)\,dx = 0.
\end{aligned} \tag{2.20}$$

We could keep going, but as a basic trigonometric identity, it is indeed the case that:

$$f(x) = \cos(\pi\,x)^3 - 3\,\cos(\pi\,x)\,\sin^2(\pi\,x) = \cos(3\,\pi\,x) = \frac{1}{\sqrt{2}}\,\mathbf{e}_3. \tag{2.21}$$

In this form, it is obvious that $\mathbf{e}_3 \cdot \mathbf{f} = 1$ and all others are zero.

Using the inner product on a function space to find the decomposition coefficients in terms of some basis is what Griffiths refers to as "Fourier's Trick". Going back to our infinite slot potential from last time, our potential, in terms of functions that satisfied all other boundary conditions, was

$$V(x,y) = \sum_{n=1}^{\infty} Q_n\,\sin\!\left(\frac{n\,\pi\,x}{d}\right)\,e^{-\frac{n\,\pi\,y}{d}}. \tag{2.22}$$

The final boundary condition we needed to set was for $y = 0$, where the above is just an infinite sum of sines. Just as there is a space of functions decomposable in terms of cosines, there is a similar vector space for sines. Now we need functions on $x \in [0, d]$, but we can still define a basis:

$$\mathbf{e}_j = \sqrt{\frac{2}{d}}\,\sin\left(\frac{j\,\pi\,x}{d}\right), \tag{2.23}$$

and an inner product $\mathbf{f} \cdot \mathbf{g} = \int_0^d f(x)\,g(x)\,dx$ such that:

$$\mathbf{e}_j \cdot \mathbf{e}_k = \frac{2}{d}\int_0^d \sin\left(\frac{j\,\pi\,x}{d}\right)\,\sin\left(\frac{k\,\pi\,x}{d}\right)\,dx = \delta_{jk}. \tag{2.24}$$

We had the boundary condition $V(x, 0) = V_0\,\sin\left(\frac{2\,\pi\,x}{d}\right)$, and this is a trivial decomposition, a pure basis function. Suppose instead we held the connecting plate at a constant potential $\mathbf{f} = V_0$ – then we need the set:

$$\begin{aligned}\mathbf{e}_j \cdot \mathbf{f} &= \sqrt{\frac{2}{d}}\int_0^d V_0\,\sin\left(\frac{j\,\pi\,x}{d}\right)\,dx = \sqrt{2\,d}\,\frac{V_0(1 - \cos(j\,\pi))}{j\,\pi} \\ &= \left\{\begin{array}{ll} 0 & j \text{ even} \\ 2\sqrt{2\,d}\,\frac{V_0}{j\,\pi} & j \text{ odd} \end{array}\right. ,\end{aligned} \tag{2.25}$$

and we have decomposed the function $\mathbf{f}$ in terms of the sine basis – then, noting (2.10), we can write:

$$\mathbf{f} \doteq V_0 = \sum_{j=1,odd}^{\infty} \frac{4\,V_0}{j\,\pi}\,\sin\left(\frac{j\,\pi\,x}{d}\right). \tag{2.26}$$

Comparing this with the value of the potential at $y = 0$, we see that the Q_n should be:

$$Q_n = \left\{\begin{array}{ll} 0 & n \text{ even} \\ \frac{4\,V_0}{n\,\pi} & n \text{ odd} \end{array}\right. \tag{2.27}$$

so that the potential solving the problem is:

$$V(x, y) = \sum_{n=1,odd}^{\infty} \frac{4\,V_0}{n\,\pi}\,\sin\left(\frac{n\,\pi\,x}{d}\right)\,e^{-\frac{n\,\pi\,y}{d}}, \tag{2.28}$$

and a contour plot of this for the first 50 non-zero components is shown in Figure 2.1.

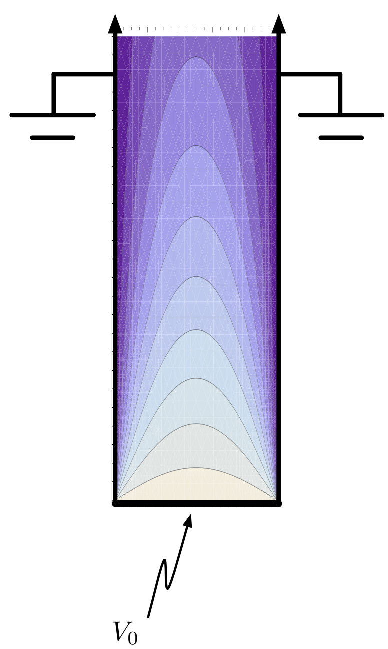

Figure 2.1: Parallel plates grounded on either side with a constant potential V_0 along the connecting plate. The first fifty odd terms in (2.28) are shown.

2.2 Notation

So far, all of this is review. Now we move on to a notational shift. Dirac's "Bra-ket" notation provides a nice, homogenous treatment of vector spaces, finite and not. But notation is *all* it is. We begin by translating the above definitions – there are two basic objects in Bra-ket notation: a Bra, denoted $\langle\alpha|$ and a ket, written $|\alpha\rangle$. The kets are elements of a vector space with scalars drawn from the complex numbers – they can be thought of as our $\mathbf{p}$ in $\mathbb{R}^N$ from (2.6) (although the entries are now complex numbers). The bras are elements of a dual space, and are something like the row form of a column vector – so they are still vectors, but a slightly different sort (you cannot, for example, add a column vector and a row vector, and similarly, you cannot add a bra to a ket). This association is motivated by the notation for inner product – we can form the bra-ket:

$$(\langle\alpha|)(|\beta\rangle) = \langle\alpha|\beta\rangle\,, \tag{2.29}$$

and this is a complex number, with the further requirement that $\langle\alpha|\beta\rangle = (\langle\beta|\alpha\rangle)^*$ (necessary to get $\langle\alpha|\alpha\rangle \geq 0$ and real).

We have basis bras and kets – suppose we have a (maximal) set of linearly independent kets: $\{|e_i\rangle\}_{i=1}^N$ that form a basis, so that any ket can be written as:

$$|\alpha\rangle = \sum_{i=1}^N a_i\;|e_i\rangle\,, \tag{2.30}$$

for a_i complex. Similarly, we have basis bras, so that

$$\langle\beta| = \sum_{i=1}^N b_i\;\langle e_i| \tag{2.31}$$

(b_i complex). As bases, we have the inner product $\langle e_i|e_j\rangle = \delta_{ij}$ (I have taken an ortho*normal* basis here – think of $\mathbf{e}_i \cdot \mathbf{e}_j$), so, using the defining properties of inner products, we can compute the bra-ket $\langle\alpha|\beta\rangle$ in terms of the coefficients of the decomposition of $|\alpha\rangle$ and $\langle\beta|$. If we take the simplest possible case, letting: $\langle\beta| = u\,\langle e_i|$ and $|\alpha\rangle = v\,|e_i\rangle$ with u, v complex, then in order to satisfy the requirement $\langle\beta|\alpha\rangle = (\langle\alpha|\beta\rangle)^*$, we must have: $\langle\alpha|\beta\rangle = u^*\,v$.

We can extend this inner product computation to the more general decomposition of bras and kets:

$$\langle \alpha | \beta \rangle = \sum_{i=1}^{N} b_i \, \langle e_i | \sum_{j=1}^{N} a_j \, | e_j \rangle = \sum_{i=1}^{N} b_i^* \, a_i, \tag{2.32}$$

and of course,

$$\langle \alpha | \alpha \rangle = \sum_{i=1}^{N} a_i^* \, a_i \, \in \mathbb{R}. \tag{2.33}$$

It is this inner product view that provides the association with row and column vector representations – in a basis, we can represent a vector in terms of its decomposition coefficients:

$$| \alpha \rangle \doteq \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \tag{2.34}$$

then if we take $\langle \alpha | \doteq \begin{pmatrix} a_1^* & a_2^* & a_3^* & \ldots \end{pmatrix}$, the bra-ket can be naturally thought of as a vector-vector multiplication of the form:

$$\langle \alpha | \alpha \rangle = \begin{pmatrix} a_1^* & a_2^* & a_3^* & \ldots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = \sum_{j=1}^{N} a_j^* \, a_j. \tag{2.35}$$

Finally, we have a handy way to represent (2.10) in this notation – the coefficient of $|\alpha\rangle$ w.r.t. the basis vector $|e_i\rangle$ is just $\langle e_i | \alpha \rangle$, so we can write any $|\alpha\rangle$ as:

$$| \alpha \rangle = \sum_{i=1}^{N} \langle e_i | \alpha \rangle \, | e_i \rangle = \sum_{i=1}^{N} | e_i \rangle \, \langle e_i | \alpha \rangle \tag{2.36}$$

where the second equality comes from the fact that $\langle e_i | \alpha \rangle$ is just a number. We sometimes define the operator:

$$\mathbf{1} = \sum_{i=1}^{N} | e_i \rangle \, \langle e_i | \, . \tag{2.37}$$

Notice that this object is neither a bra nor a ket, and we will pick up this discussion of operators, objects that act on vectors in vector spaces, next time.

Homework

Reading: Griffiths, pp. 435–440.

Problem 2.1

From "Euler's formula": $e^{i\,\alpha} = \cos(\alpha) + i\,\sin(\alpha)$, we have

$$\sin(\alpha) = \frac{1}{2\,i}\left(e^{i\,\alpha} - e^{-i\,\alpha}\right). \tag{2.38}$$

a. Using this expression, evaluate:

$$\int_0^d \sin\left(\frac{j\,\pi\,x}{d}\right)\,\sin\left(\frac{k\,\pi\,x}{d}\right)\,dx \tag{2.39}$$

for $j \neq k$.

b. Do the same thing for $j = k$, this establishes the orthonormality of the sine basis: You are showing that:

$$\frac{2}{d}\int_0^\pi \sin\left(\frac{j\,\pi\,x}{d}\right)\,\sin\left(\frac{k\,\pi\,x}{d}\right)\,dx = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{2.40}$$

Problem 2.2

In two dimensions, we have the usual basis vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, and any vector can be written as a linear combination of these two: $\mathbf{v} = v_x\,\hat{\mathbf{x}} + v_y\,\hat{\mathbf{y}}$. Suppose we introduce two vectors:

$$\mathbf{v} = 3\,\hat{\mathbf{x}} + \hat{\mathbf{y}} \qquad \mathbf{w} = -2\,\hat{\mathbf{x}} + 6\,\hat{\mathbf{y}}. \tag{2.41}$$

a. Show that $\mathbf{v}\cdot\mathbf{w} = 0$, and normalize each vector (i.e. find $\hat{\mathbf{v}}$ that points in the same direction as $\mathbf{v}$ and has $\hat{\mathbf{v}}\cdot\hat{\mathbf{v}} = 1$, and similarly for $\hat{\mathbf{w}}$) to get a new pair of orthogonal unit vectors.

b. This new pair can itself be used as the basis – find the decomposition of $\mathbf{a} = 10\,\hat{\mathbf{x}} - 3\,\hat{\mathbf{y}}$ in terms of $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$, i.e. fill in $\boxed{?}$ below:

$$\mathbf{a} = \boxed{?}\,\hat{\mathbf{v}} + \boxed{?}\,\hat{\mathbf{w}}. \tag{2.42}$$

Problem 2.3

Griffiths A.1. Note: He is using $\hat{i}$, $\hat{j}$ and $\hat{k}$ instead of $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$.