Vector Spaces

The idea of vectors dates back to the middle 1800's, but our current understanding of the concept waited until Peano's work in 1888. Even then it took many years to understand the importance and generality of the ideas involved. This one underlying idea can be used to describe the forces and accelerations in Newtonian mechanics and the potential functions of electromagnetism and the states of systems in quantum mechanics and the least-square fitting of experimental data and much more.

6.1 The Underlying Idea

What *is* a vector?

If your answer is along the lines "something with magnitude and direction" then you have something to unlearn. Maybe you heard this definition in a class that I taught. If so, I lied; sorry about that. At the very least I didn't tell the whole truth. Does an automobile have magnitude and direction? Does that make it a vector?

The idea of a vector is far more general than the picture of a line with an arrowhead attached to its end. That special case is an important one, but it doesn't tell the whole story, and the whole story is one that unites many areas of mathematics. The short answer to the question of the first paragraph is

A vector is an element of a vector space.

Roughly speaking, a vector space is some set of things for which the operation of addition is defined and the operation of multiplication by a scalar is defined. You don't necessarily have to be able to multiply two vectors by each other or even to be able to define the length of a vector, though those *are* very useful operations and will show up in most of the interesting cases. You can add two cubic polynomials together:

$$\left(2 - 3x + 4x^2 - 7x^3\right) + \left(-8 - 2x + 11x^2 + 9x^3\right)$$

makes sense, resulting in a cubic polynomial. You can multiply such a polynomial by* 17 and it's still a cubic polynomial. The set of all cubic polynomials in x forms a vector space and the vectors are the individual cubic polynomials.

The common example of directed line segments (arrows) in two or three dimensions fits this idea, because you can add such arrows by the parallelogram law and you can multiply them by numbers, changing their length (and reversing direction for negative numbers).

Another, equally important example consists of all ordinary real-valued functions of a real variable: two such functions can be added to form a third one, and you can multiply a function by a number to get another function. The example of cubic polynomials above is then a special case of this one.

A complete definition of a vector space requires pinning down these ideas and making them less vague. In the end, the way to do that is to express the definition as a set of axioms. From these axioms the general properties of vectors will follow.

A *vector space* is a set whose elements are called "vectors" and such that there are two operations defined on them: you can add vectors to each other and you can multiply them by scalars (numbers). These operations must obey certain simple rules, the axioms for a vector space.

---

* The physicist's canonical random number

6.2 Axioms

The precise definition of a vector space is given by listing a set of axioms. For this purpose, I'll denote vectors by arrows over a letter, and I'll denote scalars by Greek letters. These scalars will, for our purpose, be either real or complex numbers — it makes no difference which for now.*

1 There is a function, addition of vectors, denoted $+$, so that $\vec{v}_1 + \vec{v}_2$ is another vector.

2 There is a function, multiplication by scalars, denoted by juxtaposition, so that $\alpha\vec{v}$ is a vector.

3 $(\vec{v}_1 + \vec{v}_2) + \vec{v}_3 = \vec{v}_1 + (\vec{v}_2 + \vec{v}_3)$ (the associative law).

4 There is a zero vector, so that for each $\vec{v}$, $\vec{v} + \vec{O} = \vec{v}$.

5 There is an additive inverse for each vector, so that for each $\vec{v}$, there is another vector $\vec{v}\,'$ so that $\vec{v} + \vec{v}\,' = \vec{O}$.

6 The commutative law of addition holds: $\vec{v}_1 + \vec{v}_2 = \vec{v}_2 + \vec{v}_1$.

7 $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$.

8 $(\alpha\beta)\vec{v} = \alpha(\beta\vec{v})$.

9 $\alpha(\vec{v}_1 + \vec{v}_2) = \alpha\vec{v}_1 + \alpha\vec{v}_2$.

10 $1\vec{v} = \vec{v}$.

   In axioms 1 and 2 I called these operations "functions." Is that the right use of the word? Yes. Without going into the precise definition of the word (see section 12.1), you know it means that you have one or more independent variables and you have a single output. Addition of vectors and multiplication by scalars certainly fit that idea.

6.3 Examples of Vector Spaces

Examples of sets satisfying these axioms abound:

1 The usual picture of directed line segments in a plane, using the parallelogram law of addition.

2 The set of real-valued functions of a real variable, defined on the domain $[a \le x \le b]$. Addition is defined pointwise. If f_1 and f_2 are functions, then the value of the function $f_1 + f_2$ at the point x is the number $f_1(x) + f_2(x)$. That is, $f_1 + f_2 = f_3$ means $f_3(x) = f_1(x) + f_2(x)$. Similarly, multiplication by a scalar is defined as $(\alpha f)(x) = \alpha(f(x))$. Notice a small confusion of notation in this expression. The first multiplication, (αf), multiplies the scalar α by the vector f; the second multiplies the scalar α by the number $f(x)$.

3 Like example 2, but restricted to continuous functions. The one observation beyond the previous example is that the sum of two continuous functions is continuous.

4 Like example 2, but restricted to bounded functions. The one observation beyond the previous example is that the sum of two bounded functions is bounded.

5 The set of n-tuples of real numbers: $(a_1, a_2, \ldots, a_n)$ where addition and scalar multiplication are defined by

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n) \qquad \alpha(a_1, \ldots, a_n) = (\alpha a_1, \ldots, \alpha a_n)$$

6 The set of square-integrable real-valued functions of a real variable on the domain $[a \le x \le b]$. That is, restrict example two to those functions with $\int_a^b dx\, |f(x)|^2 < \infty$. Axiom 1 is the only one requiring more than a second to check.

7 The set of solutions to the equation $\partial^2\phi/\partial x^2 + \partial^2\phi/\partial y^2 = 0$ in any fixed domain. (Laplace's equation)

---

* For a nice introduction online see distance-ed.math.tamu.edu/Math640, chapter three.

8 Like example 5, but with $n = \infty$.

9 Like example 8, but each vector has only a finite number of non-zero entries.

10 Like example 8, but restricting the set so that $\sum_1^\infty |a_k|^2 < \infty$. Again, only axiom one takes work.

11 Like example 10, but the sum is $\sum_1^\infty |a_k| < \infty$.

12 Like example 10, but $\sum_1^\infty |a_k|^p < \infty$. $(p \geq 1)$

13 Like example 6, but $\int_a^b dx \, |f(x)|^p < \infty$.

14 Any of examples 2–13, but make the scalars complex, and the functions complex valued.

15 The set of all $n \times n$ matrices, with addition being defined element by element.

16 The set of all polynomials with the obvious laws of addition and multiplication by scalars.

17 Complex valued functions on the domain $[a \leq x \leq b]$ with $\sum_x |f(x)|^2 < \infty$. (Whatever this means. See problem 6.18)

18 $\{\vec{O}\}$, the space consisting of the zero vector alone.

19 The set of all solutions to the equations describing small motions of the surface of a drumhead.

20 The set of solutions of Maxwell's equations without charges or currents and with finite energy. That is, $\int [E^2 + B^2] d^3x < \infty$.

21 The set of all functions of a complex variable that are differentiable everywhere and satisfy

$$\int dx \, dy \, e^{-x^2 - y^2} |f(z)|^2 < \infty,$$

where $z = x + iy$.

To verify that any of these is a vector space you have to run through the ten axioms, checking each one. (Actually, in a couple of pages there's a theorem that will greatly simplify this.) To see what is involved, take the first, most familiar example, arrows that all start at one point, the origin. I'll go through the details of each of the ten axioms to show that the process of checking is very simple. There are some cases for which this checking isn't so simple, but the difficulty is usually confined to verifying axiom one.

The picture shows the definitions of addition of vectors and multiplication by scalars, the first two axioms. The commutative law, axiom 6, is clear, as the diagonal of the parallelogram doesn't depend on which side you're looking at.
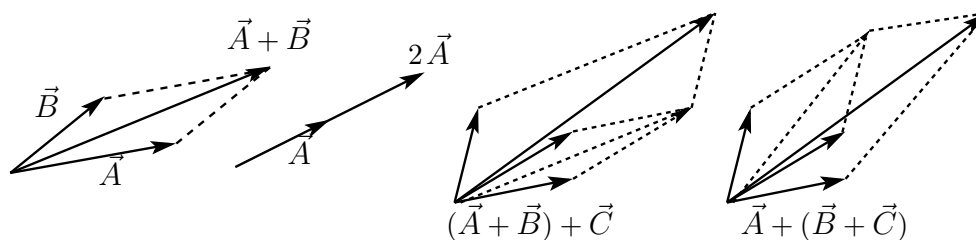
The associative law, axiom 3, is also illustrated in the picture. The zero vector, axiom 4, appears in this picture as just a point, the origin.

The definition of multiplication by a scalar is that the length of the arrow is changed (or even reversed) by the factor given by the scalar. Axioms 7 and 8 are then simply the statement that the graphical interpretation of multiplication of numbers involves adding and multiplying their lengths.

$\vec{A}$

$\vec{A'}$

$\alpha(\vec{A} + \vec{B})$
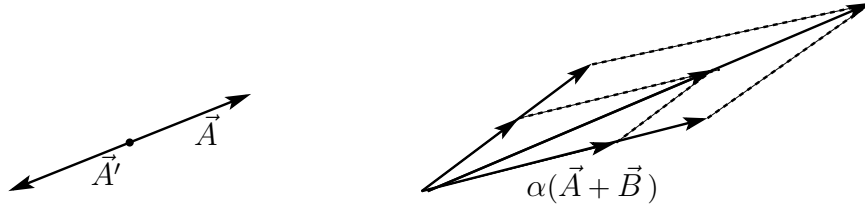
Axioms 5 and 9 appear in this picture.

Finally, axiom 10 is true because you leave the vector alone when you multiply it by one.

This process looks almost *too* easy. Some of the axioms even look as though they are trivial and unnecessary. The last one for example: why do you have to *assume* that multiplication by one leaves the vector alone? For an answer, I will show an example of something that satisfies all of axioms one through nine but *not* the tenth. These processes, addition of vectors and multiplication by scalars, are functions. I could write "$f(\vec{v}_1, \vec{v}_2)$" instead of "$\vec{v}_1 + \vec{v}_2$" and write "$g(\alpha, \vec{v})$" instead of "$\alpha\vec{v}$". The standard notation is just that — a common way to write a vector-valued function of two variables. I can define any function that I want and then see if it satisfies the required properties.

On the set of arrows just above, redefine multiplication by a scalar (the function g of the preceding paragraph) to be the zero vector for all scalars and vectors. That is, $\alpha\vec{v} = \vec{O}$ for all α and $\vec{v}$. Look back and you see that this definition satisfies all the assumptions 1–9 but not 10. For example, 9: $\alpha(\vec{v}_1 + \vec{v}_2) = \alpha\vec{v}_1 + \alpha\vec{v}_2$ because both sides of the equation are the zero vector. This observation proves that the tenth axiom is independent of the others. If you could derive the tenth axiom from the first nine, then this example couldn't exist. This construction is of course not a vector space.

Function Spaces

Is example 2 a vector space? How can a function be a vector? This comes down to your understanding of the word "function." Is $f(x)$ a function or is $f(x)$ a number? Answer: it's a number. This is a confusion caused by the conventional notation for functions. We routinely call $f(x)$ a function, but it is really the result of feeding the particular value, x, to the function f in order to get the number $f(x)$. This confusion in notation is so ingrained that it's hard to change, though in more sophisticated mathematics books it *is* changed.

In a better notation, the symbol f is the function, expressing the relation between all the possible inputs and their corresponding outputs. Then $f(1)$, or $f(\pi)$, or $f(x)$ are the results of feeding f the particular inputs, and the results are (at least for example 2) real numbers. Think of the function f as the whole graph relating input to output; the pair $(x, f(x))$ is then just one point on the graph. Adding two functions is adding their graphs. For a precise, set theoretic definition of the word function, see section 12.1. Reread the statement of example 2 in light of these comments.

f_3
f_2
f_1

$f_1 + f_2 = f_3$
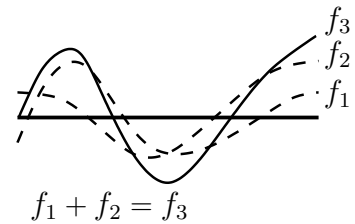
Special Function Space

Go through another of the examples of vector spaces written above. Number 6, the square-integrable real-valued functions on the interval $a \le x \le b$. The single difficulty here is the first axiom: is the sum of two square-integrable functions itself square-integrable? The other nine axioms are yours to check.

Suppose that

$$\int_a^b f(x)^2 \, dx < \infty \qquad \text{and} \qquad \int_a^b g(x)^2 \, dx < \infty.$$

simply note the combination

$$\big(f(x) + g(x)\big)^2 + \big(f(x) - g(x)\big)^2 = 2f(x)^2 + 2g(x)^2$$

The integral of the right-hand side is by assumption finite, so the same must hold for the left side. This says that the sum (and difference) of two square-integrable functions is square-integrable. For this example then, it isn't very difficult to show that it satisfies the axioms for a vector space, but it requires more than just a glance.

There are a few properties of vector spaces that seem to be missing. There is the somewhat odd notation $\vec{v}\,'$ for the additive inverse in axiom 5. Isn't that just $-\vec{v}$? Isn't the zero vector simply the number zero times a vector? Yes in both cases, but these are theorems that follow easily from the ten axioms listed. See problem 6.20. I'll do part (a) of that exercise as an example here:

Theorem: the vector $\vec{O}$ is unique.

Proof: assume it is not, then there are two such vectors, $\vec{O}_1$ and $\vec{O}_2$.

By [4], $\vec{O}_1 + \vec{O}_2 = \vec{O}_1$ ($\vec{O}_2$ is a zero vector)

By [6], the left side is $\vec{O}_2 + \vec{O}_1$

By [4], this is $\vec{O}_2$ ($\vec{O}_1$ is a zero vector)

Put these together and $\vec{O}_1 = \vec{O}_2$.

Theorem: If a subset of a vector space is closed under addition and multiplication by scalars, then it is itself a vector space. This means that if you add two elements of this subset to each other they remain in the subset and multiplying any element of the subset by a scalar leaves it in the subset. It is a "subspace."

Proof: the assumption of the theorem is that axioms 1 and 2 are satisfied as regards the subset. That axioms 3 through 10 hold follows because the elements of the subset inherit their properties from the larger vector space of which they are a part. Is this all there is to it? Not quite. Axioms 4 and 5 take a little more thought, and need the results of the problem 6.20, parts (b) and (d).

6.4 Linear Independence

A set of non-zero vectors is linearly dependent if one element of the set can be written as a linear combination of the others. The set is linearly independent if this cannot be done.

Bases, Dimension, Components

A basis for a vector space is a linearly independent set of vectors such that any vector in the space can be written as a linear combination of elements of this set. The *dimension* of the space is the number of elements in this basis.

If you take the usual vector space of arrows that start from the origin and lie in a plane, the common basis is denoted $\hat{\imath}$, $\hat{\jmath}$. If I propose a basis consisting of

$$\hat{\imath}, \qquad -\tfrac{1}{2}\hat{\imath} + \tfrac{\sqrt{3}}{2}\hat{\jmath}, \qquad -\tfrac{1}{2}\hat{\imath} - \tfrac{\sqrt{3}}{2}\hat{\jmath}$$

these will certainly span the space. Every vector can be written as a linear combination of them. They are however, redundant; the sum of all three is zero, so they aren't linearly independent and aren't a basis. If you use them as if they are a basis, the components of a given vector won't be unique. Maybe that's o.k. and you want to do it, but either be careful or look up the mathematical subject called "frames."

Beginning with the most elementary problems in physics and mathematics, it is clear that the choice of an appropriate coordinate system can provide great computational advantages. In dealing with the usual two and three dimensional vectors it is useful to express an arbitrary vector as a sum of unit vectors. Similarly, the use of Fourier series for the analysis of functions is a very powerful tool in analysis. These two ideas are essentially the same thing when you look at them as aspects of vector spaces.

If the elements of the basis are denoted $\vec{e}_i$, and a vector $\vec{a}$ is

$$\vec{a} = \sum_i a_i \vec{e}_i,$$

the numbers $\{a_i\}$ are called the *components* of $\vec{a}$ in the specified basis. Note that you don't have to talk about orthogonality or unit vectors or any other properties of the basis vectors save that they span the space and they're independent.

Example 1 is the prototype for the subject, and the basis usually chosen is the one designated $\hat{x}$, $\hat{y}$, (and $\hat{z}$ for three dimensions). Another notation for this is $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$ — I'll use $\hat{x}$-$\hat{y}$. In any case, the two (or three) arrows are at right angles to each other.

In example 5, the simplest choice of basis is

$$
\begin{aligned}
\vec{e}_1 &= (\,1 \quad 0 \quad 0 \quad \dots \quad 0 \,) \\
\vec{e}_2 &= (\,0 \quad 1 \quad 0 \quad \dots \quad 0 \,) \\
&\;\;\vdots \\
\vec{e}_n &= (\,0 \quad 0 \quad 0 \quad \dots \quad 1 \,)
\end{aligned}
\tag{6.1}
$$

In example 6, if the domain of the functions is from $-\infty$ to $+\infty$, a possible basis is the set of functions

$$
\psi_n(x) = x^n e^{-x^2/2}.
$$

The major distinction between this and the previous cases is that the dimension here is infinite. There is a basis vector corresponding to each non-negative integer. It's not obvious that this is a basis, but it's true.

If two vectors are equal to each other and you express them in the same basis, the corresponding components must be equal.

$$
\sum_i a_i \vec{e}_i = \sum_i b_i \vec{e}_i \quad \Longrightarrow \quad a_i = b_i \quad \text{for all } i
\tag{6.2}
$$

Suppose you have the relation between two functions of time

$$
A - B\omega + \gamma t = \beta t
\tag{6.3}
$$

that is, that the two *functions* are the same, think of this in terms of vectors: on the vector space of polynomials in t a basis is

$$
\vec{e}_0 = 1, \qquad \vec{e}_1 = t, \qquad \vec{e}_2 = t^2, \text{ etc.}
$$

Translate the preceding equation into this notation.

$$
(A - B\omega)\vec{e}_0 + \gamma \vec{e}_1 = \beta \vec{e}_1
\tag{6.4}
$$

For this to be valid the corresponding components must match:

$$
A - B\omega = 0, \qquad \text{and} \qquad \gamma = \beta
$$

Differential Equations

When you encounter differential equations such as

$$
m\frac{d^2 x}{dt^2} + b\frac{dx}{dt} + kx = 0, \qquad \text{or} \qquad \gamma\frac{d^3 x}{dt^3} + kt^2\frac{dx}{dt} + \alpha e^{-\beta t} x = 0,
\tag{6.5}
$$

the sets of solutions to each of these equations form vector spaces. All you have to do is to check the axioms, and because of the theorem in section 6.3 you don't even have to do all of that. The solutions

are functions, and as such they are elements of the vector space of example 2. All you need to do now is to verify that the sum of two solutions is a solution and that a constant times a solution is a solution. That's what the phrase "linear, homogeneous" means.

Another common differential equation is

$$\frac{d^2\theta}{dt^2} + \frac{g}{\ell}\sin\theta = 0$$

This describes the motion of an undamped pendulum, and the set of its solutions do *not* form a vector space. The sum of two solutions is not a solution.

The first of Eqs. (6.5) has two independent solutions,

$$x_1(t) = e^{-\gamma t}\cos\omega't, \qquad \text{and} \qquad x_2(t) = e^{-\gamma t}\sin\omega't \tag{6.6}$$

where $\gamma = -b/2m$ and $\omega' = \sqrt{\frac{k}{m} - \frac{b^2}{4m^2}}$. This is from Eq. (4.8). Any solution of this differential equation is a linear combination of these functions, and I can restate that fact in the language of this chapter by saying that x_1 and x_2 form a basis for the vector space of solutions of the damped oscillator equation. It has dimension two.

The second equation of the pair (6.5) is a third order differential equation, and as such you will need to specify three conditions to determine the solution and to determine all the three arbitrary constants. In other words, the dimension of the solution space of this equation is three.

In chapter 4 on the subject of differential equations, one of the topics was simultaneous differential equations, coupled oscillations. The simultaneous differential equations, Eq. (4.45), are

$$m_1\frac{d^2x_1}{dt^2} = -k_1x_1 - k_3(x_1 - x_2), \qquad \text{and} \qquad m_2\frac{d^2x_2}{dt^2} = -k_2x_2 - k_3(x_2 - x_1)$$

and have solutions that are pairs of functions. In the development of section 4.10 (at least for the equal mass, symmetric case), I found four pairs of functions that satisfied the equations. Now translate that into the language of this chapter, using the notation of column matrices for the functions. The solution is the vector

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

and the four basis vectors for this four-dimensional vector space are

$$\vec{e}_1 = \begin{pmatrix} e^{i\omega_1 t} \\ e^{i\omega_1 t} \end{pmatrix}, \qquad \vec{e}_2 = \begin{pmatrix} e^{-i\omega_1 t} \\ e^{-i\omega_1 t} \end{pmatrix}, \qquad \vec{e}_3 = \begin{pmatrix} e^{i\omega_2 t} \\ -e^{i\omega_2 t} \end{pmatrix}, \qquad \vec{e}_4 = \begin{pmatrix} e^{-i\omega_2 t} \\ -e^{-i\omega_2 t} \end{pmatrix}$$

Any solution of the differential equations is a linear combination of these. In the original notation, you have Eq. (4.52). In the current notation you have

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A_1\vec{e}_1 + A_2\vec{e}_2 + A_3\vec{e}_3 + A_4\vec{e}_4$$

6.5 Norms

The "norm" or length of a vector is a particularly important type of function that can be defined on a vector space. It is a function, usually denoted by $\|\ \|$, and that satisfies

1. $\|\vec{v}\| \geq 0$; $\quad \|\vec{v}\| = 0$ if and only if $\vec{v} = \vec{O}$
2. $\|\alpha\vec{v}\| = |\alpha|\,\|\vec{v}\|$
3. $\|\vec{v}_1 + \vec{v}_2\| \leq \|\vec{v}_1\| + \|\vec{v}_2\|$ (the triangle inequality) The distance between two vectors $\vec{v}_1$ and $\vec{v}_2$ is taken to be $\|\vec{v}_1 - \vec{v}_2\|$.

6.6 Scalar Product

The scalar product of two vectors is a scalar valued function of *two* vector variables. It could be denoted as $f(\vec{u}, \vec{v})$, but a standard notation for it is $\langle \vec{u}, \vec{v} \rangle$. It must satisfy the requirements

1. $\langle \vec{w}, (\vec{u} + \vec{v}) \rangle = \langle \vec{w}, \vec{u} \rangle + \langle \vec{w}, \vec{v} \rangle$
2. $\langle \vec{w}, \alpha \vec{v} \rangle = \alpha \langle \vec{w}, \vec{v} \rangle$
3. $\langle \vec{u}, \vec{v} \rangle^* = \langle \vec{v}, \vec{u} \rangle$
4. $\langle \vec{v}, \vec{v} \rangle \geq 0;$ and $\langle \vec{v}, \vec{v} \rangle = 0$ if and only if $\vec{v} = \vec{O}$

When a scalar product exists on a space, a norm naturally does too:

$$\| \vec{v} \| = \sqrt{\langle \vec{v}, \vec{v} \rangle}. \tag{6.7}$$

That this *is* a norm will follow from the Cauchy-Schwartz inequality. Not all norms come from scalar products.

Examples

Use the examples of section 6.3 to see what these are. The numbers here refer to the numbers of that section.

1 A norm is the usual picture of the length of the line segment. A scalar product is the usual product of lengths times the cosine of the angle between the vectors.

$$\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v} = u \, v \cos \vartheta. \tag{6.8}$$

4 A norm can be taken as the least upper bound of the magnitude of the function. This is distinguished from the "maximum" in that the function may not actually achieve a maximum value. Since it is bounded however, there is an upper bound (many in fact) and we take the smallest of these as the norm. On $-\infty < x < +\infty$, the function $|\tan^{-1} x|$ has $\pi/2$ for its least upper bound, though it never equals that number.

5 A possible scalar product is

$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle = \sum_{k=1}^{n} a_k^* \, b_k. \tag{6.9}$$

There are other scalar products for the same vector space, for example

$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle = \sum_{k=1}^{n} k \, a_k^* \, b_k \tag{6.10}$$

In fact any other positive function can appear as the coefficient in the sum and it still defines a valid scalar product. It's surprising how often something like this happens in real situations. In studying normal modes of oscillation the masses of different particles will appear as coefficients in a natural scalar product.

I used complex conjugation on the first factor here, but example 5 referred to real numbers only. The reason for leaving the conjugation in place is that when you jump to example 14 you want to allow for complex numbers, and it's harmless to put it in for the real case because in that instance it leaves the number alone.

For a norm, there are many possibilities:

$$
\begin{aligned}
&(1)\ \|(a_1,\ldots,a_n)\| = \sqrt{\sum_{k=1}^{n} |a_k|^2}\\
&(2)\ \|(a_1,\ldots,a_n)\| = \sum_{k=1}^{n} |a_k|\\
&(3)\ \|(a_1,\ldots,a_n)\| = \max_{k=1}^{n} |a_k|\\
&(4)\ \|(a_1,\ldots,a_n)\| = \max_{k=1}^{n} k|a_k|.
\end{aligned}
\tag{6.11}
$$

The United States Postal Service prefers a variation on the second of these norms, see problem 8.45.

6 A possible choice for a scalar product is

$$
\langle f, g \rangle = \int_a^b dx\, f(x)^*\, g(x).
\tag{6.12}
$$

9 Scalar products and norms used here are just like those used for example 5. The difference is that the sums go from 1 to infinity. The problem of convergence doesn't occur because there are only a finite number of non-zero terms.

10 Take the norm to be

$$
\|(a_1, a_2, \ldots)\| = \sqrt{\sum_{k=1}^{\infty} |a_k|^2},
\tag{6.13}
$$

and this by assumption will converge. The natural scalar product is like that of example 5, but with the sum going out to infinity. It requires a small amount of proof to show that this will converge. See problem 6.19.

11 A norm is $\|\vec{v}\| = \sum_{i=1}^{\infty} |a_i|$. There is no scalar product that will produce this norm, a fact that you can prove by using the results of problem 6.13.

13 A natural norm is

$$
\|f\| = \left[\int_a^b dx\, |f(x)|^p\right]^{1/p}.
\tag{6.14}
$$

To demonstrate that this *is* a norm requires the use of some special inequalities found in advanced calculus books.

15 If A and B are two matrices, a scalar product is $\langle A, B \rangle = \text{Tr}(A^\dagger B)$, where $\dagger$ is the transpose complex conjugate of the matrix and Tr means the trace, the sum of the diagonal elements. Several possible norms can occur. One is $\|A\| = \sqrt{\text{Tr}(A^\dagger A)}$. Another is the maximum value of $\|A\vec{u}\|$, where $\vec{u}$ is a unit vector and the norm of $\vec{u}$ is taken to be $\left[|u_1|^2 + \cdots + |u_n|^2\right]^{1/2}$.

19 A valid definition of a norm for the motions of a drumhead is its total energy, kinetic plus potential. How do you describe this mathematically? It's something like

$$
\int dx\, dy\, \frac{1}{2}\left[\left(\frac{\partial f}{\partial t}\right)^2 + (\nabla f)^2\right]
$$

I've left out all the necessary constants, such as mass density of the drumhead and tension in the drumhead. You can perhaps use dimensional analysis to surmise where they go.

There is an example in criminal law in which the distinctions between some of these norms have very practical consequences. If you're caught selling drugs in New York there is a longer sentence if your

sale is within 1000 feet of a school. If you are an attorney defending someone accused of this crime, which of the norms in Eq. (6.11) would you argue for? The legislators who wrote this law didn't know linear algebra, so they didn't specify which norm they intended. The prosecuting attorney argued for norm #1, "as the crow flies," but the defense argued that "crows don't sell drugs" and humans move along city streets, so norm #2 is more appropriate.

The New York Court of Appeals decided that the Pythagorean norm (#1) is the appropriate one and they rejected the use of the pedestrian norm that the defendant advocated (#2). www.courts.state.ny.us/ctapps/decisions/nov05/162opn05.pdf

6.7 Bases and Scalar Products

When there is a scalar product, a most useful type of basis is the orthonormal one, satisfying

$$\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{6.15}$$

The notation δ_{ij} represents the very useful Kronecker delta symbol.

In the example of Eq. (6.1) the basis vectors are orthonormal with respect to the scalar product in Eq. (6.9). It is orthogonal with respect to the other scalar product mentioned there, but it is not in that case normalized to magnitude one.

To see how the choice of even an orthonormal basis depends on the scalar product, try a different scalar product on this space. Take the special case of two dimensions. The vectors are now pairs of numbers. Think of the vectors as 2×1 matrix column and use the 2×2 matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Take the scalar product of two vectors to be

$$\langle (a_1, a_2), (b_1, b_2) \rangle = \begin{pmatrix} a_1^* & a_2^* \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = 2a_1^* b_1 + a_1^* b_2 + a_2^* b_1 + 2a_2^* b_2 \tag{6.16}$$

To show that this satisfies all the defined requirements for a scalar product takes a small amount of labor. The vectors that you may expect to be orthogonal, $(1\ 0)$ and $(0\ 1)$, are not.

In example 6, if we let the domain of the functions be $-L < x < +L$ and the scalar product is as in Eq. (6.12), then the set of trigonometric functions can be used as a basis.

$$\sin \frac{n\pi x}{L} \quad \text{and} \quad \cos \frac{m\pi x}{L}$$
$$n = 1, 2, 3, \ldots \quad \text{and} \quad m = 0, 1, 2, 3, \ldots.$$

That a function can be written as a series

$$f(x) = \sum_1^\infty a_n \sin \frac{n\pi x}{L} + \sum_0^\infty b_m \cos \frac{m\pi x}{L} \tag{6.17}$$

on the domain $-L < x < +L$ is just an example of Fourier series, and the components of f in this basis are Fourier coefficients $a_1, \ldots, b_0, \ldots$. An equally valid and more succinctly stated basis is

$$e^{n\pi ix/L}, \quad n = 0, \pm 1, \pm 2, \ldots$$

Chapter 5 on Fourier series shows many other choices of bases, all orthogonal, but not necessarily normalized.

To emphasize the relationship between Fourier series and the ideas of vector spaces, this picture represents three out of the infinite number of basis vectors and part of a function that uses these vectors to form a Fourier series.

$$f(x) = \frac{1}{2}\sin\frac{\pi x}{L} + \frac{2}{3}\sin\frac{2\pi x}{L} + \frac{1}{3}\sin\frac{3\pi x}{L} + \cdots$$

The orthogonality of the sines becomes the geometric term "perpendicular," and if you look at section 8.11, you will see that the subject of least square fitting of data to a sum of sine functions leads you right back to Fourier series, and to the same picture as here.
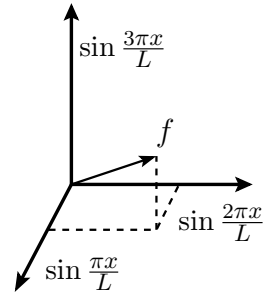
6.8 Gram-Schmidt Orthogonalization

From a basis that is not orthonormal, it is possible to construct one that is. This device is called the Gram-Schmidt procedure. Suppose that a basis is known (finite or infinite), $\vec{v}_1$, $\vec{v}_2$, …

Step 1: normalize $\vec{v}_1$. $\vec{e}_1 = \vec{v}_1 / \sqrt{\langle \vec{v}_1, \vec{v}_1 \rangle}$.

Step 2: construct a linear combination of $\vec{v}_1$ and $\vec{v}_2$ that is orthogonal to $\vec{v}_1$:
    Let $\vec{e}_{20} = \vec{v}_2 - \vec{e}_1 \langle \vec{e}_1, \vec{v}_2 \rangle$ and then normalize it.

$$\vec{e}_2 = \vec{e}_{20} / \langle \vec{e}_{20}, \vec{e}_{20} \rangle^{1/2}. \tag{6.18}$$

Step 3: Let $\vec{e}_{30} = \vec{v}_3 - \vec{e}_1 \langle \vec{e}_1, \vec{v}_3 \rangle - \vec{e}_2 \langle \vec{e}_2, \vec{v}_3 \rangle$ *etc.* repeating step 2.
What does this look like? See problem 6.3.

6.9 Cauchy-Schwartz inequality

For common three-dimensional vector geometry, it is obvious that for any real angle, $\cos^2 \theta \leq 1$. In terms of a dot product, this is $|\vec{A} \cdot \vec{B}| \leq AB$. This can be generalized to any scalar product on any vector space:

$$\left| \langle \vec{u}, \vec{v} \rangle \right| \leq \|\vec{u}\| \, \|\vec{v}\|. \tag{6.19}$$

The proof starts from a simple but not-so-obvious point. The scalar product of a vector with itself is by definition positive, so for any two vectors $\vec{u}$ and $\vec{v}$ you have the inequality

$$\langle \vec{u} - \lambda\vec{v}, \vec{u} - \lambda\vec{v} \rangle \geq 0. \tag{6.20}$$

where λ is any complex number. This expands to

$$\langle \vec{u}, \vec{u} \rangle + |\lambda|^2 \langle \vec{v}, \vec{v} \rangle - \lambda \langle \vec{u}, \vec{v} \rangle - \lambda^* \langle \vec{v}, \vec{u} \rangle \geq 0. \tag{6.21}$$

How much bigger than zero the left side is will depend on the parameter λ. To find the smallest value that the left side can have you simply differentiate. Let $\lambda = x + iy$ and differentiate with respect to x and y, setting the results to zero. This gives (see problem 6.5)

$$\lambda = \langle \vec{v}, \vec{u} \rangle / \langle \vec{v}, \vec{v} \rangle. \tag{6.22}$$

Substitute this value into the above inequality (6.21)

$$\langle \vec{u}, \vec{u} \rangle + \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\langle \vec{v}, \vec{v} \rangle} - \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\langle \vec{v}, \vec{v} \rangle} - \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\langle \vec{v}, \vec{v} \rangle} \geq 0. \tag{6.23}$$

This becomes

$$|\langle \vec{u}, \vec{v} \rangle|^2 \le \langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle \tag{6.24}$$

This isn't quite the result needed, because Eq. (6.19) is written differently. It refers to a norm and I haven't established that the square root of $\langle \vec{v}, \vec{v} \rangle$ *is* a norm. When I do, then the square root of this is the desired inequality (6.19).

For a couple of examples of this inequality, take specific scalar products. First the common directed line segments:

$$\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v} = uv \cos\theta, \qquad \text{so} \qquad |uv \cos\theta|^2 \le |u|^2 |v|^2$$

$$\left| \int_a^b dx\, f(x)^* g(x) \right|^2 \le \left[\int_a^b dx\, |f(x)|^2 \right] \left[\int_a^b dx\, |g(x)|^2 \right]$$

The first of these is familiar, but the second is not, though when you look at it from the general vector space viewpoint they are essentially the same.

Norm from a Scalar Product

The equation (6.7), $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$, defines a norm. Properties one and two for a norm are simple to check. (Do so.) The third requirement, the triangle inequality, takes a bit of work and uses the inequality Eq. (6.24).

$$
\begin{aligned}
\langle \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_2 \rangle &= \langle \vec{v}_1, \vec{v}_1 \rangle + \langle \vec{v}_2, \vec{v}_2 \rangle + \langle \vec{v}_1, \vec{v}_2 \rangle + \langle \vec{v}_2, \vec{v}_1 \rangle \\
&\le \langle \vec{v}_1, \vec{v}_1 \rangle + \langle \vec{v}_2, \vec{v}_2 \rangle + |\langle \vec{v}_1, \vec{v}_2 \rangle| + |\langle \vec{v}_2, \vec{v}_1 \rangle| \\
&= \langle \vec{v}_1, \vec{v}_1 \rangle + \langle \vec{v}_2, \vec{v}_2 \rangle + 2|\langle \vec{v}_1, \vec{v}_2 \rangle| \\
&\le \langle \vec{v}_1, \vec{v}_1 \rangle + \langle \vec{v}_2, \vec{v}_2 \rangle + 2\sqrt{\langle \vec{v}_1, \vec{v}_1 \rangle \langle \vec{v}_2, \vec{v}_2 \rangle} \\
&= \left(\sqrt{\langle \vec{v}_1, \vec{v}_1 \rangle} + \sqrt{\langle \vec{v}_2, \vec{v}_2 \rangle} \right)^2
\end{aligned}
$$

The first inequality is a property of complex numbers. The second one is Eq. (6.24). The square root of the last line is the triangle inequality, thereby justifying the use of $\sqrt{\langle \vec{v}, \vec{v} \rangle}$ as the norm of $\vec{v}$ and in the process validating Eq. (6.19).

$$\|\vec{v}_1 + \vec{v}_2\| = \sqrt{\langle \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_2 \rangle} \le \sqrt{\langle \vec{v}_1, \vec{v}_1 \rangle} + \sqrt{\langle \vec{v}_2, \vec{v}_2 \rangle} = \|\vec{v}_1\| + \|\vec{v}_2\| \tag{6.25}$$

6.10 Infinite Dimensions

Is there any real difference between the cases where the dimension of the vector space is finite and the cases where it's infinite? Yes. Most of the concepts are the same, but you have to watch out for the question of convergence. If the dimension is finite, then when you write a vector in terms of a basis $\vec{v} = \sum a_k \vec{e}_k$, the sum is finite and you don't even have to think about whether it converges or not. In the infinite-dimensional case you do.

It is even possible to have such a series converge, but not to converge to a vector. If that sounds implausible, let me take an example from a slightly different context, ordinary rational numbers. These are the number m/n where m and n are integers ($n \ne 0$). Consider the sequence

$$1, \quad 14/10, \quad 141/100, \quad 1414/1000, \quad 14142/10000, \quad 141421/100000, \ldots$$

These are quotients of integers, but the limit is $\sqrt{2}$ and that's *not** a rational number. Within the confines of rational numbers, this sequence doesn't converge. You have to expand the context to get a limit. That context is the real numbers. The same thing happens with vectors when the dimension of the space is infinite — in order to find a limit you sometimes have to expand the context and to expand what you're willing to call a vector.

Look at example 9 from section 6.3. These are sets of numbers $(a_1, a_2, \ldots)$ with just a finite number of non-zero entries. If you take a sequence of such vectors

$$(1, 0, 0, \ldots), \quad (1, 1, 0, 0, \ldots), \quad (1, 1, 1, 0, 0, \ldots), \ldots$$

Each has a finite number of non-zero elements but the limit of the sequence does not. It isn't a vector in the original vector space. Can I expand to a larger vector space? Yes, just use example 8, allowing any number of non-zero elements.

For a more useful example of the same kind, start with the same space and take the sequence

$$(1, 0, \ldots), \quad (1, \tfrac{1}{2}, 0, \ldots), \quad (1, \tfrac{1}{2}, \tfrac{1}{3}, 0, \ldots), \ldots$$

Again the limit of such a sequence doesn't have a finite number of entries, but example 10 will hold such a limit, because $\sum_1^\infty |a_k|^2 < \infty$.

How do you know when you have a vector space without holes in it? That is, one in which these problems with limits don't occur? The answer lies in the idea of a Cauchy sequence. I'll start again with the rational numbers to demonstrate the idea. The sequence of numbers that led to the square root of two has the property that even though the elements of the sequence weren't approaching a rational number, the elements were getting close *to each other*. Let $\{r_n\}$, $n = 1, 2, \ldots$ be a sequence of rational numbers.

$$\lim_{n,m\to\infty} |r_n - r_m| = 0 \qquad \text{means}$$

For any $\epsilon > 0$ there is an N so that if both n and m are $> N$ then $|r_n - r_m| < \epsilon$.

$$(6.26)$$

This property defines the sequence r_n as a Cauchy sequence. A sequence of rational numbers converges to a real number if and only if it is a Cauchy sequence; this is a theorem found in many advanced calculus texts. Still other texts will take a different approach and use the concept of a Cauchy sequence to construct the *definition* of the real numbers.

The extension of this idea to infinite dimensional vector spaces requires simply that you replace the absolute value by a norm, so that a Cauchy sequence is defined by $\lim_{n,m} \|\vec{v}_n - \vec{v}_m\| = 0$. A "complete" vector space is one in which every Cauchy sequence converges. A vector space that has a scalar product and that is also complete using the norm that this scalar product defines is called a Hilbert Space.

I don't want to imply that the differences between finite and infinite dimensional vector spaces is just a technical matter of convergence. In infinite dimensions there is far more room to move around, and the possible structures that occur are vastly more involved than in the finite dimensional case. The subject of quantum mechanics has Hilbert Spaces at the foundation of its whole structure.

---

* Proof: If it is, then express it in simplest form as $m/n = \sqrt{2} \Rightarrow m^2 = 2n^2$ where m and n have no common factor. This equation implies that m must be even: $m = 2m_1$. Substitute this value, giving $2m_1^2 = n^2$. That in turn implies that n is even, and this contradicts the assumption that the original quotient was expressed without common factors.

Exercises

1 Determine if these are vector spaces with the usual rules for addition and multiplication by scalars. If not, which axiom(s) do they violate?
(a) Quadratic polynomials of the form $ax^2 + bx$
(b) Quadratic polynomials of the form $ax^2 + bx + 1$
(c) Quadratic polynomials $ax^2 + bx + c$ with $a + b + c = 0$
(d) Quadratic polynomials $ax^2 + bx + c$ with $a + b + c = 1$

2 What is the dimension of the vector space of (up to) 5th degree polynomials having a double root at $x = 1$?

3 Starting from three dimensional vectors (the common directed line segments) and a single fixed vector $\vec{B}$, is the set of all vectors $\vec{v}$ with $\vec{v} \cdot \vec{B} = 0$ a vector space? If so, what is it's dimension? Is the set of all vectors $\vec{v}$ with $\vec{v} \times \vec{B} = 0$ a vector space? If so, what is it's dimension?

4 The set of all odd polynomials with the expected rules for addition and multiplication by scalars. Is it a vector space?

5 The set of all polynomials where the function "addition" is defined to be $f_3 = f_2 + f_1$ if the number $f_3(x) = f_1(-x) + f_2(-x)$. Is it a vector space?

6 Same as the preceding, but for (a) even polynomials, (b) odd polynomials

7 The set of directed line segments in the plane with the new rule for addition: add the vectors according to the usual rule then rotate the result by $10°$ counterclockwise. Which vector space axioms are obeyed and which not?

Problems

6.1 Fourier series represents a choice of basis for functions on an interval. For suitably smooth functions on the interval 0 to L, one basis is

$$\vec{e}_n = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}. \tag{6.27}$$

Use the scalar product $\langle f, g \rangle = \int_0^L f^*(x)g(x)\,dx$ and show that this is an orthogonal basis normalized to 1, *i.e.* it is orthonormal.

6.2 A function $F(x) = x(L - x)$ between zero and L. Use the basis of the preceding problem to write this vector in terms of its components:

$$F = \sum_1^{\infty} \alpha_n \vec{e}_n. \tag{6.28}$$

If you take the result of using this basis and write the resulting function outside the interval $0 < x < L$, graph the result.

6.3 For two dimensional real vectors with the usual parallelogram addition, interpret in pictures the first two steps of the Gram-Schmidt process, section 6.8.

6.4 For two dimensional real vectors with the usual parallelogram addition, *interpret* the vectors $\vec{u}$ and $\vec{v}$ and the parameter λ used in the proof of the Cauchy-Schwartz inequality in section 6.9. Start by considering the set of points in the plane formed by $\{\vec{u} - \lambda\vec{v}\}$ as λ ranges over the set of reals. In particular, when λ was picked to minimize the left side of the inequality (6.21), what do the vectors look like? Go through the proof and interpret it in the context of these pictures. State the idea of the whole proof geometrically.
Note: I don't mean just copy the proof. Put the geometric interpretation into words.

6.5 Start from Eq. (6.21) and show that the minimum value of the function of $\lambda = x + iy$ is given by the value stated there. Note: this derivation applies to complex vector spaces and scalar products, not just real ones. Is this a *minimum*?

6.6 For the vectors in three dimensions,

$$\vec{v}_1 = \hat{x} + \hat{y}, \qquad \vec{v}_2 = \hat{y} + \hat{z}, \qquad \vec{v}_3 = \hat{z} + \hat{x}$$

use the Gram-Schmidt procedure to construct an orthonormal basis starting from $\vec{v}_1$. Ans: $\vec{e}_3 = (\hat{x} - \hat{y} + \hat{z})/\sqrt{3}$

6.7 For the vector space of polynomials in x, use the scalar product defined as

$$\langle f, g \rangle = \int_{-1}^{1} dx\, f(x)^* g(x)$$

(Everything is real here, so the complex conjugation won't matter.) Start from the vectors

$$\vec{v}_0 = 1, \qquad \vec{v}_1 = x, \qquad \vec{v}_2 = x^2, \qquad \vec{v}_3 = x^3$$

and use the Gram-Schmidt procedure to construct an orthonormal basis starting from $\vec{v}_0$. Compare these results to the results of section 4.11. [These polynomials appear in the study of electric potentials and in the study of angular momentum in quantum mechanics: Legendre polynomials.]

6.8 Repeat the previous problem, but use a different scalar product:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} dx\, e^{-x^2}\, f(x)^* g(x)$$

[These polynomials appear in the study of the harmonic oscillator in quantum mechanics and in the study of certain waves in the upper atmosphere. With a conventional normalization they are called Hermite polynomials.]

6.9 Consider the set of all polynomials in x having degree $\leq N$. Show that this is a vector space and find its dimension.

6.10 Consider the set of all polynomials in x having degree $\leq N$ and only even powers. Show that this is a vector space and find its dimension. What about odd powers only?

6.11 Which of these are vector spaces?
(**a**) all polynomials of degree 3
(**b**) all polynomials of degree ≤ 3 [Is there a difference between (**a**) and (**b**)?]
(**c**) all functions such that $f(1) = 2f(2)$
(**d**) all functions such that $f(2) = f(1) + 1$
(**e**) all functions satisfying $f(x + 2\pi) = f(x)$
(**f**) all positive functions
(**g**) all polynomials of degree ≤ 4 satisfying $\int_{-1}^{1} dx\, x f(x) = 0$.
(**h**) all polynomials of degree ≤ 4 where the coefficient of x is zero.
   [Is there a difference between (**g**) and (**h**)?]

6.12 (**a**) For the common picture of arrows in three dimensions, prove that the subset of vectors $\vec{v}$ that satisfy $\vec{A} \cdot \vec{v} = 0$ for fixed $\vec{A}$ forms a vector space. Sketch it.
(**b**) What if the requirement is that both $\vec{A} \cdot \vec{v} = 0$ and $\vec{B} \cdot \vec{v} = 0$ hold. Describe this and sketch it.

6.13 If a norm is defined in terms of a scalar product, $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$, it satisfies the "parallelogram identity" (for real scalars),

$$\|\vec{u} + \vec{v}\|^2 + \|\vec{u} - \vec{v}\|^2 = 2\|\vec{u}\|^2 + 2\|\vec{v}\|^2. \tag{6.29}$$

6.14 If a norm satisfies the parallelogram identity, then it comes from a scalar product. Again, assume real scalars. Consider combinations of $\|\vec{u} + \vec{v}\|^2$, $\|\vec{u} - \vec{v}\|^2$ and construct what ought to be the scalar product. You then have to prove the four properties of the scalar product as stated at the start of section 6.6. Numbers four and three are easy. Number one requires that you keep plugging away, using the parallelogram identity (four times by my count).

Number two is downright tricky; leave it to the end. If you can prove it for integer and rational values of the constant α, consider it a job well done. I used induction at one point in the proof. The final step, extending α to all real values, requires some arguments about limits, and is typically the sort of reasoning you will see in an advanced calculus or mathematical analysis course.

6.15 Modify the example number 2 of section 6.3 so that $f_3 = f_1 + f_2$ means $f_3(x) = f_1(x - a) + f_2(x - b)$ for fixed a and b. Is this still a vector space?

6.16 The scalar product you use depends on the problem you're solving. The fundamental equation (5.15) started from the equation $u'' = \lambda u$ and resulted in the scalar product

$$\langle u_2, u_1 \rangle = \int_a^b dx \, u_2(x)^* u_1(x)$$

Start instead from the equation $u'' = \lambda w(x)u$ and see what identity like that of Eq. (5.15) you come to. Assume w is real. What happens if it isn't? In order to have a legitimate scalar product in the sense of section 6.6, what other requirements must you make about w?

6.17 The equation describing the motion of a string that is oscillating with frequency ω about its stretched equilibrium position is

$$\frac{d}{dx}\left(T(x)\frac{dy}{dx} \right) = -\omega^2 \mu(x)y$$

Here, $y(x)$ is the sideways displacement of the string from zero; $T(x)$ is the tension in the string (not necessarily a constant); $\mu(x)$ is the linear mass density of the string (again, it need not be a constant). The time-dependent motion is really $y(x)\cos(\omega t + \phi)$, but the time dependence does not concern us here. As in the preceding problem, derive the analog of Eq. (5.15) for this equation. For the analog of Eq. (5.16) state the boundary conditions needed on y and deduce the corresponding orthogonality equation. This scalar product has the mass density for a weight.
Ans: $\left[T(x)(y_1' y_2^* - y_1 y_2^{*\prime}) \right]_a^b = (\omega_2^{*2} - \omega_1^2) \int_a^b \mu(x) y_2^* y_1 \, dx$

6.18 The way to define the sum in example 17 is

$$\sum_x |f(x)|^2 = \lim_{c \to 0}\{\text{the sum of } |f(x)|^2 \text{ for those } x \text{ where } |f(x)|^2 > c > 0\}. \qquad (6.30)$$

This makes sense only if for each $c > 0$, $|f(x)|^2$ is greater than c for just a finite number of values of x. Show that the function

$$f(x) = \begin{cases} 1/n & \text{for } x = 1/n \\ 0 & \text{otherwise} \end{cases}$$

is in this vector space, and that the function $f(x) = x$ is not. What is a basis for this space? [Take $0 \le x \le 1$] This is an example of a vector space with non-countable dimension.

6.19 In example 10, it is assumed that $\sum_1^\infty |a_k|^2 < \infty$. Show that this implies that the sum used for the scalar product also converges: $\sum_1^\infty a_k^* b_k$. [Consider the sums $\sum |a_k + ib_k|^2$, $\sum |a_k - ib_k|^2$, $\sum |a_k + b_k|^2$, and $\sum |a_k - b_k|^2$, allowing complex scalars.]

6.20 Prove strictly from the axioms for a vector space the following four theorems. Each step in your proof must *explicitly* follow from one of the vector space axioms or from a property of scalars or from a previously proved theorem.
(a) The vector $\vec{O}$ is unique. [Assume that there are two, $\vec{O}_1$ and $\vec{O}_2$. Show that they're equal. First step: use axiom 4.]

(b) The number 0 times any vector is the zero vector: $0\vec{v} = \vec{O}$.
(c) The vector $\vec{v}\,'$ is unique.
(d) $(-1)\vec{v} = \vec{v}\,'$.

6.21 For the vector space of polynomials, are the two functions $\{1+x^2,\ x+x^3\}$ linearly independent?

6.22 Find the dimension of the space of functions that are linear combinations of
$\{1,\ \sin x,\ \cos x,\ \sin^2 x,\ \cos^2 x,\ \sin^4 x,\ \cos^4 x,\ \sin^2 x \cos^2 x\}$
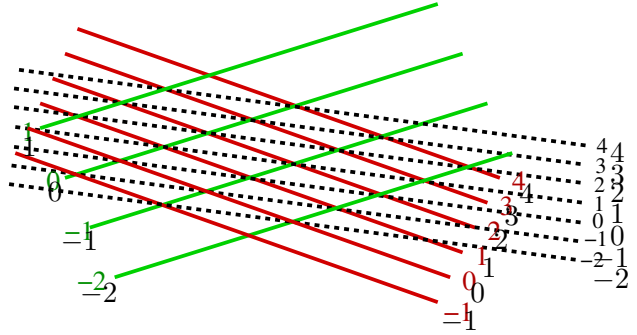
6.23 A model vector space is formed by drawing equidistant parallel lines in a plane and labelling adjacent lines by successive integers from ∞ to $+\infty$. Define multiplication by a (real) scalar so that multiplication of the vector by α means multiply the distance between the lines by $1/\alpha$. Define addition of two vectors by finding the intersections of the lines and connecting opposite corners of the parallelograms to form another set of parallel lines. The resulting lines are labelled as the sum of the two integers from the intersecting lines. (There are two choices here, if one is addition, what is the other?) Show that this construction satisfies all the requirements for a vector space. Just as a directed line segment is a good way to picture velocity, this construction is a good way to picture the gradient of a function. In the vector space of directed line segments, you pin the vectors down so that they all start from a single point. Here, you pin them down so that the lines labeled "zero" all pass through a fixed point. Did I define how to multiply by a negative scalar? If not, then you should. This picture of vectors is developed extensively in the text "Gravitation" by Misner, Wheeler, and Thorne.

6.24 In problem 6.11 **(g)**, find a basis for the space. Ans: $1,\ x,\ 3x - 5x^3$.

6.25 What is the dimension of the set of polynomials of degree less than or equal to 10 and with a triple root at $x = 1$?

6.26 Verify that Eq. (6.16) does satisfy the requirements for a scalar product.

6.27 A variation on problem 6.15: $f_3 = f_1 + f_2$ means
(a) $f_3(x) = A f_1(x - a) + B f_2(x - b)$ for fixed a, b, A, B. For what values of these constants is this a vector space?
(b) Now what about $f_3(x) = f_1(x^3) + f_2(x^3)$?

6.28 Determine if these are vector spaces:
(1) Pairs of numbers with addition defined as $(x_1, x_2) + (y_1, y_2) = (x_1 + y_2, x_2 + y_1)$ and multiplication by scalars as $c(x_1, x_2) = (cx_1, cx_2)$.
(2) Like example 2 of section 6.3, but restricted to those f such that $f(x) \geq 0$. (real scalars)
(3) Like the preceding line, but define addition as $(f + g)(x) = f(x)g(x)$ and $(cf)(x) = \big(f(x)\big)^c$.

6.29 Do the same calculation as in problem 6.7, but use the scalar product

$$\langle f, g \rangle = \int_0^1 x^2 \, dx \, f^*(x) g(x)$$

6.30 Show that the following is a scalar product.

$$\langle f, g \rangle = \int_a^b dx \left[f^*(x) g(x) + \lambda f^{*\prime}(x) g'(x) \right]$$

where λ is a constant. What restrictions if any must you place on λ? The name Sobolev is associated with this scalar product.

6.31 (a) With the scalar product of problem 6.29, find the angle between the vectors 1 and x. Here the word angle appears in the sense of $\vec{A} \cdot \vec{B} = AB \cos \theta$. (b) What is the angle if you use the scalar product of problem 6.7? (c) With the first of these scalar products, what combination of 1 and x is orthogonal to 1? Ans: $14.48°$

6.32 In the online text linked on the second page of this chapter, you will find that section two of chapter three has enough additional problems to keep you happy.

6.33 Show that the sequence of rational numbers $a_n = \sum_{k=1}^n 1/k$ is not a Cauchy sequence. What about $\sum_{k=1}^n 1/k^2$?

6.34 In the vector space of polynomials of the form $\alpha x + \beta x^3$, use the scalar product $\langle f, g \rangle = \int_0^1 dx \, f(x)^* g(x)$ and construct an orthogonal basis for this space. Ans: One pair is x, $x^3 - \frac{3}{5}x$.

6.35 You can construct the Chebyshev polynomials by starting from the successive powers, x^n, $n = 0, 1, 2, \ldots$ and applying the Gram-Schmidt process. The scalar product in this case is

$$\langle f, g \rangle = \int_{-1}^1 dx \, \frac{f(x)^* g(x)}{\sqrt{1 - x^2}}$$

The conventional normalization for these polynomials is $T_n(1) = 1$, so you should not try to make the norm of the resulting vectors one. Construct the first four of these polynomials, and show that these satisfy $T_n(\cos \theta) = \cos(n\theta)$. These polynomials are used in numerical analysis because they have the property that they oscillate uniformly between -1 and $+1$ on the domain $-1 < x < 1$. Verify that your results for the first four polynomials satisfy the recurrence relation: $T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x)$. Also show that $\cos\big((n+1)\theta\big) = 2 \cos \theta \cos(n\theta) - \cos\big((n-1)\theta\big)$.

6.36 In spherical coordinates (θ, ϕ), the angle θ is measured from the z-axis, and the function $f_1(\theta, \phi) = \cos \theta$ can be written in terms of rectangular coordinates as (section 8.8)

$$f_1(\theta, \phi) = \cos \theta = \frac{z}{r} = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$

Pick up the function f_1 and rotate it by $90°$ counterclockwise about the positive y-axis. Do this rotation in terms of rectangular coordinates, but express the result in terms of spherical coordinates: sines and cosines of θ and ϕ. Call it f_2. Draw a picture and figure out where the original and the rotated function are positive and negative and zero.

Now pick up the same f_1 and rotate it by $90°$ clockwise about the positive x-axis, again finally expressing the result in terms of spherical coordinates. Call it f_3.

If now you take the original f_1 and rotate it about some random axis by some random angle, show that the resulting function f_4 is a linear combination of the three functions f_1, f_2, and f_3. I.e., all these possible rotated functions form a three dimensional vector space. Again, calculations such as these are much easier to demonstrate in rectangular coordinates.

6.37 Take the functions f_1, f_2, and f_3 from the preceding problem and sketch the shape of the functions

$$r\, e^{-r} f_1(\theta, \phi), \qquad r\, e^{-r} f_2(\theta, \phi), \qquad r\, e^{-r} f_3(\theta, \phi)$$

To sketch these, picture them as defining some sort of density in space, ignoring the fact that they are sometimes negative. You can just take the absolute value or the square in order to visualize where they are big or small. Use dark and light shading to picture where the functions are big and small. Start by finding *where* they have the largest and smallest magnitudes. See if you can find similar pictures in an introductory chemistry text. Alternately, check out winter.group.shef.ac.uk/orbitron/

6.38 Use the results of problem 6.17 and apply it to the Legendre equation Eq. (4.55) to demonstrate that the Legendre polynomials obey $\int_{-1}^{1} dx\, P_n(x) P_m(x) = 0$ if $n \neq m$. Note: the function $T(x)$ from problem 6.17 is zero at these endpoints. That does *not* imply that there are no conditions on the functions y_1 and y_2 at those endpoints. The product of $T(x) y_1' y_2$ has to vanish there. Use the result stated just after Eq. (4.59) to show that only the Legendre polynomials and not the more general solutions of Eq. (4.58) work.

6.39 Using the result of the preceding problem that the Legendre polynomials are orthogonal, show that the equation (4.62)(a) follows from Eq. (4.62)(e). Square that equation (e) and integrate $\int_{-1}^{1} dx$. Do the integral on the left and then expand the result in an infinite series in t. On the right you have integrals of products of Legendre polynomials, and only the squared terms are non-zero. Equate like powers of t and you will have the result.

6.40 Use the scalar product of Eq. (6.16) and construct an orthogonal basis using the Gram-Schmidt process and starting from $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Verify that your answer works in at least one special case.

6.41 For the differential equation $\ddot{x} + x = 0$, pick a set of independent solutions to the differential equation — any ones you like. Use the scalar product $\langle f, g \rangle = \int_0^1 dx\, f(x)^* g(x)$ and apply the Gram-Schmidt method to find an orthogonal basis in this space of solutions. Is there another scalar product that would make this analysis simpler? Sketch the orthogonal functions that you found.