2

Complex Vector Spaces

Philosophy is written in that great book which continually lies open before us (I mean the Universe). But one cannot understand this book until one has learned to understand the language and to know the letters in which it is written. It is written in the language of mathematics, and the letters are triangles, circles and other geometric figures. Without these means it is impossible for mankind to understand a single word; without these means there is only vain stumbling in a dark labyrinth.[1]

Galileo Galilei

Quantum theory is cast in the language of complex vector spaces. These are mathematical structures that are based on complex numbers. We learned all that we need about such numbers in Chapter 1. Armed with this knowledge, we can now tackle complex vector spaces themselves.

Section 2.1 goes through the main example of a (finite-dimensional) complex vector space at tutorial pace. Section 2.2 provides formal definitions, basic properties, and more examples. Each of Section 2.3 through Section 2.7 discusses an advanced topic.

....

Reader Tip. The reader might find some of this chapter to be "just boring math." If you are eager to leap into the quantum world, we suggest reading the first two or three sections before moving on to Chapter 3. Return to Chapter 2 as a reference when needed (using the index and the table of contents to find specific topics). ♡

....

[1] *... La filosofia é scritta in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo), ma non si puo intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, ne' quali é scritto. Egli é scritto in lingua matematica, e i caratteri sono triangoli, cerchi, ed altre figure geometriche, senza i quali mezi e impossibile a intenderne umanamente parola; senza questi e un aggirarsi vanamente per un'oscuro laberinto...* (Opere Il Saggiatore p. 171).

A small disclaimer is in order. The theory of complex vector spaces is a vast and beautiful subject. Lengthy textbooks have been written on this important area of mathematics. It is impossible to provide anything more than a small glimpse into the beauty and profundity of this topic in one chapter. Rather than "teaching" our reader complex vector spaces, we aim to cover the bare minimum of concepts, terminology, and notation needed in order to start quantum computing. It is our sincere hope that reading this chapter will inspire further investigation into this remarkable subject.

2.1 $\mathbb{C}^n$ AS THE PRIMARY EXAMPLE

The primary example of a complex vector space is the set of vectors (one-dimensional arrays) of a fixed length with complex entries. These vectors will describe the states of quantum systems and quantum computers. In order to fix our ideas and to see clearly what type of structure this set has, let us carefully examine one concrete example: the set of vectors of length 4. We shall denote this set as $\mathbb{C}^4 = \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times \mathbb{C}$, which reminds us that each vector is an ordered list of four complex numbers.

A typical element of $\mathbb{C}^4$ looks like this:

$$\begin{bmatrix} 6-4i \\ 7+3i \\ 4.2-8.1i \\ -3i \end{bmatrix}. \tag{2.1}$$

We might call this vector V. We denote the jth element of V as $V[j]$. The top row is row number 0 (not 1);[2] hence, $V[1] = 7 + 3i$.

What types of operations can we carry out with such vectors? One operation that seems obvious is to form the **addition** of two vectors. For example, given two vectors of $\mathbb{C}^4$

$$V = \begin{bmatrix} 6-4i \\ 7+3i \\ 4.2-8.1i \\ -3i \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} 16+2.3i \\ -7i \\ 6 \\ -4i \end{bmatrix}, \tag{2.2}$$

[2] Computer scientists generally start indexing their rows and columns at 0. In contrast, mathematicians and physicists tend to start indexing at 1. The difference is irrelevant. We shall generally follow the computer science convention (after all, this is a computer science text).

we can add them to form $V + W \in \mathbb{C}^4$ by adding their respective entries:

$$\begin{bmatrix} 6-4i \\ 7+3i \\ 4.2-8.1i \\ -3i \end{bmatrix} + \begin{bmatrix} 16+2.3i \\ -7i \\ 6 \\ -4i \end{bmatrix} = \begin{bmatrix} (6-4i)+(16+2.3i) \\ (7+3i)+(-7i) \\ (4.2-8.1i)+(6) \\ (-3i)+(-4i) \end{bmatrix} = \begin{bmatrix} 22-1.7i \\ 7-4i \\ 10.2-8.1i \\ -7i \end{bmatrix}. \tag{2.3}$$

Formally, this operation amounts to

$$(V+W)[j] = V[j] + W[j]. \tag{2.4}$$

Exercise 2.1.1 Add the following two vectors:

$$\begin{bmatrix} 5+13i \\ 6+2i \\ 0.53-6i \\ 12 \end{bmatrix} + \begin{bmatrix} 7-8i \\ 4i \\ 2 \\ 9.4+3i \end{bmatrix}. \tag{2.5}$$

■

The addition operation satisfies certain properties. For example, because the addition of complex numbers is commutative, addition of complex vectors is also **commutative**:

$$V+W = \begin{bmatrix} (6-4i)+(16+2.3i) \\ (7+3i)+(-7i) \\ (4.2-8.1i)+(6) \\ (-3i)+(-4i) \end{bmatrix} = \begin{bmatrix} 22-1.7i \\ 7-4i \\ 10.2-8.1i \\ -7i \end{bmatrix}$$

$$= \begin{bmatrix} (16+2.3i)+(6-4i) \\ (-7i)+(7+3i) \\ (6)+(4.2-8.1i) \\ (-4i)+(-3i) \end{bmatrix} = W+V. \tag{2.6}$$

Similarly, addition of complex vectors is also **associative**, i.e., given three vectors V, W, and X, we may add them as $(V+W)+X$ or as $V+(W+X)$. Associativity states that the resulting sums are the same:

$$(V+W)+X = V+(W+X). \tag{2.7}$$

Exercise 2.1.2 Formally prove the associativity property. ■

There is also a distinguished vector called **zero**:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{2.8}$$

which satisfies the following property: for all vectors $V \in \mathbb{C}^4$, we have

$$V + \mathbf{0} = V = \mathbf{0} + V. \tag{2.9}$$

Formally, $\mathbf{0}$ is defined as $\mathbf{0}[j] = 0$.

Every vector also has an **(additive) inverse** (or **negative**). Consider

$$V = \begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix}. \tag{2.10}$$

There exists in $\mathbb{C}^4$ another vector

$$-V = \begin{bmatrix} -6 + 4i \\ -7 - 3i \\ -4.2 + 8.1i \\ 3i \end{bmatrix} \in \mathbb{C}^4 \tag{2.11}$$

such that

$$V + (-V) = \begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix} + \begin{bmatrix} -6 + 4i \\ -7 - 3i \\ -4.2 + 8.1i \\ 3i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}. \tag{2.12}$$

In general, for every vector $W \in \mathbb{C}^4$, there exists a vector $-W \in C^4$ such that $W + (-W) = (-W) + W = \mathbf{0}$. $-W$ is called the **inverse** of W. Formally,

$$(-W)[j] = -(W[j]). \tag{2.13}$$

The set $\mathbb{C}^4$ with the addition, inverse operations, and zero such that the addition is associative and commutative, form something called an **Abelian group**.

What other structure does our set $\mathbb{C}^4$ have? Take an arbitrary complex number, say, $c = 3 + 2i$. Call this number a **scalar**. Take a vector

$$V = \begin{bmatrix} 6+3i \\ 0+0i \\ 5+1i \\ 4 \end{bmatrix}. \tag{2.14}$$

We can **multiply an element by a scalar** by multiplying the scalar with each entry of the vector; i.e.,

$$(3+2i) \cdot \begin{bmatrix} 6+3i \\ 0+0i \\ 5+1i \\ 4 \end{bmatrix} = \begin{bmatrix} 12+21i \\ 0+0i \\ 13+13i \\ 12+8i \end{bmatrix}. \tag{2.15}$$

Formally, for a complex number c and a vector V, we form $c \cdot V$, which is defined as

$$(c \cdot V)[j] = c \times V[j], \tag{2.16}$$

where the $\times$ is complex multiplication. We shall omit the $\cdot$ when the scalar multiplication is understood.

Exercise 2.1.3 Scalar multiply $8 - 2i$ with $\begin{bmatrix} 16+2.3i \\ -7i \\ 6 \\ 5-4i \end{bmatrix}$. ■

Scalar multiplication satisfies the following properties: for all $c, c_1, c_2 \in \mathbb{C}$ and for all $V,\ W \in \mathbb{C}^4$,

- $1 \cdot V = V$,
- $c_1 \cdot (c_2 \cdot V) = (c_1 \times c_2) \cdot V$,
- $c \cdot (V + W) = c \cdot V + c \cdot W$,
- $(c_1 + c_2) \cdot V = c_1 \cdot V + c_2 \cdot V$.

Exercise 2.1.4 Formally prove that $(c_1 + c_2) \cdot V = c_1 \cdot V + c_2 \cdot V$. ■

An Abelian group with a scalar multiplication that satisfies these properties is called a **complex vector space**.

Notice that we have been working with vectors of size 4. However, everything that we have stated about vectors of size 4 is also true for vectors of arbitrary size. So the set $\mathbb{C}^n$ for a fixed but arbitrary n also has the structure of a complex vector space. In fact, these vector spaces will be the primary examples we will be working with for the rest of the book.

Programming Drill 2.1.1 *Write three functions that perform the addition, inverse, and scalar multiplication operations for* $\mathbb{C}^n$*, i.e., write a function that accepts the appropriate input for each of the operations and outputs the vector.*

2.2 DEFINITIONS, PROPERTIES, AND EXAMPLES

There are many other examples of complex vector spaces. We shall need to broaden our horizon and present a formal definition of a complex vector space.

Definition 2.2.1 *A* **complex vector space** *is a nonempty set* $\mathbb{V}$*, whose elements we shall call vectors, with three operations*

- *Addition:* $+ : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{V}$
- *Negation:* $- : \mathbb{V} \longrightarrow \mathbb{V}$
- *Scalar multiplication:* $\cdot : \mathbb{C} \times \mathbb{V} \longrightarrow \mathbb{V}$

and a distinguished element called the **zero vector** $\mathbf{0} \in \mathbb{V}$ *in the set. These operations and zero must satisfy the following properties: for all* $V, W, X \in \mathbb{V}$ *and for all* $c, c_1, c_2 \in \mathbb{C}$*,*

(i) *Commutativity of addition:* $V + W = W + V$,
(ii) *Associativity of addition:* $(V + W) + X = V + (W + X)$,
(iii) *Zero is an additive identity:* $V + \mathbf{0} = V = \mathbf{0} + V$,
(iv) *Every vector has an inverse:* $V + (-V) = \mathbf{0} = (-V) + V$,
(v) *Scalar multiplication has a unit:* $1 \cdot V = V$,
(vi) *Scalar multiplication respects complex multiplication:*

$$c_1 \cdot (c_2 \cdot V) = (c_1 \times c_2) \cdot V, \tag{2.17}$$

(vii) *Scalar multiplication distributes over addition:*

$$c \cdot (V + W) = c \cdot V + c \cdot W, \tag{2.18}$$

(viii) *Scalar multiplication distributes over complex addition:*

$$(c_1 + c_2) \cdot V = c_1 \cdot V + c_2 \cdot V. \tag{2.19}$$

To recap, any set that has an addition operation, an inverse operation, and a zero element that satisfies Properties (i), (ii), (iii), and (iv) is called an **Abelian group**. If, furthermore, there is a scalar multiplication operation that satisfies all the properties, then the set with the operations is called a **complex vector space**.

Although our main concern is complex vector spaces, we can gain much intuition from real vector spaces.

Definition 2.2.2 *A* **real vector space** *is a nonempty set* $\mathbb{V}$ *(whose elements we shall call vectors), along with an addition operation and a negation operation. Most important, there is a scalar multiplication that uses* $\mathbb{R}$ *and not* $\mathbb{C}$*, i.e.,*

$$\cdot : \mathbb{R} \times \mathbb{V} \longrightarrow \mathbb{V}. \tag{2.20}$$

This set and these operations must satisfy the analogous properties of a complex vector space.

The scalar multiplication of A with a complex number $c \in \mathbb{C}$ is

$$(c \cdot A)[j,k] = c \times A[j,k]. \tag{2.29}$$

Exercise 2.2.3 Let $c_1 = 2i$, $c_2 = 1 + 2i$, and $A = \begin{bmatrix} 1-i & 3 \\ 2+2i & 4+i \end{bmatrix}$. Verify Properties (vi) and (viii) in showing $\mathbb{C}^{2\times 2}$ is a complex vector space. ■

Exercise 2.2.4 Show that these operations on $\mathbb{C}^{m\times n}$ satisfy Properties (v), (vi), and (viii) of being a complex vector space. ■

Programming Drill 2.2.1 *Convert your functions from the last programming drill so that instead of accepting elements of $\mathbb{C}^n$, they accept elements of $\mathbb{C}^{m\times n}$.*

When $n = 1$, the matrices $\mathbb{C}^{m\times n} = \mathbb{C}^{m\times 1} = \mathbb{C}^m$, which we dealt with in Section 2.1. Thus, we can think of vectors as special types of matrices.

When $m = n$, the vector space $\mathbb{C}^{n\times n}$ has more operations and more structure than just a complex vector space. Here are three operations that one can perform on an $A \in \mathbb{C}^{n\times n}$:

- The **transpose** of A, denoted A^T, is defined as

$$A^T[j,k] = A[k,j]. \tag{2.30}$$

- The **conjugate** of A, denoted $\overline{A}$, is the matrix in which each element is the complex conjugate of the corresponding element of the original matrix,[3] i.e., $\overline{A}[j,k] = \overline{A[j,k]}$.
- The transpose operation and the conjugate operation are combined to form the **adjoint** or **dagger** operation. The adjoint of A, denoted as $A^\dagger$, is defined as $A^\dagger = (\overline{A})^T = \overline{(A^T)}$ or $A^\dagger[j,k] = \overline{A[k,j]}$.

Exercise 2.2.5 Find the transpose, conjugate, and adjoint of

$$\begin{bmatrix} 6-3i & 2+12i & -19i \\ 0 & 5+2.1i & 17 \\ 1 & 2+5i & 3-4.5i \end{bmatrix}. \tag{2.31}$$

■

These three operations are defined even when $m \neq n$. The transpose and adjoint are both functions from $\mathbb{C}^{m\times n}$ to $\mathbb{C}^{n\times m}$.

These operations satisfy the following properties for all $c \in \mathbb{C}$ and for all A, $B \in \mathbb{C}^{m\times n}$:

(i) Transpose is idempotent: $(A^T)^T = A$.
(ii) Transpose respects addition: $(A + B)^T = A^T + B^T$.
(iii) Transpose respects scalar multiplication: $(c \cdot A)^T = c \cdot A^T$.

[3] This notation is overloaded. It is an operation on complex numbers and complex matrices.

(iv) Conjugate is idempotent: $\overline{\overline{A}} = A$.
(v) Conjugate respects addition: $\overline{A + B} = \overline{A} + \overline{B}$.
(vi) Conjugate respects scalar multiplication: $\overline{c \cdot A} = \overline{c} \cdot \overline{A}$.
(vii) Adjoint is idempotent: $(A^\dagger)^\dagger = A$.
(viii) Adjoint respects addition: $(A + B)^\dagger = A^\dagger + B^\dagger$.
(ix) Adjoint relates to scalar multiplication: $(c \cdot A)^\dagger = \overline{c} \cdot A^\dagger$.

Exercise 2.2.6 Prove that conjugation respects scalar multiplication, i.e., $\overline{c \cdot A} = \overline{c} \cdot \overline{A}$. ■

Exercise 2.2.7 Prove Properties (vii), (viii), and (ix) using Properties (i) – (vi). ■

The transpose shall be used often in the text to save space. Rather than writing

$$\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} \tag{2.32}$$

which requires more space, we write $[c_0, c_1, \cdots, c_{n-1}]^T$.

When $m = n$, there is another binary operation that is used: **matrix multiplication**. Consider the following two 3-by-3 matrices:

$$A = \begin{bmatrix} 3+2i & 0 & 5-6i \\ 1 & 4+2i & i \\ 4-i & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 2-i & 6-4i \\ 0 & 4+5i & 2 \\ 7-4i & 2+7i & 0 \end{bmatrix}. \tag{2.33}$$

We form the matrix product of A and B, denoted $A \star B$. $A \star B$ will also be a 3-by-3 matrix. $(A \star B)[0, 0]$ will be found by multiplying each element of the 0th row of A with the corresponding element of the 0th column of B. We then sum the results:

$$\begin{aligned}(A \star B)[0, 0] &= ((3+2i) \times 5) + (0 \times 0) + ((5-6i) \times (7-4i)) \\ &= (15+10i) + (0) + (11-62i) = 26 - 52i.\end{aligned} \tag{2.34}$$

The $(A \star B)[j, k]$ entry can be found by multiplying each element of $A[j, -]$ with the appropriate element of $B[-, k]$ and summing the results. So,

$$(A \star B) = \begin{bmatrix} 26-52i & 60+24i & 26 \\ 9+7i & 1+29i & 14 \\ 48-21i & 15+22i & 20-22i \end{bmatrix}. \tag{2.35}$$

Exercise 2.2.8 Find $B \star A$. Does it equal $A \star B$? ■

Matrix multiplication is defined in a more general setting. The matrices do not have to be square. Rather, the number of columns in the first matrix must be the

Exercise 2.2.15 Show that $Func(\mathbb{N}, \mathbb{C})$ with these operations forms a complex vector space. ■

Example 2.2.12 We can generalize $Func(\mathbb{N}, \mathbb{C})$ to other sets of functions. For any $a < b$ in $\mathbb{R}$, the set of functions from the interval $[a, b] \subseteq \mathbb{R}$ to $\mathbb{C}$ denoted $Func([a, b], \mathbb{C})$ is a complex vector space. □

Exercise 2.2.16 Show that $Func(\mathbb{N}, \mathbb{R})$ and $Func([a, b], \mathbb{R})$ are real vector spaces. ■

Example 2.2.13 There are several ways of constructing new vector spaces from existing ones. Here we see one method and Section 2.7 describes another. Let $(\mathbb{V}, +, -, \mathbf{0}, \cdot)$ and $(\mathbb{V}', +', -', \mathbf{0}', \cdot')$ be two complex vector spaces. We construct a new complex vector space $(\mathbb{V} \times \mathbb{V}', +'', -'', \mathbf{0}'', \cdot'')$ called the **Cartesian product**[5] or the **direct sum** of $\mathbb{V}$ and $\mathbb{V}'$. The vectors are ordered pairs of vectors $(V, V') \in \mathbb{V} \times \mathbb{V}'$. Operations are performed pointwise:

$$(V_1, V_1') +'' (V_2, V_2') = (V_1 + V_2, V_1' +' V_2'), \tag{2.56}$$

$$-''(V, V') = (-V, -'V'), \tag{2.57}$$

$$\mathbf{0}'' = (\mathbf{0}, \mathbf{0}'), \tag{2.58}$$

$$c \cdot'' (V, V') = (c \cdot V, c \cdot' V'). \tag{2.59}$$

□

Exercise 2.2.17 Show that $\mathbb{C}^m \times \mathbb{C}^n$ is isomorphic to $\mathbb{C}^{m+n}$. ■

Exercise 2.2.18 Show that $\mathbb{C}^m$ and $\mathbb{C}^n$ are each a complex subspace of $\mathbb{C}^m \times \mathbb{C}^n$. ■

2.3 BASIS AND DIMENSION

A basis of a vector space is a set of vectors of that vector space that is special in the sense that all other vectors can be uniquely written in terms of these basis vectors.

Definition 2.3.1 *Let $\mathbb{V}$ be a complex (real) vector space. $V \in \mathbb{V}$ is a* **linear combination** *of the vectors $V_0, V_1, \ldots, V_{n-1}$ in $\mathbb{V}$ if V can be written as*

$$V = c_0 \cdot V_0 + c_1 \cdot V_1 + \cdots + c_{n-1} \cdot V_{n-1} \tag{2.60}$$

for some $c_0, c_1, \ldots, c_{n-1}$ in $\mathbb{C}$ ($\mathbb{R}$).

Let us return to $\mathbb{R}^3$ for examples.

[5] A note to the meticulous reader: Although we used $\times$ for the product of two complex numbers, here we use it for the Cartesian product of sets and the Cartesian product of vector spaces. We feel it is better to overload known symbols than to introduce a plethora of new ones.

Example 2.3.1 As

$$3\begin{bmatrix}5\\-2\\3\end{bmatrix}+5\begin{bmatrix}0\\1\\4\end{bmatrix}-4\begin{bmatrix}-6\\1\\0\end{bmatrix}+2.1\begin{bmatrix}3\\1\\1\end{bmatrix}=\begin{bmatrix}45.3\\-2.9\\31.1\end{bmatrix}, \tag{2.61}$$

we say that

$$[45.3, -2.9, 31.1]^T \tag{2.62}$$

is a linear combination of

$$\begin{bmatrix}5\\-2\\3\end{bmatrix}, \begin{bmatrix}0\\1\\4\end{bmatrix}, \begin{bmatrix}-6\\1\\0\end{bmatrix}, \quad \text{and} \begin{bmatrix}3\\1\\1\end{bmatrix}. \tag{2.63}$$

□

Definition 2.3.2 *A set* $\{V_0, V_1, \ldots, V_{n-1}\}$ *of vectors in* $\mathbb{V}$ *is called* **linearly independent** *if*

$$\mathbf{0} = c_0 \cdot V_0 + c_1 \cdot V_1 + \cdots + c_{n-1} \cdot V_{n-1} \tag{2.64}$$

implies that $c_0 = c_1 = \cdots = c_{n-1} = 0$. *This means that the only way that a linear combination of the vectors can be the zero vector is if all the* c_j *are zero.*

It can be shown that this definition is equivalent to saying that for any nonzero $V \in \mathbb{V}$, there are *unique* coefficients $c_0, c_1, \ldots, c_{n-1}$ in $\mathbb{C}$ such that

$$V = c_0 \cdot V_0 + c_1 \cdot V_1 + \cdots + c_{n-1} \cdot V_{n-1}. \tag{2.65}$$

The set of vectors are called linearly independent because each of the vectors in the set $\{V_0, V_1, \ldots, V_{n-1}\}$ cannot be written as a combination of the others in the set.

Example 2.3.2 The set of vectors

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\} \tag{2.66}$$

is linearly independent because the only way that

$$\mathbf{0} = \begin{bmatrix}0\\0\\0\end{bmatrix} = x\begin{bmatrix}1\\1\\1\end{bmatrix} + y\begin{bmatrix}0\\1\\1\end{bmatrix} + z\begin{bmatrix}0\\0\\1\end{bmatrix} \tag{2.67}$$

can occur is if $0 = x$, $0 = x + y$, and $0 = x + y + z$. By substitution, we see that $x = y = z = 0$. □

Example 2.3.3 The set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} \right\} \tag{2.68}$$

is not linearly independent (called **linearly dependent**) because

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + z \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} \tag{2.69}$$

can happen when $x = 2$, $y = -3$, and $z = -1$. □

Exercise 2.3.1 Show that the set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -4 \\ -4 \end{bmatrix} \right\} \tag{2.70}$$

is not linearly independent. ■

Definition 2.3.3 *A set $\mathcal{B} = \{V_0, V_1, \ldots, V_{n-1}\} \subseteq \mathbb{V}$ of vectors is called a* **basis** *of a (complex) vector space $\mathbb{V}$ if both*

(i) *every, $V \in \mathbb{V}$ can be written as a linear combination of vectors from $\mathcal{B}$ and*
(ii) *$\mathcal{B}$ is linearly independent.*

Example 2.3.4 $\mathbb{R}^3$ has a basis

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \tag{2.71}$$

□

Exercise 2.3.2 Verify that the preceding three vectors are in fact a basis of $\mathbb{R}^3$. ■

There may be many sets that each form a basis of a particular vector space but there is also a basis that is easier to work with called the **canonical basis** or the **standard basis**. Many of the examples that we will deal with have canonical basis. Let us look at some examples of canonical basis.

- $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \tag{2.72}$$

- $\mathbb{C}^n$ (and $\mathbb{R}^n$):

$$E_0 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, E_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \ldots, E_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{2.73}$$

Every vector $[c_0, c_1, \ldots, c_{n-1}]^T$ can be written as

$$\sum_{j=0}^{n-1} (c_j \cdot E_j). \tag{2.74}$$

- $\mathbb{C}^{m \times n}$: The canonical basis for this vector space consists of matrices of the form

$$E_{j,k} = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \vdots \\ \boldsymbol{j} \\ \vdots \\ \boldsymbol{m-1} \end{array} \begin{array}{c} \begin{array}{cccccc} \mathbf{0} & \mathbf{1} & \cdots & \boldsymbol{k} & \cdots & \boldsymbol{n-1} \end{array} \\ \begin{bmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \end{array}, \tag{2.75}$$

where $E_{j,k}$ has a 1 in row j, column k, and 0's everywhere else. There is an $E_{j,k}$ for $j = 0, 1, \ldots, m-1$ and $k = 0, 1, \ldots, n-1$. It is not hard to see that for every m-by-n matrix, A can be written as the sum:

$$A = \sum_{j=0}^{m-1} \sum_{k=0}^{n-1} A[j,k] \cdot E_{j,k}. \tag{2.76}$$

- $Poly_n$: The canonical basis is formed by the following set of monomials:

$$1, x, x^2, \ldots, x^n. \tag{2.77}$$

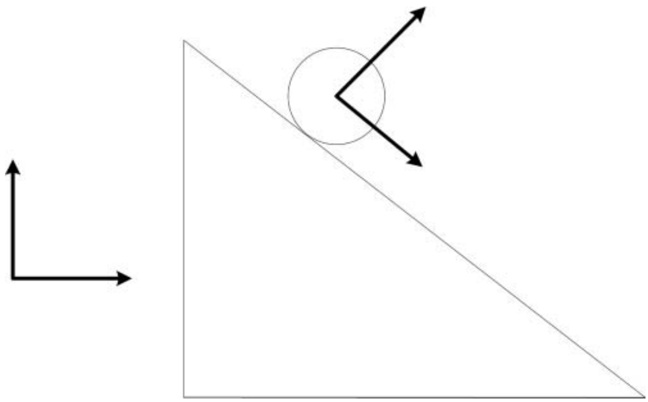

Figure 2.7. A ball rolling down a ramp and the two relevant bases.

completed the calculations, we change our results into the more understandable canonical basis and produce the desired answer. We might envision this as the flowchart shown in Figure 2.8.

Throughout this text, we shall go from one basis to another basis, perform some calculations, and finally revert to the original basis. The Hadamard matrix will frequently be the means by which we change the basis.

2.4 INNER PRODUCTS AND HILBERT SPACES

We will be interested in complex vector spaces with additional structure. Recall that a state of a quantum system corresponds to a vector in a complex vector space. A need will arise to compare different states of the system; hence, there is a need to compare corresponding vectors or measure one vector against another in a vector space.

Consider the following operation that we can perform with two vectors in $\mathbb{R}^3$:

$$\left\langle \begin{bmatrix} 5 \\ 3 \\ -7 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix} \right\rangle = [5, 3, -7] \star \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix} = (5 \times 6) + (3 \times 2) + (-7 \times 0) = 36. \tag{2.96}$$

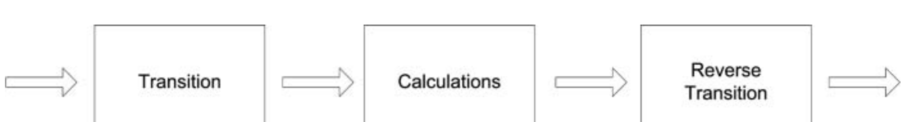

Figure 2.8. Problem-solving flowchart.

In general, for any two vectors $V_1 = [r_0, r_1, r_2]^T$ and $V_2 = [r'_0, r'_1, r'_2]^T$ in $\mathbb{R}^3$, we can form a real number by performing the following operation:

$$\langle V_1, V_2 \rangle = V_1^T \star V_2 = \sum_{j=0}^{2} r_j r'_j. \tag{2.97}$$

This is an example of an inner product of two vectors. An inner product in a complex (real) vector space is a binary operation that accepts two vectors as inputs and outputs a complex (real) number. This operation must satisfy certain properties spelled out in the following:

Definition 2.4.1 *An* **inner product** *(also called a* **dot product** *or* **scalar product***) on a complex vector space* $\mathbb{V}$ *is a function*

$$\langle -, - \rangle : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{C} \tag{2.98}$$

that satisfies the following conditions for all V, V_1, V_2, *and* V_3 *in* $\mathbb{V}$ *and for a* $c \in \mathbb{C}$*:*

(i) *Nondegenerate:*

$$\langle V, V \rangle \geq 0, \tag{2.99}$$

$$\langle V, V \rangle = 0 \textit{ if and only if } V = \mathbf{0} \tag{2.100}$$

(i.e., the only time it "degenerates" is when it is 0*).*

(ii) *Respects addition:*

$$\langle V_1 + V_2, V_3 \rangle = \langle V_1, V_3 \rangle + \langle V_2, V_3 \rangle, \tag{2.101}$$

$$\langle V_1, V_2 + V_3 \rangle = \langle V_1, V_2 \rangle + \langle V_1, V_3 \rangle. \tag{2.102}$$

(iii) *Respects scalar multiplication:*

$$\langle c \cdot V_1, V_2 \rangle = c \times \langle V_1, V_2 \rangle, \tag{2.103}$$

$$\langle V_1, c \cdot V_2 \rangle = \overline{c} \times \langle V_1, V_2 \rangle. \tag{2.104}$$

(iv) *Skew symmetric:*

$$\langle V_1, V_2 \rangle = \overline{\langle V_2, V_1 \rangle}. \tag{2.105}$$

An inner product on real vector space $\langle\ ,\ \rangle : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{R}$ must satisfy the same properties. Because any $r \in \mathbb{R}$ satisfies $\overline{r} = r$, Properties (iii) and (iv) are simpler for a real vector space.

Definition 2.4.2 *A* **(complex) inner product space** *is a (complex) vector space along with an inner product.*

Let us list some examples of inner product spaces.

- $\mathbb{R}^n$: The inner product is given as

$$\langle V_1, V_2 \rangle = V_1^T \star V_2. \tag{2.106}$$

- $\mathbb{C}^n$: The inner product is given as

$$\langle V_1, V_2 \rangle = V_1^\dagger \star V_2. \tag{2.107}$$

- $\mathbb{R}^{n\times n}$ has an inner product given for matrices $A, B \in \mathbb{R}^{n\times n}$ as

$$\langle A, B \rangle = Trace(A^T \star B), \tag{2.108}$$

where the **trace** of a square matrix C is given as the sum of the diagonal elements. That is,

$$Trace(C) = \sum_{i=0}^{n-1} C[i, i]. \tag{2.109}$$

- $\mathbb{C}^{n\times n}$ has an inner product given for matrices $A, B \in \mathbb{C}^{n\times n}$ as

$$\langle A, B \rangle = Trace(A^\dagger \star B). \tag{2.110}$$

- $Func(\mathbb{N}, \mathbb{C})$:

$$\langle f, g \rangle = \sum_{j=0}^{\infty} \overline{f(j)} g(j). \tag{2.111}$$

- $Func([a, b], \mathbb{C})$:

$$\langle f, g \rangle = \int_a^b \overline{f(t)} g(t)\, dt. \tag{2.112}$$

Exercise 2.4.1 Let $V_1 = [2, 1, 3]^T$, $V_2 = [6, 2, 4]^T$, and $V_3 = [0, -1, 2]^T$. Show that the inner product in $\mathbb{R}^3$ respects the addition, i.e., Equations (2.101) and (2.102). ■

Exercise 2.4.2 Show that the function $\langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ given in Equation (2.106) satisfies all the properties of being an inner product on $\mathbb{R}^n$. ■

Exercise 2.4.3 Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$. Show that the inner product in $\mathbb{R}^{2\times 2}$ respects addition (Equations (2.101) and (2.102)) with these matrices. ■

Exercise 2.4.4 Show that the function given for pairs of real matrices satisfies the inner product properties and converts the real vector space $\mathbb{R}^{n\times n}$ to a real inner product space. ■

Programming Drill 2.4.1 *Write a function that accepts two complex vectors of length n and calculates their inner product.*

The inner product of a complex vector with itself is a real number. We can observe this from the property that for all V_1, V_2, an inner product must satisfy

$$\langle V_1, V_2 \rangle = \overline{\langle V_2, V_1 \rangle}. \tag{2.113}$$

It follows that if $V_2 = V_1$, then we have

$$\langle V_1, V_1\rangle = \overline{\langle V_1, V_1\rangle}; \tag{2.114}$$

hence it is real.

Definition 2.4.3 *For every complex inner product space* $\mathbb{V}$, $\langle -, -\rangle$, *we can define a* **norm** *or* **length** *which is a function*

$$|\ \ | : \mathbb{V} \longrightarrow \mathbb{R} \tag{2.115}$$

defined as $|V| = \sqrt{\langle V, V\rangle}$.

Example 2.4.1 In $\mathbb{R}^3$, the norm of vector $[3, -6, 2]^T$ is

$$\left|\begin{bmatrix} 3 \\ -6 \\ 2 \end{bmatrix}\right| = \sqrt{\left\langle \begin{bmatrix} 3 \\ -6 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ 2 \end{bmatrix} \right\rangle} = \sqrt{3^2 + (-6)^2 + 2^2} = \sqrt{49} = 7. \tag{2.116}$$

□

Exercise 2.4.5 Calculate the norm of $[4 + 3i, 6 - 4i, 12 - 7i, 13i]^T$. ■

Exercise 2.4.6 Let $A = \begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix} \in \mathbb{R}^{2\times 2}$. Calculate the norm $|A| = \sqrt{\langle A, A\rangle}$. ■

In general, the norm of the vector $[x, y, z]^T$ is

$$\left|\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right| = \sqrt{\left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\rangle} = \sqrt{x^2 + y^2 + z^2}. \tag{2.117}$$

This is the Pythagorean formula for the length of a vector. The intuition one should have is that the norm of a vector in any vector space is the length of the vector.

From the properties of an inner product space, it follows that a norm has the following properties for all $V, W \in \mathbb{V}$ and $c \in \mathbb{C}$:

(i) Norm is nondegenerate: $|V| > 0$ if $V \neq \mathbf{0}$ and $|\mathbf{0}| = 0$.
(ii) Norm satisfies the **triangle inequality**: $|V + W| \leq |V| + |W|$.
(iii) Norm respects scalar multiplication: $|c \cdot V| = |c| \times |V|$.

Programming Drill 2.4.2 *Write a function that calculates the norm of a given complex vector.*

Given a norm, we can proceed and define a distance function.

Definition 2.4.4 *For every complex inner product space* $(\mathbb{V}, \langle\ ,\ \rangle)$, *we can define a* **distance function**

$$d(\ ,\ ) : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{R}, \tag{2.118}$$

where

$$d(V_1, V_2) = |V_1 - V_2| = \sqrt{\langle V_1 - V_2, V_1 - V_2 \rangle}. \tag{2.119}$$

Exercise 2.4.7 Let $V_1 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ and $V_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$. Calculate the distance between these two vectors. ■

The intuition is that $d(V_1, V_2)$ is the distance from the end of vector V_1 to the end of vector V_2. From the properties of an inner product space, it is not hard to show that a distance function has the following properties for all $U, V, W \in \mathbb{V}$:

(i) Distance is nondegenerate: $d(V, W) > 0$ if $V \neq W$ and $d(V, V) = 0$.
(ii) Distance satisfies the **triangle inequality**: $d(U, V) \leq d(U, W) + d(W, V)$.
(iii) Distance is symmetric: $d(V, W) = d(W, V)$.

Programming Drill 2.4.3 *Write a function that calculates the distance of two given complex vectors.*

Definition 2.4.5 *Two vectors* V_1 *and* V_2 *in an inner product space* $\mathbb{V}$ *are* **orthogonal** *if* $\langle V_1, V_2 \rangle = 0$.

The picture to keep in mind is that two vectors are orthogonal if they are perpendicular to each other.

Definition 2.4.6 *A basis* $\mathcal{B} = \{V_0, V_1, \ldots, V_{n-1}\}$ *for an inner product space* $\mathbb{V}$ *is called an* **orthogonal basis** *if the vectors are pairwise orthogonal to each other, i.e.,* $j \neq k$ *implies* $\langle V_j, V_k \rangle = 0$. *An orthogonal basis is called an* **orthonormal basis** *if every vector in the basis is of norm 1, i.e.,*

$$\langle V_j, V_k \rangle = \delta_{j,k} = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases} \tag{2.120}$$

$\delta_{j,k}$ *is called the* **Kronecker delta function**.

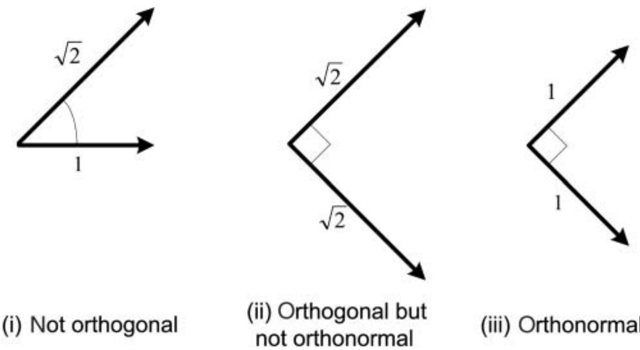

Figure 2.9. Three bases for $\mathbb{R}^2$.

Example 2.4.2 Consider the three bases for $\mathbb{R}^2$ shown in Figure 2.9.
Formally, these bases are

(i) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix},$

(ii) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix},$

(iii) $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$

□

In $\mathbb{R}^3$, the standard inner product $\langle V, V' \rangle = V^T V'$ can be shown to be equivalent to

$$\langle V, V' \rangle = |V||V'| \cos\theta, \tag{2.121}$$

where θ is the angle between V and V'. When $|V'| = 1$, this equation reduces to

$$\langle V, V' \rangle = |V| \cos\theta. \tag{2.122}$$

Exercise 2.4.8 Let $V = [3, -1, 0]^T$ and $V' = [2, -2, 1]^T$. Calculate the angle θ between these two vectors. ■

Elementary trigonometry teaches us that when $|V'| = 1$, the number $\langle V, V' \rangle$ is the length of the projection of V onto the direction of V' (Figure 2.10).

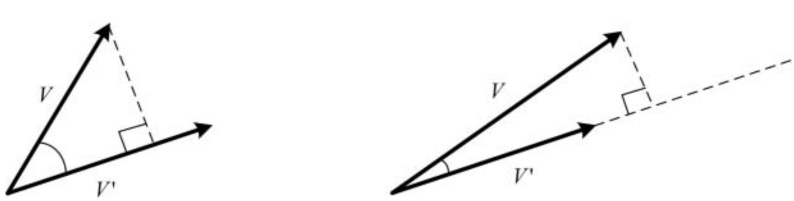

Figure 2.10. The projection of V onto V'.

Exercise 2.5.1 The following vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\} \tag{2.134}$$

are eigenvectors of the matrix

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}. \tag{2.135}$$

Find the eigenvalues. ■

If a matrix A has eigenvalue c_0 with eigenvector V_0, then for any $c \in \mathbb{C}$ we have

$$A(cV_0) = cAV_0 = cc_0V_0 = c_0(cV_0), \tag{2.136}$$

which shows that cV_0 is also an eigenvector of A with eigenvalue c_0. If cV_0 and $c'V_0$ are two such eigenvectors, then because

$$\begin{aligned} A(cV_0 + c'V_0) &= AcV_0 + Ac'V_0 = cAV_0 + c'AV_0 \\ &= c(c_0V_0) + c'(c_0V_0) = (c + c')(c_0V_0) = c_0(c + c')V_0, \end{aligned} \tag{2.137}$$

we see that the addition of two such eigenvectors is also an eigenvector. We conclude the following:

Proposition 2.5.1 Every eigenvector determines a complex subvector space of the vector space. This space is known as the **eigenspace** associated with the given eigenvector.

Some matrices have many eigenvalues and eigenvectors and some matrices have none.

2.6 HERMITIAN AND UNITARY MATRICES

We shall need certain types of important square matrices and their properties.

A matrix $A \in \mathbb{R}^{n\times n}$ is called **symmetric** if $A^T = A$. In other words, $A[j,k] = A[k,j]$. Let us generalize this notion from the real numbers to the complex numbers.

Definition 2.6.1 *An n-by-n matrix A is called* **hermitian** *if* $A^\dagger = A$. *In other words,* $A[j,k] = \overline{A[k,j]}$.

Definition 2.6.2 *If A is a hermitian matrix then the operator that it represents is called* **self-adjoint**.

Example 2.6.1 The matrix

$$\begin{bmatrix} 5 & 4+5i & 6-16i \\ 4-5i & 13 & 7 \\ 6+16i & 7 & -2.1 \end{bmatrix} \tag{2.138}$$

is hermitian. □

Exercise 2.6.1 Show that the matrix

$$\begin{bmatrix} 7 & 6+5i \\ 6-5i & -3 \end{bmatrix} \tag{2.139}$$

is hermitian. ■

Exercise 2.6.2 Show that A is hermitian if and only if $A^T = \overline{A}$. ■

Notice from the definition that the elements along the diagonal of a hermitian matrix must be real. The old notion of a symmetric matrix is a special case of hermitian that is limited to matrices with only real entries.

Proposition 2.6.1 If A is a hermitian n-by-n matrix, then for all $V, \ V' \in \mathbb{C}^n$ we have

$$\langle AV, V' \rangle = \langle V, AV' \rangle. \tag{2.140}$$

The proof is easy to see:

$$\langle AV, V' \rangle = (AV)^\dagger \star V' = V^\dagger A^\dagger V' = V^\dagger \star AV' = \langle V, AV' \rangle \tag{2.141}$$

where the first and the fourth equalities are from the definition of an inner product, the second equality is from the property of †, and the third equality is from the definition of a hermitian matrix.

Exercise 2.6.3 Prove the same proposition for symmetric real matrices. ■

Proposition 2.6.2 If A is a hermitian, then all eigenvalues are real.

To prove this, let A be a hermitian matrix with an eigenvalue $c \in \mathbb{C}$ and an eigenvector V. Consider the following sequence of equalities:

$$c\langle V, V \rangle = \langle cV, V \rangle = \langle AV, V \rangle = \langle V, AV \rangle = \langle V, cV \rangle = \overline{c}\langle V, V \rangle. \tag{2.142}$$

The first and fifth equalities are properties of the inner product. The second and fourth equalities are from the definition of eigenvalue. The third equality is from Proposition 2.6.1. Because c and V are nonzero, $c = \overline{c}$ and hence must be real.

Exercise 2.6.4 Prove that the eigenvalues of a symmetric matrix are real. ■

Proposition 2.6.3 For a given hermitian matrix, distinct eigenvectors that have distinct eigenvalues are orthogonal.

We prove this by looking at V_1 and V_2 that are distinct eigenvectors of a hermitian matrix A.

$$AV_1 = c_1 V_1 \quad \text{and} \quad AV_2 = c_2 V_2. \tag{2.143}$$

Then we have the following sequence of equalities:

$$\begin{aligned} c_1\langle V_1, V_2\rangle &= \langle c_1 V_1, V_2\rangle = \langle AV_1, V_2\rangle = \langle V_1, AV_2\rangle \\ &= \langle V_1, c_2 V_2\rangle = \overline{c_2}\langle V_1, V_2\rangle = c_2\langle V_1, V_2\rangle \end{aligned} \tag{2.144}$$

where the first and fifth equalities are from properties of inner products, the second and fourth equalities are by definition of eigenvector, the third equality follows from the fact that H is hermitian, and the last equality is from the fact that eigenvalues of hermitian matrices are real. As the left side is equal to the right side, we may subtract one from the other to get 0:

$$c_1\langle V_1, V_2\rangle - c_2\langle V_1, V_2\rangle = (c_1 - c_2)\langle V_1, V_2\rangle = 0. \tag{2.145}$$

Because c_1 and c_2 are distinct, $c_1 - c_2 \neq 0$. Hence, it follows that $\langle V_1, V_2\rangle = 0$ and they are orthogonal.

We shall need one more important proposition about self-adjoint operators.

Definition 2.6.3 *A* **diagonal matrix** *is a square matrix whose only nonzero entries are on the diagonal. All entries off the diagonal are zero.*

Proposition 2.6.4 (The Spectral Theorem for Finite-Dimensional Self-Adjoint Operators.) Every self-adjoint operator A on a finite-dimensional complex vector space $\mathbb{V}$ can be represented by a diagonal matrix whose diagonal entries are the eigenvalues of A, and whose eigenvectors form an orthonormal basis for $\mathbb{V}$ (we shall call this basis an **eigenbasis**).

Hermitian matrices and their eigenbases will play a major role in our story. We shall see in Chapter 4 that associated with every physical observable of a quantum system there is a corresponding hermitian matrix. Measurements of that observable always lead to a state that is represented by one of the eigenvectors of the associated hermitian matrix.

Programming Drill 2.6.1 *Write a function that accepts a square matrix and tells if it is hermitian.*

Another fundamental type of matrix is unitary. A matrix A is **invertible** if there exists a matrix A^{-1} such that

$$A \star A^{-1} = A^{-1} \star A = I_n. \tag{2.146}$$

Unitary matrices are a type of invertible matrix. They are invertible and their inverse is their adjoint. This fact ensures that unitary matrices "preserve the geometry" of the space on which it is acting.

Definition 2.6.4 *An n-by-n matrix U is* **unitary** *if*

$$U \star U^{\dagger} = U^{\dagger} \star U = I_n. \tag{2.147}$$

It is important to realize that not all invertible matrices are unitary.

Example 2.6.2 For any θ, the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.148}$$

is a unitary matrix. (You might have seen such a matrix when studying computer graphics. We shall see why in a few moments.) □

Exercise 2.6.5 Show that the matrix given in Equation (2.148) is unitary. ■

Example 2.6.3 The matrix

$$\begin{bmatrix} \frac{1+i}{2} & \frac{i}{\sqrt{3}} & \frac{3+i}{2\sqrt{15}} \\ \frac{-1}{2} & \frac{1}{\sqrt{3}} & \frac{4+3i}{2\sqrt{15}} \\ \frac{1}{2} & \frac{-i}{\sqrt{3}} & \frac{5i}{2\sqrt{15}} \end{bmatrix} \tag{2.149}$$

is a unitary matrix. □

Exercise 2.6.6 Show that the matrix given in Equation (2.149) is unitary. ■

Exercise 2.6.7 Show that if U and U' are unitary matrices, then so is $U \star U'$. (Hint: Use Equation (2.44)). ■

Proposition 2.6.5 Unitary matrices preserve inner products, i.e., if U is unitary, then for any $V, V' \in \mathbb{C}^n$, we have $\langle UV, UV'\rangle = \langle V, V'\rangle$.

This proposition is actually very easy to demonstrate:

$$\langle UV, UV'\rangle = (UV)^{\dagger} \star UV' = V^{\dagger}U^{\dagger} \star UV' = V^{\dagger} \star I \star V' = V^{\dagger} \star V' = \langle V, V'\rangle \tag{2.150}$$

where the first and fifth equalities are from the definition of the inner product, the second equality is from the properties of the adjoint, the third equality is from the definition of a unitary matrix, and the fourth equality is due to the face that I is the identity.

Because unitary matrices preserve inner products, they also preserve norms

$$|UV| = \sqrt{\langle UV, UV\rangle} = \sqrt{\langle V, V\rangle} = |V|. \tag{2.151}$$

In particular, if $|V| = 1$, then $|UV| = 1$. Consider the set of all vectors that have length 1. They form a ball around the origin (the zero of the vector space). We call this ball the **unit sphere** and imagine it as Figure 2.14.

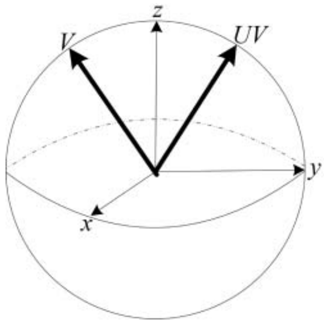

Figure 2.14. The unit sphere and the action of U on V.

If V is a vector on the unit sphere (in any dimension), then UV is also on the unit sphere. We shall see that a unitary matrix is a way of rotating the unit sphere.[7]

Exercise 2.6.8 Show that if U is a unitary matrix and V_1 and V_2 are in $\mathbb{C}^n$, then

$$d(UV_1, UV_2) = d(V_1, V_2), \tag{2.152}$$

i.e., U preserves distances. (An operator that preserves distances is called an **isometry**.) ■

What does unitary really mean? As we saw, it means that it preserves the geometry. But it also means something else: If U is unitary and $UV = V'$, then we can easily form $U^\dagger$ and multiply both sides of the equation by $U^\dagger$ to get $U^\dagger UV = U^\dagger V'$ or $V = U^\dagger V'$. In other words, because U is unitary, there is a related matrix that can "undo" the action that U performs. $U^\dagger$ takes the result of U's action and gets back the original vector. In the quantum world, all actions (that are not measurements) are "undoable" or "reversible" in such a manner.

Hermitian matrices and unitary matrices will be very important in our text. The Venn diagram shown in Figure 2.15 is helpful.

Exercise 2.6.9 Show that I_n and $-1 \cdot I_n$ are both hermitian and unitary. ■

Programming Drill 2.6.2 *Write a function that accepts a square matrix and tells if it is unitary.*

2.7 TENSOR PRODUCT OF VECTOR SPACES

At the conclusion of Section 2.2 we were introduced to the Cartesian product, which is one method of combining vector spaces. In this section, we study the tensor product, which is another, more important, method of combining vector spaces. If $\mathbb{V}$ describes one quantum system and $\mathbb{V}'$ describes another, then their tensor product describes both quantum systems as one. The tensor product is the fundamental building operation of quantum systems.

[7] These movements of the unit sphere are important in computer graphics.

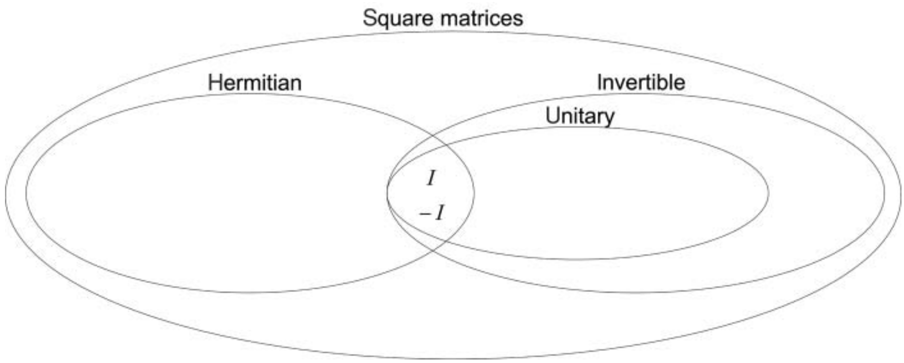

Figure 2.15. Types of matrices.

Reader Tip. A brief warning is in order. The tensor product of two vector spaces is perhaps one of the most difficult subjects in this chapter, as well as one of the most essential. Do not be intimidated if you do not understand it the first time you read it. Everyone has a hard time with tensor products. We also suggest that you read this section in conjunction with Sections 3.4 and 4.5. All these three sections deal with the tensor product from slightly different viewpoints. ♡

Given two vector spaces $\mathbb{V}$ and $\mathbb{V}'$, we shall form the **tensor product** of two vector spaces, and denote it $\mathbb{V} \otimes \mathbb{V}'$. The tensor product is generated by the set of "tensors" of all vectors:

$$\{V \otimes V' | V \in \mathbb{V} \text{ and } V' \in \mathbb{V}'\}, \tag{2.153}$$

where $\otimes$ is just a symbol. A typical element of $\mathbb{V} \otimes \mathbb{V}'$ looks like this:

$$c_0(V_0 \otimes V_0') + c_1(V_1 \otimes V_1') + \cdots + c_{p-1}(V_{p-1} \otimes V_{p-1}'), \tag{2.154}$$

where $V_0, V_1, \ldots, V_{p-1}$ are elements of $\mathbb{V}$ and $V_0', V_1', \ldots, V_{p-1}'$ are elements of $\mathbb{V}'$. We might write this as

$$\sum_{i=0}^{p-1} c_i(V_i \otimes V_i'). \tag{2.155}$$

The operations on this vector space are straightforward. For a given $\sum_{i=0}^{p-1} c_i(V_i \otimes V_i')$ and $\sum_{i=0}^{q-1} c_i'(W_i \otimes W_i')$, addition is simply the addition of summations, i.e.,

$$\sum_{i=0}^{p-1} c_i(V_i \otimes V_i') + \sum_{i=0}^{q-1} c_i'(W_i \otimes W_i'). \tag{2.156}$$

The scalar multiplication for a given $c \in \mathbb{C}$ is

$$c \cdot \sum_{i=0}^{p-1} c_i(V_i \otimes V_i') = \sum_{i=0}^{p-1}(c \times c_i)(V_i \otimes V_i'). \tag{2.157}$$

We impose the following important rewriting rules for this vector space:

(i) The tensor must respect addition in both $\mathbb{V}$ and $\mathbb{V}'$:

$$(V_i + V_j) \otimes V'_k = V_i \otimes V'_k + V_j \otimes V'_k, \tag{2.158}$$

$$V_i \otimes (V'_j + V'_k) = V_i \otimes V'_j + V_i \otimes V'_k. \tag{2.159}$$

(ii) The tensor must respect the scalar multiplication in both $\mathbb{V}$ and $\mathbb{V}'$:

$$c \cdot (V_j \otimes V'_k) = (c \cdot V_j) \otimes V'_k = V_j \otimes (c \cdot V'_k). \tag{2.160}$$

By following these rewriting rules and setting elements equal to each other, we form $\mathbb{V} \otimes \mathbb{V}'$.

Let us find a basis for $\mathbb{V} \otimes \mathbb{V}'$. Say, $\mathbb{V}$ has a basis $\mathcal{B} = \{B_0, B_1, \ldots, B_{m-1}\}$ and $\mathbb{V}'$ has a basis $\mathcal{B}' = \{B'_0, B'_1, \ldots, B'_{n-1}\}$. Given that every $V_i \in \mathbb{V}$ and $V'_i \in \mathbb{V}'$ can be written in a unique way for these bases, we can use the rewrite rules to "decompose" every element $\sum_{i=0}^{p-1} c_i (V_i \otimes V'_i)$ in the tensor product. This will give us a basis for $\mathbb{V} \otimes \mathbb{V}'$. In detail, the basis for $\mathbb{V} \otimes \mathbb{V}'$ will be the set of vectors

$$\{B_j \otimes B'_k \,|\, j = 0, 1, \ldots, m-1 \text{ and } k = 0, 1, \ldots, n-1\}. \tag{2.161}$$

Every $\sum_{i=0}^{p-1} c_i (V_i \otimes V'_i) \in \mathbb{V} \otimes \mathbb{V}'$ can be written as

$$c_{0,0}(B_0 \otimes B'_0) + c_{1,0}(B_1 \otimes B'_0) + \cdots + c_{m-1,n-1}(B_{m-1} \otimes B'_{n-1}). \tag{2.162}$$

The dimension of $\mathbb{V} \otimes \mathbb{V}'$ is the dimension of $\mathbb{V}$ times the dimension of $\mathbb{V}'$. (Remember that the dimension of $\mathbb{V} \times \mathbb{V}'$ is the dimension of $\mathbb{V}$ plus the dimension of $\mathbb{V}'$. So the tensor product of two vector spaces is usually a larger space than their Cartesian product.[8]) One should think of $\mathbb{V} \times \mathbb{V}'$ as the vector space whose states are the states of a system $\mathbb{V}$ *or* a system $\mathbb{V}'$ or both. $\mathbb{V} \otimes \mathbb{V}'$ is to be thought of as the vector space whose basic states are pairs of states, one from system $\mathbb{V}$ *and* one from the system $\mathbb{V}'$.

Given an element of $\mathbb{V}$

$$c_0 B_0 + c_1 B_1 + \cdots + c_{m-1} B_{m-1}, \tag{2.163}$$

and an element of $\mathbb{V}'$

$$c'_0 B'_0 + c'_1 B'_1 + \cdots + c'_{n-1} B'_{n-1}, \tag{2.164}$$

we can associate[9] the following element of $\mathbb{V} \otimes \mathbb{V}'$:

$$(c_0 \times c'_0)(B_0 \otimes B'_0) + (c_0 \times c'_1)(B_0 \otimes B'_1) + \cdots + (c_{m-1} \times c'_{n-1})(B_{m-1} \otimes B'_{n-1}). \tag{2.165}$$

Let us step down from the abstract highland and see what $\mathbb{C}^m \otimes \mathbb{C}^n$ actually looks like. $\mathbb{C}^m \otimes \mathbb{C}^n$ is of dimension mn and hence is isomorphic to $\mathbb{C}^{m \times n}$. What is important is how $\mathbb{C}^m \otimes \mathbb{C}^n$ is isomorphic to $\mathbb{C}^{m \times n}$. If E_j is an element of the canonical basis of each vector space, then we might identify $E_j \otimes E_k$ with $E_{j \times k}$. It is not hard to see

[8] But not always! Remember that $1 \times 1 < 1 + 1$ and $1 \times 2 < 1 + 2$, etc.

[9] It is important to notice that this "association" is not a linear map; it is something called a **bilinear map**.

from the association given in Equation (2.165) that the **tensor product of vectors** is defined as follows:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \otimes \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_0 \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \\ a_1 \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \\ a_2 \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \\ a_3 \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_0 b_0 \\ a_0 b_1 \\ a_0 b_2 \\ a_1 b_0 \\ a_1 b_1 \\ a_1 b_2 \\ a_2 b_0 \\ a_2 b_1 \\ a_2 b_2 \\ a_3 b_0 \\ a_3 b_1 \\ a_3 b_2 \end{bmatrix}. \tag{2.166}$$

In general, $\mathbb{C}^m \times \mathbb{C}^n$ is much smaller than $\mathbb{C}^m \otimes \mathbb{C}^n$.

Example 2.7.1 For example, consider $\mathbb{C}^2 \times \mathbb{C}^3$ and $\mathbb{C}^2 \otimes \mathbb{C}^3 = \mathbb{C}^6$. Consider the vector

$$\begin{bmatrix} 8 \\ 12 \\ 6 \\ 12 \\ 18 \\ 9 \end{bmatrix} \in \mathbb{C}^6 = \mathbb{C}^2 \otimes \mathbb{C}^3. \tag{2.167}$$

It is not hard to see that this is simply

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \otimes \begin{bmatrix} 4 \\ 6 \\ 3 \end{bmatrix}. \tag{2.168}$$

□

Example 2.7.2 In contrast to the above example,

$$\begin{bmatrix} 8 \\ 0 \\ 0 \\ 0 \\ 0 \\ 18 \end{bmatrix} \in \mathbb{C}^6 = \mathbb{C}^2 \otimes \mathbb{C}^3 \tag{2.169}$$

cannot be written as the tensor product of a vector from $\mathbb{C}^2$ and $\mathbb{C}^3$. In order to see this, consider the variables

$$\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} xa \\ xb \\ xc \\ ya \\ yb \\ yc \end{bmatrix}. \tag{2.170}$$

There are no solutions for the variable that will give you the required results. However, we can write the vector in Equation (2.169) as

$$\begin{bmatrix} 8 \\ 0 \\ 0 \\ 0 \\ 0 \\ 18 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 6 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}. \tag{2.171}$$

This is a summation of two vectors. □

For reasons that are made clear in Sections 3.4 and 4.5, we shall call a vector that can be written as the tensor of two vectors **separable**. In contrast, a vector that cannot be written as the tensor of two vectors (but can be written as the nontrivial sum of such tensors) shall be called **entangled**.

Exercise 2.7.1 Calculate the tensor product $\begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} \otimes \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. ■

Exercise 2.7.2 State whether $[5, 6, 3, 2, 0, 1]^T$ is a tensor product of smaller vectors from $\mathbb{C}^3$ and $\mathbb{C}^2$. ■

We will need to know not only how to take the tensor product of two vectors, but also how to determine the **tensor product of two matrices.**[10] Consider two matrices

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{0,0} & b_{0,1} & b_{0,2} \\ b_{1,0} & b_{1,1} & b_{1,2} \\ b_{2,0} & b_{2,1} & b_{2,2} \end{bmatrix}. \tag{2.172}$$

From the association given in Equation (2.165), it can be seen that the tensor product $A \otimes B$ is the matrix that has every element of A, scalar multiplied with the entire matrix B. That is,

$$A \otimes B = \begin{bmatrix} a_{0,0} \cdot \begin{bmatrix} b_{0,0} & b_{0,1} & b_{0,2} \\ b_{1,0} & b_{1,1} & b_{1,2} \\ b_{2,0} & b_{2,1} & b_{2,2} \end{bmatrix} & a_{0,1} \cdot \begin{bmatrix} b_{0,0} & b_{0,1} & b_{0,2} \\ b_{1,0} & b_{1,1} & b_{1,2} \\ b_{2,0} & b_{2,1} & b_{2,2} \end{bmatrix} \\ a_{1,0} \cdot \begin{bmatrix} b_{0,0} & b_{0,1} & b_{0,2} \\ b_{1,0} & b_{1,1} & b_{1,2} \\ b_{2,0} & b_{2,1} & b_{2,2} \end{bmatrix} & a_{1,1} \cdot \begin{bmatrix} b_{0,0} & b_{0,1} & b_{0,2} \\ b_{1,0} & b_{1,1} & b_{1,2} \\ b_{2,0} & b_{2,1} & b_{2,2} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} a_{0,0} \times b_{0,0} & a_{0,0} \times b_{0,1} & a_{0,0} \times b_{0,2} & a_{0,1} \times b_{0,0} & a_{0,1} \times b_{0,1} & a_{0,1} \times b_{0,2} \\ a_{0,0} \times b_{1,0} & a_{0,0} \times b_{1,1} & a_{0,0} \times b_{1,2} & a_{0,1} \times b_{1,0} & a_{0,1} \times b_{1,1} & a_{0,1} \times b_{1,2} \\ a_{0,0} \times b_{2,0} & a_{0,0} \times b_{2,1} & a_{0,0} \times b_{2,2} & a_{0,1} \times b_{2,0} & a_{0,1} \times b_{2,1} & a_{0,1} \times b_{2,2} \\ a_{1,0} \times b_{0,0} & a_{1,0} \times b_{0,1} & a_{1,0} \times b_{0,2} & a_{1,1} \times b_{0,0} & a_{1,1} \times b_{0,1} & a_{1,1} \times b_{0,2} \\ a_{1,0} \times b_{1,0} & a_{1,0} \times b_{1,1} & a_{1,0} \times b_{1,2} & a_{1,1} \times b_{1,0} & a_{1,1} \times b_{1,1} & a_{1,1} \times b_{1,2} \\ a_{1,0} \times b_{2,0} & a_{1,0} \times b_{2,1} & a_{1,0} \times b_{2,2} & a_{1,1} \times b_{2,0} & a_{1,1} \times b_{2,1} & a_{1,1} \times b_{2,2} \end{bmatrix}. \tag{2.173}$$

[10] It should be clear that the tensor product of two vectors is simply a special case of the tensor product of two matrices.

Formally, the tensor product of matrices is a function

$$\otimes : \mathbb{C}^{m\times m'} \times \mathbb{C}^{n\times n'} \longrightarrow \mathbb{C}^{mn\times m'n'} \tag{2.174}$$

and it is defined as

$$(A \otimes B)[j,k] = A[j/n, k/m] \times B[j \text{ Mod } n, k \text{ Mod } m]. \tag{2.175}$$

Exercise 2.7.3 Calculate

$$\begin{bmatrix} 3+2i & 5-i & 2i \\ 0 & 12 & 6-3i \\ 2 & 4+4i & 9+3i \end{bmatrix} \otimes \begin{bmatrix} 1 & 3+4i & 5-7i \\ 10+2i & 6 & 2+5i \\ 0 & 1 & 2+9i \end{bmatrix}. \tag{2.176}$$

■

Exercise 2.7.4 Prove that the tensor product is "almost" commutative. Take two 2-by-2 matrices A and B. Calculate $A \otimes B$ and $B \otimes A$. In general, although they are not equal, they do have the same entries, and one can be transformed to the other with a "nice" change of rows and columns. ■

Exercise 2.7.5 Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$, and $C = \begin{bmatrix} 6 & 5 \\ 3 & 2 \end{bmatrix}$. Calculate $A \otimes (B \otimes C)$ and $(A \otimes B) \otimes C$ and show that they are equal. ■

Exercise 2.7.6 Prove that the tensor product is associative, i.e., for arbitrary matrices A, B, and C,

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C. \tag{2.177}$$

■

Exercise 2.7.7 Let $A = \begin{bmatrix} 2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Calculate $(A \otimes B)^\dagger$ and $A^\dagger \otimes B^\dagger$ and show that they are equal. ■

Exercise 2.7.8 Prove that $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$. ■

Exercise 2.7.9 Let A, A', B, and B' be matrices of the appropriate sizes. Prove that

$$(A \star A') \otimes (B \star B') = (A \otimes B) \star (A' \otimes B'). \tag{2.178}$$

■

If A acts on V and B acts on V', then we define the action on their tensor product as

$$(A \otimes B) \star (V \otimes V') = A \star V \otimes B \star V'. \tag{2.179}$$

Such "parallel" actions will arise over and over again.