Vectors and Matrices
MANUSCRIPT

William F. Barnes
[1]

February 12, 2024

Contents

1 Vectors and Matrices **3**
1 Introduction to Vectors 3
1.1 Taxonomy of Vectors 3
1.2 Representing Vectors 4
1.3 Position Vector 4
1.4 Vector Magnitude and Direction 4
2 Vector Addition 4
2.1 Definition 4
2.2 Arrow Trick 5
2.3 Additive Inverse 5
2.4 Zero Vector 5
2.5 Vector Subtraction 5
3 Scalar Multiplication 5
3.1 Parallel Vectors 6
3.2 Straight Lines 6
3.3 Algebraic Properties of Vectors 6
4 Vector Products 6
4.1 Dot Product 6
4.2 Cross Product 7
4.3 Vector Identities 8
5 Polar Representation 9
5.1 Polar Coordinate System . . . 9
5.2 Rotated Vectors 10
5.3 Rotation Matrix 10
6 Basis Vectors 10
6.1 Unit Vectors 10
6.2 Introduction to Basis Vectors . 10
6.3 Linear Combinations 11
7 Change of Basis 11
7.1 Two Dimensions 12
7.2 N Dimensions 13
8 Vectors and Limits 14
8.1 Pi from Nested Radicals 14
9 Matrix Formalism 17
9.1 Matrix-Operator Equivalence . 17
9.2 Matrix Components 17
9.3 Projector 17
9.4 Identity Operator 18
9.5 Unified Matrix Notation 18
10 Matrix Operations 18
10.1 Matrix Addition 18
10.2 Scalar Multiplication 18
10.3 Matrix Multiplication 18
10.4 Change of Basis 19

Chapter 1

Vectors and Matrices

1 Introduction to Vectors

1.1 Taxonomy of Vectors

Definition

A *vector* is an ordered list of numbers or variables. One example of a vector is the sequence of numbers 2, 5, and 7, which can be written in *vector notation*:

$$\vec{V} = \langle 2, 5, 7 \rangle$$

Vector notation requires a *label* for the vector, $\vec{V}$ in our example, specially marked by an arrow ($\vec{\ }$). On the right, the the so-called *vector literal* is enclosed by left-and right-angle brackets $\langle\, \rangle$ as shown. Each number in the vector is separated by a comma.

Vector Components

The individual elements in a vector are formally called *components*, and the total number of components is the *dimension* of the vector. The order in which the components of a vector are listed *does* matter. For example, the three-dimensional vector $\vec{V} = \langle 2, 5, 7 \rangle$ is completely different from its reversed version $\langle 7, 5, 2 \rangle$.

Component Subscripts

In a vector, any given component is represented using the vector's symbol without the arrow, but including an *index subscript.* For instance, we could represent $\vec{V}$ as

$$\vec{V} = \langle V_a, V_b, V_c \rangle$$

with $V_a = 2$, $V_b = 5$, $V_c = 7$, but the letters a, b, c could easily have been x, y, z, or perhaps 1, 2, 3. Vector component labels are, after the dust settles, purely for bookkeeping.

1.2 Representing Vectors

Vectors of two dimensions are suited for visualization on the Cartesian plane. Given a vector $\vec{V} = \langle V_a, V_b \rangle$, we plot $\vec{V}$ with the following recipe:

- Choose any *base point* for the vector on the plane. From the base point:
- Measure V_a units horizontally, measure V_b units vertically.
- Plot the vector *tip point*. Connect base and tip with an arrow.

Plotted in Fig. 1.1 are two *equivalent* representations of the vector $\vec{A} = \langle 2, 6 \rangle$. Note that the vector doesn't 'care' about the choice of base point.

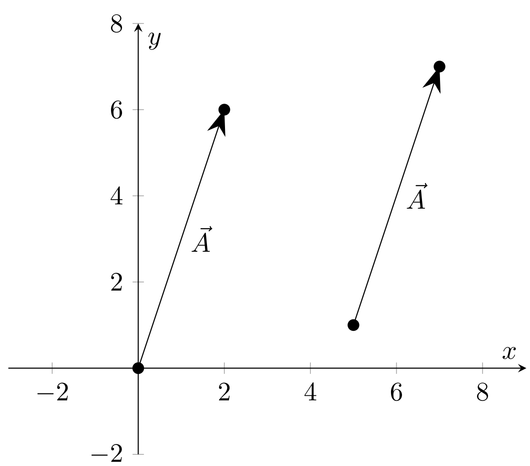

Figure 1.1: Vector $\vec{A} = \langle 2, 6 \rangle$ plotted from two different base points $(0, 0)$ and $(5, 1)$.

It should also follow that the above construction extends to dimensions beyond two. For instance, vectors of three dimensions can be visualized in a three-dimensional coordinate system, and so on.

1.3 Position Vector

A vector whose base point is the origin $(0, 0)$ is called a *position vector*, often denoted $\vec{R}$ or $\vec{X}$. A position vector $\vec{R} = \langle R_x, R_y \rangle$ is equivalent to the ordered pair (x, y), denoting a unique point in the Cartesian plane.

1.4 Vector Magnitude and Direction

Given the 'arrow' representation of a vector, we notice two important features:

- Vectors have a *magnitude*, i.e. the total arrow length.
- Vectors have a *direction*, i.e. a notion of pointing somewhere.

The 'information' in a vector is completely represented by its magnitude and its direction. (This may grant some relief as to why we can be so loose about the choice of base point.)

Calculating the Magnitude

A vector $\vec{A}$ of dimension N has a magnitude given by

$$A = \left|\vec{A}\right| = \sqrt{A_1^2 + A_2^2 + \cdots + A_N^2}\,. \tag{1.1}$$

Intuitively, the magnitude of a vector can be thought of the hypotenuse of an N-dimensional triangle. For the special case $N = 2$, the above reduces to the Pythagorean theorem.

Calculating the Direction

The direction of an N-dimensional vector $\vec{A}$ is always implied by the components A_j, but an explicit formula for the 'angle' of the vector is only trivial for small N. Working the $N = 2$ case, the direction in which a vector $\vec{A} = \langle A_x, A_y \rangle$ is pointing is given by

$$\phi = \arctan\left(\frac{A_y}{A_x}\right) \tag{1.2}$$

To justify (1.2), assume A_x and A_y are two sides of a right triangle such that

$$A_x = A\cos(\phi)$$
$$A_y = A\sin(\phi)\ ,$$

and eliminate the magnitude A.

2 Vector Addition

2.1 Definition

Two vectors $\vec{A}$, $\vec{B}$ of equal dimension N can be added by combining like components, resulting in a vector $\vec{C}$ with N components:

$$\vec{C} = \vec{A} + \vec{B} \tag{1.3}$$

The jth component is given by

$$C_j = A_j + B_j \tag{1.4}$$
$$j = 1, 2, 3, \ldots, N\,.$$

Commutativity of Vector Addition

Following immediately from (1.3)-(1.4) is the *commutativity of addition*:

$$\vec{A}+\vec{B}=\vec{B}+\vec{A} \tag{1.5}$$

In particular, (1.5) tells us that the order in which two vectors are added does not affect the result.

Associativity of Vector Addition

The sum of three vectors $\vec{A}$, $\vec{B}$, $\vec{C}$ involves two addition operations. Also following from (1.3)-(1.4) is the *associativity of addition*, telling us that the order of the two addition operations does not effect the result:

$$\vec{A}+\left(\vec{B}+\vec{C}\right)=\left(\vec{A}+\vec{B}\right)+\vec{C} \tag{1.6}$$

2.2 Arrow Trick

The 'arrow' representation of a vector avails a beautiful shortcut for vector addition. Given a pair of two-dimensional vectors $\vec{A}$, $\vec{B}$, recall that each vector can be drawn *anywhere* in the Cartesian plane. *By arranging the two vectors in tip-to-tail fashion, the vector sum goes from the tail of the first to the tip of the second.*

Fig. 1.2 demonstrates the 'arrow trick' on two example vectors $\vec{A}=\langle 2,5\rangle$, $\vec{B}=\langle 3,-3\rangle$, whose sum easily comes out to $\vec{C}=\langle 5,2\rangle$. By plotting $\vec{A}$, $\vec{B}$ as suggested, the sum $\vec{C}$ is visually represented by an arrow beginning at the base of $\vec{A}$ and ending at the tip of $\vec{B}$.

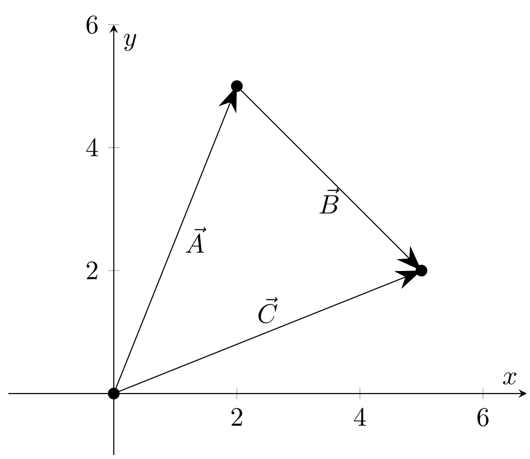

Figure 1.2: Vector addition $\vec{C}=\vec{A}+\vec{B}$.

2.3 Additive Inverse

Given any vector $\vec{A}=\langle A_1,A_2,A_3,\ldots,A_N\rangle$, the *additive inverse* is another vector that reverses the sign on all components in $\vec{A}$, denoted $-\vec{A}$, where

$$-\vec{A}=\langle -A_1,-A_2,-A_3,\ldots,-A_N\rangle\ . \tag{1.7}$$

2.4 Zero Vector

The so-called *zero vector* is the vector that contains only zeros:

$$\vec{0}=\langle 0,0,0,\ldots,0\rangle \tag{1.8}$$

For hopefully obvious reasons, turns out that the sum of any vector and its additive inverse always yields the zero vector:

$$\vec{A}+\left(-\vec{A}\right)=\vec{0}$$

In practice, the zero vector is simply written 0, omitting the arrow.

An interesting corollary to the rules of vector addition is that any *closed* sequence of vectors sums to zero. For example, drawing a triangle without lifting the pen from the surface is represented by $\vec{A}+\vec{B}+\vec{C}=0$.

2.5 Vector Subtraction

With the additive inverse established, the notion of *vector subtraction* can be framed in terms of vector addition. Given two vectors $\vec{A}$, $\vec{B}$, the difference $\vec{D}=\vec{A}-\vec{B}$ can be visualized with the same 'arrow trick', so long as we reverse the direction on $\vec{B}$ as shown in Fig. 1.3.

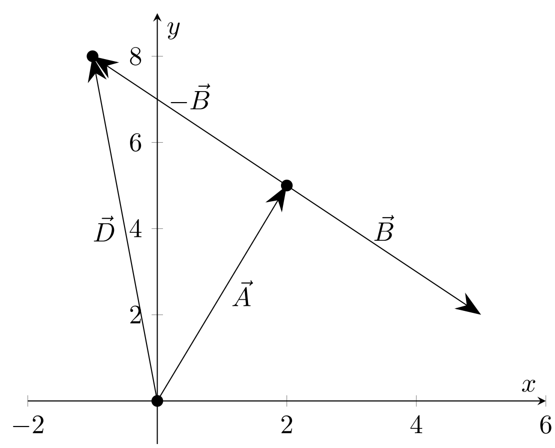

Figure 1.3: Vector subtraction $\vec{D}=\vec{A}-\vec{B}$.

3 Scalar Multiplication

A vector $\vec{A}$ can be 'scaled' by a number α called a *scalar*, which has the effect of multiplying the scalar into each component, yielding a new vector $\vec{B}$:

$$\vec{B}=\alpha\vec{A}=\langle \alpha A_1,\alpha A_2,\alpha A_3,\ldots,\alpha A_N\rangle \tag{1.9}$$

3.1 Parallel Vectors

Two vectors whose components are identical up to a scale factor α are said to be *parallel*. Somewhat like parallel lines, two parallel vectors can have different magnitudes, but point in the same direction. The vectors $\vec{A}$, $\vec{B}$ in (1.9) are necessarily parallel.

3.2 Straight Lines

Straight lines in the Cartesian plane are easily represented with vector addition and scalar multiplication. Consider the slope-intercept form of a line, namely $y = mx + b$, where m is the slope and b is the y-intercept at $(0, b)$. As a vector, the y-intercept can be written

$$\vec{b} = \langle 0, b \rangle \ .$$

Required next is a vector $\vec{m}$ that represents the slope of the line, which we capture by writing

$$\begin{aligned} \vec{m} &= \langle m_x, m_y \rangle \\ \frac{m_y}{m_x} &= m \ . \end{aligned}$$

Multiplying $\vec{m}$ by any scalar value α will lengthen, shorten, or reverse its effective placement.

Putting the two ingredients together, it follows that any point on the line $y = mx + b$ is equivalently represented as

$$\vec{r} = \alpha\, \vec{m} + \vec{b}\,, \tag{1.10}$$

where $\vec{r} = \langle x, y \rangle$ is the resulting position vector, as shown in Fig. 1.4. In case (1.10) isn't convincing, one may resolve $\vec{r}$ back into components

$$\begin{aligned} x &= \alpha\, m_x \\ y &= \alpha\, m_y + b\,, \end{aligned}$$

where eliminating α recovers the familiar $y = mx + b$.

Perpendicular Lines

Two lines in the Cartesian plane are perpendicular one line's slope is m, and the slope is $m_\perp = -1/m$. In terms of the components m_x, m_y, this means

$$\begin{aligned} m &= \frac{m_y}{m_x} \\ m_\perp &= \frac{-m_x}{m_y} \ . \end{aligned}$$

From this, the 'perpendicular slope vector' $\vec{m}_\perp$ is evidently $\vec{m}_\perp = \langle -m_y, m_x \rangle$.

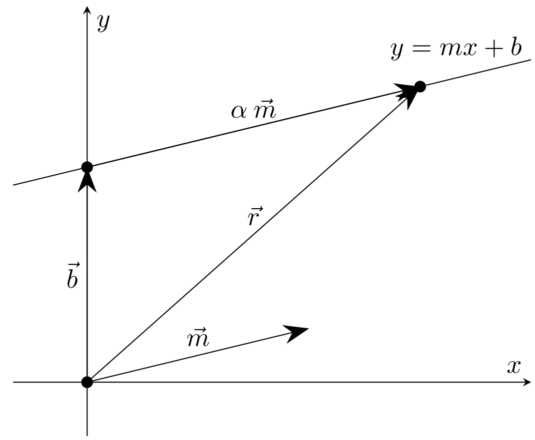

Figure 1.4: Vector construction of a straight line $y = mx + b$.

3.3 Algebraic Properties of Vectors

Associativity with Scalars

If a vector is modified by two scalars, the order in which they're applied does not matter:

$$\alpha \left(\beta \vec{A}\right) = (\alpha\beta)\, \vec{A} = (\beta\alpha)\, \vec{A} = \beta \left(\alpha \vec{A}\right) \tag{1.11}$$

Vector Distributive Properties

Readily provable from the properties of vector addition and scalar multiplication are the distributive properties involving the sum of two scalars:

$$(\alpha + \beta)\, \vec{A} = \alpha \vec{A} + \beta \vec{B} \tag{1.12}$$

$$\alpha \left(\vec{A} + \vec{B}\right) = \alpha \vec{A} + \alpha \vec{B} \tag{1.13}$$

4 Vector Products

4.1 Dot Product

Two vectors of equal dimension can be 'multiplied' to form a scalar, called the *dot product*, or the *scalar product*. The dot product is an operation that tells us how much of one vector's 'shadow' falls upon another vector, resulting in a scalar called a *projection*. For two vectors $\vec{A}$, $\vec{B}$ of dimension N, the dot product reads

$$\vec{A} \cdot \vec{B} = A_1 B_1 + A_2 B_2 + \cdots + A_N B_N \ ,$$

or, in summation notation:

$$\vec{A} \cdot \vec{B} = \sum_{j=1}^{N} A_j B_j \tag{1.14}$$

Commutativity Relation

Implicit in the definition (1.14) is the commutativity of the dot product (in any number of dimensions):

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} \tag{1.15}$$

Geometric Interpretation of Vectors

The definition (1.14) becomes more intuitive by studying the $N = 2$ case. Consider two arbitrary vectors given by

$$\vec{A} = \langle A\cos(\phi_A), A\sin(\phi_A)\rangle$$
$$\vec{B} = \langle B\cos(\phi_B), B\sin(\phi_B)\rangle \ ,$$

where A and B are the respecctive magnitudes. Calculating $\vec{A} \cdot \vec{B}$ using the formula provided results in

$$\vec{A} \cdot \vec{B} = AB\left(\cos(\phi_A)\cos(\phi_B) + \sin(\phi_A)\sin(\phi_B)\right)$$
$$\vec{A} \cdot \vec{B} = AB\cos(\phi_B - \phi_A) \ ,$$

telling us that *the dot product is equal to the product of the magnitudes and the cosine of the angle between the vectors*. In general, this result reads

$$\cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{AB} \ , \tag{1.16}$$

where θ is the angle between the vectors in any number of dimensions. The special case $N = 2$ corresponds to $\theta = \phi_B - \phi_A$.

Vector Orthogonality

From the two-dimensional dot product, note that the case $\phi_A - \phi_B = \pm\pi/2$ returns $\cos(\pm\pi/2) = 0$ on the left, telling us that *the dot product between perpendicular vectors is zero*. The formal term for 'perpendicular' is *orthogonal*, and this notion generalizes to N dimensions:

$$\vec{A} \cdot \vec{B} = 0 \tag{1.17}$$

In the Cartesian plane, recall that the slope of a line and another perpendicular line are represented by the vectors

$$\vec{m} = \langle m_x, m_y\rangle$$
$$\vec{m}_\perp = \langle -m_y, m_x\rangle \ ,$$

respectively. We verify these vectors to be orthogonal by calculating

$$\vec{m} \cdot \vec{m}_\perp = -m_x m_y + m_y m_x = 0 \, .$$

Vector Magnitude

The dot product is responsible for the formula (1.1) for calculating the magnitude of a vector. Indeed, for an N-dimensional vector $\vec{A}$, we find the dot product with itself to be

$$\vec{A} \cdot \vec{A} = A_1^2 + A_2^2 + A_3^2 + \cdots + A_N^2 \ ,$$

which is the square of the magnitude of A. More concisely:

$$A = \left|\vec{A}\right| = \sqrt{\vec{A} \cdot \vec{A}} \tag{1.18}$$

Distributive Property

For three vectors $\vec{A}$, $\vec{B}$, $\vec{C}$ of equal dimension, the dot product obeys the distributive property as one may expect:

$$\vec{A} \cdot \left(\vec{B} + \vec{C}\right) = \vec{A} \cdot \vec{B} + \vec{A} \cdot \vec{C} \tag{1.19}$$

Law of Cosines

An important relation from trigonometry called the *law of cosines* is derived using dot products. Consider the vector sum

$$\vec{A} - \vec{B} = \vec{C} \ ,$$

and then square both sides:

$$\left(\vec{A} - \vec{B}\right) \cdot \left(\vec{A} - \vec{B}\right) = \vec{C} \cdot \vec{C}$$
$$\vec{A} \cdot \vec{A} + \vec{B} \cdot \vec{B} - 2\vec{A} \cdot \vec{B} = \vec{C} \cdot \vec{C}$$

Labeling θ as the angle between vectors $\vec{A}$, $\vec{B}$, the above simplifies to the law of cosines:

$$A^2 + B^2 - 2AB\cos(\theta) = C^2 \tag{1.20}$$

Note that all right triangles have $\theta = \pi/2$, in which case (1.20) reduces to the Pythagorean theorem.

4.2 Cross Product

Two vectors of equal dimension can be 'multiplied' to form a new vector, called the *cross product*, or the *vector product*. The cross product is, for most purposes, a strictly three-dimensional operation. Consider the pair of vectors with $N = 3$:

$$\vec{A} = \langle A_x, A_y, A_z\rangle$$
$$\vec{B} = \langle B_x, B_y, B_z\rangle$$

The cross product $\vec{A} \times \vec{B}$ is defined as

$$\vec{A} \times \vec{B} = \langle C_x, C_y, C_z\rangle \ , \tag{1.21}$$

where

$$C_x = A_y B_z - A_z B_y$$
$$C_y = A_z B_x - A_x B_z$$
$$C_z = A_x B_y - A_y B_x \ ,$$

and is *orthogonal* to both $\vec{A}$ and $\vec{B}$.

Determinant Notation

The cross product formula (1.21) is tricky to memorize, and can be more transparently represented as a 'block of numbers' (*not* a matrix), sometimes called *determinant notation*:

$$\vec{A} \times \vec{B} = \begin{vmatrix} (x) & (-y) & (z) \\ A_x & A_y & A_z \\ B_z & B_y & B_z \end{vmatrix}$$

Without sweating the details of determinant notation, you can play a matching game between the determinant representation of $\vec{A} \times \vec{B}$ and the formula (1.21) to remember how it goes.

Orthogonality Check

To ensure that $\vec{A} \times \vec{B}$ is mutually orthogonal to $\vec{A}$,

$$\vec{A} \cdot \left(\vec{A} \times \vec{B} \right)$$

to see what comes out. In detail, the former case proceeds as

$$\begin{aligned} &\vec{A} \cdot \left(\vec{A} \times \vec{B} \right) \\ &\quad = A_x A_y B_z - A_x A_z B_y + A_y A_z B_x \\ &\qquad - A_y A_x B_z + A_z A_x B_y - A_z A_y B_x \\ &\quad = B_z \left(A_x A_y - A_y A_x \right) - B_y \left(A_x A_z - A_z A_x \right) \\ &\qquad + B_z \left(A_x A_y - A_y A_x \right) \\ &\quad = 0 \, . \end{aligned}$$

This also holds true for the B-case.

Null Case

The cross product of a vector with itself is identically zero:

$$\vec{A} \times \vec{A} = 0 \tag{1.22}$$

Anti-Commutativity Relation

Given the definition (1.21) of the cross product, one sees that swapping $\vec{A}$, $\vec{B}$ puts a minus sign on the result. This is known as the *anti-commutativity* of the cross product:

$$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A} \tag{1.23}$$

Right Hand Rule

There is a trick that allows one to know the direction of $\vec{A} \times \vec{B}$ known as the (oft-dreaded) *right hand rule.* To know the direction of the vector $\vec{A} \times \vec{B}$, the steps are as follows:

1. On your right hand: point your thumb, index finger, and middle finger out in perpendicular directions.

2. Let your index finger be vector $\vec{A}$, let your middle finger be vector $\vec{B}$.

3. Your thumb points along vector $\vec{A} \times \vec{B}$.

Geometric Interpretation

The definition (1.21) becomes more intuitive by studying a special case. Consider the pair of three-dimensional vectors confined to the xy-plane given by

$$\vec{A} = \langle A \cos\left(\phi_A\right), A \sin\left(\phi_A\right), 0 \rangle$$
$$\vec{B} = \langle B \cos\left(\phi_B\right), B \sin\left(\phi_B\right), 0 \rangle \ ,$$

where A and B are the respecctive magnitudes. Calculating $\vec{A} \times \vec{B}$ using the formula provided results in

$$\vec{A} \times \vec{B} = \langle 0, 0, AB \left(\cos\phi_A \sin\phi_B - \cos\phi_B \sin\phi_A \right) \rangle$$
$$\vec{A} \times \vec{B} = \langle 0, 0, AB \sin\left(\phi_B - \phi_A \right) \rangle \ ,$$

telling us that *the cross product is equal to the product of the magnitudes and the sine of the angle between the vectors.* In general, this result also tells us

$$\sin\left(\theta\right) = \frac{\left| \vec{A} \times \vec{B} \right|}{AB} \ , \tag{1.24}$$

where θ is the angle between the vectors at any relative orientation.

Area of a Parallelogram

The quantity $AB \sin\left(\theta\right)$ can be interpreted as the area of a parallelogram having base B and height $h = A \sin\phi$. For the right-angle case $\phi = \pi/2$, the parallelogram becomes a rectangle of area AB. In the language of vectors, the product $\left| \vec{A} \times \vec{B} \right|$ is the area of the parallelogram with sides A, B.

4.3 Vector Identities

Consider three vectors $\vec{A}$, $\vec{B}$, $\vec{C}$, each of three dimensions.

Triple Product

The quantity

$$V = \vec{A} \cdot \left(\vec{B} \times \vec{C}\right) \tag{1.25}$$

is a scalar called the *triple product*. Intuitively, the triple product describes the volume of the parallelpiped with sides A, B, C. One can show by brute force that (1.25) obeys the cyclic relations:

$$\vec{A} \cdot \left(\vec{B} \times \vec{C}\right) = \vec{B} \cdot \left(\vec{C} \times \vec{A}\right) = \vec{C} \cdot \left(\vec{A} \times \vec{B}\right)$$

BAC-CAB Formula

A useful equation known as the *BAC-CAB* identity, reads

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \vec{B}\left(\vec{A} \cdot \vec{C}\right) - \vec{C}\left(\vec{A} \cdot \vec{B}\right) \ . \tag{1.26}$$

The proof of (1.26) is slightly long but straightforward, using (optional) determinant notation to contain the cross product:

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \begin{vmatrix} (x) & (-y) & (z) \\ A_x & A_y & A_z \\ B_yC_z - B_zC_y & B_zC_x - B_xC_z & B_xC_y - B_yC_x \end{vmatrix}$$

$$\begin{aligned} \vec{A} \times \left(\vec{B} \times \vec{C}\right) &= \langle A_yB_xC_y - A_yB_yC_x - A_zB_zC_x + A_zB_xC_z, 0, 0\rangle + \\ &\quad \langle 0, A_zB_yC_z - A_zB_zC_y - A_xB_xC_y + A_xB_yC_x, 0\rangle + \\ &\quad \langle 0, 0, A_xB_zC_x - A_xB_xC_z - A_yB_yC_z + A_yB_zC_y\rangle \\ &= B_x\langle A_yC_y + A_zC_z, 0, 0\rangle - A_xB_x\langle 0, C_y, C_z\rangle + \\ &\quad B_y\langle 0, A_zC_z + A_xC_x, 0\rangle - A_yB_y\langle C_x, 0, C_z\rangle + \\ &\quad B_z\langle 0, 0, A_xC_x + A_yC_y\rangle - A_zB_z\langle C_x, C_y, 0\rangle \end{aligned}$$

$$\begin{aligned} \vec{A} \times \left(\vec{B} \times \vec{C}\right) &= B_x\left\langle \vec{A} \cdot \vec{C}, 0, 0\right\rangle - A_xB_x\langle C_x, C_y, C_z\rangle + \\ &\quad B_y\left\langle 0, \vec{A} \cdot \vec{C}, 0\right\rangle - A_yB_y\langle C_x, C_y, C_z\rangle + \\ &\quad B_z\left\langle 0, 0, \vec{A} \cdot \vec{C}\right\rangle - A_zB_z\langle C_x, C_y, C_z\rangle \end{aligned}$$

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \vec{B}\left(\vec{A} \cdot \vec{C}\right) - \vec{C}\left(\vec{A} \cdot \vec{B}\right)$$

5 Polar Representation

In the Cartesian plane, consider a position vector

$$\vec{r} = \langle r_x, r_y\rangle \ .$$

The 'magnitude-and-direction' interpretation of $\vec{r}$ assigns the magnitude r to the hypotenuse of a right triangle, where the adjacent and opposite sides are respectively given by

$$r_x = r\cos(\phi) \tag{1.27}$$

$$r_y = r\sin(\phi) \ , \tag{1.28}$$

congruent with equations (1.1)-(1.2). The angle parameter ϕ is also known as the *phase* of the vector, a dimensionless argument unique on the interval $[0 : 2\pi)$.

5.1 Polar Coordinate System

Equations (1.27)-(1.28) represent a mapping from system of Cartesian coordinates to the system of polar coordinates. Any point in the plane that can be represented by the ordered pair (x, y) has an equivalent representation as the ordered pair (r, ϕ). In particular, we take the position vecor in polar coordinates to be

$$\begin{aligned} \vec{r} &= \langle r\cos(\theta), r\sin(\theta)\rangle \\ &= r\langle\cos(\theta), \sin(\theta)\rangle \end{aligned}$$

5.2 Rotated Vectors

Starting with a vector $\vec{r}$ in two dimensions, particularly

$$\vec{r} = \langle r\cos(\phi), r\sin(\phi)\rangle \ ,$$

we may inquire what happens when we modify the phase such that $\phi \to \phi + \theta$, effectively rotating the vector in the plane.

Carrying this out, one writes

$$(\vec{r})' = r\,\langle \cos(\phi+\theta), \sin(\phi+\theta)\rangle \ ,$$

or, expanding the trigonometry terms,

$$\begin{aligned} r'_x &= r\left(\cos(\phi)\cos(\theta) - \sin(\phi)\sin(\theta)\right) \\ r'_y &= r\left(\sin(\phi)\cos(\theta) + \cos(\phi)\sin(\theta)\right) \ . \end{aligned}$$

Things get interesting when we keep simplifying:

$$r'_x = r_x\cos(\theta) - r_y\sin(\theta) \tag{1.29}$$

$$r'_y = r_x\sin(\theta) + r_y\cos(\theta) \tag{1.30}$$

Written this way, we see that the 'new' components r'_j are a mixture of the 'old' components r_j scaled by trigonometry terms that depend only on θ.

5.3 Rotation Matrix

Equations (1.29)-(1.30) can be packed into a single statement using *matrix notation*:

$$\begin{bmatrix} r'_1 \\ r'_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \tag{1.31}$$

Explicitly, we have made the associations

$$\begin{aligned} (\vec{r})' &= \begin{bmatrix} r'_x \\ r'_y \end{bmatrix} = \begin{bmatrix} r'_1 \\ r'_2 \end{bmatrix} \\ \vec{r} &= \begin{bmatrix} r_x \\ r_y \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \ , \end{aligned}$$

and the 'block of numbers' containing the trigonometry terms is called the *rotation* matrix, or *rotation operator*, denoted R:

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{1.32}$$

Whenever a matrix such as R occurs to the left of a vector such as $\vec{r}$ as in (1.31), there is an implied operation that 'applies' the matrix onto the vector by cross-multiplying certain components. This procedure is captured by writing (1.31) in component form:

$$r'_j = \sum_{k=1}^{2} R_{jk}\, r_k \tag{1.33}$$

6 Basis Vectors

6.1 Unit Vectors

Consider a vector $\vec{V}$ of dimension N, having magnitude V. A special vector $\hat{V}$, called a *unit vector*, is defined as $\vec{V}$ divided by its own magnitude:

$$\hat{V} = \frac{1}{V}\,\vec{V} \tag{1.34}$$

That is, a unit vector always has magnitude one, and points along the original vector. A vector of the form (1.34) is said to be *normalized.*

A more intuitive way to understand unit vectors is to rearrange (1.34) to write

$$\vec{V} = V\,\hat{V} \ ,$$

which says that a full vector $\vec{V}$ is the product of the magnitude V and the 'direction' unit vector $\hat{V}$.

Problem 1

What is the vector that bisects the angle between two vectors $\vec{U}$, $\vec{V}$?

6.2 Introduction to Basis Vectors

Consider an arbitrary vector $\vec{V}$ of dimension N. A curious way to express

$$\vec{V} = \langle V_1, V_2, V_3, \dots, V_N\rangle$$

is to fully pull apart each component so that $\vec{V}$ is the sum of N pure sub-vectors:

$$\begin{aligned} \vec{V} &= \langle V_1, 0, 0, \dots\rangle \\ &\quad + \langle 0, V_2, 0, \cdots\rangle + \langle \cdots, 0, V_3, 0, \cdots\rangle \\ &\quad + \cdots + \langle \cdots, 0, V_N\rangle \end{aligned}$$

Each sub-vector contains just one component V_j, which can be factored out of the sub-vector as a scalar. The sub-vectors that remain are called *basis vectors*, denoted $\hat{e}_j$.

$$\begin{aligned} \hat{e}_1 &= \langle 1, 0, 0, \cdots, 0\rangle \\ \hat{e}_2 &= \langle 0, 1, 0, \cdots, 0\rangle \\ \hat{e}_3 &= \langle 0, 0, 1, \cdots, 0\rangle \\ &\cdots \\ \hat{e}_N &= \langle 0, 0, 0, \cdots, 1\rangle \end{aligned} \tag{1.35}$$

There is one basis vector $\hat{e}_j$ for each of the N dimensions in which the vector is situated.

Cartesian Coordinates

In the Cartesian xy-plane, a vector is typically represented as $\vec{V} = \langle V_x, V_y \rangle$, suggesting basis vectors

$$\hat{e}_x = \langle 1, 0 \rangle$$
$$\hat{e}_y = \langle 0, 1 \rangle \ .$$

Note that the same notation extrapolates to three dimensions, in which case

$$\hat{e}_x = \langle 1, 0, 0 \rangle$$
$$\hat{e}_y = \langle 0, 1, 0 \rangle$$
$$\hat{e}_z = \langle 0, 0, 1 \rangle$$

are the basis vectors.

Orthogonality of Basis Vectors

Basis vectors are all mutually orthogonal by necessity. For two different basis vectors $\hat{e}_j$, $\hat{e}_k$, the *orthogonality relation* is

$$\hat{e}_j \cdot \hat{e}_k = 0 \ . \tag{1.36}$$

On the other hand, two of the same basis vector $\hat{e}_k$ obeys

$$\hat{e}_k \cdot \hat{e}_k = 1 \ . \tag{1.37}$$

In the general case, any set of basis vectors $\{\hat{e}_j\}$ that obeys (1.36), (1.37) is said to be *orthonormal*.

6.3 Linear Combinations

Having established the notion of basis vectors, we are free to an express arbitrary vector $\vec{V}$ as a *linear combination* of each $\hat{e}_j$, namely

$$\vec{V} = V_1 \, \hat{e}_1 + V_2 \, \hat{e}_2 + V_3 \, \hat{e}_3 + \cdots + V_N \, \hat{e}_N \ , \tag{1.38}$$

or in summation notation:

$$\vec{V} = \sum_{j=1}^{N} V_j \, \hat{e}_j \tag{1.39}$$

In the above, $\vec{V}$ can potentially point to any 'place' in the N-dimensional space in which it lives. Such a place is formally called a *vector space*.

Vector Component Isolation

One may 'solve' for the V_jth component in a vector $\vec{V}$ by exploiting the orthogonality relations (1.36), (1.37). Start with (1.38), and multiply any particular $\hat{e}_k$ into both sides:

$$\hat{e}_k \cdot \vec{V} = V_1 \, \hat{e}_k \cdot \hat{e}_1 + V_2 \, \hat{e}_k \cdot \hat{e}_2 + \cdots + V_N \, \hat{e}_k \cdot \hat{e}_N$$

Next, observe that *all except one* of the dot products on the right will cancel due to (1.37). The whole sum collapses to the term with $j = k$, namely $V_k \, \hat{e}_k \cdot \hat{e}_k$, simplifying to V_k. Formally, we have uncovered the obvious yet satisfying statement:

$$V_k = \vec{V} \cdot \hat{e}_k \tag{1.40}$$

With an explicit formula for the V_jth component of a vector, it's curious to see happens by replacing V_j in (1.39). Carrying this out, we can write a component-free way to reference a vector and its contents:

$$\vec{V} = \sum_{j=1}^{N} \left(\vec{V} \cdot \hat{e}_j \right) \hat{e}_j \tag{1.41}$$

Spanning the Vector Space

It's important to notice that a linear combination $\vec{V}$, with appropriate values of V_j, could represent *any* point in the N-dimensional *vector space* in which the vector is embedded. This is possible because the set of basis vectors $\{\hat{e}_j\}$ are said to *span* the vector space.

7 Change of Basis

Consider a two-dimensional vector $\vec{V} = \langle V_x, V_y \rangle$, naturally expressed as a linear combination in the Cartesian basis

$$\hat{e}_1 = \hat{x} = \langle 1, 0 \rangle$$
$$\hat{e}_2 = \hat{y} = \langle 0, 1 \rangle \ .$$

By convention, the Cartesian coordinate system is usually aligned with the edges of a rectangular sheet of paper or computer screen. The orientation of the coordinate system is of course arbitrary, and we must be free to *rotate* the basis vectors without 'physical' consequences.

Figure 1.5 shows a two-dimensional example with two sets of basis vectors $\{\hat{e}_j\}$, $\{\hat{u}_j\}$ embedded on the Cartesian plane. In particular, the basis vector $\hat{u}$ is rotated up from $\hat{x}$ by some arbitrary angle, and similarly $\hat{v}$ corresponds to $\hat{y}$ by the same angle. Any given linear combination $\vec{r}$ has a different representation in each basis.

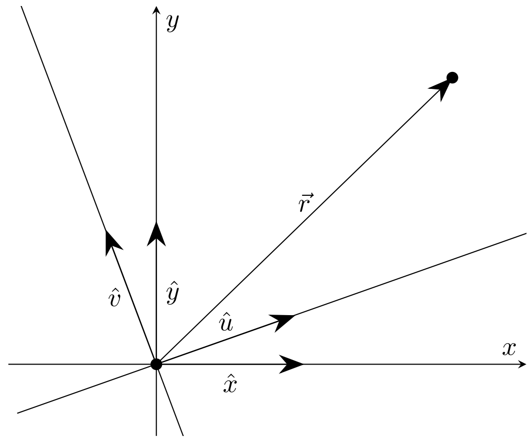

Figure 1.5: Vector $\vec{r}$ as a linear combination in two different bases.

Generalizing this idea to N diensions, we can say that linear combinations of the form (1.39) can be re-expressed in terms of a different set of orthonormal basis vectors $\{\hat{u}_j\}$:

$$(\vec{V})' = \sum_{j=1}^{N} V_j' \, \hat{u}_j \tag{1.42}$$

Analogous to (1.40), the primed components relate to the vector by:

$$V_k' = (\vec{V})' \cdot \hat{u}_k \tag{1.43}$$

7.1 Two Dimensions

Rotated Cartesian Coordinates

Suppose a different set basis vectors $\hat{u}$, $\hat{v}$ is given in terms of the original $\{\hat{e}_j\}$ basis, for instance

$$\hat{u}_1 = \hat{u} = \left\langle \frac{\sqrt{3}}{2}, \frac{1}{2} \right\rangle$$
$$\hat{u}_2 = \hat{v} = \left\langle \frac{-1}{2}, \frac{\sqrt{3}}{2} \right\rangle ,$$

or equivalently,

$$\hat{u}_1 = \frac{\sqrt{3}}{2}\, \hat{e}_1 + \frac{1}{2}\, \hat{e}_2$$
$$\hat{u}_2 = -\frac{1}{2}\, \hat{e}_1 + \frac{\sqrt{3}}{2}\, \hat{e}_2 .$$

Note that each $\hat{u}_j$ is a linear combination of each $\hat{e}_j$. The coefficients $\sqrt{3}/2$, $1/2$, etc. are carefully chosen to assure orthonormality between $\hat{u}_{1,2}$.

If a vector $\vec{r}$ is expressed in the $\{\hat{e}_j\}$ basis as the linear combination

$$\vec{r} = r_1\, \hat{e}_1 + r_2\, \hat{e}_2 ,$$

the so-called 'change of basis' occurs if we algebraically replace all $\hat{e}_j$ with $\hat{u}_j$, which first requires inverting the above relations:

$$\hat{e}_1 = \frac{\sqrt{3}}{2}\, \hat{u}_1 - \frac{1}{2}\, \hat{u}_2$$
$$\hat{e}_2 = \frac{1}{2}\, \hat{u}_1 + \frac{\sqrt{3}}{2}\, \hat{u}_2$$

Then, the vector $\vec{r}$ can be written

$$(\vec{r})' = r_1 \left(\frac{\sqrt{3}}{2}\, \hat{u}_1 - \frac{1}{2}\, \hat{u}_2 \right) + r_2 \left(\frac{1}{2}\, \hat{u}_1 + \frac{\sqrt{3}}{2}\, \hat{u}_2 \right)$$
$$(\vec{r})' = \left(r_1 \frac{\sqrt{3}}{2} + r_2 \frac{1}{2} \right) \hat{u}_1 + \left(-r_1 \frac{1}{2} + r_2 \frac{\sqrt{3}}{2} \right) \hat{u}_2 ,$$

where the components $r_{1,2}$ are finally readable as

$$r_1' = r_1 \frac{\sqrt{3}}{2} + r_2 \frac{1}{2}$$
$$r_2' = -r_1 \frac{1}{2} + r_2 \frac{\sqrt{3}}{2} ,$$

and a form like (1.42) is attained:

$$(\vec{r})' = r_1'\, \hat{u}_1 + r_2'\, \hat{u}_2$$

General Coordinate Rotations

The above example can be easily generalized such that $\hat{u}$ points anywhere in the Cartesian plane, with $\hat{v}$ appropriately perpendicular to $\hat{u}$. To achieve this, we introduce an arbitrary parameter θ such that

$$\hat{u}_1 = \cos(\theta)\, \hat{e}_1 + \sin(\theta)\, \hat{e}_2$$
$$\hat{u}_2 = -\sin(\theta)\, \hat{e}_1 + \cos(\theta)\, \hat{e}_2 .$$

By straightforward algebra, we find the inverted version to be

$$\hat{e}_1 = \cos(\theta)\, \hat{u}_1 - \sin(\theta)\, \hat{u}_2$$
$$\hat{e}_2 = \sin(\theta)\, \hat{u}_1 + \cos(\theta)\, \hat{u}_2 .$$

The pairs of equations above are suggestive of a matrix formulation, particularly

$$\begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \hat{e}_1 \\ \hat{e}_2 \end{bmatrix} , \tag{1.44}$$

and

$$\begin{bmatrix} \hat{e}_1 \\ \hat{e}_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} , \tag{1.45}$$

respectively. Comparing the above to (1.32), we see (1.45) contains the same rotation matrix R that rotates vectors in a fixed basis. Denoting the other matrix in (1.44) as $\tilde{R}$, the component version of the above reads:

$$\hat{u}_j = \sum_{k=1}^{2} \tilde{R}_{jk}\hat{e}_k$$
$$\hat{e}_j = \sum_{k=1}^{2} R_{jk}\hat{u}_k \; .$$

With a convenient representation for the basis vectors $\{\hat{e}_j\}$, an arbitrary linear combination

$$\vec{r} = r_1\,\hat{e}_1 + r_2\,\hat{e}_2$$

becomes

$$\begin{aligned}(\vec{r})' &= r_1 \sum_{k=1}^{2} R_{1k}\,\hat{u}_k + r_2 \sum_{k=1}^{2} R_{2k}\,\hat{u}_k \\ &= \sum_{k=1}^{2}\left(\sum_{j=1}^{2} R_{jk}\,r_j\right)\hat{u}_k \; ,\end{aligned}$$

telling us that the kth component of the vector $(\vec{r})'$ is given by

$$r'_k = \sum_{j=1}^{2} R_{jk}\,r_j \; , \tag{1.46}$$

which is a cousin to equation (1.33). To do a fair comparison, let us swap the j-index and the k-index in (1.46) to write

$$r'_j = \sum_{k=1}^{2} R_{kj}\,r_k \; .$$

Looking carefully, the above differs from (1.33) by the order of the subscripts on the R-term, ultimately equivalent to reversing the sign on θ. Said another way, a 'positive' rotation in the basis vectors with $\vec{r}$ fixed is equivalent to a 'negative' rotation of $\vec{r}$ with the basis fixed.

7.2 N Dimensions

Change of Basis Vectors

At the center of the change-of-basis problem is the issue of relating the two orthonormal bases $\hat{e}_j$, $\hat{u}_j$ to one another. In N dimensions, the basis vectors are related by linear combinations

$$\hat{u}_j = \sum_{k=1}^{N} \tilde{U}_{jk}\,\hat{e}_k \tag{1.47}$$
$$\hat{e}_j = \sum_{k=1}^{N} U_{jk}\,\hat{u}_k \; . \tag{1.48}$$

Having two subscripts, the terms $\tilde{U}_{jk}$, U_{jk} are not vector components, but instead *matrix* components. These typically end up being coefficients like $\sqrt{3}/2$, $1/2$, and so on.

To isolate the matrix components U_{jk} and $\tilde{U}_{jk}$, multiply (via dot product) the basis vectors $\hat{e}_m$, $\hat{u}_m$, respectively into (1.47), (1.48):

$$\hat{e}_m \cdot \hat{u}_j = \sum_{k=1}^{N} \tilde{U}_{jk}\,\hat{e}_m \cdot \hat{e}_k$$
$$\hat{u}_m \cdot \hat{e}_j = \sum_{k=1}^{N} U_{jk}\,\hat{u}_m \cdot \hat{u}_k$$

Due to orthonormality, the right side of each equation resolves to zero except for the case with $m = k$, allowing the components to be isolated:

$$\tilde{U}_{jm} = \hat{e}_m \cdot \hat{u}_j = \hat{u}_j \cdot \hat{e}_m \tag{1.49}$$
$$U_{jm} = \hat{u}_m \cdot \hat{e}_j = \hat{e}_j \cdot \hat{u}_m \tag{1.50}$$

From this, deduce also that

$$\tilde{U}_{jm} = U_{mj} \; . \tag{1.51}$$

Linear Combinations

An arbitrary linear combination

$$\vec{V} = \sum_{j=1}^{N} V_j\,\hat{e}_j$$

can be written with all $\hat{e}_j$ replaced according to (1.48):

$$(\vec{V})' = \sum_{j=1}^{N}\sum_{k=1}^{N} U_{jk}\,V_j\,\hat{u}_k = \sum_{k=1}^{N}\left(\sum_{j=1}^{N} U_{jk}\,V_j\right)\hat{u}_k$$

Comparing the above to (1.42) gives a formula for the kth component of the vector $(\vec{V})'$:

$$(V')_k = \sum_{j=1}^{N} U_{jk}\,V_j \tag{1.52}$$

Unity Condition

We can learn a bit more about the matrix components U_{jk}, $\tilde{U}_{jk}$ by eliminating $\hat{u}_k$ between (1.47)-(1.48):

$$\begin{aligned}\hat{e}_j &= \sum_{k=1}^{N} U_{jk}\left(\sum_{m=1}^{N} \tilde{U}_{km}\,\hat{e}_m\right) \\ &= \sum_{k=1}^{N}\sum_{m=1}^{N}\left(U_{jk}\tilde{U}_{km}\right)\hat{e}_m\end{aligned}$$

Next, use the symbol I to represent the quantity

$$I_{jm} = \sum_{k=1}^{N} \left(U_{jk} \tilde{U}_{km} \right) ,$$

and the above becomes

$$\hat{e}_j = \sum_{m=1}^{N} I_{jm} \, \hat{e}_m \, .$$

For the above to make sense, all terms in the right-hand sum must vanish except for that with $m = j$. Explicitly, this means I_{jm} obeys

$$I_{jm} = \begin{cases} 1 & m = j \\ 0 & m \neq j \end{cases} ,$$

reminiscent of (1.36)-(1.37).

8 Vectors and Limits

Vectors, whose components are numbers and functions, obey all of the established properties of limits. While the technical proof for this is attainable, it's worth staving off a formal effort for multivariate calculus, and for now take it on intuition that vectors and limits get along nicely.

8.1 Pi from Nested Radicals

Here we apply vectors to a curious problem that approximates the value of π by covering unit circle with triangles of known area. Figure 1.6 shows the first quadrant of the unit circle with several lines and points labeled to aid the derivation. The origin is at what would be the center of the complete circle.

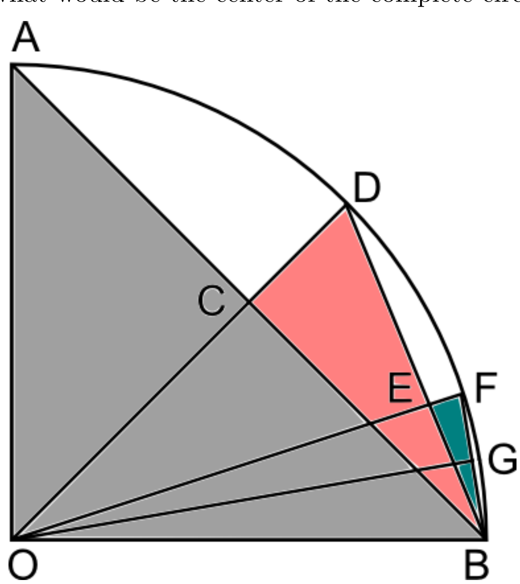

Figure 1.6: Covering a quarter circle with triangles.

To establish some notation, let the lines OA, OB define a pair of unit vectors:

$$\hat{i} = \overline{OB}$$
$$\hat{j} = \overline{OA}$$

Also, let the hypotenuse of AOB be a vector $\vec{h}_0$ such that

$$\vec{h}_0 = \hat{i} - \hat{j} \, ,$$

with magnitude

$$\left| \vec{h}_0 \right| = \sqrt{2} \, .$$

Order-Zero Triangle (1)

The largest triangle that fits in the quarter unit circle is depicted AOB, whose area is $1/2$. Using the vectors on hand, the area of AOB shall be written

$$A_0 = \frac{1}{2} \left| \hat{i} \right| \left| \hat{j} \right| \, ,$$

which is a fancy way to write $1/2$.

Order-One Triangle (2)

Next, we seek two identical triangles that cover the largest uncovered portion of the quarter circle. In Figure 1.6 these are depicted DCA, DCB, respectively.

Analyzing DCB, consider a unit vector $\hat{x}_1$ and a shorter vector $\vec{x}_1$ such that

$$\vec{x}_1 = \overline{OC} = \frac{\hat{i} + \hat{j}}{2}$$
$$\hat{x}_1 = \overline{OD} = \frac{\hat{i} + \hat{j}}{\sqrt{2}} \, ,$$

whose difference in length is CD.

Then, the area of DCB is

$$A_1 = \frac{(CD)\,(CB)}{2} = \frac{1}{2} \left| \hat{x}_1 - \vec{x}_1 \right| \frac{1}{2} \left| \vec{h}_0 \right| \, .$$

That is, the base is line CB, whose length is half the magnitude $\vec{h}_0$. The height CB is given by the difference in x-vectors.

Notice that, because $\hat{x}_1$, $\vec{x}_1$ are parallel, the following simplification can be made:

$$\left| \hat{x}_1 - \vec{x}_1 \right| = 1 - x_1$$

Calculating out A_1, one finds, after some algebra:

$$A_1 = \frac{1}{2} \left(-\frac{1}{2} + \frac{1}{\sqrt{2}} \right)$$

The hypotenuse of DCB is denoted $\vec{h}_1$ and is given by

$$\vec{h}_1 = \hat{i} - \hat{x}_1 ,$$

and has length DB. Note that the vector $\vec{h}_1$ doesn't play into the area of DCB, rather the previous $\vec{h}_0$ is used. Explicitly, the vector $\vec{h}_1$ reads

$$\vec{h}_1 = \left(1 - \frac{1}{\sqrt{2}}\right)\hat{i} + \frac{1}{\sqrt{2}}\,\hat{j} ,$$

having magnitude

$$\left|\vec{h}_1\right| = \sqrt{2 - \sqrt{2}} .$$

Order-Two Triangles (4)

To keep covering the circle, we'll take four copies of triangle FEB as shown in the Figure. To find the area of just FEB, first notice

$$\begin{aligned}\vec{x}_2 &= \overline{OE} = \hat{x}_1 + \frac{1}{2}\vec{h}_1 \\ \hat{x}_2 &= \overline{OF} = \vec{x}_2 / \left|\vec{x}_2\right| ,\end{aligned}$$

whose difference in length is EF. In detail, it's straightforward to show that

$$\vec{x}_2 = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right)\hat{i} + \frac{1}{2\sqrt{2}}\,\hat{j} ,$$

where

$$x_2 = \frac{1}{2}\sqrt{2 + \sqrt{2}} ,$$

and

$$\hat{x}_2 = \left(\frac{1 + 1/\sqrt{2}}{\sqrt{2 + \sqrt{2}}}\right)\hat{i} + \left(\frac{\sqrt{2 - \sqrt{2}}}{2}\right)\hat{j} .$$

The area of FEB is

$$\begin{aligned}A_2 &= \frac{(EF)\,(EB)}{2} \\ &= \frac{1}{2}\left|\hat{x}_2 - \vec{x}_2\right| \frac{1}{2}\left|\vec{h}_1\right| \\ &= \frac{1}{4}\left(1 - x_2\right)\left|\vec{h}_1\right| ,\end{aligned}$$

which plays much like the previous case with all indices shifted up by one. Calculating out A_2, one finds, after a lot of algebra:

$$A_2 = \frac{1}{4}\left(-\frac{1}{\sqrt{2}} + \sqrt{2 - \sqrt{2}}\right)$$

The hypotenuse of FEB is $\vec{h}_2$, given by

$$\vec{h}_2 = \hat{i} - \hat{x}_2 ,$$

which is also just like the the formula for $\vec{h}_1$ with the indices bumped by one. Explicitly, the vector $\vec{h}_2$ reads

$$\vec{h}_2 = \left(1 - \frac{\left(1 + 1/\sqrt{2}\right)}{\sqrt{2 + \sqrt{2}}}\right)\hat{i} + \left(\frac{\sqrt{2 - \sqrt{2}}}{2}\right)\hat{j} ,$$

having magnitude

$$\left|\vec{h}_2\right| = \sqrt{2 - \sqrt{2 + \sqrt{2}}} .$$

Order-Three Triangles (8)

By now we're running out of letters in the Figure, but the pattern continues. The next step has eight total triangles. Begin with

$$\begin{aligned}\vec{x}_3 &= \hat{x}_2 + \frac{1}{2}\vec{h}_2 \\ \hat{x}_3 &= \vec{x}_3 / \left|\vec{x}_3\right| ,\end{aligned}$$

implying the area to be

$$\begin{aligned}A_3 &= \frac{1}{2}\left|\hat{x}_3 - \vec{x}_3\right| \frac{1}{2}\left|\vec{h}_2\right| \\ &= \frac{1}{4}\left(1 - x_3\right)\left|\vec{h}_2\right| ,\end{aligned}$$

and furthermore:

$$\vec{h}_3 = \hat{i} - \hat{x}_3$$

Leaving the algebra to the dedicated reader, the area A_3 resolves to:

$$A_3 = \frac{1}{8}\left(-\sqrt{2 - \sqrt{2}} + 2\sqrt{2 - \sqrt{2 + \sqrt{2}}}\right)$$

Order-N Triangles

It will take an infinite number of iterations to cover the entire quarter circle with increasingly smaller triangles. At the nth step, it follows that

$$\begin{aligned}\vec{x}_n &= \hat{x}_{n-1} + \frac{1}{2}\vec{h}_{n-1} \\ \hat{x}_n &= \vec{x}_n / \left|\vec{x}_n\right| \\ \vec{h}_n &= \hat{i} - \hat{x}_n ,\end{aligned}$$

with

$$A_n = \frac{1}{4}\left(1 - x_n\right)\left|\vec{h}_{n-1}\right| .$$

While we'll take the above as a workable result, note that

$$\left|\vec{h}_{n-1}\right| = \sqrt{2\left(1 - \hat{i} \cdot \hat{x}_{n-1}\right)} .$$

With this, the area simplifies to

$$A_n = \frac{\sqrt{2}}{4}\left(1 - x_n\right)\sqrt{1 - \hat{i}\cdot\hat{x}_{n-1}}\ ,$$

which only works for $n \geq 2$.

Also, it's easy to derive

$$\hat{i}\cdot\hat{x}_{n-1} = \cos\left(\frac{\pi}{2^n}\right)$$

from the vectors on hand. It would be bad form, however, to invoke π in the midst of trying to calculate it, so we'll leave the cosine function alone.

Working out the area A_4 from the above, which is absolutely tedious without a machine, one finds:

$$A_4 = -\frac{2\sqrt{2-\sqrt{2+\sqrt{2}}}}{16} + \frac{4\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2}}}}}{16}$$

N-Sided Polygon

Tallying all areas of all triangles up to the Nth step gives one quarter the area of an N-sided polygon with equal angles and equal sides. Doing so, we write

$$P\left(N\right) = \sum_{n=0}^{N} w_n A_n\ ,$$

where A is the total area, A_n is the area of *one* triangle of order n, and w_n is the number of triangles of order n, particularly $w_n = 2^n$.

Condensing variables again, we also write, for $n \geq 2$:

$$P\left(N\right) = \sum_{n=0}^{N} \frac{2^n}{4}\left(1 - x_n\right)\left|\vec{h}_{n-1}\right| = \sum_{n=0}^{N} P_n$$

Evaluating P(N)

Now the real work begins. Starting with $N = 0$ and working up, we find

$$P_0 = 2^0 A_0 = \frac{1}{2}$$
$$P_1 = 2^1 A_1 = -\frac{1}{2} + \frac{1}{\sqrt{2}}$$
$$P_2 = 2^2 A_2 = -\frac{1}{\sqrt{2}} + \sqrt{2-\sqrt{2}}$$
$$P_3 = 2^3 A_3 = -\sqrt{2-\sqrt{2}} + 2\sqrt{2-\sqrt{2+\sqrt{2}}}\ ,$$

along with

$$P_4 = 2^4 A_4 = -2\sqrt{2-\sqrt{2+\sqrt{2}}} + 4\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2}}}}$$

Now an amazing simplification happens. Calculating the sum $P\left(N\right)$ requires adding all terms P_n up to the Nth term. However, notice that each P_n contains a positive term and a negative term. The negative term is always the exact negative of the previous positive term. The end result is, only the positive term in P_N survives the summation.

Going from the pattern on hand, we evidently have

$$P\left(N\right) = \frac{2^N}{4}\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2+\cdots}}}}$$

with N total square roots.

Area of N-Sided Polygon

Define Π (uppercase of π) such that

$$\Pi\left(N\right) = 4P\left(N\right)\ ,$$

which is the area of an N-sided polygon (all four quadrants).

For the first few orders, a calculator reveals:

$$\begin{aligned}
\Pi\left(0\right) &= 2\\
\Pi\left(1\right) &= 2.8284271248\ldots\\
\Pi\left(2\right) &= 3.0614674589\ldots\\
\Pi\left(3\right) &= 3.1214451523\ldots\\
\Pi\left(4\right) &= 3.1365484905\ldots\\
\Pi\left(5\right) &= 3.1403311570\ldots\\
\Pi\left(10\right) &= 3.1415914215\ldots\\
\Pi\left(15\right) &= 3.1415926524\ldots\\
\Pi\left(20\right) &= 3.1415926536\ldots
\end{aligned}$$

The area becomes suspiciously close to π as the number of iterations increases, and we get about ten digits of π after $N = 18$ iterations.

The number of triangles for the quarter-area is given by

$$W\left(N\right) = \sum_{n=0}^{N} w_n = \sum_{n=0}^{N} 2^n = 2^{N+1} - 1\ ,$$

and for $N = 18$, we approximately have

$$W\left(18\right) = 2^{19} - 1\ .$$

The total number of triangles is four times the above. Over two million triangles are needed to get π to ten digits:

$$4W(18) = 2^{21} - 4 = 2097148$$

Area of Unit Circle

In the limit $N \to \infty$, the triangles cover the unit circle and the area converges to π:

$$\pi = \lim_{N\to\infty} \Pi(N) \ ,$$

or

$$\pi = \lim_{N\to\infty} 2^N \sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots}}}} \ .$$

There are N square roots on the right.

This result is a bit counter-intuitive in the sense that 2^N tends to infinity while the quantity under the outermost square root goes to zero. It just happens that infinity times zero, in this particular limit, equals π.

9 Matrix Formalism

Formally, a *matrix* is a collection of numbers or variables arranged in a block with fixed rows M and columns N. Each element, i.e. *component* in the matrix requires two subscripts.

9.1 Matrix-Operator Equivalence

A primary use for a matrix is to 'operate' on a vector of dimension N, yielding a new vector of dimension M. (The term 'matrix' is often interchanged with the term 'operator'.) Symbolically, this is written

$$A\vec{x} = \vec{y} \ ,$$

and in full *block notation*, the same statement looks like

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2N} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{M1} & A_{M2} & A_{M3} & \cdots & A_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdots \\ y_M \end{bmatrix}$$

More compactly, we use *index notation* to express the same calculation:

$$\sum_{k=1}^{N} A_{jk}\, x_k = y_j \qquad (1.53)$$

$$j = 1, 2, 3, \ldots, M$$

9.2 Matrix Components

Consider two vectors $\vec{x}$, $\vec{y}$, each a linear combination in some N-dimensional basis such that

$$\vec{x} = \sum_{j=1}^{N} x_j\, \hat{e}_j$$

$$\vec{y} = \sum_{j=1}^{M} y_j\, \hat{e}_j \ .$$

While $\vec{y}$ is perfectly happy being expressed as a linear combination in the basis $\{\hat{e}_j\}$, it's instructive to re-express $\vec{y}$ in terms of its brother, $\vec{x}$. To do so, we propose an operator A such that

$$\vec{y} = A\vec{x} \ .$$

To proceed, write the above as

$$\sum_{k=1}^{M} y_k\, \hat{e}_k = \sum_{k=1}^{N} x_k\, A\hat{e}_k \ ,$$

and multiply the basis vector $\hat{e}_j$ (via dot product) into both sides:

$$\sum_{k=1}^{M} y_k\, \hat{e}_j \cdot \hat{e}_k = \sum_{k=1}^{N} x_k\, \hat{e}_j \cdot A\hat{e}_k$$

On the left, every term in the sum vanishes except that with $j = k$, and the above becomes

$$y_j = \sum_{k=1}^{N} (\hat{e}_j \cdot A\hat{e}_k)\, x_k$$

$$j = 1, 2, 3, \ldots, M \ .$$

The parenthesized quantity is what we're after:

$$A_{jk} = \hat{e}_j \cdot A\hat{e}_k \qquad (1.54)$$

The term A_{jk} is the component of the matrix A corresponding to the jth row, kth column.

9.3 Projector

Consider the curious quantity

$$P_x = \vec{x}\,\vec{x} \ , \qquad (1.55)$$

called the the *projector* of $\vec{x}$. By itself, P_x does nothing - there is no operation between the two copies of $\vec{x}$. What the projector *does* is 'wait' to be multiplied into another vector, resulting in a scaled version of $\vec{x}$. For example, applying the projector to a different vector $\vec{y}$ (of the same dimension as $\vec{x}$) goes like

$$P_x\vec{y} = \vec{x}\,(\vec{x} \cdot \vec{y}) \ .$$

9.4 Identity Operator

Consider a vector $\vec{x}$ as a linear combination in some N-dimensional basis:

$$\vec{x} = \sum_{j=1}^{N} x_j \, \hat{e}_j$$

For any one of the basis vectors $\hat{e}_k$, write the projector

$$P_{e_k} = \hat{e}_k \, \hat{e}_k \, ,$$

and then multiply $\vec{x}$ onto the right side to get

$$P_{e_k} \vec{x} = \hat{e}_k \, \hat{e}_k \cdot \vec{x} = x_k \, \hat{e}_k$$

By summing over the index k, the right side is identically $\vec{x}$:

$$\left(\sum_{k=1}^{N} P_{e_k} \right) \vec{x} = \sum_{k=1}^{N} x_k \, \hat{e}_k = \vec{x}$$

For the left side to also equal $\vec{x}$, the parenthesized quantity must be equivalent to 'multiplying by one', which we call the *identity* operator:

$$I = \sum_{k=1}^{N} P_{e_k} \tag{1.56}$$

The identity operator leaves a vector unchanged:

$$I \vec{x} = \vec{x}$$

The matrix-equivalence of I is square, has no mixed components, and has ones along the diagonal:

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{1.57}$$

9.5 Unified Matrix Notation

Recall that a matrix A relates to its components A_{jk} in a way given by (1.54), namely

$$A_{jk} = \hat{e}_j \cdot A \hat{e}_k \, .$$

To establish this directly using projectors, start with $A = IAI$ and watch what happens:

$$\begin{aligned} A = IAI &= \sum_{j=1}^{M} \sum_{k=1}^{N} P_{e_j} A P_{e_k} \\ &= \sum_{j=1}^{M} \sum_{k=1}^{N} \hat{e}_j \left(\hat{e}_j \cdot A \hat{e}_k \right) \hat{e}_k \end{aligned}$$

The parenthesized quantity is precisely A_{jk}. Evidently, the symbolic notation unifies with the index notation in the equation

$$A = \sum_{j=1}^{M} \sum_{k=1}^{N} \hat{e}_j \left(A_{jk} \right) \hat{e}_k \tag{1.58}$$

The presence of $\hat{e}_j \, \hat{e}_k$ is like a projector - it couples the component to the operator.

10 Matrix Operations

10.1 Matrix Addition

Two matrices A and B of identical dimensions, meaning M rows, N columns, can be combined to form a new matrix C such that

$$A + B = C \, ,$$

or, to elaborate:

$$A_{jk} + B_{jk} = C_{jk} \tag{1.59}$$
$$\begin{cases} j = 1, 2, 3, \ldots, M \\ k = 1, 2, 3, \ldots, N \end{cases}$$

10.2 Scalar Multiplication

A scalar α can be multiplied into each component of a matrix A to form a new matrix B such that

$$\alpha A = B \, ,$$

or:

$$\alpha A_{jk} = B_{jk} \tag{1.60}$$
$$\begin{cases} j = 1, 2, 3, \ldots, M \\ k = 1, 2, 3, \ldots, N \end{cases}$$

10.3 Matrix Multiplication

Two matrices A, B, of equal or different dimensions can be multiplied to form a new matrix C:

$$AB = C$$

The main 'rule' is that the number of *columns* in A must equal the number of *rows* in B:

$$A_{(M,K)} \times B_{(K,N)} = C_{(M,N)}$$

Matrix Non-Commutativity

If you're paying attention, the commutated product BA may violate the above, and no product is defined. In any case, we should assume that the multiplication of two matrices is not commutative:

$$AB \neq BA \tag{1.61}$$

Multiplication Formula

To derive the formula for matrix multiplication, begin with the following 'unified' representation (1.58) of the respective matrices:

$$A = \sum_{m=1}^{M} \sum_{k=1}^{K} \hat{e}_m \left(A_{mk}\right) \hat{e}_k$$

$$B = \sum_{k'=1}^{K} \sum_{n=1}^{N} \hat{e}_{k'} \left(B_{k'n}\right) \hat{e}_n$$

Then, the product AB reads

$$\begin{aligned} AB &= \sum_{m=1}^{M} \sum_{k=1}^{K} \hat{e}_m \left(A_{mk}\right) \hat{e}_k \sum_{k'=1}^{K} \sum_{n=1}^{N} \hat{e}_{k'} \left(B_{k'n}\right) \hat{e}_n \\ &= \sum_{m=1}^{M} \sum_{k=1}^{K} \sum_{k'=1}^{K} \sum_{n=1}^{N} \hat{e}_m \left(A_{mk}\right) \hat{e}_k \, \hat{e}_{k'} \left(B_{k'n}\right) \hat{e}_n \end{aligned}$$

Note that the quantity $\hat{e}_m \; \hat{e}_k \; \hat{e}_{k'} \; \hat{e}_n$ is the juxtaposition of two projectors, readily translating to $\hat{e}_m \left(\hat{e}_k \cdot \hat{e}_{k'}\right) \hat{e}_n$. Note further that the parenthesized product obeys (1.36)-(1.37), namely

$$\hat{e}_k \cdot \hat{e}_{k'} = \begin{cases} 1 & k = k' \\ 0 & k \neq k' \end{cases} ,$$

which has the effect of equating $k = k'$ in the above, eliminating one of the sums. So far then, we have

$$AB = C = \sum_{m=1}^{M} \sum_{k=1}^{K} \sum_{n=1}^{N} \left(A_{mk} B_{kn}\right) \hat{e}_m \; \hat{e}_n$$

$$C = \sum_{m=1}^{M} \sum_{n=1}^{N} \left(\sum_{k=1}^{K} A_{mk} B_{kn} \right) \hat{e}_m \; \hat{e}_n \; .$$

The symbol C has replaced the quantity AB on the left. Comparing the right side to (1.58), we conclude that the component C_{mn} of matrix C is given by the famed *matrix multiplication* formula:

$$C_{mn} = \sum_{k=1}^{K} A_{mk} B_{kn} \tag{1.62}$$

$$\begin{cases} m = 1, 2, 3, \ldots, M \\ n = 1, 2, 3, \ldots, N \end{cases}$$

Equation (1.62) reminds that it's only required that the number of columns in A match the number of rows in B. For instance, the operation $A_{(2,4)} \times B_{(4,3)} = C_{(2,3)}$, explicitly written as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

is perfectly valid, whereas the commuted product $B_{(4,3)} \times A_{(2,4)}$ is undefined.

Matrix Associativity

A direct consequence of matrix multiplication is the associativity rule:

$$\left(AB\right) C = A \left(BC\right) \tag{1.63}$$

10.4 Change of Basis

A square matrix A with components A_{jk} in the basis $\{\hat{e}_j\}$ can be represented by (1.58):

$$A = \sum_{j=1}^{N} \sum_{k=1}^{N} \hat{e}_j \left(A_{jk}\right) \hat{e}_k \; .$$

Under a change of basis $\{\hat{e}_j\} \rightarrow \{\hat{u}_j\}$, we can use (1.48)

$$\hat{e}_j = \sum_{k=1}^{N} \tilde{U}_{jk} \; \hat{u}_k$$

to replace the unit vectors, leading to

$$A' = \sum_{m=1}^{N} \sum_{n=1}^{N} \hat{u}_m \left(\sum_{j=1}^{M} \sum_{k=1}^{N} U_{mj} A_{jk} \tilde{U}_{kn} \right) \hat{u}_n \; ,$$

where the (first) term $\tilde{U}_{jm}$ has been replaced by U_{mj} due to (1.51). The parenthesized quantity is precisely the formula for the component A'_{mn} of the transformed matrix

$$A'_{mn} = \sum_{j=1}^{N} \sum_{k=1}^{N} U_{mj} A_{jk} \tilde{U}_{kn} \; , \tag{1.64}$$

or in symbolic form,

$$A' = U A \tilde{U} \; .$$

Note that the above verifies the associativity rule (1.63) for matrix multiplication. The order in which the sums are taken directly corresponds to which matrices are multiplied first. As a bonus, (1.64) tells us exactly how to take the product of three square matrices.