Appendix B

Affine Algebra

B.1 Introduction

As we saw earlier, linear algebra is the study of vectors and vector spaces. In two dimensions, a vector was treated as a quantity with direction and magnitude. It does not matter where you place the vector in the plane; it represents the same vector (see Figure A.4) since the directions and magnitudes are the same. In physics applications, among others, the location of the vector, that is, where its initial point is, may very well be significant. For example, if a particle has a certain velocity at a given instant, the velocity vector necessarily applies to the *position* of the particle at that instant. Similarly, the same force applied to two different positions on a rod have different effects on the rod. The *point* of application of the force is relevant.

Clearly, there needs to be a distinction between *points* and *vectors*. This is the essence of *affine algebra*. Let V be a vector space of dimension n. Let A be a set of elements that are called *points*. A is referred to as an *n–dimensional affine space* whenever the following conditions are met:

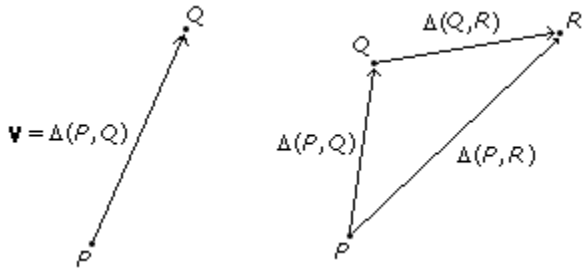
1. For each ordered pair of points $\mathcal{P}, \mathcal{Q} \in A$, there is a unique vector in V called the *difference vector* and denoted by $\Delta(\mathcal{P}, \mathcal{Q})$.

2. For each point $\mathcal{P} \in A$ and $\mathbf{v} \in V$, there is a unique point $\mathcal{Q} \in A$ such that $\mathbf{v} = \Delta(\mathcal{P}, \mathcal{Q})$.

3. For any three points $\mathcal{P}, \mathcal{Q}, \mathcal{R} \in A$, it must be that $\Delta(\mathcal{P}, \mathcal{Q}) + \Delta(\mathcal{Q}, \mathcal{R}) = \Delta(\mathcal{P}, \mathcal{R})$.

Figure B.1 illustrates these three items.

Figure B.1. *(a) A vector* $\mathbf{v}$ *connecting two points* $\mathcal{P}$ *and* $\mathcal{Q}$*. (b) The sum of vectors, each vector determined by two points.*



If $\mathcal{P}$ and $\mathcal{Q}$ are specified, $\mathbf{v}$ is uniquely determined (item 1). If $\mathcal{P}$ and $\mathbf{v}$ are specified, $\mathcal{Q}$ is uniquely determined (item 2). Figure B.1(b) illustrates item 3.

The formal definition for an affine space introduced the difference vector $\Delta(\mathcal{P}, \mathcal{Q})$. Figure B.1 gives you the geometric intuition about the difference, specifically that it appears to be a subtraction operation for two points. However, certain consequences of the definition may be proved directly without having a concrete formulation for an actual subtraction of points.

A few consequences of the definition for an affine algebra follow.

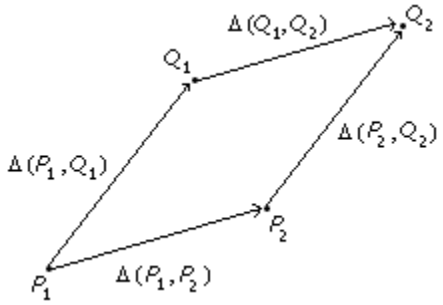
1. $\Delta(\mathcal{P}, \mathcal{P}) = \mathbf{0}$
2. $\Delta(\mathcal{Q}, \mathcal{P}) = -\Delta(\mathcal{P}, \mathcal{Q})$
3. If $\Delta(\mathcal{P}_1, \mathcal{Q}_1) = \Delta(\mathcal{P}_2, \mathcal{Q}_2)$, then $\Delta(\mathcal{P}_1, \mathcal{P}_2) = \Delta(\mathcal{Q}_1, \mathcal{Q}_2)$

The first consequence follows immediately from item 3 in the definition where $\mathcal{Q}$ is replaced by $\mathcal{P}$, $\Delta(\mathcal{P}, \mathcal{P}) + \Delta(\mathcal{P}, \mathcal{R}) = \Delta(\mathcal{P}, \mathcal{R})$. The vector $\Delta(\mathcal{P}, \mathcal{R})$ is subtracted from both sides to obtain $\Delta(\mathcal{P}, \mathcal{P}) = \mathbf{0}$.

The second consequence also follows from item 3 in the definition where $\mathcal{R}$ is replaced by $\mathcal{P}$, $\Delta(\mathcal{P}, \mathcal{Q}) + \Delta(\mathcal{Q}, \mathcal{P}) = \Delta(\mathcal{P}, \mathcal{P}) = \mathbf{0}$. The last equality is what we just proved in the previous paragraph. The first vector is subtracted from both sides to obtain $\Delta(\mathcal{Q}, \mathcal{P}) = -\Delta(\mathcal{P}, \mathcal{Q})$.

The third consequence is called the *parallelogram law*. Figure B.2 illustrates.

Figure B.2. *The parallelogram law for affine algebra.*



Item 3 in the definition can be applied in two ways:

$$\Delta(\mathcal{P}_1,\mathcal{P}_2)+\Delta(\mathcal{P}_2,\mathcal{Q}_2)=\Delta(\mathcal{P}_1,\mathcal{Q}_2) \text{ and } \Delta(\mathcal{P}_1,\mathcal{Q}_1)+\Delta(\mathcal{Q}_1,\mathcal{Q}_2)=\Delta(\mathcal{P}_1,\mathcal{Q}_2)$$

Subtracting these leads to

$$\mathbf{0}=\Delta(\mathcal{P}_1,\mathcal{P}_2)+\Delta(\mathcal{P}_2,\mathcal{Q}_2)-\Delta(\mathcal{P}_1,\mathcal{Q}_1)-\Delta(\mathcal{Q}_1,\mathcal{Q}_2)=\Delta(\mathcal{P}_1,\mathcal{P}_2)-\Delta(\mathcal{Q}_1,\mathcal{Q}_2)$$

where the last equality is valid since we assumed $\Delta(\mathcal{P}_1,\mathcal{Q}_1)=\Delta(\mathcal{P}_2,\mathcal{Q}_2)$. Therefore, $\Delta(\mathcal{P}_1,\mathcal{P}_2)=\Delta(\mathcal{Q}_1,\mathcal{Q}_2)$.

In the formal sense of affine algebra, points and vectors are distinct entities. We have already used two different fonts to help distinguish between them: $\mathcal{P}$ is a point, $\mathbf{v}$ is a vector. Even so, the following example shows how powerful the formal setting is. Given a vector space V, the points may be defined as the elements of V themselves, namely $A=V$. If $\mathcal{P}$ and $\mathcal{Q}$ are points, the corresponding vectors are labeled $\mathbf{p}$ and $\mathbf{q}$. Think of the vectors geometrically as $\mathcal{P}-\mathcal{O}$ and $\mathcal{Q}-\mathcal{O}$, respectively, where $\mathcal{O}$ is the origin. The difference vector of the points is $\Delta(\mathcal{P},\mathcal{Q})=\mathbf{q}-\mathbf{p}$, a subtraction of the two vectors. The three items in the definition for affine space can be easily verified. The example also shows that you must be steadfast in maintaining that points and vectors are different abstract quantities, even if you happen to represent them in a computer program in the same way, say as n–tuples of numbers.

Any further discussion of affine spaces in the abstract will continue to use the notation $\Delta(\mathcal{P},\mathcal{Q})$ for the vector difference of two points. However, in situations that have a computational flavor, we will instead use the more intuitive notation $\mathcal{Q}-\mathcal{P}$ with the warning that the subtraction operator is a convenient notation that is not necessarily indicative of the actual mechanism used to compute the vector difference of two points

in a particular application. We also use the suggestive notation $\mathcal{Q} = \mathcal{P} + \mathbf{v}$ when $\mathbf{v} = \Delta(\mathcal{P}, \mathcal{Q})$.

To avoid inadvertent confusion between points and vectors in a computer implementation, separate data structures for points and vectors are recommended. For example, in C++ you can define the vector class by

```
template class <T real, int n> Vector
{
public:
    // construction
    Vector ();
    Vector (const real tuple[n]);
    Vector (const Vector& v);

    // tuple access as an array
    operator const real* () const;
    operator real* ();
    real operator[] (int i) const;
    real& operator[] (int i);

    // assignment and comparison
    Vector& operator= (const Vector& v);
    bool operator== (const Vector& v) const;
    bool operator!= (const Vector& v) const;

    // arithmetic operations
    Vector operator+ (const Vector& v) const;
    Vector operator- (const Vector& v) const;
    Vector operator* (real scalar) const;
    Vector operator/ (real scalar) const;
    Vector operator- () const;
    friend Vector operator* (real scalar, const Vector& v);

private:
    real m_tuple[n];
};
```

where `real` allows you to support both `float` and `double` and where `n` is the dimension of the underlying vector space. The point class is defined by

```
template class <T real, int n> Point
{
public:
    // construction
    Point ();
    Point (const real tuple[n]);
    Point (const Point& p);

    // tuple access as an array
    operator const real* () const;
    operator real* ();
    real operator[] (int i) const;
    real& operator[] (int i);

    // assignment and comparison
    Point& operator= (const Point& p);
    bool operator== (const Point& p) const;
    bool operator!= (const Point& p) const;

    // arithmetic operations
    Point operator+ (const Vector& v) const;
    Vector operator- (const Point& p) const;

private:
    real m_tuple[n];
};
```

Of course, these are just the basic operations, but other members can be added as needed for your applications.

B.2 Coordinate Systems

Let A be an n–dimensional affine space. Let a fixed point $\mathcal{O} \in A$ be labeled as the *origin* and let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V. The set $\{\mathcal{O}; \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is called an *affine coordinate system*. Each point $P \in A$ can be uniquely represented in the coordinate system as follows. The difference $\mathcal{P} - \mathcal{O}$ is a vector and can be represented uniquely

with respect to the basis for V,

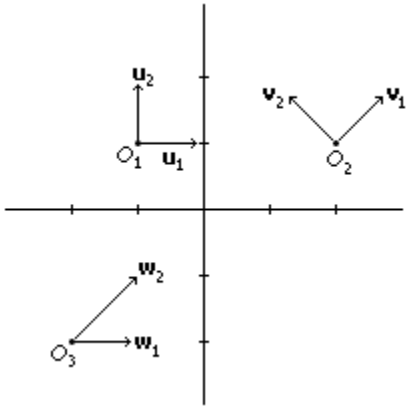
$$\mathcal{P} - \mathcal{O} = \sum_{i=1}^{n} a_i \mathbf{v}_i$$

or using our suggestive notation for sum of a point and a vector,

$$\mathcal{P} = \mathcal{O} + \sum_{i=1}^{n} a_i \mathbf{v}_i$$

The numbers $(a_1, \ldots, a_n)$ are called the *affine coordinates* of $\mathcal{P}$ relative to the specified coordinate system. The origin $\mathcal{O}$ has coordinates $(0, \ldots, 0)$. Figure B.3 shows three coordinate systems in the plane.

Figure B.3. *Three coordinate systems in the plane. Observe that the vectors in the coordinate system are not required to be unit length or perpendicular in pairs.*



A natural question is how to change coordinates from one system to another. Let $\{\mathcal{O}_1; \mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\{\mathcal{O}_2; \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be two affine coordinate systems for A. A point $\mathcal{P} \in A$ has affine coordinates $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$, respectively. The origin $\mathcal{O}_2$ has affine coordinates $(c_1, \ldots, c_n)$ in the first coordinate system. The relationship

$$\mathcal{P} = \mathcal{O}_1 + \sum_{i=1}^{n} a_i \mathbf{u}_i = \mathcal{O}_2 + \sum_{i=1}^{n} b_i \mathbf{v}_i = \mathcal{O}_1 + \sum_{i=1}^{n} c_i \mathbf{u}_i + \sum_{i=1}^{n} b_i \mathbf{v}_i$$

implies

$$\sum_{i=1}^{n} b_i \mathbf{v}_i = \sum_{i=1}^{n} (a_i - c_i) \mathbf{u}_i$$

The two bases are related by a change of basis transformation, $\mathbf{u}_i = \sum_{j=1}^{n} m_{ji}\mathbf{v}_j$ for $1 \le i \le n$, so

$$\sum_{i=1}^{n} b_i\mathbf{v}_i = \sum_{i=1}^{n}(a_i - c_i)\sum_{j=1}^{n} m_{ji}\mathbf{v}_j$$

Renaming the index i to j on the left–hand side and rearranging terms on the right–hand side,

$$\sum_{j=1}^{n} b_j\mathbf{v}_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} m_{ji}(a_i - c_i)\right)\mathbf{v}_j$$

The uniqueness of the representation for a vector requires

$$b_j = \sum_{i=1}^{n} m_{ji}(a_i - c_i), \quad 1 \le j \le n$$

which is the desired relationship between the coordinate systems.

Example B.1. Here are two coordinate systems in the plane, the ones shown in Figure B.3. The origin of the first is $\mathcal{O}_1 = (-1, 2)$ and the basis vectors (coordinate axis directions) are $\mathbf{u}_1 = (1, 0)$ and $\mathbf{u}_2 = (0, 1)$. The origin of the second is $\mathcal{O}_2 = (2, 1)$ and the basis vectors are $\mathbf{v}_1 = (1, 1)/\sqrt{2}$ and $\mathbf{v}_2 = (-1, 1)/\sqrt{2}$. Let us represent the origin of the second system in terms of the first:

$$(3, -1) = (2, 1) - (-1, 2) = \mathcal{O}_2 - \mathcal{O}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 = c_1(1, 0) + c_2(0, 1) = (c_1, c_2)$$

The change of basis matrix is determined by $\mathbf{u}_1 = m_{11}\mathbf{v}_1 + m_{21}\mathbf{v}_2$ and $\mathbf{u}_2 = m_{12}\mathbf{v}_1 + m_{22}\mathbf{v}_2$. In block matrix form where the basis vectors are written as columns,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$$

The solution is

$$\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

If $\mathcal{P} = \mathcal{O}_1 + a_1\mathbf{u}_1 + a_2\mathbf{u}_2 = \mathcal{O}_2 + b_1\mathbf{v}_1 + b_2\mathbf{v}_2$, then

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} a_1 - c_1 \\ a_2 - c_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} a_1 + a_2 - 2 \\ -a_1 + a_2 + 4 \end{bmatrix}$$

The point $(0,0)$ is represented as $\mathcal{O}_1 + (1)\mathbf{u}_1 + (-2)\mathbf{u}_2$ in the first coordinate system. Using the change of coordinate relationship above, verify that the representation in the second coordinate system is $\mathcal{O}_2 + (-3/\sqrt{2})\mathbf{v}_1 + (1/\sqrt{2})\mathbf{v}_2$. Carefully sketch a picture to convince yourself that the coordinates are accurate. $\bowtie$

Exercise B.1. How does the analysis of Example B.1 change when you have the common origin $\mathcal{O}_1 = \mathcal{O}_2 = (0,0)$? In the first coordinate system, let R be the matrix that represents a counterclockwise rotation about the origin by an angle $\pi/4$. What is the matrix R? How is it related to the change of basis matrix? $\bowtie$

Exercise B.2. Repeat the construction in Example B.1, but use the coordinate system $\{\mathcal{O}_3; \mathbf{w}_1, \mathbf{w}_2\}$ instead of $\{\mathcal{O}_2; \mathbf{v}_1, \mathbf{v}_2\}$, where $\mathcal{O}_3 = (-2,-2)$, $\mathbf{w}_1 = (1,0)$, and $\mathbf{w}_2 = (1,1)$. $\bowtie$

B.3 Subspaces

Let A be an affine space. An *affine subspace* of A is a set $A_1 \subseteq A$ such that $V_1 = \{\Delta(\mathcal{P},\mathcal{Q}) \in V : \mathcal{P},\mathcal{Q} \in S\}$ is a subspace of V. The geometric motivation for this definition is quite simple. If $V = \mathbb{R}^3$, the 1–dimensional subspaces are lines through the origin and the 2–dimensional subspaces are planes through the origin. If A is the set of points in space, the 1–dimensional subspaces are lines (not necessarily through the origin) and the 2–dimensional subspaces are planes (not necessarily through the origin). That is, the affine subspaces are just translations of the vector subspaces.

Let A_1 and A_2 be affine subspaces of A with corresponding vector subspaces V_1 and V_2 of V, respectively. The subspaces are said to be *parallel* if $V_1 \subseteq V_2$ or if $V_2 \subseteq V_1$. If it is the case that $V_1 \subseteq V_2$, then as sets either A_1 and A_2 are disjoint ($A_1 \cap A_2 = \emptyset$) or A_1 is contained in A_2 ($A_1 \subseteq A_2$).

Example B.2. Let A be the set of points in space and $V = \mathbb{R}^3$. Let $\mathcal{P} = (0,0,1)$, $\mathbf{u}_1 = (1,0,0)$, and $\mathbf{u}_2 = (0,1,0)$. Let $\mathcal{Q} = (0,0,2)$ and $\mathbf{v}_1 = (1,1,0)$. The sets $A_1 = \{\mathcal{P} + s_1\mathbf{u}_1 + s_2\mathbf{u}_2 : s_1, s_2 \in \mathbb{R}\}$ and $A_2 = \{\mathcal{Q} + t_1\mathbf{v}_1 : t_1 \in \mathbb{R}\}$ are parallel, affine subspaces. In this case, A_1 and A_2 are disjoint. If instead $A_2 = \{\mathcal{P} + t_1\mathbf{v}_1 : t_1 \in \mathbb{R}\}$, A_1 and A_2 are still parallel subspaces, but A_2 is a proper subset of A_1. $\bowtie$

Exercise B.3. Let A be the set of points in space and $V = \mathbb{R}^3$. Let $\mathcal{P} = (0,0,0)$, $\mathbf{u} = (1,0,0)$, $\mathcal{Q} = (0,0,1)$, and $\mathbf{v} = (0,1,0)$. Are the sets $A_1 = \{\mathcal{P} + s\mathbf{u} : s \in \mathbb{R}\}$ and $A_2 = \{\mathcal{Q} + t\mathbf{v} : t \in \mathbb{R}\}$ parallel subspaces? $\bowtie$

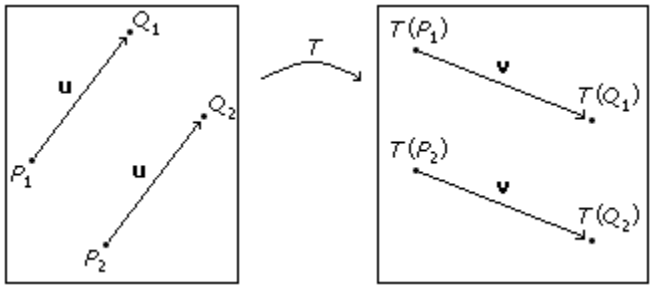
B.4 Transformations

Just as we defined linear transformations for vector spaces, we can define *affine transformations* for affine spaces. Let A be an affine space with vector space V and vector difference operator Δ_A. Let B be an affine space with vector space W and vector difference operator Δ_B. An affine transformation is a function $T : A \to B$ such that the following are true:

1. $\Delta_A(\mathcal{P}_1, \mathcal{Q}_1) = \Delta_A(\mathcal{P}_2, \mathcal{Q}_2)$ implies that $\Delta_B(T(\mathcal{P}_1), T(\mathcal{Q}_1)) = \Delta_B(T(\mathcal{P}_2), T(\mathcal{Q}_2))$.
2. The function $L : V \to W$ defined by $L(\Delta_A(\mathcal{P}, \mathcal{Q})) = \Delta_B(T(\mathcal{P}), T(\mathcal{Q}))$ is a linear transformation.

Figure B.4 illustrates condition 1 in the definition. Setting $\mathbf{u} = \Delta_A(\mathcal{P}_1, \mathcal{Q}_1)$ and $\mathbf{v} = \Delta_B(T(\mathcal{P}_1), T(\mathcal{Q}_1))$, in geometric terms the condition states that no matter what point $\mathcal{P}$ is used as the initial point for $\mathbf{u}$, the initial point for $\mathbf{v}$ must be $T(\mathcal{P})$. Condition 2 just states that the vectors themselves must be transformed linearly.

Figure B.4. *An illustration of condition 1 of the definition for affine transformation.*



If $\mathcal{O}_A$ is selected as the origin for A and if $\mathcal{O}_B = T(\mathcal{O}_A)$ is selected as the origin for B, then the affine transformation is of the form

$$T(\mathcal{O}_A + \mathbf{x}) = T(\mathcal{O}_A) + L(\mathbf{x}) = \mathcal{O}_B + L(\mathbf{x})$$

Consider the special case when the two affine spaces are the same $B = A$, $W = V$, and $\Delta_B = \Delta_A$. Define $\mathcal{O}_B - \mathcal{O}_A = \mathbf{b}$. The affine transformation is now of the form

$$T(\mathcal{O}_A + \mathbf{x}) = \mathcal{O}_A + \mathbf{b} + L(\mathbf{x})$$

Thus, for a fixed origin $\mathcal{O}_A$ and for a specific matrix representation M of the linear transformation L, the induced action of the affine transformation on the vector space elements is

$$\mathbf{y} = M\mathbf{x} + \mathbf{b} \tag{B.1}$$

Hopefully this form should look familiar! If M is the identity matrix, then the affine transformation is a *translation* by $\mathbf{b}$. If M is an orthogonal matrix with determinant 1 and $\mathbf{b}$ is the zero vector, the affine transformation is a pure *rotation* about the origin. If M is an orthogonal matrix of determinant 1 and $\mathbf{b}$ is any vector, the affine transformation is called a *rigid motion*, quite an important concept for physics applications.

B.5 Barycentric Coordinates

The definition of an affine space allows for computing the difference of two points and for computing the sum of a point and a vector. The latter operation is the natural way you "move" from one point to another. The sum of two points is just not defined. However, there is an operation on two points that does make intuitive sense, that of a weighted average of two points. If $\mathcal{P}$ and $\mathcal{Q}$ are two points and $\mathbf{v} = \mathcal{Q} - \mathcal{P}$, then $\mathcal{Q} = \mathcal{P} + \mathbf{v}$. For each $t \in \mathbb{R}$, the quantity $t\mathbf{v}$ is, of course, a vector, so $\mathcal{P} + t\mathbf{v}$ is a point itself. Using our suggestive notation for subtraction of points, we have

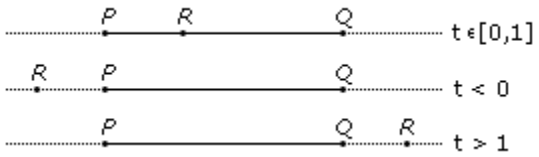
$$\mathcal{P} + t\mathbf{v} = \mathcal{P} + t(\mathcal{Q} - \mathcal{P})$$

It is an error to distribute the multiplication by t across the difference of points, because the definition of affine algebra does not allow the operation of a scalar times a point. That is, $t(\mathcal{Q} - \mathcal{P})$ is well–defined since t is a scalar and $\mathcal{Q} - \mathcal{P}$ is a vector, but $t\mathcal{Q} - t\mathcal{P}$ is ill–defined. But let's go ahead and distribute anyway, then combine the $\mathcal{P}$ terms to obtain

$$\mathcal{R} = (1 - t)\mathcal{P} + t\mathcal{Q} \tag{B.2}$$

$\mathcal{R}$ is said to be a *barycentric combination* of $\mathcal{P}$ and $\mathcal{Q}$ with *barycentric coordinates* $1 - t$ and t, respectively. Observe that the sum of the coordinates is 1, a necessity for a pair of numbers to be barycentric coordinates. For $t \in [0, 1]$, $\mathcal{R}$ is a point on the line segment connecting $\mathcal{P}$ and $\mathcal{Q}$. For $t < 0$, $\mathcal{R}$ is on the line through $\mathcal{P}$ and $\mathcal{Q}$ with $\mathcal{P}$ between $\mathcal{R}$ and $\mathcal{Q}$. Similarly, for $t > 1$, $\mathcal{R}$ is on the line with $\mathcal{Q}$ between $\mathcal{R}$ and $\mathcal{P}$. Figure B.5 illustrates these cases.

Figure B.5. *Various barycentric combinations of two points $\mathcal{P}$ and $\mathcal{Q}$.*



To support barycentric combinations, the C++ template code for points has one additional member function. The other two arithmetic operations are shown just for comparison.

```
template class <T real, int n> Point
{
public:
    // return_point = this + v
    Point operator+ (const Vector& v) const;

    // return_vector = this - p
    Vector operator- (const Point& p) const;

    // return_point = (1 - t) * this + t * p (barycentric combination)
    Point operator+ (real t, const Point& p) const;
}
```

B.5.1 Triangles

The concept of barycentric coordinates extends to three noncolinear points $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{R}$. The points, of course, form a triangle. Two vectors located at $\mathcal{P}$ are $\mathbf{u} = \mathcal{Q} - \mathcal{P}$ and $\mathbf{v} = \mathcal{R} - \mathcal{P}$. For any scalars s and t, $s\mathbf{u}$ and $t\mathbf{v}$ are vectors and may be added to any point to obtain another point. In particular,

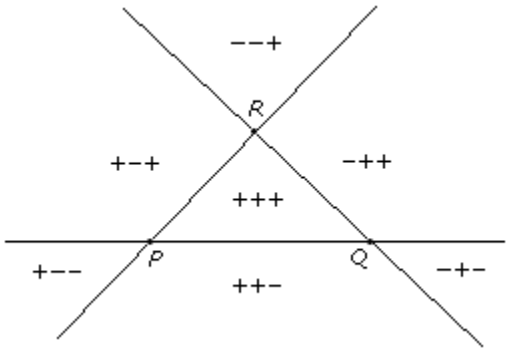
$$\mathcal{P} + s\mathbf{u} + t\mathbf{v} = \mathcal{P} + s(\mathcal{Q} - \mathcal{P}) + t(\mathcal{R} - \mathcal{P})$$

is a point. Just as equation (B.2) was motivated by distributing the scalar product across the differences and collected terms, we do the same here to obtain

$$\mathcal{B} = (1 - s - t)\mathcal{P} + s\mathcal{Q} + t\mathcal{R} \tag{B.3}$$

$\mathcal{B}$ is said to be a *barycentric combination* of $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{R}$ with *barycentric coordinates* $c_1 = 1 - s - t$, $c_2 = s$, and $c_3 = t$, respectively. As with a barycentric combination of two points, the barycentric coordinates must sum to one: $c_1 + c_2 + c_3 = 1$. The location of $\mathcal{B}$ relative to the triangle formed by the three points is illustrated in Figure B.6.

Figure B.6. *The triangle partitions the plane into seven regions. The signs of c_1, c_2, and c_3 are listed as ordered triples.*



The signs of c_1, c_2, and c_3 are listed as ordered triples. On a boundary between two regions, not including the vertices, one of the c_i is 0. At a vertex, two of the c_i are 0, the other coordinate necessarily 1. The coordinates cannot be simultaneously negative since the sum of three negative numbers cannot be 1.

Exercise B.4. *Linear interpolation over a triangle.* A real–valued function $f(x, y)$, unknown to you, has been sampled at three noncolinear points (x_i, y_i) with function values f_i, $0 \le i \le 2$. If you are given the information that $f : \mathbb{R}^2 \to \mathbb{R}$ is a linear transformation, construct an explicit formula for f. $\bowtie$

B.5.2 Tetrahedra

The concept of barycentric coordinates also extends to four noncoplanar points $\mathcal{P}_i$, $0 \le i \le 3$. The points form a tetrahedron. Using a construction similar to that for a segment and a triangle, a *barycentric combination* of the points is

$$\mathcal{B} = (1 - c_1 - c_2 - c_3)\mathcal{P}_0 + c_1\mathcal{P}_1 + c_2\mathcal{P}_2 + c_3\mathcal{P}_3 \tag{B.4}$$

The values $c_0 = 1 - c_1 - c_2 - c_3$, c_1, c_2, and c_3 are the *barycentric coordinates* of $\mathcal{B}$ and they sum to 1. The tetrahedron partitions space into 15 regions, each region labeled with an ordered quadruple of signs for the four coefficients. The only invalid combination of signs is all negative since the sum of four negative numbers cannot equal 1.

Exercise B.5. *Linear interpolation over a tetrahedron.* A real–valued function $f(x, y, z)$, unknown to you, has been sampled at four noncoplanar points (x_i, y_i, z_i) with function values f_i, $0 \le i \le 3$. If you are given the information that $f : \mathbb{R}^3 \to \mathbb{R}$ is a linear transformation, construct an explicit formula for f. $\bowtie$

B.5.3 Simplices

We have seen barycentric coordinates relative to a segment, a triangle, and a tetrahedron. The concept extends to affine spaces of n–tuples for any $n \ge 2$. The name of the object that generalizes triangle and tetrahedron is *simplex* (plural *simplices*). A simplex is formed by $n + 1$ points $\mathcal{P}_i$, $0 \le i \le n$, such that the set of vectors $\{\mathcal{P}_i - \mathcal{P}_0\}_{i=1}^n$ are linearly independent. A *barycentric combination of the points is*

$$\mathcal{B} = \sum_{i=0}^{n} c_i \mathcal{P}_i \tag{B.5}$$

and the c_i are the *barycentric coordinates* of $\mathcal{B}$ with respect to the given points. As before, the coefficients sum to 1, $\sum_{i=0}^n c_i = 1$.

Although we tend to work in 2D or 3D, let us be abstract for a moment and ask the same question we did for segments, triangles, and tetrahedra. A segment, a simplex in $\mathbb{R}$, partitioned $\mathbb{R}$ into 3 regions. A triangle, a simplex in $\mathbb{R}^2$, partitioned $\mathbb{R}^2$ into 7 regions. A tetrahedron, a simplex in $\mathbb{R}^3$, partitioned $\mathbb{R}^3$ into 15 regions. How many regions in $\mathbb{R}^n$ are obtained by a partitioning of a simplex? The sequence with increasing dimension is 3, 7, 15, so you might guess that the answer is $2^{n+1} - 1$. This is correct. An intuitive reason is supported by looking at the signs of the $n + 1$ coefficients. Each region is labeled with an ordered $(n + 1)$–tuple of signs, each sign positive or negative. There are two choices of sign for each of the $n + 1$ components, leading to 2^{n+1} possibilities. As in the cases we've already looked at, all negative signs are not allowed since the sum would be negative and you cannot obtain a sum of 1. This means only $(2^{n+1} - 1)$ tuples of signs are possible.

A more geometric approach to counting the regions is based on an analysis of the components of the simplices. A segment has 2 vertices. The interior of the segment is formed by both vertices, so you can think of that region as occuring as the only

possibility when choosing 2 vertices from a set of 2 vertices. That is, the number of interior regions is $C(2,2) = 1$, where

$$C(n,k) = \frac{n!}{k!(n-k)!}$$

is the number of combinations of n items choosing k at a time. The segment has two exterior regions, each formed by a single vertex. The number of such regions is $C(2,1) = 2$ since you have 2 vertices, but choose only 1 at a time. The total number of regions in the partition of the line is $C(2,1) + C(2,2) = 2 + 1 = 3$.

A triangle has 3 vertices. The interior is formed by choosing all three vertices. The number of interior regions is $C(3,3) = 1$. Figure B.6 shows that each edge has a corresponding exterior region. An edge is formed by 2 vertices and you have 3 vertices to choose from. The total number of edges is $C(3,2) = 3$. The figure also shows that each vertex has a corresponding exterior region. The total number of vertices is $C(3,1) = 3$. The total number of regions in the partition of the plane is $C(3,1) + C(3,2) + C(3,3) = 3 + 3 + 1 = 7$.

The same argument applies to a tetrahedron with 4 vertices. The interior is formed by all four vertices; the number of interior regions is $C(4,4) = 1$. A face is formed by 3 vertices; there are $C(4,3) = 4$ such possibilities. An edge is formed by 2 vertices; there are $C(4,2) = 6$ such possibilities. Finally, there are $C(4,1) = 4$ vertices. An exterior region is associated with each vertex, each edge, and each face. The total number of regions is $C(4,1) + C(4,2) + C(4,3) + C(4,4) = 4 + 6 + 4 + 1$.

Consider now the general case, a simplex formed by $n+1$ points. The components, so to speak, of the simplex are classified by how many vertices form them. If a component uses k vertices, let's call that a k–vertex component. The interior of the simplex is the only $(n+1)$–vertex component. Each k–vertex component where $1 \le k < n+1$ has a single exterior region associated with it. The total number of regions is therefore

$$C(n+1,n+1) + C(n+1,n) + \cdots + C(n+1,1) = \sum_{k=1}^{n+1} C(n+1,k) = 2^{n+1} - 1$$

The term $C(n,k)$ indicates the number of k–vertex components, the number of combinations of n vertices choosing k at a time. Where did that last equality come from? Recall the binomal expansion for a power of a sum of two values:

$$\begin{aligned}(x+y)^m &= C(m,0)x^m + C(m,1)x^{m-1}y + \cdots + C(m,m-1)xy^{m-1} + C(m,m)y^m \\ &= \textstyle\sum_{k=0}^{m} C(m,k)x^{m-k}y^k\end{aligned}$$

Setting $x = 1$, $y = 1$, and $m = n + 1$, we arrive at

$$2^{n+1} = (1+1)^{n+1} = \sum_{k=0}^{n+1} C(n+1,k) = 1 + \sum_{k=1}^{n+1} C(n+1,k)$$

B.5.4 Length, Area, Volume, and Hypervolume

The constructions in this section are designed to show that the area of a triangle (simplex in 2D) can be computed as a sum (in an integration sense) of lengths of line segments (simplices in 1D). I also show that the volume of a tetrahedron (simplex in 3D) can be computed as a sum of areas of triangles (simplices in 2D). These two facts indicate the process is recursive in dimension. Intuitively, a simplex in 4D has a "volume," so to speak, that can be computed as a sum of volumes of tetrahedra (simplices in 3D). Sometimes this is called a *hypervolume*, but that leads us to wanting names for the similar concept in yet higher dimensions. I will use the term *hypervolume* for any dimension and note that hypervolume in 1D is length, hypervolume in 2D is area, and hypervolume in 3D is volume. Given a simplex formed by points $\mathcal{P}_i$ for $0 \le i \le n$, the hypervolume is denoted $H(\mathcal{P}_0, \ldots, \mathcal{P}_n)$.

Length of a Segment

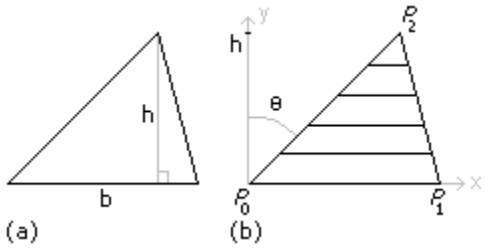
A line segment with end points $\mathcal{P}_0$ and $\mathcal{P}_1$ is a simplex with two vertices. The length of the simplex is $|\mathcal{P}_1 - \mathcal{P}_0|$. Using our notation for hypervolume,

$$H(\mathcal{P}_0, \mathcal{P}_1) = |\mathcal{P}_1 - \mathcal{P}_0| \tag{B.6}$$

Area of a Triangle

Recall that the area of a triangle with base length b and height h is $A = bh/2$. Figure B.7(a) shows the standard drawing one normally uses to show b and h. Figurel B.7(b) shows a triangle viewed as the union of line segments of varying lengths. The drawing shows the bottom edge aligned with the x–axis, but the ensuing arguments do not depend on this. The area of the triangle may be thought of as the sum of the lengths of all the line segments in the union. This is an informal and mathematically nonrigorous view of the situation, but intuitively it works quite well. The number of line segments is infinite, one segment per y–value in the interval $[0, h]$. The sum cannot be computed in the usual sense. In this setting it becomes an integration of the lengths as a function of y.

Figure B.7. *(a) A triangle with base length b and height h marked. The area of the triangle is bh/2. (b) A triangle viewed as a union of an infinite number of line segments of varying lengths (only a few are shown). The area of the triangle is the sum of the lengths of those line segments.*



Select a value of $y \in [0, h]$. The line segment has two end points, one on the triangle edge from $\mathcal{P}_0$ to $\mathcal{P}_2$ and one on the triangle edge from $\mathcal{P}_1$ to $\mathcal{P}_2$. The fraction of the distance along each edge on which the end points lie is $y/h \in [0, 1]$. That is, the end points are the barycentric combinations $(1 - y/h)\mathcal{P}_0 + (y/h)\mathcal{P}_2$ and $(1 - y/h)\mathcal{P}_1 + (y/h)\mathcal{P}_2$. The length of the segment is the length of the vector connecting the end points,

$$\mathbf{v} = ((1 - y/h)\mathcal{P}_1 + (y/h)\mathcal{P}_2) - (1 - y/h)\mathcal{P}_0 + (y/h)\mathcal{P}_2 = (1 - y/h)(\mathcal{P}_1 - \mathcal{P}_0)$$

The length of the segment as a function of y is

$$L(y) = |\mathbf{v}| = (1 - y/h)|\mathcal{P}_1 - \mathcal{P}_0| = (1 - y/h)b$$

where b is the length of the base of the triangle. To "add" all the values $L(y)$ for $y \in [0, h]$ in order to obtain the area A, we need to integrate

$$A = \int_0^h L(y)\,dy = \int_0^h (1 - y/h)b\,dy = b\,\frac{-h(1 - y/h)^2}{2}\bigg|_0^h = \frac{bh}{2}$$

In terms of our hypervolume notation,

$$H(\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2) = \frac{h}{2} H(\mathcal{P}_0, \mathcal{P}_1) \tag{B.7}$$

where $h = |\mathcal{P}_2 - \mathcal{P}_0| \cos\theta$ with θ the angle as shown in Figure B.7.

The height h also may be computed as the length of the projection of $\mathcal{P}_2 - \mathcal{P}_0$ onto the vertical axis. A convenient vector to use, whose direction is that of the vertical

axis, is the following. Let $\mathcal{P}_1 - \mathcal{P}_0 = \mathbf{v} = (v_1, v_1)$. A perpendicular vector in the direction of the positive vertical axis as shown in Figure B.7 is $-\text{Perp}(\mathcal{P}_0, \mathcal{P}_1)$ where

$$\text{Perp}(\mathcal{P}_0, \mathcal{P}_1) = (v_2, -v_1) \tag{B.8}$$

is called the *perp* vector. You will notice that the perp vector itself is in the direction of the *negative* vertical axis as shown in Figure B.7. Using the coordinate–free definition of dot product, equation (A.5),

$$(\mathcal{P}_2 - \mathcal{P}_0) \cdot (-\text{Perp}(\mathcal{P}_0, \mathcal{P}_1)) = |\mathcal{P}_2 - \mathcal{P}_0||\text{Perp}(\mathcal{P}_0, \mathcal{P}_1)| \cos\theta = h|\text{Perp}(\mathcal{P}_0, \mathcal{P}_1)|$$

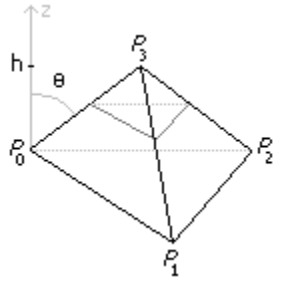
The perp vector has the same length as $\mathcal{P}_1 - \mathcal{P}_0$, so the dot product in the previous equation is twice the area of the triangle, bh. This gives us the nonrecursive formula for the area of a triangle,

$$H(\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2) = -\frac{1}{2}(\mathcal{P}_2 - \mathcal{P}_0) \cdot \text{Perp}(\mathcal{P}_0, \mathcal{P}_1) \tag{B.9}$$

Volume of a Tetrahedra

The volume of a tetrahedron may be computed in the same intuitive way that was used for the area of a triangle. Let $\mathcal{P}_i$, $0 \leq i \leq 3$, denote the vertices of the tetrahedron. The base of the tetrahedron will be selected as the triangle formed by $\mathcal{P}_i$, $0 \leq i \leq 2$. The tetrahedron may be informally thought of as an infinite union of triangles that are parallel to its base. Figure B.8 shows a tetrahedron and one of its triangle slices.

Figure B.8. *A tetrahedron with base formed by $\mathcal{P}_0$, $\mathcal{P}_1$, and $\mathcal{P}_2$. A triangle slice parallel to the base is shown. The direction perpendicular to the base is marked as the positive z-axis.*



Select a value of $z \in [0,h]$. The corresponding triangle has three vertices, one on each tetrahedron edge with end points $\mathcal{P}_i$ and $\mathcal{P}_3$ for $0 \le i \le 2$. The fraction of the distance along each edge on which the end points lie is $z/h \in [0,1]$. The end points are the barycentric combinations $\mathcal{Q}_i = (1-z/h)\mathcal{P}_i + (z/h)\mathcal{P}_3$ for $0 \le i \le 2$. If $A(z)$ denotes the area of this triangle, the volume of the tetrahedron is

$$V = \int_0^h A(z)\,dz$$

We would like to use equation (B.9) to compute $A(z)$, but there is a problem. Even though that equation is written in a coordinate–free manner, it is implicitly tied to 2D points via the definition of the perp vector in equation (B.8) that requires points in two dimensions. For the time being, let's instead use the area formula from equation (A.10). We will later return to the issue of the perp vector and provide a definition that is valid in any dimension.

Set $\mathbf{v}_i = \mathcal{Q}_i - \mathcal{Q}_0 = (1-z/h)(\mathcal{P}_i - \mathcal{P}_0)$ for $1 \le i \le 2$. The triangle slice is half of a parallelogram formed by $\mathbf{v}_1$ and $\mathbf{v}_2$, so using equation (A.10) the area of the triangle is

$$A(z) = \frac{1}{2}|\mathbf{v}_1 \times \mathbf{v}_2| = \frac{(1-z/h)^2}{2}|(\mathcal{P}_1 - \mathcal{P}_0) \times (\mathcal{P}_2 - \mathcal{P}_0)| = (1-z/h)^2 b$$

where b is the area of the base of the tetrahedron. The volume of the tetrahedron is

$$V = \int_0^h A(z)\,dz = \int_0^h (1-z/h)^2 b\,dz = b\,\frac{-h(1-z/h)^3}{3}\bigg|_0^h = \frac{bh}{3}$$

In our hypervolume notation, the volume is

$$H(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2,\mathcal{P}_3) = \frac{h}{3}H(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2) \tag{B.10}$$

where $h = |\mathcal{P}_3 - \mathcal{P}_0|\cos\theta$ with θ the angle shown in Figure B.8.

The height may also be computed as the length of the projection of $\mathcal{P}_3 - \mathcal{P}_0$ onto the vertical axis. We already know a vector with that direction, the cross product $(\mathcal{P}_1 - \mathcal{P}_0) \times (\mathcal{P}_2 - \mathcal{P}_0)$. Using equation (B.8) as motivation, define

$$\text{Perp}(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2) = (\mathcal{P}_1 - \mathcal{P}_0) \times (\mathcal{P}_2 - \mathcal{P}_0) \tag{B.11}$$

Using the coordinate–free definition of dot product, equation (A.5),

$$(\mathcal{P}_3-\mathcal{P}_0)\cdot\text{Perp}(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2) = |\mathcal{P}_3-\mathcal{P}_0||\text{Perp}(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2)|\cos\theta = h|\text{Perp}(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2)| = 2hb$$

where b is the area of the triangle base of the tetrahedron. Dividing by 6 we have a nonrecursive formula for the volume,

$$H(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2,\mathcal{P}_3) = \frac{1}{6}(\mathcal{P}_3 - \mathcal{P}_0)\cdot\text{Perp}(\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2) \tag{B.12}$$

Hypervolume of a Simplex

Notice the strong similarities between equations (B.7) and (B.10) and between equations (B.9) and (B.12). The first pair of equations suggests that for a simplex formed from $n+1$ points $\mathcal{P}_i$, $0 \le i \le n$, the recursive formula for the hypervolume is

$$H(\mathcal{P}_0, \ldots, \mathcal{P}_n) = \frac{h}{n} H(\mathcal{P}_0, \ldots, \mathcal{P}_{n-1}) \tag{B.13}$$

where $h = |\mathcal{P}_n - \mathcal{P}_0| \cos\theta$ with θ the angle between $\mathcal{P}_n - \mathcal{P}_0$ and a vector that is perpendicular to the base of the simplex. The base is itself a simplex but formed by n points $\mathcal{P}_i$ for $0 \le i \le n-1$. As we saw in the cases $n = 2$ and $n = 3$, we want the perpendicular vector chosen so that θ is an acute angle. In 2D we used $-\text{Perp}(\mathcal{P}_0, \mathcal{P}_1)$ and in 3D we used $\text{Perp}(\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2)$. This suggests that in general dimension n, we will use a vector $(-1)^{n+1}\text{Perp}(\mathcal{P}_0, \ldots, \mathcal{P}_{n-1})$, where the perp vector $\text{Perp}(\mathcal{P}_0, \ldots, \mathcal{P}_{n-1})$ is appropriately defined. The second pair of equations suggests the nonrecursive formula

$$H(\mathcal{P}_0, \ldots, \mathcal{P}_n) = \frac{(-1)^{n+1}}{n!} (\mathcal{P}_n - \mathcal{P}_1) \cdot \text{Perp}(\mathcal{P}_0, \ldots, \mathcal{P}_{n-1}) \tag{B.14}$$

This leaves us with the task of finding a formula for $\text{Perp}(\mathcal{P}_0, \ldots, \mathcal{P}_{n-1})$, hopefully in a way that applies to any dimensional input points. We accomplish this last goal by introducing an indexed quantity that stores information about permutations. In tensor calculus, this is called the *Levi–Civita permutation tensor*, but the name and tensor calculus are not really important in our context.

Let's look at the 2D problem first. The doubly–indexed quantity e_{ij} for $1 \le i \le 2$ and $1 \le j \le 2$ represents four numbers, each number in the set $\{-1, 0, 1\}$. The value is 0 if the two indices are the same: $e_{11} = 0$ and $e_{22} = 0$. The other values are $e_{12} = 1$ and $e_{21} = -1$. The choice of 1 or -1 is based on whether the ordered index pair (i, j) is an even or odd permutation of $(1, 2)$. In the current case, $(1, 2)$ is an even permutation of $(1, 2)$ (zero transpositions, zero is an even number) so $e_{12} = 1$. The pair $(2, 1)$ is an odd permutation of $(1, 2)$ (one transposition, one is an odd number) so $e_{21} = -1$. You should notice that $e_{ji} = -e_{ij}$. Treating e_{ij} as the elements of a 2×2 matrix E, that matrix is skew–symmetric: $E^{\mathrm{T}} = -E$.

The arguments of the area function and perp operation are points. Since the area and perp vector are invariant under translations, we may as well replace the arguments by vectors $\mathbf{v}^{(i)} = \mathcal{P}_i - \mathcal{P}_0$ for $i \ge 1$. In 2D, $\mathbf{v}^{(i)} = (v_1^{(i)}, v_2^{(i)})$. Thus,

$$(u_1, u_2) = \mathbf{u} = \text{Perp}(\mathbf{v}^{(1)}) = (v_2^{(1)}, -v_1^{(1)}) = (e_{12} v_2^{(1)}, e_{21} v_1^{(1)})$$

In summation notation the components of $\mathbf{u}$ are

$$u_j = \sum_{i=1}^{2} e_{ji} v_i^{(1)} \tag{B.15}$$

The area of the triangle is

$$\begin{aligned} H(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) &= -\tfrac{1}{2}\mathbf{v}^{(2)} \cdot \text{Perp}(\mathbf{v}^{(1)}) && \text{Using equation (B.9)} \\ &= -\tfrac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} e_{ji} v_j^{(2)} v_i^{(1)} && \text{Using equation (B.15)} \\ &= \tfrac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} e_{ij} v_i^{(1)} v_j^{(2)} && e_{ij} = -e_{ji} \text{ and swapping terms} \end{aligned} \tag{B.16}$$

In 3D the triply–indexed permutation tensor is e_{ijk} for $1 \le i \le 3$, $1 \le j \le 3$, and $1 \le k \le 3$. Each value is in the set $\{-1, 0, 1\}$. If any pair of indices is the same, the value is zero; for example, $e_{111} = 0$, $e_{112} = 0$, and $e_{233} = 0$ (there are 21 zero elements). Otherwise, $e_{ijk} = 1$ if (i, j, k) is an even permutation of $(1, 2, 3)$ or $e_{ijk} = -1$ if (i, j, k) is an odd permutation of $(1, 2, 3)$. Under these conditions only six elements are nonzero: $e_{123} = e_{231} = e_{312} = 1$ and $e_{132} = e_{321} = e_{213} = -1$. As in the 2D case, if a pair of indices is swapped, the sign is changed: $e_{jik} = e_{ikj} = e_{kji} = -e_{ijk}$. Define $\mathbf{v}^{(i)} = \mathcal{P}_i - \mathcal{P}_0 = (v_1^{(i)}, v_2^{(i)}, v_3^{(i)})$, $1 \le i \le 3$; then

$$\begin{aligned} (u_1, u_2, u_3) &= \mathbf{u} \\ &= \text{Perp}(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) \\ &= \mathbf{v}^{(1)} \times \mathbf{v}^{(2)} \\ &= \left(v_2^{(1)} v_3^{(2)} - v_3^{(1)} v_2^{(2)}, v_3^{(1)} v_1^{(2)} - v_1^{(1)} v_3^{(2)}, v_1^{(1)} v_2^{(2)} - v_2^{(1)} v_1^{(2)}\right) \\ &= \left(e_{123} v_2^{(1)} v_3^{(2)} + e_{132} v_3^{(1)} v_2^{(2)}, e_{231} v_3^{(1)} v_1^{(2)} + e_{213} v_1^{(1)} v_3^{(2)}, e_{312} v_1^{(1)} v_2^{(2)} + e_{321} v_2^{(1)} v_1^{(2)}\right) \end{aligned}$$

In summation notation the components of $\mathbf{u}$ are

$$u_k = \sum_{i=1}^{3}\sum_{j=1}^{3} e_{kij} v_i^{(1)} v_j^{(2)} \tag{B.17}$$

The volume of the tetrahedron is

$$\begin{aligned} H(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}) &= \tfrac{1}{6}\mathbf{v}^{(3)} \cdot \text{Perp}(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) && \text{Using equation (B.12)} \\ &= \tfrac{1}{6}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} e_{kij} v_k^{(3)} v_i^{(1)} v_j^{(2)} && \text{Using equation (B.17)} \\ &= \tfrac{1}{6}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} -e_{ikj} v_i^{(1)} v_j^{(2)} v_k^{(3)} && e_{kij} = -e_{ikj} \text{ and swapping terms} \\ &= \tfrac{1}{6}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} e_{ijk} v_i^{(1)} v_j^{(2)} v_k^{(3)} && e_{ikj} = -e_{ijk} \end{aligned} \tag{B.18}$$

The last couple of steps are just to swap the k into the last position. In 2D one swap was required. In 3D two swaps were required.

The pattern holds for general dimension n. The permutation tensor is an n–indexed quantity $e_{i_1\cdots i_n}$ that is zero if any pair of indices is repeated, is 1 if $(i_1,\ldots,i_n)$ is an even permutation of $(1,\ldots,n)$, or is -1 if $(i_1,\ldots,i_n)$ is an odd permutation of $(1,\ldots,n)$. Only $n!$ values are nonzero where $n!$ is the number of permutations of n numbers. The vectors of interest are $\mathbf{v}^{(i)} = \mathcal{P}_i - \mathcal{P}_0$ for $1 \le i \le n$. The vector $(u_1,\ldots,u_n) = \mathbf{u} = \text{Perp}(\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n)})$ has components

$$u_j = \sum_{i_1=1}^{n} \cdots \sum_{i_{n-1}=1}^{n} e_{i_n i_1 \cdots i_{n-1}} v_{i_1}^{(1)} \cdots v_{i_{n-1}}^{(n-1)} \tag{B.19}$$

The hypervolume of the simplex is

$$\begin{aligned} H(\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n)}) &= \tfrac{(-1)^{n+1}}{n!}\mathbf{v}^{(n)} \cdot \text{Perp}(\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n-1)}) && \text{Using equation (B.14)} \\ &= \tfrac{(-1)^{n+1}}{n!} \textstyle\sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} e_{i_n i_1 \cdots i_{n-1}} v_{i_n}^{(n)} v_{i_1}^{(1)} \cdots v_{i_{n-1}}^{(n-1)} && \text{Using equation (B.19)} \\ &= \tfrac{1}{n!} \textstyle\sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} e_{i_n i_1 \cdots i_{n-1}} v_{i_1}^{(1)} \cdots v_{i_{n-1}}^{(n-1)} v_{i_n}^{(n)} \end{aligned} \tag{B.20}$$

The last equation is the result of swapping i_n with each of the $n-1$ other indices. Each swap introduces a factor of -1, so the total swaps introduces the factor $(-1)^{n-1}$. Combining with the other sign factor $(-1)^{n+1}$ results in a factor of $(-1)^{n-1}(-1^{n+1} = (-1)^{2n} = 1$.

The final interesting observation is that the summations in the hypervolume equation (B.20) are just the determinant of a matrix. The n–indexed quantity $e_{i_1\cdots i_n}$ is the same quantity as e_σ introduced in the section on determinants. If the vectors $\mathbf{v}^{(i)}$ are written as the columns of an $n \times n$, say $[\mathbf{v}^{(1)} \mid \cdots \mid \mathbf{v}^{(n)}]$, then the hypervolume of the simplex is

$$H(\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n)}) = \frac{1}{n!} \det \left[\begin{array}{c|c|c} \mathbf{v}^{(1)} & \cdots & \mathbf{v}^{(n)} \end{array} \right] \tag{B.21}$$

The formula was developed with a certain ordering in mind for the input vectors. For general ordering, equation (B.21) can produce negative numbers, in which case the formula generates the *signed hypervolume*. To be sure you have the nonnegative hypervolume, just take the absolute value of the right–hand side of the equation.