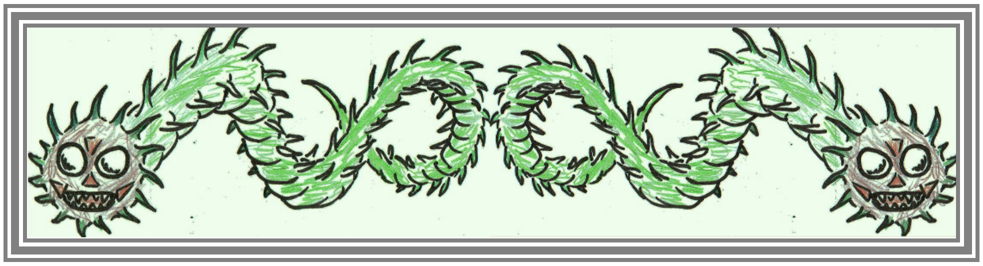

Vector Algebra Is Actually a Subset Of Quaternion Algebra?

Rev 0.02
Jim Satterwhite

James Satterwhite PE – NC Ret.

5/31/2025

Table of Contents

1. Introduction

In my study of Maxwell's Treatise and his resultant equations I have become embedded in what I will call the Vector-Quaternion controversy. There are many elements of this controversy, as well as a number of personalities and agendas involved. From this I have come to believe that Vector Algebra is *necessarily* a subset of Quaternions, although many have worked diligently to deny that fact. Let's look at what is involved and afterward see what you think.

2. Mathematical Notation

There has been considerable evolution of mathematical notation since Maxwell's and Hamilton's time, circa mid 1800s. The mathematical notation conventions of their time are difficult to comprehend sometimes when we see them and try to work with them today. I have to remember that they were doing the best they could with the existing conventions and at the same time breaking barriers and advancing the state of the art. The part that disturbs me is that some folks want to hang on the old notation as if it were some spiritual artifact. Let's give credit where it is due and move on.

A notation that particularly irritates me in the quaternion community are the $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ unit vectors that stand for the x, y, z coordinates and to me are best notated by the $\hat{x}, \hat{y}, \hat{z}$ unit vector notation.

3. Vector Algebra

Vector algebra consists of addition and subtraction of vectors; however, no direct multiplication. Instead the dot and cross products are introduced. What is the product of two vectors? Let's take a look. Assume two vectors $\vec{A}$ and $\vec{B}$

$$\vec{A} = A_x\hat{x} + A_y\hat{y} + A_z\hat{z}$$
$$\vec{B} = B_x\hat{x} + B_y\hat{y} + B_z\hat{z}$$

then multiply them directly

$$\begin{aligned}
\vec{A}\vec{B} &= \left(A_x\hat{x} + A_y\hat{y} + A_z\hat{z}\right)\left(B_x\hat{x} + B_y\hat{y} + B_z\hat{z}\right) \\
&= A_xB_x\hat{x}^2 + A_xB_y\hat{x}\hat{y} + A_xB_z\hat{x}\hat{z} + A_yB_x\hat{y}\hat{x} \\
&\quad + A_yB_y\hat{y}^2 + A_yB_z\hat{y}\hat{z} + A_zB_x\hat{z}\hat{x} + A_zB_y\hat{z}\hat{y} + A_zB_z\hat{z}^2 \\
&= -\underbrace{\left(A_xB_x + A_yB_y + A_zB_z\right)}_{Scalar} \\
&\quad + \underbrace{A_xB_y\hat{x}\hat{y} + A_xB_z\hat{x}\hat{z} + A_yB_x\hat{y}\hat{x} + A_yB_z\hat{y}\hat{z} + A_zB_x\hat{z}\hat{x} + A_zB_y\hat{z}\hat{y}}_{Vector}.
\end{aligned} \tag{1}$$

The result is a vector and a scalar; however, how do we resolve the vector as it contains products of unit vectors ($\hat{x}\hat{y}$, $\hat{x}\hat{z}$, $\hat{y}\hat{x}$, $\hat{y}\hat{z}$, $\hat{z}\hat{x}$, $\hat{z}\hat{y}$, $\hat{x}^2$, $\hat{y}^2$, $\hat{z}^2$) and vector algebra as taught does not give us a way to resolve this except for the unit vector squared ie, $\hat{x}^2 = \hat{y}^2 = \hat{z}^2 = -1$. This is important because Equation 1 yields the Dot and Cross products when you take the definition of these unit vector products from Hamilton's definition of his Quaternions and apply them. This is what Maxwell saw in Hamilton's quaternions and it gave him a formalism for many of the equations he had derived. He could see the form of the equations; however he had no name or definition for them. We will see how this evolves below.

In vector algebra, the dot and cross products $\vec{\nabla} \circ \vec{A}$ & $\vec{\nabla} \times \vec{A}$ are just introduced as the available products for vectors without a definition of where they come from. Yes, there are some round about derivations that prove the validity of the construct; however, they don't evolve naturally as they do from Hamilton.

4. Maxwell's Treatise

I have been translating the section of Maxwell's Treatise related to his equations into modern notation with interest in the controversy concerning whether his treatise was written in quaternion form as he claims or not.

I know from working with Maxwell's Treatise that he derived the *form* of the dot and cross product in his work; however did not have the notation for them until he came to be familiar with Hamilton's quaternions. He seized upon Hamilton's ∇ operator on a vector function. He defined his own version notation for what Gibbs and Heaviside defined to be the dot and cross products involving the del ∇ operator, as I understand it, by Hamilton.

Let's see it in Maxwell's own words. From Article 25 of the Preliminary section of Volume 1 of Maxwell's Treatise. Here we will see his words and notation and my modernization of his notation in blue. This presentation is what I used when I transcribed and translated the notation the Maxwell's Equations section of his Treatise into modern notation. First presented is a copy of the original text paragraph by paragraph followed by transcription of the text and translation into modern notation. This allows the reader to see the evolution of his work into modern notation. I make no apology for my modern notation not necessarily being the present convention. (Shades of Heaviside).

25. On the Effect of the Operator ∇ $\left(\vec{\nabla}\right)$ on a Vector Function

On the effect of the operator ∇ *on a vector function.*

25.] We have seen that the operation denoted by ∇ is that by which a vector quantity is deduced from its potential. The same operation, however, when applied to a vector function, produces results which enter into the two theorems we have just proved (III and IV). The extension of this operator to vector displacements, and most of its further development, is due to Professor Tait †.

We have seen that the operation denoted by ∇ $\left(\vec{\nabla}\right)$ is that by which a vector quantity is deduced from its potential. The same operation; however, when applied to a vector function, produces results which enter into the two theorems we have just proved (III and IV). The extension of this operator to vector displacements, and most of its further development, is due to Professor Tait.

Let σ be a vector function of ρ, the vector of a variable point. Let us suppose, as usual, that

$$\rho = ix + jy + kz,$$

and

$$\sigma = iX + jY + kZ;$$

where X, Y, Z are the components of σ in the directions of the axes.

We have to perform on σ the operation

$$\nabla = i\frac{d}{dx} + j\frac{d}{dy} + k\frac{d}{dz}.$$

Let σ $\left(\vec{\sigma}\right)$ be a vector function of ρ $\left(\vec{\rho}\right)$, the vector of a variable point. Let us suppose, as usual, that

$$\rho = ix + -jy + kz$$

and

$$\sigma = iX + -jY + kZ;$$

$$\left\{\begin{array}{ll} & \vec{\rho} = \rho_x \hat{i} + -\rho_y \hat{j} + \rho_z \hat{k} \\ \text{and} & \vec{\sigma} = \sigma_X \hat{i} + -\sigma_y \hat{j} + \sigma_z \hat{k}; \end{array}\right\}$$

where X, Y, Z $\left(\sigma_X, \sigma_y, \sigma_z\right)$ are the components of σ $\left(\vec{\sigma}\right)$ in the directions of the axes.

We have to perform on $\vec{\sigma}$ the operation

$$\nabla = \frac{d}{dx}i + \frac{d}{dy}j + \frac{d}{dz}k,$$

$$\left\{\vec{\nabla} = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} + \frac{\partial}{\partial z}\hat{k},\right\}$$

$$\left\{\text{or equally} \quad \vec{\nabla} = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z},\right\}$$

Performing this operation, and remembering the rules for the multiplication of i, j, k, we find that $\nabla\sigma$ consists of two parts, one scalar and the other vector.

The scalar part is

$$S\nabla\sigma = -\left(\frac{dX}{dx} + \frac{dY}{dy} + \frac{dZ}{dz}\right), \text{ see Theorem III,}$$

and the vector part is

$$V\nabla\sigma = i\left(\frac{dZ}{dy} - \frac{dY}{dz}\right) + j\left(\frac{dX}{dz} - \frac{dZ}{dx}\right) + k\left(\frac{dY}{dx} - \frac{dX}{dy}\right).$$

Performing this operation, and remembering the rules for the multiplication of i, j, k $\left(\hat{i}, \hat{j}, \hat{k} \text{ or } \hat{x}, \hat{y}, \hat{z}\right)$, we find that $\nabla\sigma$ $\left(\vec{\nabla}\vec{\sigma}\right)$ consists of two parts, one scalar and the other a vector.

The scalar part is

$$S\nabla\sigma = -\left(\frac{dX}{dx} + \frac{dY}{dy} + \frac{dZ}{dz}\right), \quad \text{see Theorem III}$$

$$\left\{S\vec{\nabla}\vec{\sigma} = -\left(\frac{\partial\sigma_x}{\partial x} + \frac{\partial\sigma_y}{\partial y} + \frac{\partial\sigma_z}{\partial z}\right) = -\vec{\nabla}\circ\vec{\sigma},\right\}$$

and the vector part is

$$V\nabla\sigma = i\left(\frac{\partial Z}{\partial y}-\frac{\partial Y}{\partial z}\right)+j\left(\frac{\partial X}{\partial z}-\frac{\partial Z}{\partial x}\right)+k\left(\frac{\partial Y}{\partial x}-\frac{\partial X}{\partial y}\right),$$

$$\left\{V\vec{\nabla}\vec{\sigma}=\left(\frac{\partial\sigma_z}{\partial y}-\frac{\partial\sigma_y}{\partial z}\right)\hat{i}+\left(\frac{\partial\sigma_x}{\partial z}-\frac{\partial\sigma_z}{\partial x}\right)\hat{j}+\left(\frac{\partial\sigma_y}{\partial x}-\frac{\partial\sigma_x}{\partial y}\right)\hat{k}=\vec{\nabla}\times\vec{\sigma},\right\}$$

It is interesting that he did not show the multiplication that lead to his definitions as it is interesting.

$$\vec{\nabla}\vec{\sigma}=\underbrace{-\left(\frac{\partial\sigma_x}{\partial x}+\frac{\partial\sigma_y}{\partial y}+\frac{\partial\sigma_z}{\partial z}\right)}_{Scalar\ \ S\vec{\nabla}\vec{\sigma}\quad -\left(\vec{\nabla}\circ\vec{\sigma}\ \ Divergence\right)}$$

$$+\underbrace{\left(\frac{\partial\sigma_z}{\partial y}-\frac{\partial\sigma_y}{\partial z}\right)\hat{i}+\left(\frac{\partial\sigma_x}{\partial z}-\frac{\partial\sigma_z}{\partial x}\right)\hat{j}+\left(\frac{\partial\sigma_y}{\partial x}-\frac{\partial\sigma_x}{\partial y}\right)\hat{k}}_{Vector\ V\vec{\nabla}\vec{\sigma}\quad \left(\vec{\nabla}\times\vec{\sigma}\ \ Curl\right)}$$

So he defines the scalar part of $\vec{\nabla}\vec{\sigma}$ as $S\vec{\nabla}\vec{\sigma}$ $\left(-\vec{\nabla}\circ\vec{\sigma}\right)$ and the vector part as $V\vec{\nabla}\vec{\sigma}$ $\left(-\vec{\nabla}\times\vec{\sigma}\right)$. He called $S\vec{\nabla}\vec{\sigma}$ the convergence. I don't remember his name for $V\vec{\nabla}\vec{\sigma}$.

This notation $\left(S \,\&\, V\right)$ was very confusing to me on my first pass at translating his Treatise notation. Heaviside and Gibbs made up for this in their subsequent work, leading to their vector algebra from which we get the dot and cross products and the divergence and curl.

5. Hamilton's Quaternions

In 1843 Hamilton introduced his quaternion theory and expanded the field of vector analysis. Hamilton’s original notation was quite innovative at the time; however, in my opinion, still muddies the water in current usage.

$$\boxed{\underbrace{a=a_0+a_1i+a_2j+a_3k}_{\text{Hamilton's Quaternion Form}}}$$

and where

$$\boxed{\begin{aligned} &i^2 = j^2 = k^2 = ijk = -1 \\ &ij = k = -ji \\ &jk = i = -kj \\ &ki = j = -ik \end{aligned}}$$

First, I will use the four dot over-stroke to denote a quaternion, for example $\ddddot{A}$.

$$\boxed{\begin{aligned} \ddddot{A} &= A_0 + A_x\hat{x} + A_y\hat{y} + A_z\hat{z} \\ &= A_0 + \vec{A} \end{aligned}}$$

where

$$\boxed{\begin{aligned} &\hat{x}^2 = \hat{y}^2 = \hat{z}^2 = \hat{x}\hat{y}\hat{z} = -1 \\ &\hat{x}\hat{y} = \hat{z} = -\hat{y}\hat{x} \\ &\hat{y}\hat{z} = \hat{x} = -\hat{z}\hat{y} \\ &\hat{z}\hat{x} = \hat{y} = \hat{x}\hat{z} \end{aligned}}$$

Hamilton used i, j, k for his unit vectors. I imagine that was because the notation of the over-caret for a unit vector had not evolved at that time. Something like $f(x)x$ to express the product of a function of x times the unit vector x is confusing and would suggest that $f(x)i$ would be much clearer ***until*** you could write $f(x)\hat{x}$. The i unit vector is actually for the x axis, so in my mind, $\hat{x}$ is a much better choice for the unit vector for the x axis than the unit vector i.

At last, it was becoming clear to me what Maxwell means by the scalar and vector products and it comes from quaternion multiplication. If you take two quaternion vectors $\ddddot{A}$ **and** $\ddddot{B}$ (In my notation the four over-dots define a quaternion) whose components are

$A_0, A_x, A_y, A_z,$ and B_0, B_x, B_y, B_z or simply $\left(A_0 + \vec{A}\right)$ and $\left(B_0 + \vec{B}\right)$

are multiplied as follows

$$
\begin{aligned}
\overset{....}{A}\,\overset{....}{B} &= \left(A_0+A_x\hat{\mathrm{i}}+A_y\hat{\mathrm{j}}+A_z\hat{k}\right)\left(B_0+B_x\hat{i}+B_y\hat{j}+B_z\hat{k}\right)\\
&=\left(A_0B_0-A_xB_x-A_yB_y-A_zB_z\right)\\
&\quad+A_0\left(B_x\hat{i}+B_y\hat{j}+B_z\hat{k}\right)+B_0\left(A_x\hat{\mathrm{i}}+A_y\hat{\mathrm{j}}+A_z\hat{k}\right)\\
&\quad+\left(A_yB_z-A_zB_y\right)\hat{\mathrm{i}}+\left(A_zB_x-A_xB_z\right)\hat{\mathrm{j}}+\left(A_xB_y-A_yB_x\right)\hat{k}\\
&\\
&=A_0B_0-\left(A_xB_x+A_yB_y+A_zB_z\right)+A_0\vec{B}+B_0\vec{A}\\
&\quad+\left(A_yB_z-A_zB_y\right)\hat{\mathrm{i}}+\left(A_zB_x-A_xB_z\right)\hat{\mathrm{j}}+\left(A_xB_y-A_yB_x\right)\hat{k}\\
&\\
&=\boxed{\underbrace{A_0B_0-\vec{A}\circ\vec{B}}_{\text{Scalar}}+\underbrace{A_0\vec{B}+B_0\vec{A}+\vec{A}\times\vec{B}}_{\text{Vector}}}
\end{aligned}
$$

If we make $\overset{....}{A}$ **and** $\overset{....}{B}$ "pure", as I have seen in some literature, quaternions or vectors $\vec{A}$ **and** $\vec{B}$ by assigning the scalars to zero value, we have

$$
\begin{aligned}
\overset{....}{A}\,\overset{....}{B} &= \vec{A}\vec{B}\\
&=\left(A_x\hat{\mathrm{i}}+A_y\hat{\mathrm{j}}+A_z\hat{k}\right)\left(B_x\hat{i}+B_y\hat{j}+B_z\hat{k}\right)\\
&=\left(-A_xB_x-A_yB_y-A_zB_z\right)\\
&\quad+\left(A_yB_z-A_zB_y\right)\hat{\mathrm{i}}+\left(A_zB_x-A_xB_z\right)\hat{\mathrm{j}}+\left(A_xB_y-A_yB_x\right)\hat{k}\\
&\\
&=-\left(A_xB_x+A_yB_y+A_zB_z\right)\\
&\quad+\left(A_yB_z-A_zB_y\right)\hat{\mathrm{i}}+\left(A_zB_x-A_xB_z\right)\hat{\mathrm{j}}+\left(A_xB_y-A_yB_x\right)\hat{k}\\
&\\
&=\boxed{\underbrace{\vec{A}\circ\vec{B}}_{\text{Scalar}}+\underbrace{\vec{A}\times\vec{B}}_{\text{Vector}}}
\end{aligned}
\tag{2}
$$

Here we see the dot and cross products evolve naturally and Maxwell saw a formalization for the equations he had derived in his work and he could give them a name and symbol. "Ah Ha". And, he shared that in Article 25.

6. Vector Algebra Again

Now we can see that Vector Algebra as taught today cannot deal with equation (1); however, if we allow that vector algebra is a subset of quaternions, equation (1) falls out naturally as shown in equation (2) above. note that this comes without collateral damage to vector algebra.

This reminds me that the many presentations of "Maxwell's Equations" I received in my education and career were not actually Maxwell's Equations, but the presentation and reduction of Maxwell's Equations by Gibbs and Heaviside.

Both Gibbs and Heaviside felt that quaternions were, I will say, too awkward and complicated for general use decided to create vector algebra and created "vector algebra" that we all enjoy today. They felt the need to distance vector algebra from the complexities of quaternions. I feel that that was probably appropriate at the time. However, that legacy is important, even though, it denies the fact that vector algebra is a subset of quaternions and was derived from quaternions. I can't imagine the difficulties in getting Maxwell's ideas accepted if this "divorce" had not been done. It was difficult enough as it was. As a side note, I often wonder what might have happened if Maxwell had lived long enough to collaborate with Heaviside and Gibbs. It is something to ponder.

7. On Quaternions

From what I have been able to tell, quaternions have not found much real usage except in the field of rotations involving things like graphics and inertial platforms. For example, I implemented quaternion rotation in the position solution for the dynamics of a strapdown inertial platform in the handle. That was my introduction to quaternions and it was difficult for me; however, it was really a "cookbook" implementation and did not get into the depth of quaternions. It is a well developed art built on quaternions, yet, has little meaning to other applications.

The question that sticks in my mind is what physical property can be represented by a quaternion where it's definition is defined by a scalar and vector?

8. Vector and Quaternion Products

I am an engineer and not a mathematician and I have wondered for sometime why the direct product of two vectors or quaternions have no meaning and are ignored. We use the dot and cross products instead. It is interesting that the product of two normal vectors produces a quaternion, that is a scalar and a vector; however, it is never spoken of, not in my education anyway. It seems to me that the mathematician's could not make sense of it, ignored it and got around it by devising the dot and cross products by proving that they are valid constructs.

As an engineer I am less interested in a mathematical proof rather than that it works and can predict the performance of a physical entity. The mathematical proof is the icing on the cake.

9. Conclusion

It seems to me that it is evident that vector algebra is *necessarily* a subset of quaternion theory and we don't contaminate vector algebra by acknowledging that.

End of Document

D:\!\Maxwell\Useful\OnQuaternionsAndVectorAlgebra R03.doc