AN AMAZING SYMMETRY PARADOX OF THE "SIMPLE" VECTOR PRODUCT

CIBELLE CELESTINO SILVA and
ROBERTO DE ANDRADE MARTINS

Group of History and Theory of Science, DRCC, Instituto de Física "Gleb Wataghin", P.O. Box 6165, UNICAMP, 13083-970 Campinas, SP, Brazil

ABSTRACT: This paper analyzes some difficulties related to the use of vector product in the teaching of electromagnetic theory and shows that these difficulties have deep historical roots, arising from deficient conceptual elucidation in the works of the founders of electromagnetism and vector analysis.

The direction of the vector $\vec{\mathbf{F}} = q\vec{\mathbf{v}} \times \vec{\mathbf{B}}$ is perpendicular to both vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{B}}$, and its direction is determined by the "right hand rule". Why does the force $\vec{\mathbf{F}}$ have one direction and not the opposite one? The usual answer to this question is that this is a "convention". However, behind this convention, there stand non-arbitrary symmetry proprieties of the vectors. It is also necessary to stress the difference between a vector quantity and its graphical representation. Both the right hand rule and graphical representation of a vector had been discussed by the founders of the vector calculus.

At the beginning of 20$^{\text{th}}$ century, some authors such as Paul Langevin and Woldemar Voigt discussed conceptual issues and the better notation to represent the different kinds of vectors. They were not successful, however, in creating a new vector system. In this specific case, the historical study allowed us to unravel the root of conceptual difficulties underlying the current teaching practice of the vector product. However, the solution of those difficulties is not purely historical: it requires a deeper discussion of the symmetry properties of the two types of vectors, and the use of a new notation.

1. Introduction

The teaching of electromagnetic theory is quite difficult. Besides the abstract character of the electromagnetic fields, there is a lot of new mathematical concepts to be taught together with the new physical concepts. This paper will analyze some difficulties related to the use of vector products in electromagnetism and will show that these difficulties have deep historical roots, arising from deficient conceptual elucidation in the works of the founders of electromagnetism and vector analysis.

This vector operation is usually introduced in the fundamental physics course in the chapters on dynamics of rotation. However, it is within electromagnetism that the vector product acquires an essential role, when the magnetic force $\vec{\mathbf{F}}$ acting upon a moving charge *q* is introduced and computed using the equation $\vec{\mathbf{F}} = q\vec{\mathbf{v}} \times \vec{\mathbf{B}}$, the vector $\vec{\mathbf{F}}$ is perpendicular to both vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{B}}$, and its direction is determined by the "right hand rule".

The right hand rule is just a mnemonic device. Why does the force $\vec{\mathbf{F}}$ have one direction and not the opposite one? The usual answer to this question is that this is a "convention". As a matter of fact, the direction of the vector product is indeed a convention. However, behind this convention, there stand non-arbitrary symmetry proprieties of the vectors. There are *polar vectors* and *axial vectors*, and they have quite different symmetry properties.

It is also necessary to stress the difference between a vector quantity and its graphical representation. In order to represent both polar and axial vector quantities graphically we usually use arrows, that is, we use the same symbol to represent two different things. Both the right hand rule and graphical representation of a vector had been discussed by the founders of the vector calculus.

2. The symmetry proprieties of vectors

There are two kinds of directional physical magnitudes: polar vectors (such as those that correspond to displacement, velocity, acceleration, force and electric field) and axial vectors (such as those that correspond to angular velocity, torque, angular momentum and magnetic field). By multiplying two vectors it is possible to generate new objects, and the vector product of two polar vectors is an axial vector.

When one represents a *polar* vector $\vec{\mathbf{A}}$ as $\vec{\mathbf{A}} = A_x \vec{\mathbf{i}} + A_y \vec{\mathbf{j}} + A_z \vec{\mathbf{k}}$, the symbols $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ are understood to be unit *polar* vectors and the components of $\vec{\mathbf{A}}$ (that is, A_x, A_y and A_z) are scalar quantities. Now, if one attempts to represent an *axial* vector $\vec{\mathbf{C}}$ as $\vec{\mathbf{C}} = C_x \vec{\mathbf{i}} + C_y \vec{\mathbf{j}} + C_z \vec{\mathbf{k}}$, a problem arises: if the symbols $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ stand for unit *polar* vectors, then $\vec{\mathbf{C}}$ should also be a polar vector, because the addition of polar vectors produces polar vectors.

There is a way out of the problem, however. The answer is in C_x, C_y, C_z. It is possible to interpret them as *pseudoscalars*. The product of a polar vector by a pseudoscalar is an axial vector, and vice-versa. Therefore, if one assumes that $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ are *polar* vectors, it is possible to represent an *axial* vector as $\vec{\mathbf{C}} = C_x \vec{\mathbf{i}} + C_y \vec{\mathbf{j}} + C_z \vec{\mathbf{k}}$ provided that C_x, C_y, C_z are pseudoscalars. This is implicitly presupposed in the usual vector algebra used nowadays by physicists, although elementary textbooks do not introduce this concept.

Conversely, it would be possible to regard $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ as unit *axial* vectors. In that case, it would be possible to represent a polar vector $\vec{\mathbf{A}}$ as $\vec{\mathbf{A}} = A_x \vec{\mathbf{i}} + A_y \vec{\mathbf{j}} + A_z \vec{\mathbf{k}}$, if A_x, A_y, A_z are pseudoscalar quantities.

3. An amazing paradox

Among the various physical laws associated with the vector product, let us consider the expression $\vec{\mathbf{F}} = q\vec{\mathbf{v}} \times \vec{\mathbf{B}}$. One may wonder how could it be possible that the vector product between the polar vector $\vec{\mathbf{v}}$ and the axial vector $\vec{\mathbf{B}}$ produces a polar vector $\vec{\mathbf{F}}$. This is not so obvious and we should look more closely at the nature of those vectors to get the answer.

Let us adopt the usual interpretation of the symbols $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ as unit *polar* vectors. As discussed above, the components of $\vec{\mathbf{F}}$ and $\vec{\mathbf{v}}$ must be scalars and the components of $\vec{\mathbf{B}}$ must be pseudoscalars. The distinction between the quantities involved is usually hidden by the use of the same notation to describe them. We shall use a bent arrow to identify axial vectors (as suggested by Paul Langevin in 1912) and a bent bar to identify pseudoscalars.

The force $\vec{\mathbf{F}}$ is obtained by the vector product $q\,\vec{\mathbf{v}} \times \overset{\circlearrowleft}{\mathrm{B}}$ and can be written as

$$\vec{\mathbf{F}} = q\,\vec{\mathbf{v}} \times \overset{\circlearrowleft}{\mathrm{B}} = q.(v_x \vec{\mathbf{i}} + v_y \vec{\mathbf{j}} + v_z \vec{\mathbf{k}}) \times (\breve{B}_x \vec{\mathbf{i}} + \breve{B}_y \vec{\mathbf{j}} + \breve{B}_z \vec{\mathbf{k}}) \therefore$$
$$\therefore \vec{\mathbf{F}} = q(v_y \breve{B}_z - v_z \breve{B}_y)\vec{\mathbf{i}} + q(v_z \breve{B}_x - v_x \breve{B}_z)\vec{\mathbf{j}} + q(v_x \breve{B}_y - v_y \breve{B}_x)\vec{\mathbf{k}}\,. \tag{1}$$

The electric charge **q** is a scalar. In the final expression, each quantity between parentheses is a pseudoscalar. Therefore, the vector product $\vec{\mathbf{v}} \times \vec{\mathbf{B}}$ should be an axial vector. A paradox arises, since the force $\vec{\mathbf{F}}$ is a polar vector and the right side of the equation seems to be an axial vector.

The problem, now, is the interpretation of $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ in the final expression. It is necessary to regard them as axial vectors in the above equation because they are the result of vector products: $\vec{\mathbf{i}} = \vec{\mathbf{j}} \times \vec{\mathbf{k}}$, $\vec{\mathbf{j}} = \vec{\mathbf{k}} \times \vec{\mathbf{i}}$, $\vec{\mathbf{k}} = \vec{\mathbf{i}} \times \vec{\mathbf{j}}$. However, to stress that $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ are now axial vectors, we should use a different notation, such as bent arrows. Therefore, we will have $\overset{\circlearrowleft}{\mathrm{i}} = \vec{\mathbf{j}} \times \vec{\mathbf{k}}$, etc., and instead of equation (1) we should have:

$$\vec{\mathbf{F}} = q(v_y \breve{B}_z - v_z \breve{B}_y)\,\overset{\circlearrowleft}{\mathrm{i}} + q(v_z \breve{B}_x - v_x \breve{B}_z)\,\overset{\circlearrowleft}{\mathrm{j}} + q(v_x \breve{B}_y - v_y \breve{B}_x)\,\overset{\circlearrowleft}{\mathrm{k}}. \tag{2}$$

Notice, however, that now we have a new set of unit vectors $\overset{\circlearrowleft}{\mathrm{i}}, \overset{\circlearrowleft}{\mathrm{j}}, \overset{\circlearrowleft}{\mathrm{k}}$ and this is inconsistent with the usual interpretation of $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ as polar vectors.

If one wants to retain only the old set $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$ of unit polar vectors, it is necessary to use an unit pseudoscalar $\breve{1}$ and to introduce the new rules $\vec{\mathbf{j}} \times \vec{\mathbf{k}} = \breve{1}\vec{\mathbf{i}}$, etc. In that case, instead of Equation (2) we can write:

$$\vec{\mathbf{F}} = \breve{1}\,[q(v_y \breve{B}_z - v_z \breve{B}_y)\vec{\mathbf{i}} + q(v_z \breve{B}_x - v_x \breve{B}_z)\vec{\mathbf{j}} + q(v_x \breve{B}_y - v_y \breve{B}_x)\vec{\mathbf{k}}] \tag{3}$$

It is impossible to build a closed vector algebra using only polar (or axial) vectors and scalars, because the vector product will generate axial vectors from polar vectors (and vice-versa). Therefore, when one represents vectors by their components, there are two possibilities: either one uses both polar and axial unit vectors, together with scalars and suitable multiplication rules (such as $\overset{\circlearrowleft}{\mathrm{i}} = \vec{\mathbf{j}} \times \vec{\mathbf{k}}$ etc.) or one uses only polar or axial unit vectors together with both scalars and pseudoscalars and suitable multiplication rules (such as $\vec{\mathbf{j}} \times \vec{\mathbf{k}} = \breve{1}\vec{\mathbf{i}}$, etc.). Whatever the choice, it is useful to use different symbols to represent scalars, pseudoscalars, polar vectors and axial vectors.

Suppose, for instance, that we choose to use only unit axial vectors, that is $\overset{\circlearrowleft}{\mathrm{i}}$, $\overset{\circlearrowleft}{\mathrm{j}}$, $\overset{\circlearrowleft}{\mathrm{k}}$. In this case, the components of the velocity $\breve{v}_x, \breve{v}_y, \breve{v}_z$ should be pseudoscalars and the components of the magnetic field B_x, B_y, B_z should be scalars in order to have the velocity $\vec{\mathbf{v}}$ as a polar vector and the magnetic field $\overset{\circlearrowleft}{\mathrm{B}}$ as an axial vector. The force $\vec{\mathbf{F}}$ may be calculated as $\vec{\mathbf{F}} = q\vec{\mathbf{v}} \times \overset{\circlearrowleft}{\mathrm{B}}$. This force is a sum of polar vector terms such as $q(\breve{v}_y B_z - \breve{v}_z B_y)\overset{\circlearrowleft}{\mathrm{i}}$.

4. The quaternion system and the rise of contemporary vector system

A quaternion is a special mathematical entity containing four components. William Hamilton (1805-1865) invented it in 1843. A quaternion can be written as q = a + bi + cj + zk, where i, j, k are imaginary units which Hamilton interpreted as "versors" (meaning "rotators") that produce a $\pi/2$ rotation when applied to a vector. He gave the name "pure quaternion" or "vector" to a quaternion without a scalar part, of the form q = bi + cj + zk.

Within the quaternion theory, the units *i, j, k* are *axial* vectors. However, as Hamilton called pure quaternions "vectors" and stated that vectors can represent entities such as position and displacement (that is, *polar* vectors), there was from the very beginning a confusion between the two types of vectors. This was the root of misunderstandings that still afflict the teaching of vector calculus.

The traditional notation of arrows to represent both polar and axial vectors makes it difficult to realize that the force and velocity are physical quantities with polar symmetry whereas the magnetic field has axial symmetry. This tradition began in late 19th century with the invention of the present day vector system by Gibbs and Heaviside from quaternion system. Our full paper will discuss some of the historical features of this development.

At the beginning of 20th century, some authors such as Paul Langevin and Woldemar Voigt discussed conceptual issues and the better notation to represent the different kinds of vectors. They were not successful, however, in creating a new vector system.

The subject discussed in this paper arose from an historical study of the foundations of vector analysis. In this specific case, the historical study allowed us to unravel the root of conceptual difficulties underlying the current teaching practice of the vector product. However, the solution of those difficulties is not purely historical: it requires a deeper discussion of the symmetry properties of the two types of vectors, and the use of a new notation.