A Historical Study of Vector Analysis

C.T. Tai
Radiation Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, Michigan
48109-2122

Technical Report RL 915

THE UNIVERSITY OF MICHIGAN

Radiation Laboratory
Department of Electrical Engineering
and Computer Science
Ann Arbor, Michigan 48109-2122
USA

May, 1995

1. Introduction

Vector analysis is an indispensable tool in the teaching and learning of electromagnetics, hydrodynamics, and mechanics. In a book on the history of vector analysis [1], Michael J. Crowe made a very thorough investigation of the decline of quaternion analysis and the evolution of vector analysis during the 19th century until the beginning of this century. The topics covered are mostly vector algebra and quaternion analysis. He did not comment much on the technical aspects of the subject from the point of view of a mathematician or theoretical physicist. For example, the difference between the presentations of Gibbs and Heaviside, considered to be two founders of modern vector analysis, is not discussed in Crowe's book, and less attention is paid to the history of vector differentiation and integration, and to the role played by the del operator, ∇.

A short history of vector analysis is also found in several other books. For example, in a book by Burali-Forti and Marcolongo [2] published in 1920, there are four historical notes in the appendix entitled: On the definition of abstraction, On vectors, On vector and scalar (interior) products, and On grad, rot, div. In another book published in 1965 by Moon and Spencer [3] there is a brief but very critical review of the history of vector analysis from a technical perspective. Many of the assertions in that book will be discussed later. One important reason for these two authors to present vector analysis by way of tensor analysis is stated very firmly in the introduction of that book [3, p. 9]:

> The present book differs from the customary textbook on vectors in stressing the idea of invariance under groups of transformations. In other words, elementary tensor technique is introduced, and in this way, the subject is placed on the firm, logical foundation which vector textbooks have previously lacked.

In Appendix C of that book [3, p. 323] they make the following comment about the del operator:

> In reading the foregoing book [referring to their book], one may wonder why nothing has been said about the operator ∇, which is usually considered such an important part of vector analysis. The truth is that ∇, though providing the subject

> with fluency, is an unreliable device because it often gives incorrect results. For this reason - and because it is not necessary - we have omitted it in the body of the book. Here, however, we shall indicate briefly the use of the operator ∇

These two quotations are sufficient to indicate that after decades of application of vector analysis there seems to be no systematic treatment of the subject that could be considered satisfactory according to these two authors. This observation is also supported by the fact that we have so far no standard notations in vector analysis. Many books on electromagnetics, for example, use the linguistic notations for the gradient, divergence, and curl - namely, grad u, div $\overline{f}$, and curl $\overline{f}$, while many others prefer Gibbs' notations for these functions, namely ∇u, $\nabla \cdot \overline{f}$, and $\nabla \times \overline{f}$. Is there a good explanation to the students why we do not yet have a universally accepted standard notation besides saying: "It is a matter of personal choice."? In regard to Moon and Spencer's comments about the lack of a firm, logical foundation in previous books on vector analysis, there has been no elaboration. They do give an example of an incorrect result from using ∇ to find the expression for divergence in an orthogonal curvilinear coordinate system, but no explanation was given as to the cause of such a wrong result. In fact, the views expressed by these two authors are also found in many books treating vector analysis. These will be reviewed and commented upon later.

In writing this essay, we have in mind the reader who already has an acquaintance with the subject matter of vector analysis, and who feels the need for a critical scrutiny of what he or she has already learned. Many students must have felt such a need, because the conventional curriculum avoids thorough critical examination of many topics. The primary objective of many schools of physical and engineering sciences is to teach the students how to <u>use</u> particular tools (such as vector analysis) to formulate and to solve problems. It is usually felt that students in the hard sciences, particularly at the undergraduate level, do not have the luxury of time to deal with the logic and many of the fine points of subject fundamentals. Many of these subtle details are overlooked in favor of developing skills in applying results, sometimes bluntly. This essay attempts to point out those inadequate or illogical treatments of some basic aspects of the subject which have arisen in the past, and offers a more logical and systematic alternative for the reader to consider.

2. Notations and Operators

2.1 Past and Present Notations in Vector Analysis

In a book on advanced vector analysis published in 1924, Weatherburn [4] compiled a table of notations in vector analysis which had been used up to that time. The names of the authors in that table are: Gibbs/Wilson, Heaviside, Abraham, Ignatowsky, Lorentz, and Burali-Forti/Marcolongo. In Moon and Spencer's *Vectors* (quoted in the previous chapter), published in 1965, there is also a table of notations. The names of the authors in that table are: Maxwell, Gibbs, Gibbs/Wilson, Heaviside, Gans, Lagally, Burali-Forti, Marcolongo, Phillips, and Moon/Spencer. Among these authors, Gibbs, Wilson, Phillips, Moon, and Spencer are American. Maxwell and Heaviside belong to the English schools. Abraham, Ignatowsky, Gans, and Lagally belong to the German schools. Lorentz was a Dutch physicist and Burali-Forti and Marcolongo were Italians. Ignatowsky was a native of Russia but was trained in Germany. For our study, we prepare another list which contains several contemporary authors and some more notations in Table I. The dyadic notation is added because we need it to characterize the gradient of a vector, which is a dyadic function. A rudimentary introduction to dyadic analysis will be given after we present the list of notations given below. In looking at this list, most readers will recognize the linguistic notations grad u, div $\bar{a}$, curl $\bar{a}$, or rot $\bar{a}$ for the three key functions. They are probably accustomed to Gibbs' notations ∇u, $\nabla.\alpha$, and $\nabla \times \alpha$ except that the period '.' in $\nabla.\alpha$ is now replaced by a dot '·' as in Wilson's notations, and his Greek letters for vectors are now commonly replaced by boldface, Clarendon or equivalent fonts while the linguistic notations are used by many authors in Europe and a few in the U. S. A. There is no doubt that Gibbs' notations have been adopted in many books published in the U. S. A. We quote here two very well known books in electromagnetic theory, one by Stratton, and another by Jackson Their treatises are well known to many electrical engineers as well as physicists.

Historically, vector analysis was developed a few years after Maxwell formulated his monumental work in electromagnetic theory. When he wrote his treatise on electricity and magnetism [5] in

Table I: Notations

Author(s)	Vectors	Scalar Product	Vector Product	Dyadic in 3-space	Tensor in 3-space
Maxwell [5]	σ,ρ $\mathcal{A},\mathcal{B}$*	$S\sigma\rho$	$V\sigma\rho$	–	–
Gibbs [6]	α,β	$\alpha.\beta$	$\alpha\times\beta$	$\alpha\beta$	–
Wilson [7]	$\bar{a},\bar{b}$	$\bar{a}\cdot\bar{b}$	$\bar{a}\times\bar{b}$	$\bar{a}\bar{b}$	–
Heaviside [8]	$\bar{a},\bar{b}$	$\bar{a}\bar{b}$	$V\bar{a}\bar{b}$	–	–
Gans [9]	$\mathcal{A},\mathcal{B}$	$(\mathcal{A},\mathcal{B})$	$[\mathcal{A},\mathcal{B}]$	–	–
Burati-Forti/ Marcolongo [2]	$\bar{a},\bar{b}$	$\bar{a}\times\bar{b}$	$\bar{a}\wedge\bar{b}$	–	–
Stratton [10]	$\bar{a},\bar{b}$	$\bar{a}\cdot\bar{b}$	$\bar{a}\times\bar{b}$	–	T_{ij}
Jackson [11]	$\bar{a},\bar{b}$	$\bar{a}\cdot\bar{b}$	$\bar{a}\times\bar{b}$	$\overleftrightarrow{T}$	T_{ij}
Moon/ Spencer [12]	$\bar{a},\bar{b}$	$\bar{a}\cdot\bar{b}$	$\bar{a}\times\bar{b}$	–	T_{ij}

Author(s)	gradient of a scalar	gradient of a vector	divergence of a vector	curl or rot of a vector	Laplacian of a scalar	Laplacian of a vector
Maxwell [5]	∇u	–	$-S\nabla\rho$	$V\nabla\rho$	$\nabla^2 u$	–
Gibbs [6]	∇u	$\nabla\alpha$	$\nabla\cdot\alpha$	$\nabla\times\alpha$	$\nabla\cdot\nabla u$	$\nabla\cdot\nabla\alpha$
Wilson [7]	∇u	$\nabla\bar{a}$	$\nabla\cdot\bar{a}$	$\nabla\times\bar{a}$	$\nabla\cdot\nabla u$	$\nabla\cdot\nabla\bar{a}$
Heaviside [8]	∇u	$\nabla.\bar{a}$	$\nabla\bar{a}$; div $\bar{a}$	$V\nabla\bar{a}$; curl $\bar{a}$	$\nabla^2 u$	$\nabla^2\bar{a}$
Gans [9]	∇u; grad u	–	$\nabla\cdot\bar{a}$; div $\bar{a}$	$\nabla\times\bar{a}$; rot $\bar{a}$	Δu	$\Delta\bar{a}$
Burati-Forti/ Marcolongo [2]	grad u	–	div $\bar{a}$	rot $\bar{a}$	$\Delta_2 u$	$\Delta'_2\bar{a}$
Stratton [10]	∇u	$\nabla\bar{a}$	$\nabla\cdot\bar{a}$	$\nabla\times\bar{a}$	$\nabla^2 u$	$\nabla^2\bar{a}$
Jackson [11]	∇u	$\nabla\bar{a}$	$\nabla\cdot\bar{a}$	$\nabla\times\bar{a}$	$\nabla^2 u$	$\nabla^2\bar{a}$
Moon/ Spencer [12]	grad u	–	div $\bar{a}$	curl $\bar{a}$	$\nabla^2 u$	✡ $\bar{a}$

* Upper case script symbols are used here in place of capital German letters originally used by Maxwell and Gans.

1873, vector analysis was not yet available. Its forerunner, quaternion analysis, developed by Hamilton (1805 - 1865) in 1843, was then advocated by many of Hamilton's followers. It is probably for this reason that Maxwell wrote an article in his book (Article 618) entitled "Quaternion Expressions for the Electromagnetic Equations." Maxwell's notations on our list are based on this document. Actually, he used very little of these notations in the entire book and in his papers published elsewhere.

The notation used by Heaviside is not conventional from the present point of view. His notation for the scalar product and the divergence does not have a dot and his notation for the curl is of quaternion form like Maxwell's. The notations used by Burati-Forti and Marcolongo are obsolete now. Occasionally we still see the notation $\bar{a} \wedge \bar{b}$ for the cross product in European books. As a whole, we now have basically two sets of notations in current use: the linguistic notation and Gibbs' notation. The names of Moon and Spencer are included on our list primarily because these two authors considered the use of ∇ to be unreliable and they frequently emphasize their view that the rigorous method of formulating vector analysis is to follow the route of tensor analysis. In addition, their new notation for the Laplacian of a vector function will be a subject of detailed examination in the chapter on orthogonal curvilinear systems.

2.2 Quaternion Analysis

The rise of vector analysis as a distinct branch of applied mathematics has its origin in quaternion analysis. It is therefore necessary to review briefly the laws of quaternion analysis to show its influence upon the development of vector analysis and also explain the notations in the previous list. Quaternions are complex numbers of the form

$$q = w + ix + jy + kz \tag{2.1}$$

where w, x, y, and z are real numbers, and i, j, and k are unit vectors, directed along the x, y, and z axes respectively. These unit vectors obey the following laws of multiplication:

$$\begin{aligned} ij &= k,\ jk = i,\ ki = j \\ ji &= -k,\ kj = -i,\ ik = -j \\ ii &= jj = kk = -1 \end{aligned} \tag{2.2}$$

We must not at this stage associate the above relations with our current laws in vector analysis. We consider the subject as a new algebra, which is indeed the case. The product of the multiplication of two quaternions σ and ρ in which the scalar parts w and w' are zero is obtained as follows:

We let

$$\begin{aligned} \sigma &= iD_1 + jD_2 + kD_3 \\ \rho &= iX + jY + kZ \end{aligned}$$

then

$$\begin{aligned} \sigma\rho = &-(D_1X + D_2Y + D_3Z) \\ &+ i(D_2Z - D_3Y) \\ &+ j(D_3X - D_1Z) \\ &+ k(D_1Y - D_2X) \end{aligned} \tag{2.3}$$

The resultant quaternion, $\sigma\rho$, has two parts, one scalar and one vector. In Hamilton's original notation they are:

$$S.\sigma\rho = -(D_1X + D_2Y + D_3Z) \tag{2.4}$$

$$\begin{aligned} V.\sigma\rho = &\ i(D_2Z - D_3Y) \\ &+ j(D_3X - D_1Z) \\ &+ k(D_1Y - D_2X) \end{aligned} \tag{2.5}$$

The period between S or V and $\sigma\rho$ can be omitted without causing any ambiguity. When one identifies σ as ∇, Hamilton's del operator, that is:

$$\sigma = \nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \tag{2.6}$$

then

$$S\nabla\rho = -\left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}\right) \tag{2.7}$$

$$\begin{aligned} V\nabla\rho = {} & i\left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}\right) \\ & + j\left(\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}\right) \\ & + k\left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right) \end{aligned} \tag{2.8}$$

$S\nabla\rho$ is the quaternion notation used by Maxwell for the negative of the divergence of ρ and he called it the convergence of ρ. He used the quaternion notation $V\nabla\rho$ for the curl of ρ. The term 'curl' was his creation, and it is now a standard name. According to Crowe [1, p. 142] the term 'divergence' was originally due to William Kingdom Clifford (1845 - 1879) who was also the first person to define the modern notations for the scalar and vector products. However, his original definition of the scalar product is the negative of the modern scalar product. In the list of notations, we notice that Heaviside used the quaternion notation for the curl even though he was opposed to quaternion analysis. In one of his writings [8, p. 35] he concurred with Gibbs' treatment of vector analysis but criticized Gibbs' notations without offering a reason. Heaviside's remark will be quoted and discussed in Sec. 3.1.

We should mention that the long controversy between quaternionists led by Tait and the proponents of the then new vector analysis led by Gibbs was covered in great detail by Crowe [1]. Such stories are very educational to young scientists and engineers. The other topic which needs to be reviewed deals with the dyadic notation which was used by Gibbs, Wilson, and Jackson but not the other authors in the previous list. An understanding of this notation is necessary in order to explain the Laplacian of a vector function, particularly in the general curvilinear coordinate system. Additionally, dyadic analysis is a natural extension of vector analysis. Problems formulated using tensor analysis in a three-dimensional Euclidean space can be handled by dyadic analysis in a relatively simpler format.

2.3 Dyadic Analysis

For the time being, we will provide the basic formulas in dyadic analysis to be used in our investigation of the past works. In vector algebra, a vector function denoted by $\overline{F}$ is represented in a Cartesian coordinate system by:

$$\overline{F} = \sum F_i \hat{x}_i, \ i = 1,2,3 \tag{2.9}$$

where $\hat{x}_i$, with $i = 1,2,3$ denotes the unit vector in the x,y,z direction and F_i corresponds to the components of $\overline{F}$ in the x,y, and z direction. It is understood that the summation goes from i=1 to 3. A dyadic function or a dyadic for short, denoted by $\overline{\overline{F}}$ is defined in the same Cartesian coordinate system by

$$\overline{\overline{F}} = \sum_j \overline{F}_j \hat{x}_j, \ j = 1,2,3 \tag{2.10}$$

where

$$\overline{F}_j = \sum_i F_{ij} \hat{x}_i, \ i = 1,2,3 \tag{2.11}$$

denotes three independent or distinct vector functions. The relative position of $\overline{F}_j$ and $\hat{x}_j$ in (2.10) must be maintained in that order, and one is not supposed to interchange the ordering of these two vectors. In other words, the commutative rule does not apply to (2.10). When (2.11) is substituted into (2.10), we obtain:

$$\overline{\overline{F}} = \sum_i \sum_j F_{ij} \hat{x}_i \hat{x}_j \tag{2.12}$$

Equations (2.10) - (2.12) contain the definition of a dyadic in a Cartesian coordinate system. There are nine scalar components of $\overline{\overline{F}}$. The doublets $\hat{x}_i \hat{x}_j$, juxtaposed together $\hat{x}_i$ and $\hat{x}_j$ with i,j = (1,2,3), are called dyads, and there are nine of them too. The dyads are not commutative, that is:

$$\hat{x}_i \hat{x}_j \neq \hat{x}_j \hat{x}_i, \ i \neq j$$

The transpose of a dyadic $\overline{\overline{F}}$, denoted by $\left[\overline{\overline{F}}\right]^T$ is defined by

$$\begin{aligned}\left[\overline{\overline{F}}\right]^T &= \sum_j \hat{x}_j \overline{F_j} \\ &= \sum_i \sum_j F_{ij} \hat{x}_j \hat{x}_i \\ &= \sum_i \sum_j F_{ji} \hat{x}_i \hat{x}_j\end{aligned} \tag{2.13}$$

There are two scalar products between a dyadic and a vector. The anterior scalar product between a vector $\overline{a}$ and a dyadic $\overline{\overline{F}}$, denoted by $\overline{a} \cdot \overline{\overline{F}}$, is defined by

$$\overline{a} \cdot \overline{\overline{F}} = \sum_j \left(\overline{a} \cdot \overline{F_j}\right) \hat{x}_j = \sum_i \sum_j a_i F_{ij} \hat{x}_j \tag{2.14}$$

which is a vector. The posterior scalar product between $\overline{a}$ and $\overline{\overline{F}}$, denoted by $\overline{\overline{F}} \cdot \overline{a}$, is defined by

$$\overline{\overline{F}} \cdot \overline{a} = \sum_j \overline{F_j} \left(\hat{x}_j \cdot \overline{a}\right) = \sum_i \sum_j a_j F_{ij} \hat{x}_i \tag{2.15}$$

which is also a vector. In general,

$$\overline{a} \cdot \overline{\overline{F}} \neq \overline{\overline{F}} \cdot \overline{a}$$

The two products are equal when $\overline{\overline{F}}$ is a symmetric dyadic characterized by $F_{ij} = F_{ji}$. There are two vector products possible between $\overline{a}$ and $\overline{\overline{F}}$. The anterior vector product is defined by

$$\overline{a} \times \overline{\overline{F}} = \sum_j \left(\overline{a} \times \overline{F_j}\right) \hat{x}_j = \sum_i \sum_j F_{ij} \left(\overline{a} \times \hat{x}_i\right) \hat{x}_j \tag{2.16}$$

which is a dyadic. The posterior vector product is defined by

$$\overline{\overline{F}} \times \overline{a} = \sum_j \overline{F_j} \left(\hat{x}_j \times \overline{a}\right) = \sum_i \sum_j F_{ij} \hat{x}_i \left(\hat{x}_j \times \overline{a}\right) \tag{2.17}$$

which is another dyadic.

The gradient of a vector function in a Cartesian system, denoted by $\nabla \overline{F}$, is defined by

$$\nabla\overline{F} = \sum_i \hat{x}_i \frac{\partial\overline{F}}{\partial x_i} = \sum_i \sum_j \hat{x}_i \frac{\partial F_j}{\partial x_i}\hat{x}_j$$
$$= \sum_i \sum_j \frac{\partial F_j}{\partial x_i}\hat{x}_i\hat{x}_j \qquad (2.18)$$

which is a dyadic. We will derive the expression for the gradient of $\overline{F}$ in general curvilinear systems later. This introduction of dyadic analysis is merely presented to show that the gradient of a vector function as tabulated in the list of notations in Table I is a dyadic. It should be mentioned that a dyadic can also be written in the form

$$\overline{\overline{F}} = \overline{A}\,\overline{B} \qquad (2.19)$$

Knowing $\overline{\overline{F}}$ and with a specified $\overline{A}$ we can find $\overline{B}$ as follows:

$$\overline{A}\cdot\overline{\overline{F}} = \overline{A}\cdot\overline{A}\,\overline{B} = \left|\overline{A}\right|^2\overline{B}$$

Hence,

$$\overline{B} = \frac{1}{\left|\overline{A}\right|^2}\overline{A}\cdot\overline{\overline{F}} \qquad (2.20)$$

where $\overline{A}\cdot\overline{\overline{F}}$ is the anterior scalar product between $\overline{A}$ and $\overline{\overline{F}}$. If $\overline{B}$ is specified, we can find $\overline{A}$ as follows:

$$\overline{\overline{F}}\cdot\overline{B} = \overline{A}\,\overline{B}\cdot\overline{B} = \overline{A}\left|\overline{B}\right|^2$$

hence

$$\overline{A} = \frac{1}{\left|\overline{B}\right|^2}\overline{\overline{F}}\cdot\overline{B} \qquad (2.21)$$

where $\overline{\overline{F}}\cdot\overline{B}$ is the posterior scalar product between $\overline{B}$ and $\overline{\overline{F}}$. In the list of notations, Gibbs and Wilson use the $\overline{A}\,\overline{B}$ form for the dyadics. The next topic to be reviewed deals with operators.

2.4 Operators

For our convenience we would like to discuss in sufficient detail, the classification and the characteristics of a number of operators appearing in this study. We will focus on unary and binary operators and will consider such operators in cascade or compound arrangements as the complexity of the case at hand requires.

A unary operator involves only one operand. A binary operator needs two operands, one anterior, and another posterior. A cascade operator could be unary or binary. As an example we consider the derivative symbol $\frac{\partial}{\partial x}$ to be a unary operator. When it operates on an operand P, it produces the derivative, $\frac{\partial P}{\partial x}$. In some writings, the operator $\frac{\partial}{\partial x}$ is denoted by D_x. The operand under consideration can be a scalar function of x and other independent variables or a vector function, or a dyadic function, that is,

$$\frac{\partial P}{\partial x}; \quad \frac{\partial a}{\partial x}, \frac{\partial \bar{a}}{\partial x}, \frac{\partial \bar{\bar{F}}}{\partial x}$$

are all valid applications of the unary differential operator.

The partial derivative of a dyadic function in a Cartesian system is defined by

$$\begin{aligned} \frac{\partial \bar{\bar{F}}}{\partial x} &= \sum_j \frac{\partial \bar{F}_j}{\partial x} \hat{x}_j \\ &= \sum_i \sum_j \frac{\partial F_{ij}}{\partial x} \hat{x}_i \hat{x}_j \end{aligned} \tag{2.22}$$

We list in Table II below, several commonly used unary operators and their possible operands. The function a in the weighted differential operator $a\frac{\partial}{\partial x}$ is assumed to be a scalar function. A vector operator such as $\bar{a}\frac{\partial}{\partial x}$ can operate on a dyadic that would yield a 'tridic' - a quantity which is not included in this study. The last operator in the table is the del operator or the gradient operator. It can be applied to an operand which is either a scalar or a vector.

Table II: Valid Application of Some Unary Differential Operators

Operator	$\frac{\partial}{\partial x}$	$a\frac{\partial}{\partial x}$	$\bar{a}\frac{\partial}{\partial x}$	$\nabla = \sum_i \hat{x}_i \frac{\partial}{\partial x_i}$ (2.23)
Type of Operand	$b,\ \bar{b},\ \bar{\bar{b}}$	$b,\ \bar{b},\ \bar{\bar{b}}$	$b,\ \bar{b}$	$b,\ \bar{b}$
Result	$\frac{\partial b}{\partial x},\ \frac{\partial \bar{b}}{\partial x},\ \frac{\partial \bar{\bar{b}}}{\partial x};$	$a\frac{\partial b}{\partial x},\ a\frac{\partial \bar{b}}{\partial x},\ a\frac{\partial \bar{\bar{b}}}{\partial x};$	$\bar{a}\frac{\partial b}{\partial x},\ \bar{a}\frac{\partial \bar{b}}{\partial x};$	$\nabla b,\ \nabla \bar{b}$

A binary operator requires two operands. In arithmetic and algebra we have four binary operators: $+$ (addition), $-$ (subtraction), $\times$ (multiplication), and $\div$ (division). In these cases, we need two operands, one anterior and another posterior, as in $2+3$, $4-3$, 5×3, and $6\div 3$. It should be remarked that the symbols + and − are also used to denote 'plus' and 'minus' signs. For example, $-a = |a|$ when a is negative. In this case the minus sign is not considered to be a binary operator in our classification, but rather as a unary 'sign change' operator. The two binary operators involved frequently in our work are the dot ($\cdot$) and the cross ($\times$). They appear in Gibbs' notations for the scalar and vector products, that is, $\bar{a}\cdot\bar{b}$ and $\bar{a}\times\bar{b}$. We consider the dot and the cross as two binary operators, and their operands, one anterior and one posterior, must be vectors, that is

$$\bar{A}\cdot\bar{B} \text{ and } \bar{A}\times\bar{B}$$

The dot operator is not the same as the multiplication operator in arithmetic and the cross operator is not the same as the multiplication operator, although we use the same symbol.

According to the definitions of the scalar and vector products,

$$\bar{A}\cdot\bar{B} = \bar{B}\cdot\bar{A} = |\bar{A}||\bar{B}|\cos\theta \tag{2.24}$$

$$\bar{A}\times\bar{B} = -\bar{B}\times\bar{A} = |\bar{A}||\bar{B}|\sin\theta\,\hat{c} \tag{2.25}$$

where θ is the angle measured from $\bar{A}$ to $\bar{B}$ in the plane containing these two vectors and $\hat{c}$ is the unit vector $\perp$ to both $\bar{A}$ and $\bar{B}$ and is pointed in the right-screw advancing direction when $\bar{A}$ turns into $\bar{B}$. The dot and the cross can also be applied to operands where one of them or both are dyadics. Thus, we have

$$\overline{A}\cdot\overline{\overline{B}},\overline{\overline{B}}\cdot\overline{A},$$
$$\overline{A}\times\overline{\overline{B}},\overline{\overline{B}}\times\overline{A},\overline{\overline{A}}\cdot\overline{\overline{B}},\overline{\overline{B}}\cdot\overline{\overline{A}} \quad (2.26)$$

The first two entities are vectors and the remaining four are dyadics.

The last group of operators are called cascade or compound operators. Of particular concern in this study is the proper treatment of a pair of operators of different types, which are applied sequentially. When one of the operators is a scalar differential unary operator, and the other is a vector binary operator, there arise a number of hazards in their application which, if not properly treated, could lead to invalid results. Several commonly used cascade operators are of the form:

$$\cdot\frac{\partial}{\partial y},\ \cdot\nabla,\ \times\frac{\partial}{\partial y},\ \times\nabla \quad (2.27)$$

These operators also require two operands; the anterior operand must be a vector and the posterior operand must be compatible with the part in front. Thus we can have

$$\overline{A}\cdot\frac{\partial\overline{B}}{\partial y},\ \overline{A}\cdot\frac{\partial\overline{\overline{B}}}{\partial y};$$
$$\overline{A}\cdot\nabla u,\ \overline{A}\cdot\nabla\overline{B};$$
$$\overline{A}\times\frac{\partial\overline{B}}{\partial y},\ \overline{A}\times\frac{\partial\overline{\overline{B}}}{\partial y};$$
$$\overline{A}\times\nabla u,\ \overline{A}\times\nabla\overline{B} \quad (2.28)$$

In (2.27) the unary operator $\frac{\partial}{\partial y}$ and ∇, and the binary operators $\cdot$ and $\times$ are not commutative; hence, the following combinations or assemblies are not valid cascade operators.

$$\frac{\partial}{\partial y}\cdot,\ \nabla\cdot,\ \frac{\partial}{\partial y}\times,\ \nabla\times \quad (2.29)$$

These assemblies are formed by interchanging the positions of the symbols in (2.27). They are not operators in the sense that we cannot find an operand to form a meaningful entity. For example,

$$\frac{\partial}{\partial y}\cdot\overline{A},\ \frac{\partial}{\partial y}\cdot\overline{\overline{B}};\ \nabla\cdot\overline{A},\ \nabla\cdot\overline{\overline{B}};$$
$$\frac{\partial}{\partial y}\times\overline{A},\ \frac{\partial}{\partial y}\times\overline{\overline{B}};\ \nabla\times\overline{A},\ \nabla\times\overline{\overline{B}} \quad (2.30)$$

do not have any meaningful interpretation.

The reader has probably noticed that there are two assemblies, $\nabla \cdot \overline{A}$ and $\nabla \times \overline{A}$ in (2.30) which correspond to Gibbs' notation for the divergence and curl. This is very true, but that does not mean that $\nabla \cdot \overline{A}$ is a scalar product between ∇ and $\overline{A}$, nor is $\nabla \times \overline{A}$ a vector product between ∇ and $\overline{A}$. In fact, this is a central issue in this study to be examined very critically in the following chapters. We now have the necessary tools to investigate many of the past presentations of vector analysis.

3. The Pioneer Works of J. Willard Gibbs (1839 - 1903)

3.1 Two Pamphlets Printed in 1881 and 1884

Gibbs' original works on vector analysis are found in two pamphlets entitled *Elements of Vector Analysis* [6], privately printed in New Haven. The first consists of 33 pages published in 1881 and the second of 40 pages published in 1884. These pamphlets were distributed to his students at Yale University and also to many scientists and mathematicians including Heaviside, Helmholtz, Kirchhoff, Lorentz, Rayleigh (Lord), Stokes, Tait, and Thomson (J. J.) [12, Appendix IV]. The contents are divided into five chapters and a note on bivectors:

Chapter I. Concerning the algebra of vectors

Chapter II. Concerning the differential and integral calculus of vectors

Chapter III. Concerning linear vector functions

Chapter IV. Concerning the differential and integral calculus of vectors (Supplement to Chapter II)

Chapter V. Concerning transcendental functions of dyadics

A Note on bivector analysis

The most important formulations for our immediate discussions are covered in Articles 50 - 54 and 68 - 71 which are reproduced below:

Functions of Positions in Space

50. Def. - If u is any scalar function of position in space (i.e., any scalar quantity having continuously varying values in space), ∇u is the vector function of position in space which has everywhere the direction of the most rapid increase of u, and a magnitude equal to the rate of that increase per unit of length. ∇u may be called the derivative of u, and u, the primitive of ∇u.

We may also take any one of the Nos. 51, 52, 53 for the definition of ∇u.

51. If ρ is the vector defining the position of a point in space,

$$du = \nabla u \cdot d\rho$$

52.
$$\nabla u = i\frac{du}{dx} + j\frac{du}{dy} + k\frac{du}{dz} \tag{3.1}$$

53.
$$\frac{du}{dx} = i \cdot \nabla u, \quad \frac{\mathrm{du}}{\mathrm{dy}} = j \cdot \nabla u, \quad \frac{du}{dz} = k \cdot \nabla u$$

54. Def. - If ω is a vector having continuously varying values in space,

$$\nabla \cdot \omega = i \cdot \frac{d\omega}{dx} + j \cdot \frac{d\omega}{dy} + k \cdot \frac{d\omega}{dz} \tag{3.2}$$

$$\nabla \times \omega = i \times \frac{d\omega}{dx} + j \times \frac{d\omega}{dy} + k \times \frac{d\omega}{dz} \tag{3.3}$$

$\nabla \cdot \omega$ is called the divergence of ω and $\nabla \times \omega$ its curl.

If we set

$$\omega = Xi + Yj + Zk,$$

we obtain by substitution the equation

$$\nabla \cdot \omega = \frac{dX}{dx} + \frac{dY}{dy} + \frac{dZ}{dz} \tag{3.4}$$

and

$$\nabla \times \omega = i\left(\frac{dZ}{dy} - \frac{dY}{dz}\right) + j\left(\frac{dX}{dz} - \frac{dZ}{dx}\right) + k\left(\frac{dY}{dx} - \frac{dX}{dy}\right) \tag{3.5}$$

which may also be regarded as defining $\nabla \cdot \omega$ and $\nabla \times \omega$.

Combinations of the Operators ∇, $\nabla \cdot$, and $\nabla \times$

68. If ω is any vector function of space, $\nabla \cdot \nabla \times \omega = 0$. This may be deduced directly from the definition of No. 54.

The converse of this proposition will be proved hereafter.

69. If u is any scalar function of position in space, we have by Nos. 52 and 54

$$\nabla \cdot \nabla u = \left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} \right) u \qquad (3.6)$$

70. Def. - If ω is any vector function of position in space, we may define $\nabla \cdot \nabla \omega$ by the equation

$$\nabla \cdot \nabla \omega = \left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} \right) \omega, \qquad (3.7)$$

the expression $\nabla \cdot \nabla$ being regarded, for the present at least, as a single operator when applied to a vector. (It will be remembered that no meaning has been attributed to ∇ before a vector.) It should be noticed that if

$$\omega = iX + jY + kZ,$$

$$\nabla \cdot \nabla \omega = i\nabla \cdot \nabla X + j\nabla \cdot \nabla Y + k\nabla \cdot \nabla Z, \qquad (3.8)$$

that is, the operator $\nabla \cdot \nabla$ applied to a vector affects separately its scalar components.

71. From the above definition with those of Nos. 52 and 54 we may easily obtain

$$\nabla \cdot \nabla \omega = \nabla \nabla \cdot \omega - \nabla \times \nabla \times \omega \qquad (3.9)$$

The effect of the operator $\nabla \cdot \nabla$ is therefore independent of the direction of the axes used in its definition.

In quoting these sections we have changed Gibbs' original notation for the divergence from $\nabla.\omega$ to $\nabla\cdot\omega$, i.e., the period has been replaced by a dot. In addition, some equation numbers have been added for our reference later on.

After Gibbs revealed his new work on vector analysis he was attacked fiercely by Tait, a chief advocate of the quaternion analysis, who stated [13, Preface]:

> Even Prof. Willard Gibbs must be ranked as one of the retarders of quaternion progress, in virtue of his pamphlet on vector analysis; a sort of hermaphrodite monster, compounded by the notations of Hamilton and Grassman.

This infamous statement has been quoted by many authors in the past. Gibbs' gentlemanly but firm response to Tait's attack was [14]:

> The merit or demerits of a pamphlet printed for private distribution a good many years ago do not constitute a subject of any great importance, but the assumption implied in the sentence quoted are suggestive of certain reflections and inquiries which are of broad interest; and seem not untimely at a period when the methods and results of the various forms of multiple algebra are attracting so much attention. It seems to be assumed that a departure from quaternionic usage in the treatment of vectors is an enormity. If this assumption is true, it is an important truth; if not, it would be unfortunate if it should remain unchallenged, especially when supported by so high an authority. The criticism relates particularly to notations, but I believe that there is a deeper question of notions underlying that of notations. Indeed, if my offense had been solely in the matter of notation, it would have been less accurate to describe my production as a monstrosity, than to characterize its dress as uncouth.

Gibbs then continued on to explain the advantage of his treatment of vector analysis in comparison to quaternion analysis. In the final part of that paper he stated:

> The particular form of signs we adopt is a matter of minor consequence. In order to keep within the resources of an ordinary printing office, I have used a dot and a cross, which are already associated with multiplication, which is best denoted by the simple juxtaposition of factors. I have no special predilection for these particular signs. The use of the dot is indeed liable to the objection that it interferes with its use as a separatrix, or instead of a parenthesis.

Although Gibbs considered his choice of the signs or notations a matter of minor importance, actually it had a tremendous consequence as will be shown in this study. Before we discuss it, a comment from Heaviside, generally considered by the scientific community as a co-founder with Gibbs of the modern vector analysis, should be quoted. During the peak of the controversy between Tait and Gibbs, Heaviside made the following remark [8, p. 35]:

> Prof. W. Gibbs is well able to take care of himself. I may, however, remark that the modifications referred to are evidence of modifications felt to be needed, and that Prof. Gibbs' pamphlet (Not published, New Haven, 1881-4, p.83), is not a quaternionic treatise, but an able and in some respects original little treatise on vector analysis, though too condensed and also too advanced for learners' use, and that Prof. Gibbs, being no doubt a little touched by Prof. Tait's condemnation, has recently (in the pages of Nature) made a powerful defense of his position. He has by a long way the best of the argument, unless Prof. Tait's rejoinder has still to appear. Prof. Gibbs clearly separates the quaternionic question from the question of a suitable notation, and argues strongly against the quaternionic establishment of vector analysis. I am able (and am happy) to express a general concurrence of opinion with him about the quaternion and its comparative uselessness in practical vector analysis. As regards his notation, however, I do not like it. Mine is Tait's, but simplified, and made to harmonize with Cartesians.

There are two implications in Heaviside's remark which are of interest to us. When he considered Gibbs' pamphlet to be too condensed it implies that some of the treatments may not have been obvious to him (or may not even have been comprehended by him). Secondly, he stated dislike for Gibbs' notations but without giving his reason(s). The fact that Heaviside used some of Tait's quaternionic notations seems to indicate that he did not approve of Gibbs' notations at all. We now believe that many workers, including Heaviside, did not appreciate the most eloquent and complete theory of vector analysis formulated by Gibbs. For this reason, we would like to offer a digest of Gibbs' work so we can have a clear understanding of his formulation.

3.2 Divergence and Curl Operators and Their New Notations

The basic definitions of the gradient, divergence, and the curl formulated by Gibbs are given by (3.1), (3.2), and (3.3). For convenience, we will make some changes in symbols to allow the convenience of using the summation sign. These changes are:

$$x,\ y,\ z \text{ to } x_1,\ x_2,\ x_3$$
$$i,\ j,\ k, \text{ to } \hat{x}_1,\ \hat{x}_2,\ \hat{x}_3.$$

The old total derivative symbols will be replaced by partial derivatives and the Greek letters for vectors by boldface letters. Thus, Eqs. (3.1) - (3.3) become:

$$\nabla u = \sum_i \hat{x}_i \frac{\partial u}{\partial x_i} \tag{3.10}$$

$$\nabla \cdot \boldsymbol{F} = \sum_i \hat{x}_i \cdot \frac{\partial \boldsymbol{F}}{\partial x_i} \tag{3.11}$$

$$\nabla \times \boldsymbol{F} = \sum_i \hat{x}_i \times \frac{\partial \boldsymbol{F}}{\partial \hat{x}_i} \tag{3.12}$$

It is understood that the summation goes from $i = 1$ to 3

The most important information passed to us by Gibbs concerns the nomenclature for the notations in these expressions. In the title preceding Article 68 quoted previously, he designated ∇, $\nabla\cdot$ and $\nabla\times$ as operators. If we examine the expressions given by (3.10), (3.11), and (3.12) it is quite obvious that the gradient operator or the del operator is unmistakably given by

$$\nabla = \sum_i \hat{x}_i \frac{\partial}{\partial x_i} \tag{3.13}$$

For the divergence, Gibbs used two symbols, a del followed by a dot, to denote his divergence operator. For the curl, he used a del followed by a cross to denote the curl operator. If we examine the expressions for the divergence and the curl defined by (3.11) and (3.12) it is clear that his two notations mean:

$$(\nabla \cdot)_G \rightarrow \sum_i \hat{x}_i \cdot \frac{\partial}{\partial x_i} \tag{3.14}$$

$$(\nabla \times)_G \rightarrow \sum_i \hat{x}_i \times \frac{\partial}{\partial x_i} \tag{3.15}$$

We emphasize this point by labeling his two notations with a subscript 'G', and we use an arrow instead of an equal sign to denote 'a notation for.'

According to our classification of the operators in Chapter 2, Gibbs' $(\nabla \cdot)_G$ and $(\nabla \times)_G$ are not compound operators; they are assemblies used by Gibbs as the notations for the divergence and curl. On the other hand, the terms at the right side of (3.14) and (3.15) are indeed compound operators according to our classification. Since these operators are distinct from the gradient operator we will introduce two notations for them. They are:

$$\nabla\!\!\!\!\cdot\;\; = \sum_i \hat{x}_i \cdot \frac{\partial}{\partial x_i} \tag{3.16}$$

$$\nabla\!\!\!\!\times = \sum_i \hat{x}_i \times \frac{\partial}{\partial x_i} \tag{3.17}$$

They are called, respectively, the divergence operator and the curl operator. Although these operators are so far defined in the Cartesian coordinate system we will demonstrate later that they are invariant to the choice of coordinate system. One important feature of $\nabla\!\!\!\!\cdot\;\;$ and $\nabla\!\!\!\!\times$ is that both these operators are independent of the gradient operator ∇. In other words, <u>∇ is not a constituent of the divergence operator nor of the curl operator</u>. These two symbols are suggested by the appearance of the dot or the cross in between the unit vectors $\hat{x}_i$ and the partial derivatives $\frac{\partial}{\partial x_i}$ of the ∇ operator as defined by (3.13). In Gibbs' notations, $(\nabla \cdot)_G$ and $(\nabla \times)_G$, ∇ is a part of his notations for the divergence and the curl that leads to a very serious misinterpretation by many later users and which is a key issue in our study. With the introduction of these two new notations, Eqs. (3.1) to (3.9) become:

$$\nabla u = \sum_i \hat{x}_i \frac{\partial u}{\partial x_i} \tag{3.18}$$

$$\nabla_{\cdot} F = \sum_i \hat{x}_i \cdot \frac{\partial F}{\partial x_i} \qquad (3.19)$$

$$\nabla_{\cdot} F = \sum_i \frac{\partial F_i}{\partial x_i} \qquad (3.20)$$

$$\boldsymbol{\nabla}_{\times} F = \sum_i \hat{x}_i \times \frac{\partial F}{\partial x_i} \qquad (3.21)$$

$$\boldsymbol{\nabla}_{\times} F = \sum_i \hat{x}_i \left(\frac{\partial F_k}{\partial x_j} - \frac{\partial F_j}{\partial x_k} \right) \qquad (3.22)$$

$(i, j, k) = (1,2,3)$ in cyclic order

$$\nabla_{\cdot} \nabla u = \sum_i \frac{\partial^2 u}{\partial x_i^2} \qquad (3.23)$$

$$\nabla_{\cdot} \nabla F = \sum_i \frac{\partial^2 F}{\partial x_i^2} \qquad (3.24)$$

$$\nabla_{\cdot} \nabla F = \sum_i \hat{x}_i \nabla_{\cdot} \nabla F_i \qquad (3.25)$$

$$\nabla_{\cdot} \nabla F = \nabla \nabla_{\cdot} F - \boldsymbol{\nabla}_{\times} \boldsymbol{\nabla}_{\times} F \qquad (3.26)$$

In these formulas the del operator only enters in the gradient of a scalar, (3.18), or of a vector, (3.24) - (3.26). Except for the notations for the divergence and the curl, we have not changed the content of Gibbs' work at all. These equations will be used later in our study of other people's presentations.

4. Book by Edwin Bidwell Wilson Founded upon the Lectures of J. Willard Gibbs

4.1 Gibbs' Lecture Notes

In 1901 the first book on vector analysis by an American author was published. The book was written by Wilson [7], then an instructor at Yale University, and founded upon the lectures of Gibbs. According to the general preface of that book, the greater part of the material has been taken from the course of lectures on Vector Analysis delivered annually at Yale University by Professor Gibbs. There is one historical document well-kept at the Sterling Memorial Library of Yale University which is the record of the lectures [15]. It is a cloth-bound book of notes, hand-written in ink on 8-1/2" by 11" ruled paper, consisting of fifteen chapters covering 289 - plus pages. The title page and the table of contents are:

Lectures Delivered upon

Vector Analysis

and its

Applications to Geometry and Physics

by

Professor J. Willard Gibbs 1899-90

reported by Mr. E. B. Wilson

Table of contents page

4.2 Wilson's Book

Presumably, Wilson's book (436 pages) is mainly based on these notes. It was mentioned in the preface of Wilson's book that some use, however, has been made of the chapters on vector analysis in Heaviside's *Electromagnetic Theory* (1893) and in Föppl's lectures on Maxwell's *Theory of Electricity* (1894). Apparently, Gibbs himself was not involved in the preparation of this book. We quote here two paragraphs in the preface by Professor Gibbs:

> I was very glad to have one of the hearers of my course on Vector Analysis in the year 1899-1900 undertake the preparation of a text-book on the subject.
>
> I have not desired that Dr. Wilson should aim simply at the reproduction of my lectures, but rather that he should use his own judgment in all respects for the production of a text-book in which the subject should be so illustrated by an

> adequate number of examples as to meet the wants of students of geometry and physics.

In the general preface, Wilson stated:

> When I undertook to adapt the lectures of Professor Gibbs on Vector Analysis for publication in the Yale Bicentennial Series, Professor Gibbs himself was already so fully engaged in his work to appear in the same series, *Elementary Principles in Statistical Mechanics*, that it was understood no material assistance in the composition of this book could be expected from him. For this reason he wished me to feel entirely free to use my own discretion alike in the selection of the topics to be treated and in the mode of treatment. It has been my endeavor to use the freedom thus granted only in so far as was necessary for presenting his method in text-book form.

One very important remark by Wilson is found in the preface:

> It has been the aim here to give also an exposition of scalar and vector products of the operator ∇, of divergence and curl which have gained such universal recognition since the appearance of Maxwell's *Treatise on Electricity and Magnetism*, slope, potential, linear vector functions, etc. such as shall be adequate for the needs of students of physics at the present day and adapted to them.

We would like to point out here that in Gibbs' pamphlets and in the lecture notes reported by Wilson, there is no mention of the scalar and vector products of the operator ∇. We believe this concept or interpretation was created by Wilson and unfortunately, it has had a tremendously detrimental effect upon the learning of vector analysis within the framework of Gibbs' original contributions.

In explaining the meaning of the divergence of a vector function Wilson misinterpreted Gibbs' notation for this function, namely $\nabla \cdot \boldsymbol{F}$. After defining the ∇ operator for the gradient in a Cartesian system as

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \tag{4.1}$$

he stated in Sec. 70, p. 150 of Wilson's book [7]:

> Although the operation ∇V has not been defined and cannot be at present, two formal combinations of the vector operator ∇ and a vector function V may be treated. These are the (formal) scalar product and the (formal) vector product of ∇ into V. They are:
>
> $$\nabla \cdot V = \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \right) \cdot V \tag{4.2}$$
>
> $$\nabla \times V = \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \right) \times V \tag{4.3}$$
>
> The differentiations $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$, being scalar operators, pass by the dot and the cross, that is
>
> $$\nabla \cdot V = \left(i \cdot \frac{\partial V}{\partial x} + j \cdot \frac{\partial V}{\partial y} + k \cdot \frac{\partial V}{\partial z} \right) \tag{4.4}$$
>
> $$\nabla \times V = \left(i \times \frac{\partial V}{\partial x} + j \times \frac{\partial V}{\partial y} + k \times \frac{\partial V}{\partial z} \right) \tag{4.5}$$
>
> They may be expressed in terms of the components V_1, V_2, V_3 of V

We have identified the equations with our own numbers. In order to compare these expressions with Gibbs' expressions now described by (3.19) to (3.22), we again, will change the notations for V; x, y, z; i, j, k; to $\boldsymbol{F}$; x_1, x_2, x_3; $\hat{x}_1, \hat{x}_2, \hat{x}_3$; and $\nabla \cdot V$ and $\nabla \times V$ to $\nabla \boldsymbol{F}$ and $\nabla \boldsymbol{F}$. Eq. (4.1) to (4.5) become:

$$\nabla = \sum_i \hat{x}_i \frac{\partial}{\partial x_i} \tag{4.6}$$

$$\nabla F = \left(\sum_i \hat{x}_i \frac{\partial}{\partial x_i} \right) \cdot F \tag{4.7}$$

$$\nabla F = \left(\sum_i \hat{x}_i \frac{\partial}{\partial x_i} \right) \times F \tag{4.8}$$

$$\nabla F = \sum_i \hat{x}_i \cdot \frac{\partial F}{\partial x_i} \tag{4.9}$$

$$\nabla F = \sum_i \hat{x}_i \times \frac{\partial F}{\partial x_i} \tag{4.10}$$

Equations (4.6), (4.9), and (4.10) are identical to Gibbs' (3.13), (3.19), and (3.21). However, (4.7) and (4.8) are not found in Gibbs' works. Wilson obtained or derived (4.9) and (4.10) from (4.7) and (4.8). The derivation involves two crucial steps or assumptions. First, he considers Gibbs' notations $\nabla \cdot \boldsymbol{F}$ and $\nabla \times \boldsymbol{F}$ as 'formal' scalar and vector products between ∇ and $\boldsymbol{F}$. In the following we will refer to this model as the FSP (formal scalar product) and FVP (formal vector product). He did not explain the meaning of the word 'formal'. Secondly, after he formed the FSP and FVP he let the differentiation $\frac{\partial}{\partial x_i}$ pass by the dot and the cross with the argument that the differentiations $\frac{\partial}{\partial x_i}$, $(i = 1,2,3)$ are scalar operators. The statement appears to be quite firm. But standard books on mathematical analysis do not have such a theorem. Later on, [7, p. 152] Wilson attempts to soften his attitude by saying:

> From some standpoints objections may be brought forward against treating ∇ as a symbolic vector and introducing $\nabla \cdot V$ and $\nabla \times V$ as the symbolic scalar and vector products of ∇ into V respectively. These objections may be avoided by simply laying down the definition that the symbol $\nabla\cdot$ and $\nabla\times$, which may be looked upon as entirely new operators quite distinct from ∇, shall be
>
> $$\nabla \cdot V = i \cdot \frac{\partial V}{\partial x} + j \cdot \frac{\partial V}{\partial y} + k \cdot \frac{\partial V}{\partial z} \tag{4.11}$$
>
> and

$$\nabla \times V = i \times \frac{\partial V}{\partial x} + j \times \frac{\partial V}{\partial y} + k \times \frac{\partial V}{\partial z} \tag{4.12}$$

But for practical purposes and for remembering formulas, it seems by all means advisable to regard

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}$$

as a symbolic vector differentiator. This symbol obeys the same laws as a vector just in so far as the differentiations $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$, obey the same laws as ordinary scalar quantities.

The contradictions between the above statement and the FSP and FVP assertion seems quite evident. Equations (4.11) and (4.12), of course, are the same as Gibbs' (3.11) and (3.12) with V replaced by $\boldsymbol{F}$ and $x, y, z;$ and $i, j, k;$ by x_1, x_2, x_3 and $\hat{x}_1, \hat{x}_2, \hat{x}_3$. The difference is that Gibbs never spoke of a FSP and FVP but Wilson introduced these concepts to derive the expressions for div $\boldsymbol{F}$ and curl $\boldsymbol{F}$ by imposing some non-valid manipulations. What is the consequence? Many later authors followed his practice and encountered difficulties when the same treatment was applied to orthogonal curvilinear coordinate systems. Before we discuss this topic, Heaviside's treatment of vector analysis, particularly his handling of ∇ should be reviewed and commented upon.

We have pointed out that Gibbs' pamphlets were communicated to Heaviside. On the other hand, Wilson also mentioned some use of Heaviside's treatment of vector analysis in his book *Electromagnetic Theory* (1893) in his preface. The exchange between Heaviside and Wilson was therefore, mutual. However, Heaviside goes his own way in presenting the same topics. Before we turn to the next chapter, Wilson's FSP and FVP model will be analytically examined.

If we start with one of Gibbs' definitions of divergence, without using his notation but rather by using the linguistic notation, i.e.,

$$\text{div}\, \boldsymbol{F} = \sum_i \frac{\partial F_i}{\partial x_i} \tag{4.13}$$

then by substituting $F_i = \hat{\boldsymbol{x}}_i \cdot \boldsymbol{F}$ into (4.13) we find

$$\begin{aligned}\text{div } \boldsymbol{F} &= \sum_i \frac{\partial(\hat{\boldsymbol{x}}_i \cdot \boldsymbol{F})}{\partial x_i} \\ &= \sum_i \left[\hat{\boldsymbol{x}}_i \cdot \frac{\partial \boldsymbol{F}}{\partial x_i} + \frac{\partial \hat{\boldsymbol{x}}_i}{\partial x_i} \cdot \boldsymbol{F} \right]\end{aligned} \tag{4.14}$$

Since $\frac{\partial \hat{\boldsymbol{x}}_i}{\partial x_i} = 0$ (4.14) reduces to

$$\text{div } \boldsymbol{F} = \sum_i \hat{\boldsymbol{x}}_i \cdot \frac{\partial \boldsymbol{F}}{\partial x_i} \tag{4.15}$$

which is obviously not equal to

$$\left(\sum \hat{\boldsymbol{x}}_i \frac{\partial}{\partial x_i} \right) \cdot \boldsymbol{F}$$

or $\nabla \cdot \boldsymbol{F}$. This is a proof of the lack of validity of the FSP. A similar proof can be executed with respect to the FVP. Another demonstration of the fallacy of a FSP is to consider a 'twisted' differential operator of the form

$$\nabla_t = \hat{\boldsymbol{x}}_2 \frac{\partial}{\partial x_1} + \hat{\boldsymbol{x}}_3 \frac{\partial}{\partial x_2} + \hat{\boldsymbol{x}}_1 \frac{\partial}{\partial x_3} \tag{4.16}$$

and a 'twisted' vector function defined by

$$\boldsymbol{F}_t = \hat{\boldsymbol{x}}_2 F_1 + \hat{\boldsymbol{x}}_3 F_2 + \hat{\boldsymbol{x}}_1 F_3 \tag{4.17}$$

If the FSP were a valid product then by following Wilson's pass-by procedure we obtain

$$\nabla_t \cdot \boldsymbol{F}_t = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} \tag{4.18}$$

In other words, div $\boldsymbol{F}$ is now treated as the formal scalar product between ∇_t and $\boldsymbol{F}_t$. The result is the same as Wilson's FSP between ∇ and $\boldsymbol{F}$. Such a manipulation is, of course, not a valid mathematical procedure. We have now refuted Wilson's treatment of div $\boldsymbol{F}$ and curl $\boldsymbol{F}$ based on the FSP and FVP. The legitimate compound differential operators for the divergence and the curl are, respectively, $\nabla\!\!\!\!\cdot\;$ and $\nabla\!\!\!\!\times$ defined by (3.16) and (3.17). $(\nabla \cdot)_G$ and $(\nabla \times)_G$ are merely Gibbs' notations suggested for the divergence and the curl. They are not operators.

4.3 The Spread of the Formal Scalar Product (FSP) and Formal Vector Product (FVP)

Being the first book on vector analysis published in 1901 in the U. S. A., Wilson's book became very popular. The 8th reprinting was made in 1943 and a paperback reprint by Dover Publications appeared in 1960. Many later authors freely adopted Wilson's presentation using the FSP and FVP to derive the expressions for divergence and curl in the Cartesian coordinate system. We have found over fifty books [16] containing such a treatment. We now quote herein a few examples to show Wilson's influence.

1.) In the book *Advanced Vector Analysis* by Weatherburn [4] published in 1924, we find the following statement:

> To justify the notation ∇·, we have only to expand the formal products according to the distributive law, then

$$\nabla \cdot \bar{f} = \left[\sum_i \left(\bar{a}_i \frac{\partial}{\partial x_i}\right)\right] \cdot \bar{f} = \sum_i \frac{\partial f_i}{\partial x_i} = \text{div } \bar{f}$$

We shall remark here that any distributive law in mathematics should be proved. In this case, there is no distributive law to speak of because the author is dealing with an assembly of mathematical symbols and ∇· is not a compound operator. Incidentally, Weatherburn's book appears to be the first book published in England wherein Gibbs' notations, but not Heaviside's, have been used in addition to the linguistic notations, namely, grad u, div $\bar{f}$, and curl $\bar{f}$.

2) A German book by Lagally [17] published in 1928 contains the following statement on p. 123:

> The rotation (curl) of $\bar{f}$ is denoted by the vector product between ∇ with field function $\bar{f}$ ... and

$$\text{div grad} f = \nabla \cdot \nabla f = \left(\sum_i \hat{x}_i \frac{\partial}{\partial x_i}\right) \cdot \left(\sum_j \hat{x}_j \frac{\partial f}{\partial x_j}\right) = \nabla^2 f$$

The above is an English translation of the original text in German. It is seen that a term like $\hat{x}_1 \frac{\partial}{\partial x_1} \cdot \hat{x}_1 \frac{\partial}{\partial x_1}$ is an assembly of symbols. It is not a compound operator.

3.) In a book by Mason and Weaver [18, p. 336] we find the following statement:

> The differential operator ∇ can be considered formally as a vector of components $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$, so that its scalar and vector products with another vector may be taken.

In comparison with Wilson's treatment Mason and Weaver have used the word 'formally' to be associated with ∇ and then speak of scalar and vector products with vector functions.

4.) In a book *Applied Mathematics* by Schelkunoff [19, p. 126], the author first derived the differential expression for the divergence based on the flux model; then he added:

> In Sec. 6 the vector operator del was introduced. If we treat it as a vector and multiply it by a vector $\overline{F}$, we find
>
> $$\nabla \cdot \overline{F} = \left(\sum_i \hat{x}_i \frac{\partial}{\partial x_i} \right) \cdot \left(\sum_j \hat{x}_j F_j \right) = \sum_i \frac{\partial F_i}{\partial x_i} = \text{div } \overline{F}$$
>
> For this reason $\nabla\cdot$ may be used as an alternative for div; however, the notation is tied too specifically to Cartesian coordinates.

There are two messages in this statement: the first one is his acceptance of the FSP as a valid entity. The second one is his implication that FSP only applies to the Cartesian system. Actually, the divergence operator, ∇, is invariant with respect to the choice of the coordinate system, a property to be demonstrated later, but '$\nabla\cdot$' is an assembly, not an operator. Only by means of an illegitimate manipulation does it yield the differential expression for the divergence in the Cartesian coordinate system.

5.) From a well-known book by Feynman, Leighton and Sands [20, p. 2-7] we find the following statement:

Let us try the dot product between ∇ with a vector field that we know, say $\overline{f}$; we write

$$\nabla \cdot \overline{f} = \nabla_x f_x + \nabla_y f_y + \nabla_z f_z$$

or

$$\nabla \cdot \overline{f} = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$$

The authors remarked on the same page before the above statement:

> With operators we must always keep the sequence right, so that the operations make the proper sense ...

This remark is very important. Our discussion and use of the operators in chapter 2, particularly that related to the compound operators, closely adheres to this principle. In the case of Gibbs' notation, $\nabla \cdot \overline{f}$, we are faced with a dot symbol after ∇, so that the differentiation cannot be applied to $\overline{f}$; it is blocked by a dot in the assembly. Thus, the authors seem to have violated their own rule by trying to form a dot product or FSP.

6.) In the English translation of a Russian book by Borisenko and Tarapov [21, p. 157] we find the following statement:

> The expression (4.29) $\left(\nabla = \sum i_k \frac{\partial}{\partial x_k}\right)$ for the operator ∇ implies the following representation for the divergence of $\overline{A}$:
>
> $$\text{div } \overline{A} = \frac{\partial A_k}{\partial x_k} = i_k \frac{\partial}{\partial x_k} \cdot \overline{A} = \nabla \cdot \overline{A}$$
>
> A coordinate-free symbolic representation of the operator ∇ is
>
> $$\nabla(\cdots) = \lim_{V \to 0} \frac{1}{V} \oiint_S \overline{n}(\cdots) dS \qquad (4.19)$$
>
> where $(\cdots)$ is some expression (possibly preceded by a dot or a cross) on which the given operator acts. In fact, according to (4.31) and (4.29),

$$\text{grad } \varphi = \lim_{V \to 0} \frac{1}{V} \oint\!\!\oint_S \varphi \bar{n}\, dS, \tag{4.20}$$

$$\text{div } \bar{A} = \lim_{V \to 0} \frac{1}{V} \oint\!\!\oint_S \bar{A} \cdot \bar{n} dS . \tag{4.21}$$

From the above statement we see that the two authors believe the validity of the FSP. Their (4.19) also implies that they consider ∇ as a constituent of the divergence and the curl in addition to comprising the gradient operator. The formula described by (4.19) appeared earlier in the book by Gans [22, p. 49, Sixth Edition] who used both Gibbs' notations and the linguistic notations in this edition.

There are several authors presenting ∇ as defined by $\sum (\partial/\partial x_i)\hat{x}_i$ instead of $\sum \hat{x}_i(\partial/\partial x_i)$, and the Laplacian, defined by div grad, is often treated as the scalar product between two nablas, presumably because Gibbs used $\nabla \cdot \nabla$ as the notation for this compound operator. These practices, including the use of a FSP and FVP, go beyond the boundary of the U. S. A. and continental Europe. There are books in Chinese and Japanese doing the same.

5. ∇ in the Hands of Oliver Heaviside (1850 - 1925)

Although we have traced the concept of the FSP and FVP as due to Wilson, the same practice is found in the works of Heaviside. In Vol. I of his book *Electromagnetic Theory* [8, § 127] published in 1893, Heavside stated:

> When the operand of ∇ is a vector, say $\overline{D}$, we have both the scalar product and the vector product to consider. Taking the formula along first, we have
>
> $$\operatorname{div}\overline{D} = \nabla_1 D_1 + \nabla_2 D_2 + \nabla_3 D_3.$$
>
> This function of $\overline{D}$ is called the divergence and is a very important function in physical mathematics.

He then considered the curl of a vector function as the vector product between ∇ and that vector. At the time of his writing he was already aware of Gibbs' pamphlets on vector analysis but Wilson's book was not yet published. It seems, therefore, that Heaviside and Wilson independently introduced the misleading concept for the scalar and vector products between ∇ and a vector function. Both were, perhaps, induced by Gibbs' notations for the divergence and the curl. Heaviside did not even include the word 'formal' in his description of the products. We should mention that Heaviside's notations for these two products and the gradient are not the same as Gibbs' (See the table of notations in Sec. 2.1). His notation for the divergence of $\overline{f}$ is $\nabla\overline{f}$ and his notation for the curl of $\overline{f}$ is $V\nabla\overline{f}$ (a quaternion notation) while his notation for the gradient of a scalar function f is $\nabla.f$. Having treated $\nabla\cdot\overline{f}$ and $\nabla\times\overline{f}$ (Gibbs' notations for the divergence and the curl) as two 'products', Heaviside simply considered ∇ as a vector in deriving various differential identities. One of them was presented as follows [8, § 132]:

> The examples relate principally to the modification introduced by the differentiating functions of ∇.
>
> (a) We have the parallelopiped property

$$\overline{N}\,V\nabla\overline{E} = \nabla V\overline{E}\,\overline{N} = \overline{E}\,V\overline{N}\nabla \qquad (176)$$

where ∇ is a common vector. The equations remain true when ∇ is vex, provided we consistently employ the differentiating power in the three forms. Thus, the first form, expressing $\overline{N}$ component of curl $\overline{E}$, is not open to misconception. But in the second form, expressing the divergence of $V\overline{E}\,\overline{N}$, since $\overline{N}$ follows ∇, we must understand that $\overline{N}$ is supposed to remain constant. In the third form, again, the operand $\overline{E}$ precedes the differentiator; we must either, then, assume that ∇ acts backwards, or else, which is preferable, change the third form to $V\overline{N}\nabla.\overline{E}$, the scalar product of $V\overline{N}\nabla$ and $\overline{E}$, or $\left(V\overline{N}\nabla\right)\overline{E}$ if that is plainer.

(b) Suppose, however, that both vectors in the vector product are variable. Thus, required the divergence of $V\overline{E}\overline{H}$, expanded vectorially. We have

$$\nabla V\overline{E}\,\overline{H} = \overline{E}\,V\overline{H}\nabla = \overline{H}\,V\nabla\overline{E}, \qquad (177)$$

where the first form alone is entirely unambiguous. But we may use either of the others, provided that the differentiating power of ∇ is made to act on both $\overline{E}$ and $\overline{H}$. But if we keep to the plainer and more usual convention that the operand is to follow the operator, then the third term, in which $\overline{E}$ alone is differentiated, gives one part of the result, whilst the second form, or rather its equivalent, $-\overline{E}V\nabla\overline{H}$, wherein $\overline{H}$ alone is differentiated, gives the rest. So we have, complete, and without ambiguity

$$\text{div}\; V\overline{E}\,\overline{H} = \overline{H}\,\text{Curl}\;\overline{E} - \overline{E}\;\text{Curl}\;\overline{H} \qquad (178)$$

a very important transformation.

First of all, in terms of Gibbs' notations, Heaviside's Eqs. (176), (177), and (178) would be written in the form

$$\overline{N}\cdot\nabla\times\overline{E} = \nabla\cdot\left(\overline{E}\times\overline{N}\right) = \overline{E}\cdot\left(\overline{N}\times\nabla\right) \qquad (5.1)$$

$$\nabla\cdot\left(\overline{E}\times\overline{H}\right)=\overline{E}\cdot\left(\overline{H}\times\nabla\right)=\overline{H}\cdot\left(\nabla\times\overline{E}\right) \tag{5.2}$$

$$\nabla\cdot\left(\overline{E}\times\overline{H}\right)=\overline{H}\cdot\nabla\times\overline{E}-\overline{E}\cdot\nabla\times\overline{H} \tag{5.3}$$

And in terms of the new notations for the divergence operator and the curl operator, Heaviside's three equations will be written in the form

$$\overline{N}\cdot\boldsymbol{\nabla}\overline{E}=\boldsymbol{\nabla}\left(\overline{E}\times\overline{N}\right)=\overline{E}\cdot\left(\overline{N}\times\nabla\right) \tag{5.4}$$

$$\boldsymbol{\nabla}\left(\overline{E}\times\overline{H}\right)=\overline{E}\cdot\left(\overline{H}\times\nabla\right)=\overline{H}\cdot(\boldsymbol{\nabla}\overline{E}) \tag{5.5}$$

$$\boldsymbol{\nabla}\left(\overline{E}\times\overline{H}\right)=\overline{H}\cdot\boldsymbol{\nabla}\overline{E}-\overline{E}\cdot\boldsymbol{\nabla}\overline{H} \tag{5.6}$$

According to the established mathematical rules, Heaviside's logic in arriving at his (178) or our (5.1) or (5.4) is entirely unnacceptable; in particular, present day students would never write an equation (177) or (5.2) or (5.5) with ∇ being the ∇ operator. The second term in (5.2) or (5.5) is a weighted operator while the first and the third are functions and they are not equal to each other. His Eq. (178) or (5.3) in Gibbs' notation or (5.6) in our notation is a valid vector identity but his derivation of this identity is not based on established mathematical rules. It is obtained by a manipulation of mathematical symbols and selecting the desired terms. The most important message passed to us is his practice of considering $\nabla\cdot\overline{f}$ and $\nabla\times\overline{f}$ as two legitimate products, the same as Wilson's FSP and FVP. Heaviside's 'equations' will be examined again in a later section and will be cast in proper form in terms of the symbolic vector and/or a partial symbolic vector.

6. Shilov's Formulation of Vector Analysis

A book in Russian on vector analysis was written by Shilov [23] in 1954, who advocated a new formulation with the intent of providing a rather broader treatment of vector analysis. Shilov's work was adopted by Fang [24] who studied in the U. S. S. R. We were informed of Shilov's work through Fang. After a careful examination of the English translation of the two key chapters in that book we found the contradictions as described below:

Shilov defined an 'expression' for ∇ denoted by $T(\nabla)$ as:

$$T(\nabla) = \frac{\partial}{\partial x}T(i) + \frac{\partial}{\partial y}T(j) + \frac{\partial}{\partial z}T(k) \tag{6.1}$$

where i, j, k denote the Cartesian unit vectors and ∇ (nabla) is identified as the Hamilton differential operator, that is,

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}.$$

Equation (6.1) is the same as Shilov's Eq. (18) on p. 18 of [24]. We want to emphatically call attention to the fact that the only meaningful expression for $T(\nabla)$ involving ∇ is ∇f, the gradient of f. In that case, (6.1) is an identity because the right side of (6.1) yields

$$\frac{\partial}{\partial x}(if) + \frac{\partial}{\partial y}(jf) + \frac{\partial}{\partial z}(kf) = i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}$$

which is ∇f.

The most serious contradiction in Shilov's work is his derivation of the expression for the divergence and the curl by letting $T(\nabla)$ equal to $\nabla \cdot \overline{f}$ and $\nabla \times \overline{f}$ respectively. We have pointed out before that these two products do not exist. Shilov is defining a meaningless assembly to make it meaningful. It is like defining $2 + \times 3$ to be equal to $2 \times +3$ $(= +6)$.

7. Orthogonal Curvilinear Systems

After having revealed a number of 'historical' confusions and contradictions in vector analysis so far presented in the Cartesian system, we now examine several presentations in curvilinear coordinate systems. We will show even more clearly the sources of the various misrepresentations.

In this section we limit ourselves to problems in orthogonal curvilinear coordinate systems, leaving the discussion of non-orthogonal curvilinear systems to a later section. In an orthogonal curvilinear system with coordinate variables v_i, $i = 1,2,3$, the total differential of the position vector $\boldsymbol{r}$ of a point in 3-dimensional space can be written in the form

$$d\boldsymbol{r} = \frac{\partial \boldsymbol{r}}{\partial v_1} dv_1 + \frac{\partial \boldsymbol{r}}{\partial v_2} dv_2 + \frac{\partial \boldsymbol{r}}{\partial v_3} dv_3 \tag{7.1}$$

If we denote the metric coefficients in an orthogonal curvilinear system by h_i with $i = (1,2,3)$, then

$$\frac{\partial \boldsymbol{r}}{\partial v_i} = h_i \hat{\boldsymbol{u}}_i,\ i = (1,2,3) \tag{7.2}$$

where $\hat{\boldsymbol{u}}_i$ denote the unit vectors in that system. An orthogonal system is characterized by

$$\hat{\boldsymbol{u}}_i \cdot \hat{\boldsymbol{u}}_j = 0,\ i \neq j \tag{7.3}$$

and

$$\hat{\boldsymbol{u}}_i \times \hat{\boldsymbol{u}}_j = \hat{\boldsymbol{u}}_k,\ \ i \neq j \neq k \tag{7.4}$$

with (i, j, k) in cyclic order of $(1,2,3)$. Thus $\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2$, and $\hat{\boldsymbol{u}}_3$ form a right-hand system.

For our discussion, a review of the invariance of the three differential operators is in order.

7.1 Invariance of the Differential Operators ∇, $\nabla\!\!\!\cdot$ and $\nabla\!\!\!\times$ in Orthogonal Curvilinear Systems

When (7.2) is substituted into (7.1) we obtain

$$d\boldsymbol{r} = \sum_i h_i \hat{\boldsymbol{u}}_i dv_i \tag{7.5}$$

It is understood that the summation in (7.5) goes from $i = 1$ to 3. The total differential of a scalar function is defined by

$$df = \sum_i \frac{\partial f}{\partial v_i} dv_i \tag{7.6}$$

which can be written in the form

$$\begin{aligned} df &= \sum \frac{1}{h_i}\frac{\partial f}{\partial v_i} h_i dv_i \\ &= \sum \frac{\hat{\boldsymbol{u}}_i}{h_i}\frac{\partial f}{\partial v_i} \cdot \hat{\boldsymbol{u}}_i h_i dv_i \end{aligned} \tag{7.7}$$

The last line of (7.7) represents the scalar product of the gradient of f and $\boldsymbol{dr}$ given by (7.5). Thus

$$df = \nabla f \cdot \boldsymbol{dr} \tag{7.8}$$

where

$$\nabla f = \sum \frac{\hat{\boldsymbol{u}}_i}{h_i}\frac{\partial f}{\partial v_i} \tag{7.9}$$

If we write $\boldsymbol{dr} = \hat{\boldsymbol{r}}dr$ in (7.8), then

$$\frac{\partial f}{\partial r} = \hat{\boldsymbol{r}} \cdot \nabla f \tag{7.10}$$

The name for the gradient used by Maxwell is space-variation, presumably because of the relationship described by (7.10). In view of (7.9), the gradient operator in an orthogonal curvilinear coordinate system is therefore given by

$$\nabla = \sum \frac{\hat{\boldsymbol{u}}_i}{h_i}\frac{\partial}{\partial v_i} \tag{7.11}$$

For the special case of the Cartesian coordinate system, $h_i = 1$, $\hat{\boldsymbol{u}}_i = \hat{\boldsymbol{x}}_i$, $v_i = x_i$, we recover the del operator of Hamilton. Although our derivation of (7.11) is independent of the choice of the coordinate system, it is desirable to show analytically that the gradient operator is indeed invariant to that choice. If we have another orthogonal curvilinear coordinate system with coordinate variables v_i', unit vectors $\hat{\boldsymbol{u}}_i'$ and metric coefficients h_i' and we denote the gradient operator in that system by ∇', we want to show that

$$\nabla = \nabla' \tag{7.12}$$

Since

$$d\boldsymbol{r} = \sum_i h_i \hat{\boldsymbol{u}}_i dv_i = \sum_j h'_j \boldsymbol{u}'_j dv'_j \tag{7.13}$$

hence

$$h'_j dv'_j = \sum h_i \left(\hat{\boldsymbol{u}}'_j \cdot \hat{\boldsymbol{u}}_i\right) dv_i \tag{7.14}$$

Thus,

$$h'_j \frac{\partial v'_j}{\partial v_i} = h_i \left(\hat{\boldsymbol{u}}'_j \cdot \hat{\boldsymbol{u}}_i\right)$$

or

$$\hat{\boldsymbol{u}}'_j \cdot \hat{\boldsymbol{u}}_i = \frac{h'_j}{h_i} \frac{\partial v'_j}{\partial v_i} \tag{7.15}$$

Now

$$\hat{\boldsymbol{u}}'_j = \sum_i \left(\hat{\boldsymbol{u}}'_j \cdot \hat{\boldsymbol{u}}_i\right) \hat{\boldsymbol{u}}_i \tag{7.16}$$

Substituting (7.15) into (7.16), we obtain

$$\hat{\boldsymbol{u}}'_j = \sum_i \frac{h'_j}{h_i} \frac{\partial v'_j}{\partial v_i} \hat{\boldsymbol{u}}_i$$

or

$$\frac{\hat{\boldsymbol{u}}'_j}{h'_j} = \sum_i \frac{1}{h_i} \frac{\partial v'_j}{\partial v_i} \hat{\boldsymbol{u}}_i \tag{7.17}$$

By definition,

$$\nabla' = \sum_j \frac{\hat{\boldsymbol{u}}'_j}{h'_j} \frac{\partial}{\partial v'_j} \tag{7.18}$$

and by the chain rule of differentiation, we have

$$\nabla' = \sum_j \frac{\hat{\boldsymbol{u}}'_j}{h'_j} \sum_k \frac{\partial v_k}{\partial v'_j} \frac{\partial}{\partial v_k} \tag{7.19}$$

Eliminating $\hat{\boldsymbol{u}}'_j / h'_j$ between (7.17) and (7.19) we obtain

$$\nabla' = \sum_i \sum_j \sum_k \frac{\hat{\boldsymbol{u}}_i}{h_i} \frac{\partial v'_j}{\partial v_i} \frac{\partial v_k}{\partial v'_j} \frac{\partial}{\partial v_k} \tag{7.20}$$

Since

$$\frac{\partial v_k}{\partial v_i} = \sum_j \frac{\partial v_k}{\partial v'_j} \frac{\partial v'_j}{\partial v_i} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$$

eq. (7.20) reduces to

$$\nabla' = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i} \frac{\partial}{\partial v_i} = \nabla \tag{7.21}$$

This completes our proof. By following the same procedure it is not difficult to show that the divergence operator and the curl operator are also invariant with respect to the choice of the orthogonal curvilinear coordinate system, i.e.,

$$\nabla \cdot = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i} \cdot \frac{\partial}{\partial v_i} = \sum_j \frac{\hat{\boldsymbol{u}}'_j}{h'_j} \cdot \frac{\partial}{\partial v'_j} = \nabla' \cdot \tag{7.22}$$

and

$$\nabla \times = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i} \times \frac{\partial}{\partial v_i} = \sum_j \frac{\hat{\boldsymbol{u}}'_j}{h'_j} \times \frac{\partial}{\partial v'_j} = \nabla' \times \tag{7.23}$$

With these expressions at our disposal, let us look at some of the treatments of the FSP and FVP in orthogonal curvilinear coordinate systems and some of the presentations of vector identities involving the del operator.

7.2 Two Examples from the Book by Moon and Spencer

In the book by Moon and Spencer [3, p. 325] they stated:

> Let me apply the definition, Eq. (1.4), (of ∇ in the orthogonal curvilinear system, our 7.20) to divergence. By the usual definition of a scalar product,

$$\nabla \cdot V = \frac{1}{(g_{11})^{\frac{1}{2}}} \frac{\partial (V)_1}{\partial x^1} + \frac{1}{(g_{22})^{\frac{1}{2}}} \frac{\partial (V)_2}{\partial x^2} + \frac{1}{(g_{33})^{\frac{1}{2}}} \frac{\partial (V)_3}{\partial x^3} \tag{7.24}$$

But this is <u>not</u> divergence, which is found to be ....

Similar inconsistencies are obtained with other applications of Eq. (14).

In (7.24), their $(g_{ii})^{\frac{1}{2}}$ correspond to our metric coefficients h_i and their x^i to our variables v_i.

In the first place, they have now applied the FSP to ∇ and V in an orthogonal curvilinear coordinate system without realizing that the FSP is a non-valid entity in any coordinate system including the Cartesian system. After obtaining a wrong formula for the divergence, (7.24), they did not offer an explanation of the reason for the failure. It is seen in our discussion of the invariance of ∇ that ∇f is equal to $\nabla' f'$ and ∇f or $\nabla' f'$ is not a scalar product between ∇ and f nor ∇' and f'.

In discussing the Laplacian of a vector function, the two authors stated [3, p. 235]:

> Section (7.08) showed that there are three meaningful combinations of differential operators: div grad, grad div, and curl curl. Of these, the first is the scalar Laplacian, ∇^2. It is convenient to combine the other two operators to form the <u>vector Laplacian</u>, ✡:
>
> $$✡ = \text{grad div} - \text{curl curl} \qquad (7.25)$$
>
> Evidently the vector Laplacian can operate only on a vector, so
>
> $$✡\boldsymbol{E} = \text{grad div}\,\boldsymbol{E} - \text{curl curl}\,\boldsymbol{E} \qquad (7.26)$$
>
> Since the quantities on the right are vectors, ✡$\boldsymbol{E}$ transforms as a univalent tensor or vector.
>
> As noted in Table 1.01 (their table of notations on page 10), the scalar and vector Laplacians are often represented by the same symbol. This is poor practice, however, since the two are basically quite different:
>
> $$\nabla^2 = \text{div grad} \qquad (7.27)$$
>
> $$✡ \equiv \text{grad div} - \text{curl curl} \qquad (7.28)$$

> This difference is evident also when the expression for the vector Laplacian is expanded. ...

Analytically it can be proved [25, pp. 124 - 126] that in any orthogonal curvilinear system,

$$\text{div grad}\, f = \text{grad div}\, f - \text{curl curl}\, f$$

or

$$\nabla\cdot\nabla f = \nabla\,\nabla\cdot f - \nabla\times\nabla\times f \tag{7.29}$$

where ∇f denotes the gradient of a vector function that is a dyadic function. The divergence of a dyadic function is a vector function. The use of ∇^2 to denote the Laplacian is an old practice, but the use of $\nabla\cdot\nabla$ is preferred because it shows the structure of the Laplacian when it is applied to either a scalar function or a vector function. By treating (7.25) as the definition for the Laplacian applied to a vector function, the two authors have probably been influenced by a remark made by Stratton [10, p. 50]:

> The vector $\nabla\cdot\nabla \boldsymbol{F}$ may now be obtained by subtraction of (85) [an expansion of $\nabla\times\nabla\times\boldsymbol{F}$ in an orthogonal curvilinear system] from the expansion of $\nabla\nabla\cdot\boldsymbol{F}$, and the result differs from that which follows a direct application of the Laplacian to the curvilinear components of $\boldsymbol{F}$.

As shown in our proof [25, pp. 124 - 126] $\nabla\boldsymbol{F}$ is a dyadic, where the gradient operator must apply to the entire vector function containing both the components and the unit vectors. When this is done, we find that (7.29) is indeed an identity. In view of our analysis, it is clear that a special symbol for the Laplacian is not necessary when it is operating on a vector function. The same remark holds true for the two different notations for the Laplacian introduced by Burati-Forti and Marcolongo as shown in Table I.

These two examples also show why Moon and Spencer thought that ∇ is an unreliable device. The past history of vector analysis seems to have led them to make such a conclusion. ∇ is a reliable device when it is used in the gradient of a scalar or vector function, but not in any other application. For the divergence and the curl, the divergence operator, $\dot{\nabla}$, and the curl operator, $\boldsymbol{\nabla}$, are the proper operators. They are distinctly different from ∇.

7.3 A Search for the Divergence Operator in Orthogonal Curvilinear Coordinate Systems

In a very well known book on the methods of theoretical physics [26, p. 44] the authors try to find the differential operators for the three key functions in an orthogonal curvilinear coordinate system. They state:

> The vector operator must have different forms for its different uses:
>
> $$\nabla = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i}\frac{\partial}{\partial v_i} \quad \text{for the gradient}$$
>
> $$= \frac{1}{\Omega}\sum_i \hat{\boldsymbol{u}}_i \frac{\partial}{\partial v_i}\left(\frac{\Omega}{h_i}\right) \quad \text{for the divergence}$$
>
> and no form which can be written for the curl.

We have used Ω to represent $h_1h_2h_3$ and have changed their coordinate variables ξ_i to v_i and their symbols $\boldsymbol{a}_i$ to $\hat{\boldsymbol{u}}_i$. It is obvious that the 'operator' introduced by these two authors for the divergence can produce the correct expression for the divergence only if the operation is interpreted as

$$\left[\frac{1}{\Omega}\sum_i \hat{\boldsymbol{u}}_i \frac{\partial}{\partial v_i}\left(\frac{\Omega}{h_i}\right)\right]\cdot \boldsymbol{f} \rightarrow \frac{1}{\Omega}\sum_i \frac{\partial}{\partial v_i}\left(\frac{\Omega}{h_i}\hat{\boldsymbol{u}}_i \cdot \boldsymbol{f}\right) \qquad (7.25)$$

Such an interpretation is quite arbitrary, and it does not follow the accepted rule of a differential operator because the first term within the bracket is a function so the entire expression represents the scalar product of $[\cdots]$ and $\boldsymbol{f}$. One is not supposed to move the unit vector $\hat{\boldsymbol{u}}_i$ to the right side of Ω / h and then combine $\hat{\boldsymbol{u}}_i$ with $\cdot \boldsymbol{f}$ as shown in the right term of (7.25). It is a matter of creating a desired expression by arbitrarily rearranging the terms in a function and the position of the dot operator. A reader must recognize now that ∇ can never be a part of the divergence operator nor the curl operator. The proper operators for the divergence and the curl are $\nabla\!\!\!\cdot\,$ and $\nabla\!\!\!\!\times$ respectively. We could have used any two symbols for that matter such as D and C.

7.4 The Use of ∇ to Derive Vector Identities

There are many authors who have tried to apply identities in vector algebra to 'derive' vector identities involving the differential functions ∇f, $\nabla\!\!\!\cdot\, f$, and $\nabla\!\!\!\!\times f$. We quote here two examples. The first example is from the book by Borisenko and Tarapov [21, p. 180] where a problem is posed and 'solved':

> Prob. 7. Find $\nabla(\boldsymbol{A}\cdot\boldsymbol{B})$
>
> Solution. Clearly
>
> $$\nabla(\boldsymbol{A}\cdot\boldsymbol{B}) = \nabla(\boldsymbol{A}_c\cdot\boldsymbol{B}) + \nabla(\boldsymbol{A}\cdot\boldsymbol{B}_c) \tag{7.26}$$
>
> where the subscript c has the same meaning as on p. 170 (the subscript c denotes that the quantity to which it is attached is momentarily being held fixed). According to formula (1.30)
>
> $$\boldsymbol{c}(\boldsymbol{a}\cdot\boldsymbol{b}) = (\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b} - \boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) \tag{7.27}$$
>
> Hence, setting
>
> $$\boldsymbol{a} = \boldsymbol{A}_c,\ \boldsymbol{b} = \boldsymbol{B},\ \boldsymbol{c} = \nabla,$$
>
> we have
>
> $$\nabla(\boldsymbol{A}_c\cdot\boldsymbol{B}) = (\boldsymbol{A}_c\cdot\nabla)\boldsymbol{B} + \boldsymbol{A}_c\times(\nabla\times\boldsymbol{B}) \tag{7.28}$$
>
> and similarly,
>
> $$\nabla(\boldsymbol{A}\cdot\boldsymbol{B}_c) = \nabla(\boldsymbol{B}_c\cdot\boldsymbol{A}) = (\boldsymbol{B}_c\cdot\nabla)\boldsymbol{A} + \boldsymbol{B}_c\times(\nabla\times\boldsymbol{A}) \tag{7.29}$$
>
> Thus, finally,
>
> $$\nabla(\boldsymbol{A}\cdot\boldsymbol{B}) = (\boldsymbol{A}\cdot\nabla)\boldsymbol{B} + (\boldsymbol{B}\cdot\nabla)\boldsymbol{A} + \boldsymbol{A}\times\operatorname{curl}\boldsymbol{B} + \boldsymbol{B}\times\operatorname{curl}\boldsymbol{A} \tag{7.30}$$

As far as the final result, (7.30), is concerned, they have indeed obtained a correct answer. But what is the justification of applying (7.27) with c replaced by ∇ and why cannot the same formula be applied directly to $\nabla(\boldsymbol{A}\cdot\boldsymbol{B})$? There is no way to provide the answers to these questions. A reader has to accept blindly such a treatment.

The second example is found in the book by Panofsky and Phillips [27, p. 470]. They wrote:

$$\begin{aligned}\nabla\times(\boldsymbol{A}\times\boldsymbol{B}) &= (\nabla\cdot\boldsymbol{B})\boldsymbol{A}-(\nabla\cdot\boldsymbol{A})\boldsymbol{B}\\ &= (\nabla\cdot\boldsymbol{B}_c)\boldsymbol{A}+(\nabla\cdot\boldsymbol{B})\boldsymbol{A}_c-(\nabla\cdot\boldsymbol{A}_c)\boldsymbol{B}-(\nabla\cdot\boldsymbol{A})\boldsymbol{B}_c\end{aligned} \tag{7.31}$$

> where the subscript c indicates that the function is constant and may be permuted with the vector operator, with due regard to sign changes if such changes are indicated by the ordinary vector relations.

It is seen that their $(\nabla\cdot\boldsymbol{B})\boldsymbol{A}$ in the first line is not $(\mathrm{div}\,\boldsymbol{B})\boldsymbol{A}$. Rather it is equal to $(\nabla\cdot\boldsymbol{B}_c)\boldsymbol{A}+(\nabla\cdot\boldsymbol{B})\boldsymbol{A}_c$. Secondly, if $\boldsymbol{B}_c$ is constant, the established rule in differential calculus would consider their $\nabla\cdot\boldsymbol{B}_c$ (i.e., div $\boldsymbol{B}_c$) $=0$. The use of algebraic identities to derive differential identities by replacing a vector by ∇ has no foundation - the first line of (7.31). For the exercise in consideration, one way to find the identity is to prove first that

$$\bar{\nabla}(\boldsymbol{A}\times\boldsymbol{B})=\bar{\nabla}(\boldsymbol{BA}-\boldsymbol{AB})$$

or

$$\mathrm{Curl}\,(\boldsymbol{A}\times\boldsymbol{B})=\mathrm{div}\,(\boldsymbol{BA}-\boldsymbol{AB}) \tag{7.32}$$

where $\boldsymbol{AB}$ is a dyadic and $\boldsymbol{BA}$ its transpose (see Sec. 2.3). Then by means of dyadic analysis one finds

$$\bar{\nabla}(\boldsymbol{BA})=(\bar{\nabla}\,\boldsymbol{B})\boldsymbol{A}+\boldsymbol{B}\cdot\nabla\boldsymbol{A} \tag{7.33}$$

$$\bar{\nabla}(\boldsymbol{AB})=(\bar{\nabla}\,\boldsymbol{A})\boldsymbol{B}+\boldsymbol{A}\cdot\nabla\boldsymbol{B} \tag{7.34}$$

Hence

$$\bar{\nabla}(\boldsymbol{A}\times\boldsymbol{B})=(\,\bar{\nabla}\,\boldsymbol{B})\boldsymbol{A}+\boldsymbol{B}\cdot\nabla\boldsymbol{A}$$

$$-(\nabla A)B - A\cdot\nabla B \tag{7.35}$$

where ∇A and ∇B are two dyadic functions. A simpler method of deriving (7.35) will be shown in later section. It should be emphasized that one cannot legitimately write

$$\nabla\times(A\times B)=(\nabla\cdot B)A-(\nabla\cdot A)B$$

as the two authors did and then change $(\nabla\cdot B)A$ to $\nabla\cdot(BA)$, and similarly for $(\nabla\cdot A)B$ in order to create a desired identity.

A general comment on the analogy and no analogy between algebraic vector identities and differential vector identities was made by Milne [28]. It was stated on p. 77:

> The above examples [referring to nine differential vector identities expressed in linguistic notations such as [grad $(X\cdot Y)=(\text{grad}\,X)\cdot Y+(\text{grad}\,Y)\cdot X$ etc.] whilst exhibiting the relations between the symbols in vector or tensor form, conceal the nature of the identities. A little gain of insight is obtained occasionally if the symbol ∇ is employed. E.g., Example (9) [$\text{curl}\,\text{curl}\,X=\text{grad}\,\text{div}\,X-\nabla^2X$] may be written
>
> $$\nabla\times(\nabla\times X)=\nabla(\nabla\cdot X)-\nabla^2X \tag{7.36}$$
>
> which bears an obvious analogy to
>
> $$Q\times(Q\times X)=Q(Q\cdot X)-Q^2X \tag{7.37}$$
>
> where Q denotes a vector function.
>
> On the other hand Example (5)
>
> [Curl $(X\times Y)=Y\cdot\text{grad}\,X-X\cdot\text{grad}\,Y+X\,\text{div}\,Y-Y\,\text{div}\,X$]
>
> may be written
>
> $$\nabla\times(X\times Y)=Y\cdot\nabla X-X\cdot\nabla Y+X(\nabla\cdot Y)-Y(\nabla\cdot X) \tag{7.38}$$
>
> which bears no obvious analogy to

$$Q \times (X \times Y) = X(Q \cdot Y) - Y(Q \cdot X) \tag{7.39}$$

To obtain a better analogy, one would have to write

$$Q \times (X \times Y) = Q \cdot (YX - XY) \tag{7.40}$$

and replace Q by ∇.

We do not understand why (7.40) is a better analogy than (7.39) because as algebraic vector identities, they are equivalent. There is only one interpretation of (7.40), namely,

$$Q \times (X \times Y) = (Q \cdot Y)X - (Q \cdot X)Y \tag{7.41}$$

By replacing Q by ∇ in (7.40), and treating the resultant expression as the divergence of the dyadic $XY - YX$ the manipulation is identical to the one used by Panofsky and Phillips. This short paragraph on the role played by del in an authoritative book on vectorial mechanics shows the consequence of treating Gibbs' notations for the divergence and the curl as two products, one scalar and and one vector.

We have now shown the failures by several prominent authors in trying to invoke ∇ as an operator, not only for the gradient but also for the divergence and the curl. The role is now filled in by the symbolic vector to be discussed in the next section. Many of the ambiguities which have occurred in the past presentations covered in this paper will be recast correctly and unambiguously by our new method utilizing the symbolic vector. In fact, the entire subject of vector analysis can be developed from one single defining equation including the derivation of vector identities and integral theorems. The method can also be extended to non-orthogonal curvilinear systems.

8. The Method of Symbolic Vector

The method of symbolic vector was first disclosed in an article published in 1991 [29]. The entire subject was then expanded and compiled in a book [25] published in 1992. We are not going to give any derivations nor proofs of the theorems to be used here. Rather, we will just use some important formulas in that method in order to clarify the historical presentations covered in the previous sections.

In the first place, a symbolic expression, denoted by $T(\nabla)$, involves a symbolic vector, or a dummy vector denoted by ∇. The symbolic expression is defined by

$$T(\nabla) = \lim_{\Delta V \to 0} \frac{\oiint_S T(\hat{\boldsymbol{n}})dS}{\Delta V} \tag{8.1}$$

where $\hat{\boldsymbol{n}}$ denotes an outward unit normal vector of a surface, S, enclosing the volume ΔV The symbolic expression is created by replacing a vector in a well-defined vector expression by ∇ Some of the well-defined vector expressions are:

$$\boldsymbol{ad},\ \boldsymbol{a}\cdot\boldsymbol{d},\ \boldsymbol{a}\times\boldsymbol{d},\ \boldsymbol{a}\cdot(\boldsymbol{d}\times\boldsymbol{b}),\ \boldsymbol{d}(\boldsymbol{a}\cdot\boldsymbol{b}),\ \boldsymbol{d}\cdot(\boldsymbol{ab}),\ \boldsymbol{d}\times(\boldsymbol{ab}),\ \text{and}\ \boldsymbol{d}\times(\boldsymbol{a}\times\boldsymbol{b}) \tag{8.2}$$

where the dot and the cross represent the two binary operators, namely, the scalar product operator and the vector product operator.

When the vector $\boldsymbol{d}$ in (8.2) is replaced by ∇, the symbolic vector or the dummy vector, we obtain the following symbolic expressions:

$$a\nabla,\ \boldsymbol{a}\cdot\nabla,\ \boldsymbol{a}\times\nabla,\ \boldsymbol{a}\cdot(\nabla\times\boldsymbol{b}),\ \nabla(\boldsymbol{a}\cdot\boldsymbol{b}),\ \nabla\cdot(\boldsymbol{ab}),\ \nabla\times(\boldsymbol{ab}),\ \text{and}\ \nabla\times(\boldsymbol{a}\times\boldsymbol{b}) \tag{8.3}$$

To be more specific, if $T(\nabla)$ contains only one function besides ∇, like the first three expressions in (8.3), we sometimes would use the notation $T(\nabla, a)$ where a may be a scalar or a vector. When $T(\nabla)$ contains two functions like the last five in (8.3), the notation $T(\nabla, a, b)$ will be used for clarity if necessary, where a and b may be scalars or vectors or one of each. As far as the notation is concerned, we may use any other symbol to denote the symbolic vector such as $\boldsymbol{S}$

(symbolic) or $\boldsymbol{D}$ (dummy). We adopt the symbol $\tilde{\nabla}$ because by a proper choice of the symbolic expression, we can produce the three key functions, ∇f, $\overset{\cdot}{\nabla} \boldsymbol{F}$, and $\overset{\times}{\nabla} \boldsymbol{F}$ that makes $\tilde{\nabla}$ the creator of the operators ∇, $\overset{\cdot}{\nabla}$, and $\overset{\times}{\nabla}$.

The definition of $T(\tilde{\nabla})$ given by (8.1) is the most important formula in our new method. We would like to consider three simple symbolic expressions to derive the expressions for the gradient, the divergence, and the curl. Let us consider an expression $T(\tilde{\nabla})$ given by

$$T(\tilde{\nabla}) = \boldsymbol{a} \cdot \tilde{\nabla} \tag{8.4}$$

$T(\hat{\boldsymbol{n}})$ is then given by

$$T(\hat{\boldsymbol{n}}) = \boldsymbol{a} \cdot \hat{\boldsymbol{n}} \tag{8.5}$$

which is a function of both $\boldsymbol{a}$ and $\hat{\boldsymbol{n}}$. By substituting (8.4) and (8.5) into (8.1), we obtain

$$\boldsymbol{a} \cdot \tilde{\nabla} = \lim_{\Delta V \to 0} \frac{\oint\!\!\oint_S \boldsymbol{a} \cdot \hat{\boldsymbol{n}}\, dS}{\Delta V} \tag{8.6}$$

For simplicity, let us evaluate the limit of the integral-differential expression in (8.6) in the Cartesian coordinate system; we obtain:

$$\boldsymbol{a} \cdot \tilde{\nabla} = \sum_i \hat{\boldsymbol{x}}_i \cdot \frac{\partial \boldsymbol{a}}{\partial x_i} \tag{8.7}$$

which is the expression for the divergence of $\boldsymbol{a}$ now denoted by $\overset{\cdot}{\nabla} \boldsymbol{a}$. The chain of events is recapitulated in the following line:

$$\boldsymbol{a} \cdot \tilde{\nabla} = \lim_{\Delta V \to 0} \frac{\oint\!\!\oint_S \boldsymbol{a} \cdot \hat{\boldsymbol{n}}\, dS}{\Delta V} = \sum_i \hat{\boldsymbol{x}}_i \cdot \frac{\partial \boldsymbol{a}}{\partial x_i}$$

$$= \overset{\cdot}{\nabla} \boldsymbol{a} \text{ (divergence of } \boldsymbol{a}) \tag{8.8}$$

If we start with a $T(\tilde{\nabla})$ represented by $\tilde{\nabla} \cdot \boldsymbol{a}$ we would obtain the same result because $\hat{\boldsymbol{n}} \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \hat{\boldsymbol{n}}$, i.e.,

$$\tilde{\nabla} \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \tilde{\nabla} = \overset{\cdot}{\nabla} \boldsymbol{a} \tag{8.9}$$

We also want to remark that the dot in $\boldsymbol{a}\cdot\boldsymbol{\nabla}$ or $\boldsymbol{\nabla}\cdot\boldsymbol{a}$ is a sign or a binary operator for the scalar product, but the product is executed in $\hat{\boldsymbol{n}}\cdot\boldsymbol{a}$ or $\boldsymbol{a}\cdot\hat{\boldsymbol{n}}$. This is one of the most important and delicate concepts in the method of symbolic vector. A reader must grasp firmly this concept in order to understand and to use this method freely without uncertainty.

The symbolic vector $\boldsymbol{\nabla}$ and the function $\boldsymbol{a}$ in the symbolic expression $\boldsymbol{\nabla}\cdot\boldsymbol{a}$ is therefore commutative. At this stage we request the reader to leave aside Gibbs' notations for the divergence and the curl. The use of Gibbs' notation now would bring lots of confusion into understanding the method of symbolic vector. On the other hand, if we had used $\boldsymbol{S}$ as the notation for the symbolic vector we would obtain

$$\boldsymbol{S}\cdot\boldsymbol{a}=\boldsymbol{a}\cdot\boldsymbol{S}=\boldsymbol{\nabla}\,\boldsymbol{a} \tag{8.10}$$

that does not change the result at all.

We consider now a symbolic expression given by $f\,\boldsymbol{\nabla}$ or $\boldsymbol{\nabla}\,f$. Then, an application of (8.1) yields the expression for the gradient, i.e.,

$$f\,\boldsymbol{\nabla}=\boldsymbol{\nabla}\,f=\nabla f \tag{8.11}$$

Finally, if we let $T(\boldsymbol{\nabla})$ equal to $\boldsymbol{\nabla}\times\boldsymbol{f}$ or $-\boldsymbol{f}\times\boldsymbol{\nabla}$ we obtain the expression for the curl:

$$\boldsymbol{\nabla}\times\boldsymbol{f}=-\boldsymbol{f}\times\boldsymbol{\nabla}=\boldsymbol{\nabla}\,\boldsymbol{f} \tag{8.12}$$

The three sample examples show very clearly that by means of the definition of $T(\boldsymbol{\nabla})$ we can readily derive the differential expressions for the three key functions in vector analysis. In general, the integral/differential limit for any $T(\boldsymbol{\nabla})$ defined by (8.1) can be evaluated in an orthogonal curvilinear system that yields

$$T(\boldsymbol{\nabla})=\frac{1}{\Omega}\sum\frac{\partial}{\partial v_i}\left[\frac{\Omega}{h_i}T(\hat{\boldsymbol{u}}_i)\right] \tag{8.13}$$

where h_i and $\hat{\boldsymbol{u}}_i$ denote, respectively, the metric coefficients and the unit vectors in that system, and $\Omega=h_1h_2h_3$. The derivation of (8.13) is found in [25, p. 38].

There are two lemmas associated with the definition of $T(\nabla)$:

Lemma 1: For any symbolic expression $T(\nabla)$, which is generated from a valid vector expression, we can treat ∇ in that expression as a vector, and all of the algebraic identities in vector algebra are applicable. Thus we have

$$a\nabla = \nabla a \tag{8.14}$$

$$\nabla \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \nabla \tag{8.15}$$

$$\nabla \times \boldsymbol{a} = -\boldsymbol{a} \times \nabla \tag{8.16}$$

$$\boldsymbol{b} \cdot (\boldsymbol{a} \times \nabla) = \nabla \cdot (\boldsymbol{b} \times \boldsymbol{a}) = \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b}) \tag{8.17}$$

$$\nabla \times (\boldsymbol{a} \times \boldsymbol{b}) = (\nabla \cdot \boldsymbol{b})\boldsymbol{a} - (\nabla \cdot \boldsymbol{a})\boldsymbol{b} \tag{8.18}$$

When $T(\nabla)$ contains more than one function besides ∇, say, two functions a and b which could be both scalars or vectors or one of each, we will specifically use the form $T(\nabla, a, b)$ to write such a symbolic expression. In this case we have

Lemma 2: For a symbolic expression containing two functions, the following relation holds true:

$$T(\nabla, a, b) = T(\nabla_a, a, b) + T(\nabla_b, a, b) \tag{8.19}$$

where ∇_a and ∇_b denote two partial symbolic vectors.

A symbolic expression containing a partial symbolic vector is defined by

$$T(\nabla_a, a, b) = \lim_{\Delta V \to 0} \frac{\left[\oiint_S T(\hat{\boldsymbol{n}}, a, b)\, dS \right]_{b=\text{constant}}}{\Delta V} \tag{8.20}$$

In an orthogonal curvilinear coordinate system, the differential form of $T(\nabla_a, a, b)$ is given by

$$T(\nabla_a, a, b) = \frac{1}{\Omega} \sum_i \frac{\partial}{\partial v_i} \left[\frac{\Omega}{h_i} T(\hat{\boldsymbol{u}}_i, a, b) \right]_{b=\text{constant}} \tag{8.21}$$

Similarly for $T(\nabla_b, a, b)$. Lemma 1 also applies to expressions containing a partial symbolic vector. There is no need for us to consider symbolic expressions containing more than two functions. For example, to determine an expression like

$$(\boldsymbol{a}\times\nabla)\cdot(\boldsymbol{b}\times\boldsymbol{c})$$

we can treat $(\boldsymbol{a}\times\nabla)\cdot\boldsymbol{d}$ first. After we obtain the result we let $\boldsymbol{d}=\boldsymbol{b}\times\boldsymbol{c}$ and simplify the resultant. With (8.13) - (8.21) and the two lemmas at our disposal, we can re-examine the presentations by these authors discussed previously.

To find the expressions for the three key functions in orthogonal curvilinear systems, we let $T(\nabla)$ equal to ∇f or $f\nabla$, $f\cdot\nabla$ or $\nabla\cdot f$, and $\nabla\times f$ or $-f\times\nabla$ respectively in (8.13), one finds

$$\nabla f = f\nabla = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i}\frac{\partial f}{\partial v_i} = \nabla f\,(\text{gradient of} f) \tag{8.22}$$

$$\nabla\cdot f = f\cdot\nabla = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i}\cdot\frac{\partial f}{\partial v_i} = \nabla f\,(\text{divergence of } f) \tag{8.23}$$

$$\nabla\times f = -f\times\nabla = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i}\times\frac{\partial f}{\partial v_i} = \nabla f\,(\text{curl of } f) \tag{8.24}$$

Equations (8.23) and (8.24) can be converted into the form

$$\nabla f = \frac{1}{\Omega}\sum_i \frac{\partial}{\partial v_i}\left(\frac{\Omega}{h_i}f_i\right) \tag{8.25}$$

$$\nabla f = \frac{1}{\Omega}\sum_{i,j,k} h_i\hat{\boldsymbol{u}}_i\left[\frac{\partial(h_k f_k)}{\partial v_j}-\frac{\partial(h_j f_j)}{\partial v_k}\right] \tag{8.26}$$

In (8.26) the summation is taken in the cyclic order of $i,j,k=(1,2,3)$. In obtaining these results we made use of the following identities [25, p. 13 - 15]:

$$\sum_i \frac{\partial}{\partial v_i}\left(\frac{\Omega}{h_i}\hat{\boldsymbol{u}}_i\right)=0 \tag{8.27}$$

$$\frac{\partial\hat{\boldsymbol{u}}_i}{\partial v_i} = -\left(\frac{1}{h_j}\frac{\partial h_i}{\partial v_j}\hat{\boldsymbol{u}}_j+\frac{1}{h_k}\frac{\partial h_i}{\partial v_k}\hat{\boldsymbol{u}}_k\right) \tag{8.28}$$

$$\frac{\partial \hat{\boldsymbol{u}}_i}{\partial v_j} = \frac{1}{h_i}\frac{\partial h_j}{\partial v_i}\hat{\boldsymbol{u}}_j \;\; (i \neq j) \tag{8.29}$$

In (8.28) $i, j, k = (1,2,3)$ is taken in cyclic order. It is very clear that the method of symbolic vector yields directly the operational expressions for the divergence and the curl as well as the gradient. Morse and Feshbach's failure to find the divergence operator and the curl operator is remedied in this analysis. The curl operator evolves just as readily as the divergence operator. From the structure of (8.23) we see clearly that $\boldsymbol{\nabla} f$ is not the scalar product between ∇ and f, an assumption made by Moon and Spencer for the orthogonal curvilinear system.

In regard to Heaviside's treatment the proper substitutes for his 'equation' (5.1) to (5.3) or (5.4) to (5.6) are:

$$\boldsymbol{N} \cdot \boldsymbol{\nabla} \times \boldsymbol{E} = \boldsymbol{\nabla}(\boldsymbol{E} \times \boldsymbol{N}) = \boldsymbol{E} \cdot (\boldsymbol{N} \times \boldsymbol{\nabla}) \tag{8.30}$$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{E} \times \boldsymbol{H}) = \boldsymbol{E} \cdot (\boldsymbol{H} \times \boldsymbol{\nabla}) = \boldsymbol{H} \cdot (\boldsymbol{\nabla} \times \boldsymbol{E}) \tag{8.31}$$

$$\begin{aligned} \boldsymbol{\nabla} \cdot (\boldsymbol{E} \times \boldsymbol{H}) &= \boldsymbol{\nabla}_E \cdot (\boldsymbol{E} \times \boldsymbol{H}) - \boldsymbol{\nabla}_H \cdot (\boldsymbol{H} \times \boldsymbol{E}) \\ &= \boldsymbol{H} \cdot (\boldsymbol{\nabla}_E \times \boldsymbol{E}) - \boldsymbol{E} \cdot (\boldsymbol{\nabla}_H \times \boldsymbol{H}) \end{aligned} \tag{8.32}$$

Equation (8.32) yields

$$\boldsymbol{\nabla}(\boldsymbol{E} \times \boldsymbol{H}) = \boldsymbol{H} \cdot \boldsymbol{\nabla} \boldsymbol{E} - \boldsymbol{E} \cdot \boldsymbol{\nabla} \boldsymbol{H} \tag{8.33}$$

Although our equations have the same form as Heaviside's except that his ∇ has been replaced by $\boldsymbol{\nabla}$, the symbolic vector, yet there is a vast difference in meaning of the two sets. For example, his $\boldsymbol{H} \cdot \nabla \times \boldsymbol{E}$ in (5.2) is interpreted as $\boldsymbol{H} \cdot \mathrm{Curl}\, \boldsymbol{E}$ but our $\boldsymbol{H} \cdot \boldsymbol{\nabla} \times \boldsymbol{E}$ is the same as $\boldsymbol{\nabla} \cdot (\boldsymbol{E} \times \boldsymbol{H})$ because of Lemma 2 and it is equal to $\boldsymbol{\nabla}(\boldsymbol{E} \times \boldsymbol{H})$.

Every term in (8.30) to (8.33) is well defined. Both Lemma 1 and Lemma 2 are used to obtain the vector identity stated by (8.33).

Return now to the problems posed by Borinsenko and Parapov we start with the symbolic expression $\forall(\boldsymbol{A}\cdot\boldsymbol{B})$ for $\nabla(\boldsymbol{A}\cdot\boldsymbol{B})$; then by applying Lemma 2 we have

$$\forall(\boldsymbol{A}\cdot\boldsymbol{B})=\forall_A(\boldsymbol{A}\cdot\boldsymbol{B})+\forall_B(\boldsymbol{A}\cdot\boldsymbol{B}) \tag{8.34}$$

Applying Lemma 1 we have

$$\forall_A(\boldsymbol{A}\cdot\boldsymbol{B})=(\boldsymbol{B}\cdot\forall_A)\boldsymbol{A}-\boldsymbol{B}\times(\boldsymbol{A}\times\forall_A) \tag{8.35}$$

and

$$\forall_B(\boldsymbol{A}\cdot\boldsymbol{B})=(\boldsymbol{A}\cdot\forall_B)\boldsymbol{B}-\boldsymbol{A}\times(\boldsymbol{B}\times\forall_B) \tag{8.36}$$

hence,

$$\forall_A(\boldsymbol{A}\cdot\boldsymbol{B})=\boldsymbol{B}\cdot\nabla\boldsymbol{A}+\boldsymbol{B}\times\forall\boldsymbol{A} \tag{8.37}$$

and

$$\forall_B(\boldsymbol{A}\cdot\boldsymbol{B})=\boldsymbol{A}\cdot\nabla\boldsymbol{B}+\boldsymbol{A}\times\forall\boldsymbol{B} \tag{8.38}$$

Thus,

$$\nabla(\boldsymbol{A}\cdot\boldsymbol{B})=\boldsymbol{A}\cdot\nabla\boldsymbol{B}+\boldsymbol{B}\cdot\nabla\boldsymbol{A}+\boldsymbol{A}\times\forall\boldsymbol{B}+\boldsymbol{B}\times\forall\boldsymbol{A} \tag{8.39}$$

Our derivation of (8.39) appears to be similar to the derivation by Borisenko and Tarapov in form, but the use of the FSP and FVP in their formulation and the treatment of (7.26) as an algebraic identity is entirely unacceptable while each of our steps are supported by the basic principle in the method of symbolic vector, particularly the two Lemmas therein.

The exercise posed by Panofsky and Phillips can be formulated correctly by our new method. The steps are outlined below:

We start with $\forall\times(\boldsymbol{A}\times\boldsymbol{B})$ which is the symbolic expression of $\forall(\boldsymbol{A}\times\boldsymbol{B})$; then by means of Lemma 2

$$\forall\times(\boldsymbol{A}\times\boldsymbol{B})=\forall_A\times(\boldsymbol{A}\times\boldsymbol{B})+\forall_B\times(\boldsymbol{A}\times\boldsymbol{B}) \tag{8.40}$$

By means of Lemma 1, we have

$$\boldsymbol{\nabla}_A \times (A \times B) = (B \cdot \boldsymbol{\nabla}_A)A - (\boldsymbol{\nabla}_A \cdot A)B$$

$$= B \cdot \nabla A - B \boldsymbol{\nabla} A$$

Similarly

$$\boldsymbol{\nabla}_B \times (A \times B) = (B \cdot \boldsymbol{\nabla}_B)A - (A \cdot \boldsymbol{\nabla}_B)B$$

$$= A \boldsymbol{\nabla} B - A \cdot \nabla B$$

hence,

$$\boldsymbol{\nabla}(A \times B) = A \boldsymbol{\nabla} B - A \cdot \nabla B$$

$$-B \boldsymbol{\nabla} A + B \cdot \nabla A \tag{8.41}$$

which is the same as (7.35) obtained previously in Sec. 7 by a more complicated analysis. The convenience and the simplicity of the method of symbolic vector to derive vector identities hopefully has been demonstrated very clearly in the last two examples. All commonly used vector identities have been derived in this way as shown in [25, pp. 52 - 54]. The method of symbolic vector has so far been applied only to orthogonal curvilinear systems. The method can be applied equally well to general or non-orthogonal curvilinear systems. The formulation is shown in the next section.

9. General Curvilinear Coordinate Systems

Expressions for the three key functions in vector analysis have been derived previously by Stratton [10, pp. 38 - 47] in the general curvilinear coordinate system or non-orthogonal coordinate system. The expression for the gradient is obtained by means of the relation $d\phi = \nabla\phi \cdot d\boldsymbol{r}$ using the contravarient components of the displacement vector $d\boldsymbol{r}$. The expression for the divergence is found by applying Gauss's theorem to an infinitesimal region and the expression for the curl is obtained by applying Stokes' theorem to an infinitesimal contour. We will derive these expressions by the method of symbolic vector by applying the basic definition of $T(\forall)$ to all three functions. To carry out the analysis it is necessary to reveal the concept of unitary vectors and reciprocal vectors. However, we will not follow the usual treatment based on tensor analysis and notations. Rather we will treat the subject entirely within the framework of vector analysis except to share some technical nomenclature commonly used in tensor analysis. Vector analysis in general curvilinear coordinate systems is not covered in [25]; the material to be presented is new.

9.1 Unitary Vectors and Reciprocal Vectors

In the general curvilinear coordinate system, henceforth to be abbreviated as GCS, the total differential of a displacement vector will be written in the form

$$d\boldsymbol{r} = \sum_i \boldsymbol{a}_i dv_i \tag{9.1}$$

where $\boldsymbol{a}_i$, with $i = 1,2,3$, are called unitary vectors and v_i the coordinate variables. The unitary vectors are not necessarily of unit length, nor of the dimension of length. The three base vectors define a differential volume given by

$$\begin{aligned} dV &= \boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) dv_1 dv_2 dv_3 \\ &= \Lambda dv_1 dv_2 dv_3 \end{aligned} \tag{9.2}$$

where

$$\Lambda = \boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = \boldsymbol{a}_2 \cdot (\boldsymbol{a}_3 \times \boldsymbol{a}_1) = \boldsymbol{a}_3 \cdot (\boldsymbol{a}_1 \times \boldsymbol{a}_2)$$

$\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ are in general, not orthogonal to each other. Three reciprocal vectors, denoted by $\boldsymbol{b}_j$ with $j = 1,2,3$, are defined by

$$\boldsymbol{b}_1 = \frac{1}{\Lambda}\boldsymbol{a}_2 \times \boldsymbol{a}_3,\ \boldsymbol{b}_2 = \frac{1}{\Lambda}\boldsymbol{a}_3 \times \boldsymbol{a}_1,\ \boldsymbol{b}_3 = \frac{1}{\Lambda}\boldsymbol{a}_1 \times \boldsymbol{a}_2 \qquad (9.3)$$

They are called reciprocal vectors because

$$\boldsymbol{b}_j \cdot \boldsymbol{a}_i = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \qquad (9.4)$$

The unitary vectors can be expressed in terms of the reciprocal vectors in the form:

$$\boldsymbol{a}_1 = \Lambda(\boldsymbol{b}_2 \times \boldsymbol{b}_3),\ \boldsymbol{a}_2 = \Lambda(\boldsymbol{b}_3 \times \boldsymbol{b}_1),\ \boldsymbol{a}_3 = \Lambda(\boldsymbol{b}_1 \times \boldsymbol{b}_2) \qquad (9.4)^*$$

The total differential of the same displacement vector defined in (9.1) can be written in the form

$$d\boldsymbol{r} = \sum_j \boldsymbol{b}_j dw_j \qquad (9.5)$$

where w_i, with $j = 1,2,3$, denote the coordinate variables measured along the reciprocal base vectors, so

$$\frac{\partial \boldsymbol{r}}{\partial w_j} = \boldsymbol{b}_j \qquad (9.6)$$

while

$$\frac{\partial \boldsymbol{r}}{\partial v_i} = \boldsymbol{a}_i \qquad (9.7)$$

It can be readily shown that

$$\boldsymbol{b}_1 \cdot (\boldsymbol{b}_2 \times \boldsymbol{b}_3) = \boldsymbol{b}_2 \cdot (\boldsymbol{b}_3 \times \boldsymbol{b}_1) = \boldsymbol{b}_3 \cdot (\boldsymbol{b}_1 \times \boldsymbol{b}_2) = \frac{1}{\Lambda} \qquad (9.8)$$

---

* Stratton [10, p. 39] had inadvertently written, in our notation, $\boldsymbol{a}_1 = \frac{1}{\Lambda}(\boldsymbol{b}_2 \times \boldsymbol{b}_3)$, etc.

A vector function $\boldsymbol{F}$ can be expressed either in terms of the unitary vectors or their reciprocal vectors. We write:

$$\boldsymbol{F} = \sum_i f_i \boldsymbol{a}_i = \sum_j g_j \boldsymbol{b}_j \tag{9.9}$$

The notation F_i will be reserved to denote the components of $\boldsymbol{F}$ in an orthogonal curvilinear system. Our f_i and g_i correspond, respectively, to the contravariant and covariant components of $\boldsymbol{F}$ designated by Stratton [26], two names commonly used in tensor analysis. We merely call g_i the components of $\boldsymbol{F}$ in the unitary (vector) system, or the unitary components, and f_i the components of $\boldsymbol{F}$ in the reciprocal (vector) system, or the reciprocal components.

If we denote

$$\alpha_{ij} = \boldsymbol{a}_i \cdot \boldsymbol{a}_j = \alpha_{ji} \tag{9.10}$$

and

$$\beta_{ij} = \boldsymbol{b}_i \cdot \boldsymbol{b}_j = \beta_{ji} \tag{9.11}$$

then the relations between f_i and g_i are:

$$f_i = \sum_j \beta_{ij} g_j \tag{9.12}$$

$$g_i = \sum_j \alpha_{ij} f_j \tag{9.13}$$

We have purposely avoided the superscript notations for g_i and β_{ij} commonly used in tensor analysis for these quantities, mainly to show that vector analysis can be treated properly without the aid of tensor analysis.

On account of the orthogonal relations between $\boldsymbol{a}_i$ and $\boldsymbol{b}_j$ as stated by (9.4) the following relations can be derived from (9.9):

$$f_i = \boldsymbol{F} \cdot \boldsymbol{b}_i \tag{9.14}$$

$$g_j = \boldsymbol{F} \cdot \boldsymbol{a}_j \tag{9.15}$$

hence

$$F = \sum_i (F \cdot b_i) a_i = \sum_j (F \cdot a_j) b_j \tag{9.16}$$

In the language of dyadic analysis

$$\sum_i b_i a_i = \sum_j a_j b_j = \overline{\overline{I}} \tag{9.17}$$

where $\overline{\overline{I}}$ is called the idemfactor such that

$$F \cdot \overline{\overline{I}} = \overline{\overline{I}} \cdot F = F \tag{9.18}$$

We have now sufficient materials to apply the method of symbolic vector to the general curvilinear system.

9.2 Gradient, Divergence and Curl in a General Curvilinear System

In a GCS, the differential length along the coordinates v_i in the direction of the unitary vector a_i is:

$$ds_i = a_i dv_i ,\ i = 1,2,3 \tag{9.19}$$

hence

$$\begin{aligned} ds_i = |ds_i| &= (a_i \cdot a_i)^{\frac{1}{2}} dv_i \\ &= (\alpha_{ii})^{\frac{1}{2}} dv_i \end{aligned} \tag{9.20}$$

A differential area bounded by ds_2 and ds_3 is given by

$$\begin{aligned} dS_1 &= ds_2 \times ds_3 \\ &= a_2 \times a_3 dv_2 dv_3 \end{aligned} \tag{9.21}$$

In general,

$$\begin{aligned} dS_i &= ds_j \times ds_k \\ &= (a_j \times a_k) dv_j dv_k \end{aligned} \tag{9.22}$$

where i, j, k are carried out in cyclic order of $(1,2,3)$.

Then

$$\begin{aligned} dS_i &= \left[(\boldsymbol{a}_j \times \boldsymbol{a}_k)\cdot(\boldsymbol{a}_j \times \boldsymbol{a}_k)\right]^{\frac{1}{2}} dv_j dv_k \\ &= \left[(\boldsymbol{a}_j \cdot \boldsymbol{a}_j)(\boldsymbol{a}_k \cdot \boldsymbol{a}_k) - (\boldsymbol{a}_j \cdot \boldsymbol{a}_k)(\boldsymbol{a}_k \cdot \boldsymbol{a}_j)\right]^{\frac{1}{2}} dv_j dv_k \\ &= \left(\alpha_{jj}\alpha_{kk} - \alpha_{jk}^2\right)^{\frac{1}{2}} dv_j dv_k \end{aligned} \qquad (9.23)$$

where the α_{ij}'s are defined in (9.10).

The differential volume is given by

$$dV = \boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) dv_1 dv_2 dv_3 = \Lambda dv_1 dv_2 dv_3 \qquad (9.24)$$

As shown by Stratton, [10, p. 43]

$$\Lambda = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}^{\frac{1}{2}} \qquad (9.25)$$

Because our notations are different from Stratton's we have repeated some of his presentations mainly for the readers to get accustomed to our notations. We must mention that much of the basic works on unitary and reciprocal vectors are the original contributions of Gibbs found in his first pamphlet [6, Vol. 1, Chapter 1].

In the subsequent analysis we need a theorem which states:

$$\sum_i \frac{\partial}{\partial v_i}(\boldsymbol{a}_j \times \boldsymbol{a}_k) = 0 \qquad (9.26)$$

To prove this theorem we have, in view of (9.7),

$$\frac{\partial \boldsymbol{a}_i}{\partial v_j} = \frac{\partial \boldsymbol{a}_j}{\partial v_i} \qquad (9.27)$$

hence,

$$\sum_i \frac{\partial}{\partial v_i}\left(\boldsymbol{a}_j \times \boldsymbol{a}_k\right) = \boldsymbol{a}_2 \times \frac{\partial \boldsymbol{a}_3}{\partial v_1} + \frac{\partial \boldsymbol{a}_2}{\partial v_1} \times \boldsymbol{a}_3$$
$$+ \boldsymbol{a}_3 \times \frac{\partial \boldsymbol{a}_1}{\partial v_2} + \frac{\partial \boldsymbol{a}_3}{\partial v_2} \times \boldsymbol{a}_1 \qquad (9.28)$$
$$+ \boldsymbol{a}_1 \times \frac{\partial \boldsymbol{a}_2}{\partial v_3} + \frac{\partial \boldsymbol{a}_1}{\partial v_3} \times \boldsymbol{a}_2 .$$

Since

$$\frac{\partial \boldsymbol{a}_3}{\partial v_1} = \frac{\partial \boldsymbol{a}_1}{\partial v_3}, \quad \frac{\partial \boldsymbol{a}_2}{\partial v_1} = \frac{\partial \boldsymbol{a}_1}{\partial v_2}, \quad \frac{\partial \boldsymbol{a}_3}{\partial v_2} = \frac{\partial \boldsymbol{a}_2}{\partial v_3}$$

the six terms in (9.28) cancel each other so (9.26) holds true. The geometrical interpretation of this theorem is that the total vector area of a closed surface vanishes. The theorem for an orthogonal curvilinear system was proved in [26, p. 15]. The proof therein appears more complicated than the present proof using unitary vectors.

We consider the symbolic expression introduced previously by (8.1) but now write it in the form

$$T(\nabla) = \operatorname*{Lim}_{\nabla V \to 0} \frac{\sum_i T(\hat{\boldsymbol{n}}_i)\Delta S_i}{\Delta V} \qquad (9.29)$$

with

$$\hat{\boldsymbol{n}}_i \Delta S_i = \Delta \boldsymbol{S}_i = \left(\boldsymbol{a}_j \times \boldsymbol{a}_k\right)\Delta v_j \Delta v_k$$

with

$$\Delta V = \Lambda \Delta v_1 \Delta v_2 \Delta v_3 .$$

Because of the linearity of $T(\hat{\boldsymbol{n}}_i)$ with respect to $\hat{\boldsymbol{n}}_i$ we find that the differential expression for $T(\nabla)$ in a general curvilinear coordinate system is given by

$$T(\nabla) = \frac{1}{\Lambda}\sum_i \frac{\partial}{\partial v_i} T\left(\boldsymbol{a}_j \times \boldsymbol{a}_k\right) \qquad (9.30)$$

where $i, j, k = (1,2,3)$ taken in cyclic order.

To find the expression for the gradient, we let

$$T(\forall) = \forall f = f \forall$$

so

$$T(\boldsymbol{a}_j \times \boldsymbol{a}_k) = (\boldsymbol{a}_j \times \boldsymbol{a}_k) f$$

that yields

$$\forall f = f \forall = \frac{1}{\Lambda}\sum_i \left(\boldsymbol{a}_j \times \boldsymbol{a}_k\right)\frac{\partial f}{\partial v_i} = \nabla f = \operatorname{grad} f \tag{9.31}$$

where we have made use of (9.26) to eliminate the sum of the derivatives of the cross products. In terms of the reciprocal vectors $\boldsymbol{b}_i$

$$\boldsymbol{a}_j \times \boldsymbol{a}_k = \Lambda \boldsymbol{b}_i;$$

hence

$$\nabla f = \sum_i \boldsymbol{b}_i \frac{\partial f}{\partial v_i} \tag{9.32}$$

The gradient operator in GCS is therefore represented by

$$\nabla = \sum_i \boldsymbol{b}_i \frac{\partial}{\partial v_i} \tag{9.33}$$

To find the expression for the divergence, we let

$$T(\forall) = \forall \cdot \boldsymbol{F} = \boldsymbol{F} \cdot \forall$$

so

$$T(\boldsymbol{a}_j \times \boldsymbol{a}_k) = (\boldsymbol{a}_j \times \boldsymbol{a}_k) \cdot \boldsymbol{F} = \boldsymbol{F} \cdot (\boldsymbol{a}_j \times \boldsymbol{a}_k).$$

Substituting them into (9.30) we obtain

$$\begin{aligned} \forall \cdot \boldsymbol{F} = \boldsymbol{F} \cdot \forall &= \frac{1}{\Lambda}\sum_i \frac{\partial}{\partial v_i}\left[\left(\boldsymbol{a}_j \times \boldsymbol{a}_k\right) \cdot \boldsymbol{F}\right] \\ &= \frac{1}{\Lambda}\sum_i \left[\left(\boldsymbol{a}_j \times \boldsymbol{a}_k\right) \cdot \frac{\partial \boldsymbol{F}}{\partial v_i}\right] \\ &= \sum_i \boldsymbol{b}_i \cdot \frac{\partial \boldsymbol{F}}{\partial v_i} = \forall \boldsymbol{F} = \operatorname{div} \boldsymbol{F} \end{aligned} \tag{9.34}$$

Equation (9.34) shows that the divergence operator in the general curvilinear system is represented by

$$\nabla = \sum_i b_i \cdot \frac{\partial}{\partial v_i} \quad \text{(divergence operator)} \tag{9.35}$$

Comparing it with (9.33), we see that the dot in the right term of (9.34) lies between b_i and $\frac{\partial}{\partial v_i}$, and '$\cdot\frac{\partial}{\partial v_i}$' is a compound operator which is now applied to a vector as the posterior operand. By changing $a_j \times a_k$ to Λb_i in the first line of (9.34) we obtain

$$\nabla F = \frac{1}{\Lambda}\sum_i \frac{\partial}{\partial v_i}(\Lambda \mathbf{b}_i \cdot \mathbf{F}) \tag{9.36}$$

Since

$$b_i \cdot F = f_i$$

according to (9.14) where f_i denotes the reciprocal components of F or the contravariant components of F, (9.36) can be written in the form

$$\nabla F = \frac{1}{\Lambda}\sum_i \frac{\partial}{\partial v_i}(\Lambda f_i) \tag{9.37}$$

which gives directly, the differential form of ∇F without the need to evaluate the derivatives of the unitary vectors or the reciprocal vectors.

By letting

$$T(\nabla) = \nabla \times F = -F \times \nabla \tag{9.38}$$

we can obtain the operational form and the direct differential form of ∇F. They are given by

$$\nabla F = \sum_i b_i \times \frac{\partial F}{\partial v_i} \quad \text{(operational form)} \tag{9.39}$$

with

$$\nabla = \sum_i b_i \times \frac{\partial}{\partial v_i} \quad \text{(Curl operator)} \tag{9.40}$$

and

$$\nabla \mathbf{F} = \frac{1}{\Lambda}\sum_i \mathbf{a}_i \left(\frac{\partial g_k}{\partial v_j} - \frac{\partial g_j}{\partial v_k} \right) \quad \text{(differential form)} \tag{9.41}$$

where g_i denotes the unitary components or the covariant components of F.

Equations (9.32), (9.37), and (9.41) are the same as Stratton's (49), (55), and (63). on pp. 44 - 47 of his book [10]. His formulation does not yield directly our equations (9.34) and (9.39) although they can be obtained from (9.37) and (9.41), or Stratton's (55) and (63) by a proper transformation of variables and with the aid of (9.26). However, it is a rather complicated exercise. Presumably it is for this reason that the proper forms of the divergence operator and the curl operator in general curvilinear coordinate systems are not treated in the literature, including orthoginal curvilinear systems except the Cartesian system as found in Gibbs' classic.

In comparing the present method of deriving (9.37) and (9.42) with that of Stratton, we see that Stratton applies the divergence theorem to obtain the differential expression for the divergence and Stokes' theorem to obtain the differential expression for the curl while our method is based upon one single expression, namely (9.30), from which the expressions for the three key functions can be derived.

Before we apply (9.32), (9.36), and (9.41) to the orthogonal curvilinear systems as special cases, it is desirable to show the invariance of the three operatorsin the general curvilinear system. We consider first the gradient operator:

$$\nabla = \sum_i b_i \frac{\partial}{\partial v_i} \tag{9.42}$$

In a primed system with coordinate variables v_i' and reciprocal vectors b_j' the gradient operator will be denoted by ∇' and given by

$$\nabla' = \sum_j b_j' \frac{\partial}{\partial v_j'} \tag{9.43}$$

We want to show that $\nabla = \nabla'$.

Let

$$b_i = \sum_j c_{ij} b'_j \tag{9.44}$$

By taking the scalar product of (9.44) with a'_k, a typical unitary vector in the primed system, we obtain

$$c_{ik} = b_i \cdot a'_k$$

or

$$c_{ij} = b_i \cdot a'_j \tag{9.45}$$

Thus

$$b_i = \sum_j \left(b_i \cdot a'_j\right) b'_j \tag{9.46}$$

By definition,

$$dr = \sum_i a_i dr_i = \sum_j a'_j dv'_j \tag{9.47}$$

By taking the scalar product of (9.47) with b_k, a typical reciprocal vectorin the unprimed system, we obtain

$$dv_k = \sum_j \left(b_k \cdot a'_j\right) dv'_j$$

or

$$dv_i = \sum_j \left(b_i \cdot a'_j\right) dv'_j \tag{9.48}$$

Hence,

$$\frac{\partial v_i}{\partial v'_j} = b_i \cdot a'_j \tag{9.49}$$

Substituting (9.49) into (9.46), and in view of (9.42), we have

$$\begin{aligned} \nabla &= \sum_i \sum_j b'_j \frac{\partial v_i}{\partial v'_j} \frac{\partial}{\partial v'_i} \\ &= \sum_j b'_j \frac{\partial}{\partial v'_j} = \nabla' \end{aligned} \tag{9.50}$$

Equation (9.50) describes the invariance property of the gradient operator in any two GCS's. Similar proofs apply to the invariance of ∇ and ∇.

To obtain the formulas for the three key functions in an orthogonal curvilinear system which is a special case of the GCS, we let

$$a_i = h_i \hat{u}_i,\ i = 1,2,3 \tag{9.51}$$

where the h_i's denote the metric coefficients, and

$$\hat{u}_i \times \hat{u}_j = \begin{cases} 0, & i = j \\ \hat{u}_k, & i \neq j \neq k \end{cases} \tag{9.52}$$

with $i, j, k = (1,2,3)$ in cyclic order. Then

$$\Lambda = \Omega = h_1 h_2 h_3 \tag{9.53}$$

$$b_i = \hat{u}_i / h_i \tag{9.54}$$

In an orthogonal system the components of F have been denoted by F_i, i.e.,

$$F = \sum_i F_i \hat{u}_i \tag{9.55}$$

Thus, in view of (9.14), (9.15), (9.51), and (9.54) we have

$$f_i = F_i \cdot b_i = F_i / h_i \tag{9.56}$$

$$g_i = F_i \cdot a_i = h_i F_i \tag{9.57}$$

Equations (9.32) to (9.37) and (9.39) to (9.41) become

$$\nabla f = \sum_i \frac{\hat{u}_i}{h_i} \frac{\partial f}{\partial v_k} \tag{9.58}$$

$$\nabla F = \sum_i \frac{\hat{u}_i}{h_i} \cdot \frac{\partial F}{\partial v_i} \tag{9.59}$$

$$\nabla = \sum_i \frac{\hat{u}_i}{h_i} \cdot \frac{\partial}{\partial v_i} \tag{9.60}$$

$$\nabla F = \frac{1}{\Omega} \sum_i \frac{\partial}{\partial v_i} \left(\frac{\Omega}{h_i} \hat{u}_i \cdot F \right) \tag{9.61}$$

$$\nabla F = \frac{1}{\Omega} \sum_i \frac{\partial}{\partial v_i} \left(\frac{\Omega}{h_i} F_i \right) \tag{9.62}$$

$$\boldsymbol{\nabla} \boldsymbol{F} = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i} \times \frac{\partial \boldsymbol{F}}{\partial v_i} \tag{9.63}$$

$$\boldsymbol{\nabla} = \sum_i \frac{\hat{\boldsymbol{u}}_i}{h_i} \times \frac{\partial}{\partial v_i} \tag{9.64}$$

$$\boldsymbol{\nabla} \boldsymbol{F} = \frac{1}{\Omega} \sum_i h_i \hat{\boldsymbol{u}}_i \left[\frac{\partial (h_k F_k)}{\partial v_j} - \frac{\partial (h_j F_j)}{\partial v_k} \right]. \tag{9.65}$$

As shown before, these formulas can be derived by using Gibbs' formulas for the three key functions in the Cartesian system and invoking the invariance of the three distinct differential operators ∇, $\bar{\nabla}$, and $\boldsymbol{\nabla}$. We list these expressions at the end of the main body of this essay to point out that all these expressions are now derived from one single defining equation, namely the symbolic expression $T(\boldsymbol{\nabla})$ given by (8.1).

10. Retrospect

In this work we have examined critically some practices of presenting vector analysis in several early works and in a few contemporary writings. It should be pointed out emphatically that the whole subject of vector analysis was formulated by the great American scientist J. Willard Gibbs in a very precise and elegant fashion. Although his original works are confined to formulations in a Cartesian coordinate system, they can be extended to curvilinear systems as a result of the invariance of the differential operators, as reviewed in this paper, without the necessity of resorting to the aid of tensor analysis.

In spite of the richness of Gibbs' theory of vector analysis, his notations for the divergence and the curl, in the opinion of this author, have induced several later workers, including one of his students, Wilson, to make some inappropriate interpretations. The adoption of these interpretations is world-wide. We have selected a few examples from the works of several seasoned scientists and engineers to illustrate the prevalence of the improper use of ∇.

Many authors in the past have considered Heaviside to be a co-founder with Gibbs of the modern vector analysis. We do not share this view. In Heaviside's treatment of vector analysis, he spoke freely of the scalar product and the vector product between ∇ and a vector function $\boldsymbol{F}$ and he used ∇ as a vector in deriving algebraic vector identities which incorporate differential entities. In view of these mathematically insupportable treatments, Heaviside's status as a pioneer in vector analysis is not of the same level as Gibbs'. In the historical introduction of a 1950 edition of Heaviside's book on *Electromagnetic Theory* [8], Ernst Weber stated:

> Chap. III of the *Electromagnetic Theory* dealing with "The Elements of Vectorial Algebra and Analysis" is practically the model of modern treatises on vector analysis. Considerable moral assistance came from a pamphlet by J. W. Gibbs who independently developed vector analysis during 1881-4 in Heaviside sense - but

> using the less attractive notation of Tait; however, Gibbs deferred publication until 1901.

The above statement contains, unfortunately, several misleading messages. In the first place, in view of our detailed study of Heaviside's works, his treatment would be a poor model if it were used to teach vector calculus. Secondly, if Heaviside truly received moral assistance from Gibbs' pamphlet, he would not have committed himself to the improper use of ∇, and would have restricted his use of it to the expression for the gradient. Most important of all, Gibbs did not develop his theory in the Heaviside sense. His development is completely different from that of Heaviside. Finally, the book published in 1901 was written by Wilson, not by Gibbs himself. Even though it was founded upon the lectures of Gibbs, it contained some of Wilson's own interpretations which are not found in Gibbs' original pamphlets nor in his lecture notes reported by Wilson. The two prefaces, one by Gibbs and another by Wilson, which we quote in Sec. 1, are proofs of our assertion. We were reluctant to criticize a scientist of Heaviside's status and the opinion expressed by Prof. Weber. After all, Heaviside had contributed very much to electromagnetic theory and had been recognized as a rare genius. However, in the field of vector analysis we must set the record straight and call attention to the outstanding contribution of Gibbs who stood above all his contemporaries in the last century. For the sake of future generations of students, we have the obligation to remove unsound arguments and arbitrary manipulations in an otherwise precise branch of mathematical science.

The recently published symbolic method of treating vector analysis, which has been introduced briefly in this paper, shows that some of the treatments can be remedied by using the technique in this relatively new method. It also shows that the entire subject of vector calculus can be conveniently developed based on one single defining expression that includes all the integral theorems which are not mentioned in this paper, but can be found in [25]. As far as notations are concerned, in addition to the long established notations of Gibbs and the linguistic notations, we have presented new symbols for two distinct differential operators for the divergence and the curl to accompany the existing gradient operator. For a decision on whether these new notations will be of interest and use to students, we leave the matter in the hands of future generations.

We have examined a history covering a period of over one century. It represents a very interesting period in the development of the mathematical foundations of electromagnetic theory. However, in view of the long-entrenched and widespread mis-use of the gradient operator, ∇, as a component of the divergence and curl operators, the obligation of sharing the insight presented here with many of our colleagues in this field has been a labour fraught with frustration.

We hope that this presentation is clear enough that the issue(s) will be understood by the serious workers in this subject, and that future students will not have to ponder over contradictions and misrepresentations.

Acknowledgement

I am grateful to Prof. Philip H. Alexander of the University of Windsor, Canada, currently a visiting scholar at the University of Michigan, for his most valuable suggestions and assistance during my preparation of this manuscript. Dr. John Bryant has given me constant advice and encouragement ever since this research was started. Prof. W. Jack Cunningham of Yale University was very kind in helping me to search for the Lecture Notes of Prof. Gibbs recorded by Dr. Wilson, currently stored at Sterling Library of Yale University. It is a very valuable document in my study. The support which I have received from Dr. Thomas Phipps, Jr. is very much appreciated. I want to thank Prof. Fawwaz Ulaby, Director of the Radiation Laboratory at the University of Michigan for his encouragement, and Dr. Shou-Zhong Wang for his technical assistance.

References

[1] Crowe, Michael J., *A History of Vector Analysis*, Dover Publications, N. Y., 1985, an unabridged and corrected republication of the work first published by the University of Notre Dame Press in 1967.

[2] Burati-Forti, C., and Marcolongo, R., *Elements de Calcul Vectoriel*, French edition of the original in Italian by S. Lattès, Librairie Scientifique, A. Hermann et Fils, Paris, 1910.

[3] Moon, Parry and Spencer, Domina Eberle, *Vectors*, D. Van Nostrand, N. J., 1965.

[4] Weatherburn, *Advanced Vector Analysis*, G. Bell and Sons, London, 1924.

[5] Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Oxford University Press, Oxford, 1873.

[6] Gibbs, J. W., *Elements of Vector Analysis*, privately printed, New Haven (first part in 1881, and second part in 1884), reproduced in *The Scientific Papers of J. Willard Gibbs*, Vol. II, pp. 17 - 50, 50 - 90, Longmans, Green & Co., London, 1906.

[7] Wilson, Edwin Bidwell, *Vector Analysis*, Charles Scribner's Sons, N. Y., 1901.

[8] Heaviside, Oliver, *Electromagnetic Theory*, Vol. I completed in 1893, Vol. II in 1898, Vol. III in 1912. Complete and unabridged edition of the volumes reproduced by Dover Publications in 1950 with a critical and historical introduction by Ernst Weber.

[9] Gans, Richard, *Einfürhrung in die vektor-analysis mit Anwendungen auf die Mathematische Physik*, B. G. Teubner, Leipzig, 1905. English translation of the Sixth Edition by W. M. Deans, Blakie, London, 1932.

[10] Stratton, Julius Adams, *Electromagnetic Theory*, McGraw-Hill, N. Y., 1941.

[11] Jackson, John David, *Classical Electrodynamics*, John Wiley, N. Y., 1962.

[12] Wheeler, Lynde Phelps, *Josiah Willard Gibbs*, Yale University Press, New Haven, 1952.

[13] Tait, P. G., *An Elementary Treatise on Quaternions*, Cambridge University Press, 1890.

[14] Gibbs, J. W., "On the role of quaternions in the algebra of vectors," *Nature*, Vol. XLIII, pp. 511 - 513, 1891; reproduced in *The Scientific Papers of J. Willard Gibbs* [6, pp. 155 - 160]

[15] Wilson, E. B., "Lectures delivered upon vector analysis and its applications to geometry and physics by Prof. J. Willard Gibbs, 1899-90," a report of 289 - plus pages stored at Sterling Memorial Library, Yale University, New Haven.

[16] Tai, C. T., "A Survey of the Improper Uses of ∇ in Vector Analysis," Technical Report RL 909, Radiation Laboratory, Department of Electrical Engineering and Computer Science, the University of Michigan, Ann Arbor, MI 48109, November 1994.

[17] Lagally, Max, *Vorlesungen über Vektor-Rechnung*, Akademische Verlagsgesellschaft, Leipzig, 1928.

[18] Mason, M., and Weaver, W., *The Electromagnetic Field*, The University of Chicago Press, Chicago, 1929.

[19] Schelkunoff, S. A., *Applied Mathematics for Engineers and Scientists*, Van Nostrand, New York, 1948.

[20] Feynman, R. P., Leighton, R. B., and Sands, M., *The Feynman Lectures on Physics*, Vol. II, Addison-Wesley, Reading, 1964.

[21] Borisenko, A. I., and Tarapov, I. E., *Vector and Tensor Analysis*, Russian original 1966; English translation by Richard A. Silverman, Dover Publications, New York, 1979.

[22] Gans, Richard, *Einführung in die Vektor Analysis*, B. G. Teubner, Leipzig, 1905. English translation of the Sixth Edition by W. M. Deans, Blakie, London, 1932; the Seventh German Edition was published in 1950 by the same company.

[23] Shilov, G. E., *Lectures on Vector Analysis* (in Russian), The Government Publishing House of Technical Theoretical Literature, Moscow, 1954; an English translation of pp. 18 - 21 of

Chapter 2, and pp. 85 - 92 of Chapter 8 containing the essential formulas of Shilov's formulation was communicated to this author by Mr. N. H. Fang of the Nanjing Institute of Electronics.

[24] Fang, N. H., *Introduction to Electromagnetic Theory* (in Chinese), Science Press, Beijing, China, 1986.

[25] Tai, C. T., *Generalized Vector and Dyadic Analysis*, IEEE Press, Piscataway, N. J., 1992.

[26] Morse, P. M., and Feshbach, H., *Methods of Theoretical Physics*, Vol. I, McGraw-Hill, N. Y., 1953.

[27] Panofsky, W. K. H., and Phillips, M., *Classical Electricity and Magnetism*, Second Edition, Addison-Wesley, Reading, MA, 1962.

[28] Milne, E. A., *Vectorial Mechanics*, Methuen, London, 1948.

[29] Tai, C. T., and Fang, N. H., "A Systematic Treatment of Vector Analysis," *IEEE Trans. on Education*, Vol. 34, No. 2, pp. 167 - 174, 1991.