

REMARKS ON STANDARD TEXTBOOK DEFINITIONS OF VECTOR OPERATORS

Burak POLAT (*burakpolat@beykent.edu.tr*)

Beykent University, Dept. of Electronics and Communications Engineering, Istanbul

ABSTRACT

The common definitions of gradient, curl and divergence operators in standard textbooks on vector analysis and electromagnetic theory in the space of continuous functions are reviewed, criticized and suggestions are done for their clear and precise comprehension with quotations from literature

Keywords : *Vector calculus, vector operators, electromagnetic fields, electromagnetics education.*

VEKTÖR OPERATÖRLERİN STANDART DERS KİTABI TANIMLAMALARI ÜZERİNE YORUMLAR

Burak POLAT (*burakpolat@beykent.edu.tr*)

Beykent Üniversitesi, Elektronik ve Haberleşme Mühendisliği Bölümü, İstanbul

ÖZET

Bu çalışmada tüm dünyada yaygın şekilde kullanılan temel vektörel analiz ve elektromanyetizma kitaplarında karşımıza çıkan vektör operatörlerin (gradyan, diverjans ve rotasyonel) sürekli fonksiyonlar uzayındaki geleneksel tanımları incelenmiş, bu operatörlerin daha net ve doğru şekilde açıklanmaları amacıyla literatürden atıflarla beraber eleştiriler ve öneriler sunulmuştur

Anahtar Kelimeler: *Vektörel Analiz, Vektör Operatörler, Elektromanyetik Alanlar, Elektromanyetizma Eğitimi.*

1. INTRODUCTION

The importance of clear and precise comprehension of vector operators is of vital importance in classical electromagnetic theory. The proper physical interpretation of field equations rests on nothing but the mathematical descriptions of the gradient, curl and divergence operators in particular. It is for this reason that most introductory textbooks on electromagnetic theory start with a chapter or has an appendix that reviews vector calculus with emphasis on these vector operators. It can also be observed that the definitions, descriptions and style are generally alike in such review chapters, where the main concern is to provide a physical picture of a certain field equation in the first place by using most suitable (differential or integral) representations of the vector operators that serve to the purpose. However, such an introduction should be criticized since lack of proper and exact descriptions of the vector operators may yield confusion and misinterpretation. With regard to the importance of the concepts of vector calculus, in this paper I provide examples for the common definitions of vector (gradient, curl and divergence) operators in C^1 , the space of continuous functions with continuous derivatives, raise critics and suggest approaches with quotations from literature in search of the most scientific introduction to the topic.

2. The Gradient Operator

In this section we shall first focus on the differential form definitions of the gradient operator. There are two alternative representations to start which always yield another. The first approach is to introduce the gradient operator directly in Cartesian coordinates and generalize it by a directional derivative representation supported with physical interpretations involving the rate of change in altitude when walking up an incline or temperature distribution in a building. Among many such books written since mid 20th century some well known ones include [1, Sect.4.4], [2, Sect.33], [3, Sect.1.19], [4, Sec.1.8], [5, Sect.AII], [6, Sect.1.03], [7, Sect.2.3], [8, Sect.A.2], [9, Sect.1.7], [10, Sect.1.1], [11,

Sect.1.3], [12, Sect.1.22], [13, Sect.5.1], [14, Sect.V.14], [15, Sect.2.2], [16, pg.37,38], [17, Ch.1], [18, Sect.3.4]. And the second is vice versa (cf. , [19, Sect.2], [20, Sect.1.2], [21, Sect.2.5], [22, Sect.2.13], [23, Sect.3.2.3], [24, Sect.3.7], [25, Sect.4.4], [26, Sect.4.6]). Let us quote from [16, pg.37,38] with equation numbers modified, which utilizes the first approach:

“...Let us consider the change in the value of an arbitrary scalar field f as we move from (x, y, z) to $(x + dx, y + dy, z + dz)$. We will denote this change as df . From ordinary multivariable calculus,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \quad (1)$$

This expression can be written as the following dot product between two vectors:

$$df = \left(\frac{\partial f}{\partial x} \hat{\mathbf{a}}_x + \frac{\partial f}{\partial y} \hat{\mathbf{a}}_y + \frac{\partial f}{\partial z} \hat{\mathbf{a}}_z \right) \cdot (dx \hat{\mathbf{a}}_x + dy \hat{\mathbf{a}}_y + dz \hat{\mathbf{a}}_z) \quad (2)$$

The vector on the far right is simply the differential displacement vector $d\mathbf{l}$ along the path of movement, so we can write Equation (2) in the form

$$df = \left(\frac{\partial f}{\partial x} \hat{\mathbf{a}}_x + \frac{\partial f}{\partial y} \hat{\mathbf{a}}_y + \frac{\partial f}{\partial z} \hat{\mathbf{a}}_z \right) \cdot d\mathbf{l} \quad (3)$$

The vector quantity in paranthesis is called the **gradient** of f and is denoted symbolically by $\mathbf{grad}f$. Hence

$$df = \mathbf{grad}f \cdot d\mathbf{l} \quad (4)$$

where, in the Cartesian coordinate system,

$$\mathbf{grad}f = \frac{\partial f}{\partial x} \hat{\mathbf{a}}_x + \frac{\partial f}{\partial y} \hat{\mathbf{a}}_y + \frac{\partial f}{\partial z} \hat{\mathbf{a}}_z \quad (5)$$

“... We can write Equation (4) in the form

$$df = \mathbf{grad}f \cdot d\mathbf{l} = |\mathbf{grad}f| dl \cos \theta \quad (6)$$

where θ is the angle between $\mathbf{grad}f$ and $d\mathbf{l}$. Dividing both sides by dl yields

$$\frac{df}{dl} = |\mathbf{grad}f| \cos \theta \quad (7)$$

When the direction of the path is parallel to $\mathbf{grad}f$, $\cos \theta = 1$. Along such a path, df/dl attains its maximum value. Thus,

$$\left. \frac{df}{dl} \right|_{\max} = |\mathbf{grad}f| \quad (8)$$

Using Equation (8), we can define $\mathbf{grad}f$ as

$$\mathbf{grad}f = \left. \frac{\partial f}{\partial l} \right|_{\max} \hat{\mathbf{a}}_n \quad (9)$$

where $\hat{\mathbf{a}}_n$ points in the direction of maximum increase in f . This definition is valid in all coordinate systems. Thus, the gradient $\mathbf{grad}f$ is a vector that points in the direction of maximum increase of the function f . “

Comments: The main concern here is to provide the simplest, i.e. Cartesian representation of the gradient operator -among others- directly for a smooth introduction. Technically it should not be treated as a *definition* since a *general* definition of a mathematical operator should always be coordinate-free. It should be realized that the more general (coordinate-free) definition (9) is derived from another definition (5). Of course, it is a minor issue since it could also be introduced in the first place which would yield (5) as a special case (cf.[18,Sect.3.6], [21, Sect.2.5], [27, Sect.2.5]). By the way, the change of notation from (8) to (9) in differentiation from d to ∂ operator is unnecessary and unjustified. However, what primarily concerns us is

the *ambiguity* of the definition (9). A proper mathematical definition must also serve as a description while the *definition* (9) involves no

information for determining neither $\left. \frac{\partial f}{\partial l} \right|_{\max}$ nor $\hat{\mathbf{a}}_n$. It certainly makes

sense for geometrical and physical interpretations of the gradient operator associated with level (equivalue) surfaces or field lines of a scalar field but should not be considered as a definition in that regard. To summarize so far, in author's view, neither of the differential form definitions (5) or (9) serves as a proper definition of the gradient operator for different reasons.

The integral form representation of the gradient operator is also introduced – however only in textbooks on vector calculus- and is always derived as an application of divergence theorem for a vector field as $\vec{A} = \vec{C}f$, where \vec{C} is fixed but arbitrary vector, assuming that the gradient and divergence operators, divergence and mean value theorems are already introduced (cf.[28, Ch.6,pr.24], [29, Sect.4.3.4]).

“At any point P

$$\mathbf{grad}f = \lim_{\Delta\mathcal{G} \rightarrow 0} \frac{1}{\Delta\mathcal{G}} \oiint_{\Delta\Sigma} f \hat{n} d\Sigma \tag{10}$$

where $\Delta\mathcal{G}$ is an arbitrary volume enclosed by the surface $\Delta\Sigma$, \hat{n} is the unit outward normal of $\Delta\Sigma$ and the limit is obtained by shrinking $\Delta\mathcal{G}$ to the point P .” (modified from [Spiegel, Ch.6,pr.24])

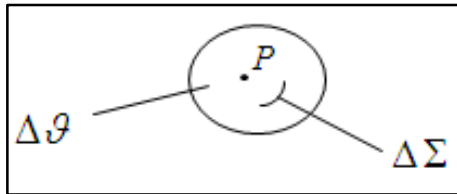


Figure 1. An infinitely small region $\Delta\mathcal{G}$ with enclosure $\Delta\Sigma$

Comments: The integral definition of gradient operator is certainly its unique proper mathematical description since it is coordinate-free and

also yields the differential form representations in any curvilinear coordinate system provided that the arbitrary volume of integration $\Delta\mathcal{G}$ is chosen properly. It also does not need to introduce the concept of differential derivative, gradient or divergence operators, divergence theorem or mean value theorems in the first place. Then the relation (4) can be seen to be satisfied immediately which also yields (9) suitable for a physical interpretation. A careful eye can realize a minor inconsistency between the two sides of (10). The left hand side is a *vector field*, while the right hand side is not a function of space variables and only yields numerical values for a given observation point. For any algebraic operation on the integral definition of the gradient operator it needs to be modified by introducing the characteristic function of region $\Delta\mathcal{G}$,

$$H[\Delta\mathcal{G}] = \begin{cases} 1, \vec{r} \in \Delta\mathcal{G} \\ 0, \vec{r} \notin \Delta\mathcal{G} \end{cases} \quad (11)$$

into (10) as

$$\mathbf{grad}f(\vec{r}) = \lim_{\Delta\mathcal{G} \rightarrow 0} \frac{H[\Delta\mathcal{G}]}{\Delta\mathcal{G}} \oint_{\Delta\Sigma} f(\vec{r}) \hat{n} d\Sigma \quad (12)$$

To sum up, for a proper mathematical definition of the gradient operator, (12) should be the starting point. A brief introduction to generalized curvilinear coordinate systems should have been given earlier so that the differential representation of the gradient operator in orthogonal curvilinear coordinate system (u_1, u_2, u_3) with metric coefficients (h_1, h_2, h_3) can be achieved by choosing the volume element as $\Delta\mathcal{G} = h_1\Delta u_1 h_2\Delta u_2 h_3\Delta u_3$ and taking the limit as $\Delta u_1, \Delta u_2, \Delta u_3$ approach zero to get

$$\mathbf{grad}f = \frac{\partial f}{h_1 \partial u_1} \hat{\mathbf{a}}_{u_1} + \frac{\partial f}{h_2 \partial u_2} \hat{\mathbf{a}}_{u_2} + \frac{\partial f}{h_3 \partial u_3} \hat{\mathbf{a}}_{u_3} \quad (13)$$

Then, based on the properties $df = \frac{\partial f}{\partial u_1} du_1 + \frac{\partial f}{\partial u_2} du_2 + \frac{\partial f}{\partial u_3} du_3$,

$d\mathbf{l} = h_1 du_1 \hat{\mathbf{a}}_{u_1} + h_2 du_2 \hat{\mathbf{a}}_{u_2} + h_3 du_3 \hat{\mathbf{a}}_{u_3}$ it can be seen that the relation

(4) is satisfied in any orthogonal coordinate system, from which one can also reach to (9) suitable for physical interpretations related to level (equivalue) surfaces or field lines.

3. The Curl Operator

The curl of a vector field is generally introduced by a line integral as a measure of its “circulation” around a closed path that encircles the point of interest and shrinks onto it in the limit. It is generally introduced either in determinant form in Cartesian coordinates (cf. [1, Sect.4.6], [4, Sect.1.11], [8, Sect.A.2], [9, Sect.1.8], [12, Sect.1.25], [17, Ch.1], [23, Sect.3.2.5], [25, Sect.4.1], [30, Sect.7]) or via its component in an arbitrary direction which yields an entire representation (cf. [16, pg.46], [25, Sect.4.3], [31, Sect.1.12], [21, Sect.2.7], [24, Sect.3.8.1], [14, Sect.V.18], [6, Sect.1.03], [7, Sect.2.15], [5, Sect.AII], [22, Sect.5.18], [26, Sect.8.4], [10, Sect.1.6], [11, Sect.1.5], [32, Appx.C], [15, Sect.2.4]) or vice versa (cf. [27, Sect.2.8]).

“...We define the component of $\mathbf{curl} \mathbf{A}$ in the direction $\hat{\mathbf{a}}_i$ by

$$(\mathbf{curl} \mathbf{A})_i = \hat{\mathbf{a}}_i \cdot \mathbf{curl} \mathbf{A} \equiv \lim_{\Delta S_i \rightarrow 0} \frac{1}{\Delta S_i} \oint_{C_i} \mathbf{A} \cdot d\mathbf{l} \quad (14)$$

where ΔS_i is a small surface that is bounded by the contour (i.e., path) C_i and has unit normal $\hat{\mathbf{a}}_i$. The direction of C_i is governed by the right-hand rule, which says that when the right-hand thumb is placed along the path, the remaining fingers “poke” through the surface ΔS_i in the direction of $\hat{\mathbf{a}}_i$.

Since $\mathbf{curl} \mathbf{A}$ is a vector, we can represent it by its magnitude and direction, which we will denote as $|\mathbf{curl} \mathbf{A}|$ and $\hat{\mathbf{a}}_n$, respectively. To

find $|\mathbf{curl} \mathbf{A}|$, we notice from equation (14) that the values of the components of $|\mathbf{curl} \mathbf{A}|$ vary with the orientations of the integration path C_i . Since the maximum value that any component of a vector can attain equals the vector's magnitude, we can conclude that

$$|\mathbf{curl} \mathbf{A}| \equiv \left[\lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \oint_C \mathbf{A} \cdot d\mathbf{l} \right]_{\max} \quad (15)$$

where C is the differential path that maximizes the circulation integral. Thus we can write $\mathbf{curl} \mathbf{A}$ as

$$\mathbf{curl} \mathbf{A} \equiv \hat{\mathbf{a}}_n \left[\lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \oint_C \mathbf{A} \cdot d\mathbf{l} \right]_{\max} \quad (16)$$

where $\hat{\mathbf{a}}_n$ is perpendicular to the surface bounded by C and points in the direction determined by the right-hand rule. (modified from [16, pg.46])

Comments: Regarding the “definition” (14) it should be questioned how meaningful it is to introduce a component of $\mathbf{curl} \mathbf{A}$ rather than its entire representation (16) in the first place. In electromagnetics textbooks it is preferred for a quick introduction since it is easily deduced from Stokes theorem and helps visualize surface magnetic currents (cf.[33, Ch.1, Sect.19]) or the rotation of a paddle wheel in a fluid with a simple device called curl meter (cf. [3, Sect.1.12], [7, Sect.2.18], [24, Sect.3.9.2], [25, Sect.4.3]). This is explained in [32, Appx.C] as follows: “...*Since the curl of \mathbf{A} is a vector, a prescription for calculating it at a point implies a prescription for calculating its components along an arbitrary direction*”. However, starting with (14) does not make much sense unless the analytical relations between $(\hat{\mathbf{a}}_i, \Delta s_i, C_i)$ and $(\hat{\mathbf{a}}_n, \Delta s, C)$ are provided. Besides, (16) is an *ambiguous* definition (as was (9)) since it involves no information for determining neither $\hat{\mathbf{a}}_n$ or

$$\left[\lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \oint_C \mathbf{A} \cdot d\mathbf{l} \right]_{\max}$$

For a precise definition of the curl operator one needs to introduce its integral definition. This is available in textbooks on rather vector calculus than electromagnetic theory. Similar to the case with the gradient operator, it is derived as an application of divergence theorem assuming that the divergence operator, divergence and mean value theorems are already introduced (cf. [28, Ch.6,pr.24]).

“At any point P

$$\mathbf{curl} \mathbf{A} = \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{1}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} \hat{\mathbf{n}} \times \mathbf{A} d\Sigma \quad (17)$$

where $\Delta \mathcal{G}$ is an arbitrary volume enclosed by the surface $\Delta \Sigma$, $\hat{\mathbf{n}}$ is the unit outward normal of $\Delta \Sigma$ and the limit is obtained by shrinking $\Delta \mathcal{G}$ to the point P .” (see Fig.1)

Comments: For algebraic purposes this needs to be improved as

$$\mathbf{curl} \mathbf{A} = \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{H[\Delta \mathcal{G}]}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} \hat{\mathbf{n}} \times \mathbf{A} d\Sigma \quad (18)$$

When the definition (18) is taken as the starting point, the property (14) can be obtained as an application in an infinitely small cylinder pointed in the direction $\hat{\mathbf{a}}_i$ with cross area Δs_i (cf.[29, Sect.4.4.5], [2, Sect.3.6]), which also yields (16). Similarly, by choosing the volume element as $\Delta \mathcal{G} = h_1 \Delta u_1 h_2 \Delta u_2 h_3 \Delta u_3$, the differential representation of the curl operator

$$\mathbf{curl} \mathbf{A} = \begin{vmatrix} h_1 \hat{\mathbf{u}}_1 & h_2 \hat{\mathbf{u}}_2 & h_3 \hat{\mathbf{u}}_3 \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix} \quad (19)$$

can also be obtained simultaneously.

4. The Divergence Operator

The divergence operator is mostly defined in integral form

$$\mathbf{div} \mathbf{A} = \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{1}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} \mathbf{A} \cdot \hat{\mathbf{n}} d\Sigma \quad (20)$$

in textbooks either directly or by deriving from the divergence theorem. The reason that it is generally introduced in integral form in the first place is that the divergence operator is always associated with the laws of conservation of electric and magnetic flux and the representation (20) is most suitable for physical interpretation. The integral form definition (20) also yields the differential form representation

$$\mathbf{div} \mathbf{A} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u_1} (A_1 h_2 h_3) + \frac{\partial}{\partial u_2} (A_2 h_3 h_1) + \frac{\partial}{\partial u_3} (A_3 h_1 h_2) \right] \quad (21)$$

The definition (20) is compatible with (12) and (18) except that it needs to be modified as

$$\mathbf{div} \mathbf{A} = \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{H[\Delta \mathcal{G}]}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} \mathbf{A} \cdot \hat{\mathbf{n}} d\Sigma \quad (22)$$

for algebraic purposes. An exception is observed in [8, Sect.A.2] where the divergence operator is introduced in differential form in Cartesian coordinates from which the integral form (20) is derived.

5. On Nabla Operator

The introduction of nabla (aka del) operator is prone to critical misinterpretation in vector calculus. In almost every textbook the differential form representations of gradient, divergence and curl operators in Cartesian coordinates are introduced as ordinary multiplication, scalar and vector products of the nabla operator with the associated scalar/vector function, respectively (cf.[28, Ch.4]). First let us quote from [13, Sect.5.1]:

“The operator $\vec{\nabla}$ is a symbolic vector with components $\frac{\partial}{\partial x_1}$, $\frac{\partial}{\partial x_2}$,

$\frac{\partial}{\partial x_3}$ It may be applied to a scalar function $F(x_1, x_2, x_3)$ or to a vector

function $\vec{y}(x_1, x_2, x_3)$ with components $y_i(x_1, x_2, x_3)$, $i = 1, 2, 3$. In the latter case, we have either the scalar product $\vec{\nabla} \cdot \vec{y}$ or the vector product $\vec{\nabla} \times \vec{y}$. The usefulness of the symbol lies in the fact that it makes many physical formulas more compact”

Next, let us quote from [12, Sect.1.2.3] on the description and use of nabla operator:

“The gradient has the formal appearance of a vector, ∇ , “multiplying” as scalar T :

$$\nabla T = \left(\hat{\mathbf{a}}_x \frac{\partial}{\partial x} + \hat{\mathbf{a}}_y \frac{\partial}{\partial y} + \hat{\mathbf{a}}_z \frac{\partial}{\partial z} \right) T \quad (23)$$

(I write the unit vectors to the left, just so no one will think this means $\frac{\partial \hat{\mathbf{a}}_x}{\partial x}$, and so on – which would be zero, since $\hat{\mathbf{a}}_x$ is constant). The term in paranthesis is called ‘del’:

$$\nabla = \hat{\mathbf{a}}_x \frac{\partial}{\partial x} + \hat{\mathbf{a}}_y \frac{\partial}{\partial y} + \hat{\mathbf{a}}_z \frac{\partial}{\partial z} \quad (24)$$

Of course, del is not a vector, in the usual sense. Indeed, it is without specific meaning until we provide it with a function to act upon. Furthermore, it does not “multiply” T ; rather, it is an instruction to differentiate that function. To be precise, then, we would say that ∇ is a vector operator that acts upon T , not a vector that multiplies T .

With this qualification, though, ∇ mimics the behavior of an ordinary vector in virtually every way; almost anything that can be done with

other vectors can also be done with ∇ , if we merely translate “multiply” by “act upon”. So by all means take the vector appearance of ∇ seriously: it is a marvellous piece of notational simplification, as you will appreciate if you ever consult Maxwell’s original work on electromagnetism, written without the benefit of ∇

Now an ordinary vector \mathbf{A} can multiply in three ways:

Multiply a scalar a : $\mathbf{A}a$

Multiply another vector \mathbf{B} , via the dot product: $\mathbf{A} \cdot \mathbf{B}$

Multiply another vector via the cross product: $\mathbf{A} \times \mathbf{B}$

Correspondingly, there are three ways the operator ∇ can act:

On a scalar function T : ∇T (the gradient)

On a vector function \mathbf{v} , via the dot product: $\nabla \cdot \mathbf{v}$ (the “divergence”)

On a vector function, via the cross product: $\nabla \times \mathbf{v}$ (the “curl”)

Comments: The history of nabla operator and its use in vector calculus have been investigated in detail in the works of Chen-To Tai. Regarding the use of scalar and cross product operations between the nabla operator and a vector function, the notation was introduced by Gibbs [34] while a scalar or vector product operation was never pronounced. It has been Wilson [35, Sect.70] who misinterpreted the notation as formal scalar and vector product operations between nabla operator and a vector function. Let us quote from [36]:

“...Unfortunately, the ‘scalar product’ model is not a valid method at all. The interpretation was forced on the notation for the divergence introduced by Gibbs, namely $\nabla \cdot \mathbf{f}$, who also introduced the notation for the curl as $\nabla \times \mathbf{f}$. The fact that both the ‘scalar product’ and the ‘vector product’ of ∇ and \mathbf{f} do not exist can be illustrated by a simple arithmetic analogy. For example, an assembly of numbers and signs, in the form of $2 + \times 3$, has no meaning in arithmetic. But if we move the plus sign to the front, we create a well-defined number, $+6$, and if we

move the plus sign to the back, we create a numerical operator, $\nabla \cdot$. Neither of these two expressions is equivalent to the original assembly.

Now if one considers Gibbs' notation for the divergence in a Cartesian system as the 'scalar product' between ∇ and \mathbf{f} , then

$$\nabla \cdot \mathbf{f} = \left(\sum_i \hat{\mathbf{a}}_i \frac{\partial}{\partial x_i} \right) \cdot \left(\sum_j f_j \hat{\mathbf{a}}_j \right) \quad (25)$$

The right member of (25) is meaningless, because it consists of an assembly of functions and symbols. Let us assume for the time being that the distributive rule is applicable to the two groups in (25). Then, one member of the assembly has the form

$$\hat{\mathbf{a}}_1 \frac{\partial}{\partial x_1} \cdot f_1 \hat{\mathbf{a}}_1 \quad (26)$$

Analogous to the arithmetical example, (26) is also an assembly. It is not a meaningful expression. We cannot arbitrarily move the dot sign to the front of the differential sign to create an expression of our liking, viz.

$$\hat{\mathbf{a}}_1 \cdot \frac{\partial}{\partial x_1} f_1 \hat{\mathbf{a}}_1 = \frac{\partial f_1}{\partial x_1} \quad (27)$$

nor can we move the front unit vector behind the differential sign and put two brackets around the remaining functions to create the same partial derivative as in (27), viz.,

$$\frac{\partial}{\partial x_1} (\hat{\mathbf{a}}_1 \cdot f_1 \hat{\mathbf{a}}_1) = \frac{\partial f_1}{\partial x_1} \quad (28)$$

Neither (27) nor (28) is equivalent to the original assembly, (25). This is not a matter of interpretation; it is a manipulation which is not allowed in mathematics."

The second reason why the scalar/vector product interpretation fails is given in [36] as follows:

“...The fact that the expression so arbitrarily created from (25) does not represent the correct expression for the divergence in a Cartesian system has fooled many people about the true nature of the ‘scalar product’ model. When the same model is applied to a curvilinear coordinate system, people found that it does not work [37, Appx.C, p.325], but they never questioned the meaning of the model itself. In other words, they did not realize that they were dealing with an assembly to ‘derive’ the differential expression of the divergence in the Cartesian system. The amazing story is that mathematicians, physicists, and engineers who used Gibbs’ notation have practiced this manipulation for generations, and it has reached every corner of the world”.

6. Concluding Remarks

In this review paper I focused on various representations of the vector operators in standard undergraduate textbooks on electromagnetics and vector calculus, raised critics with quotations from literature and suggested approaches for their clear and precise comprehension. To conclude, the critics raised in the present paper can be summarized as follows:

1. With so many misconceptions and misinterpretations embedded in literature for over a century regarding the nabla operator, its introduction in basic textbooks should be handled very carefully. With this in mind, the author has preferred to avoid the nabla operator and always adopted the old notation (‘grad’, ‘div’, ‘curl’) throughout his career and scientific works following the teachings of Prof. Mithat İdemem with the concern that it makes no sense mathematically except for being practical for memorizing the differential Cartesian representations of these operators.
2. Regarding the most proper introduction to vector operators, the integral form uniform and coordinate-free definitions (12), (18), (22) should be introduced in the first place in defining the vector operators simultaneously without any reference to their physical interpretations (cf.[2, Sect.3.6], [38]). An introduction to general

curvilinear coordinate systems (u_1, u_2, u_3) should have been given earlier. As a second stage the integrations (12), (18) and (21) should be performed in a volume $\Delta \mathcal{G} = h_1 \Delta u_1 h_2 \Delta u_2 h_3 \Delta u_3$ so that the differential forms of the vector operators (13), (19) and (21) are also obtained simultaneously. It should be noted that one can also start with (13), (19) and (21) and integrate over $\Delta \mathcal{G} = h_1 \Delta u_1 h_2 \Delta u_2 h_3 \Delta u_3$. However, in that case one only obtains (12), (18) and (22) as expressed in (u_1, u_2, u_3) system, not the general coordinate-free definitions over an *arbitrary* domain. This view is also supported in the Introduction section of [39] as follows:

“Vector analysis is a classical subject in applied mathematics and is an important tool in various branches of physical science and engineering. In many early books on vector analysis [40], [41], the differential forms of three key functions, namely the divergence, curl, and gradient, are usually introduced in the Cartesian coordinate system with aid of the del operator (some authors call it Nabla operator or Hamilton operator). Although the differential expressions of these functions in other orthogonal coordinate systems can be found by a transformation of coordinate variables, it is simpler to start with the definition of these functions in a form independent of the choice of the coordinate system and then derive their differential expressions in a general curvilinear orthogonal system. Several different approaches are available to execute the derivation. They are found in [30], [38], [42-44].”

To overcome the deficiencies of the nabla operator Chen-To Tai has introduced a new symbolic (functional) S-operator and its associated operators in [39] (see also [45]) which do not have explicit differential forms and satisfy commutative property. With these operators he has been able to present a systematic treatment of the entire subject of vector analysis, while it can also be extended to dyadic analysis.

Moreover, the integral form definitions (12), (18) and (22) can also be extended from the space of continuous functions to the space of Schwartz-Sobolev distributions directly when point, space curve or

surface type singularities are involved, which have also been subjects of previous investigations by the author (see [46-48]).

- Regarding the physical interpretation of the gradient operator, (13) directly requires the identity (4), from which one finally obtains (9), which is most suitable associated with the level (equivalency) surfaces or field lines of a scalar field.

$$\mathbf{grad} f \triangleq \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{H[\Delta \mathcal{G}]}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} f \hat{n} d\Sigma \Rightarrow \mathbf{grad} f = \sum_{i=1}^3 \frac{\partial f}{h_i \partial u_i} \hat{\mathbf{a}}_{u_i}$$

$$\Rightarrow df = \mathbf{grad} f \cdot d\mathbf{l} \Rightarrow \mathbf{grad} f = \left. \frac{\partial f}{\partial l} \right|_{\max} \hat{\mathbf{a}}_n$$

- Regarding the physical interpretation of the curl operator on a surface, the relations (19), (14) and eventually (16) can be obtained respectively from the definition (18) in a straightforward manner as shown in [2, Sect.3.6] or [29, Sect.4.4.5].

$$\mathbf{curl} \mathbf{A} \triangleq \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{H[\Delta \mathcal{G}]}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} \hat{n} \times \mathbf{A} d\Sigma \Rightarrow$$

$$\left\{ \begin{array}{l} \mathbf{curl} \mathbf{A} = \begin{vmatrix} h_1 \hat{\mathbf{u}}_1 & h_2 \hat{\mathbf{u}}_2 & h_3 \hat{\mathbf{u}}_3 \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix}, \text{ and simultaneously} \\ \hat{\mathbf{a}}_i \cdot \mathbf{curl} \mathbf{A} = \lim_{\Delta s_i \rightarrow 0} \frac{1}{\Delta s_i} \oint_{C_i} \mathbf{A} \cdot d\mathbf{l} \Rightarrow \mathbf{curl} \mathbf{A} = \hat{\mathbf{a}}_n \left[\lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \oint_C \mathbf{A} \cdot d\mathbf{l} \right]_{\max} \end{array} \right.$$

- The integral form definition (22) of the divergence operator is already its most suitable form for a physical interpretation associated with conservation laws. Integration over a differential volume element for an orthogonal curvilinear coordinate system directly yields (21).

$$\mathbf{div} \mathbf{A} \triangleq \lim_{\Delta \mathcal{G} \rightarrow 0} \frac{H[\Delta \mathcal{G}]}{\Delta \mathcal{G}} \oiint_{\Delta \Sigma} \mathbf{A} \cdot \hat{n} d\Sigma \Rightarrow \mathbf{div} \mathbf{A} = \frac{1}{\Omega} \sum_{i=1}^3 \frac{\partial}{\partial u_i} \left(\frac{\Omega A_i}{h_i} \right)$$

$$\Omega = h_1 h_2 h_3$$

6. Regarding a systematic treatment of vector analysis, differential forms, dyadics, bivectors, tensors, quaternions, and Clifford algebras stand out as advanced analytical tools. Among them the calculus of differential forms, developed by H.G. Grassmann [49] and E. Cartan [50], establishes a direct connection to geometrical images, which provides an ability to visualize vector operators. They have been applied to electromagnetic theory among other branches of physical sciences since [51]. Their advantages over traditional methods are outlined in [52] as follows: *“First, forms clarify the relationship between field intensity and flux density, by providing distinct mathematical and graphical representations for the two types of fields. Second, Ampere’s and Faraday’s laws obtain graphical representations that are as intuitive as the representation of Gauss’ law. Third, the vector Stokes theorem and the divergence theorem become special cases of a single relationship that is easier for the student to remember, apply, and visualize than their vector formulations. Fourth, computational simplifications result from the use of forms. Derivatives are easier to employ in curvilinear coordinates, integration becomes more straightforward, and families of vector identities are replaced by algebraic rules”*. For an introduction to differential forms with applications in electromagnetic theory one may refer to the recent survey paper [53] and the references cited therein.
7. Finally, a history of notations in Vector Analysis can be reached at [54, Vol.II, pp.495-501]. On the other hand, for a review of the contributions of Gibbs and Heaviside on the development of the modern system of Vector Analysis the reader may refer to [55, Ch.5].

REFERENCES

- [1] G.E. Hay, Vector and Tensor Analysis, Dover Pub. 1953.
- [2] A.P. Wills, Vector Analysis with an Introduction to Tensor Analysis, Dover Pub. 1958.
- [3] R.E. Collin, Principles and Applications of Electromagnetic Fields,

- McGraw-Hill Book Company, Inc. 1961.
- [4] N. Tralli, *Classical Electromagnetic Theory*, McGraw-Hill Book Company, Inc. 1963.
- [5] R. Becker, *Electromagnetic Fields and Interactions*, Dover Publ. 1964.
- [6] E.C. Jordan, *Electromagnetic Waves and Radiating Systems*, Prentice Hall, Inc. 1964.
- [7] E.M. Purcell, *Electricity and Magnetism*, Berkeley Physics Course. Vol.2, McGraw-Hill Book Company, Inc. 1965.
- [8] W.T. Scott, *The Physics of Electricity and Magnetism*, John Wiley & Sons, Inc. 1966.
- [9] E.V. Bohn, *Introduction to Electromagnetic Fields and Waves*, Addison-Wesley Pub. 1968.
- [10] A. Shadowitz, *The Electromagnetic Field*, Dover Publ. 1975.
- [11] M. Schwartz, *Principles of Electrodynamics*, Dover Publ. 1987.
- [12] D.J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, Inc. 1989.
- [13] D.V. Widder, *Advanced Calculus*, 2nd Ed. Dover Pub. 1989.
- [14] R.S. Elliot, *Electromagnetics. History, Theory, and Applications*, IEEE Press, 1993.
- [15] R.E. DuBroff, S.V. Marshall and G.G. Skitek, *Electromagnetic Concepts and Applications*, 4th Ed. Prentice Hall, Inc. 1996.
- [16] K.R. Demarest, *Engineering Electromagnetics*, Prentice Hall, Inc. 1998.
- [17] B. Di Bartolo, *Classical Theory of Electromagnetism*, World Scientific Publ. 2004
- [18] F.T. Ulaby, *Fundamentals of Applied Electromagnetics*, 6th Ed. Prentice Hall, Inc. 2010.
- [19] C.E. Weatherburn, *Advanced Vector Analysis*, Bell's Mathematical Series, 1949.

- [20] S.A. Schelkunoff, *Electromagnetic Waves*, D. Van Nostrand Company, Inc. 1956.
- [21] R.M. Fano, L.J.Chu and R.B. Adler, *Electromagnetic Fields, Energy, and Forces*, John Wiley & Sons, Inc. 1960.
- [22] J.D. Kraus, *Electromagnetics*, McGraw-Hill, Inc. 1968.
- [23] P. Hammond, *Applied Electromagnetism*, Pergamon Press, 1971.
- [24] L.M. Magid, *Electromagnetic Fields, Energy and Waves*, John Wiley and Sons, Inc., 1972.
- [25] N.N. Rao, *Elements of Engineering Electromagnetics*, Prentice Hall, Inc. 1991.
- [26] W.H. Hayt, *Engineering Electromagnetics*, McGraw-Hill, Inc. 1989.
- [27] D.K. Cheng, *Field and Wave Electromagnetics*, Addison-Wesley Publ. Company, Inc. 1983
- [28] M.R. Spiegel, *Theory and Problems of Vector Analysis*, Schaum's Outline Series, McGraw-Hill Book Company, Inc. 1974.
- [29] A.I. Borisenko and I.E. Tarapov, *Vector and Tensor Analysis with Applications*, (Translated and edited by R.A. Silverman), Prentice Hall Inc., 1968.
- [30] C.E. Weatherburn, *Differential Geometry of Three Dimensions*, Cambridge University Press, 1927.
- [31] R.W.P. King, *The Theory of Electromagnetism*, Pergamon Press, 1964.
- [32] L. Eyges, *The Classical Electromagnetic Field*, Dover Publ. 1980.
- [33] R.W.P. King, *Electromagnetic Engineering*, McGraw-Hill Book Company, Inc. 1945.
- [34] J.W. Gibbs, *The Scientific Papers of J. Willard Gibbs*, Vol. II, Ox Bow Pr Publ. 1994.
- [35] E.B. Wilson, *Vector Analysis*, N.Y., Charles Scribner's Sons, 1901.
- [36] C.T. Tai, "Another Matter of History", *IEEE Antennas and Propagation Magazine*, 21-26, Feb. 1991

- [37] P. Moon and D.E. Spencer, *Vectors*, Princeton, N.J., D. Van Nostrand, 1965.
- [38] C.T. Tai, "Unified Definition of Divergence, Curl, and Gradient", *Appl. Math. Mech.* Vol.7, No.1, 1986.
- [39] C.T. Tai, "A Systematic Treatment of Vector Analysis", *IEEE Trans. On Education*, Vol.34, No.2 167-174, 1991.
- [40] J.W. Gibbs, *Vector Analysis*, New Haven, CT: Yale University Press, 1948.
- [41] H.B. Philips, *Vector Analysis*, New York: Wiley, 1933.
- [42] H.E. Newell, *Vector Analysis*, New York McGraw Hill, 1955
- [43] M. Javid and P.M. Brown, *Field Analysis and Electromagnetics*, New York: McGraw Hill, 1963.
- [44] N.H. Fang and T.Y. Zhu, "Unified Representation of Vector Functions and Dyadic Functions in Field Theory in Orthogonal Systems", *Acta Antennae Sinica*, Vol.3, No.1, 1987 (in Chinese)
- [45] C.T. Tai, *Generalized Vector and Dyadic Analysis*, IEEE Press 1992.
- [46] B. Polat, "A Distributional Investigation of Point Sources in Classical Electromagnetism", *Revue Roumaine des Sciences Techniques - Série Électrotechnique et Énergétique*, Vol.57 No.2, 123-133 (2012)
- [47] B. Polat, "Space Curve Types of Distributions and Sources", *IEEE Antennas and Propagation Magazine*, Vol.54 No.3, 68-85, (2012).
- [48] B. Polat, "Distributional Derivatives on a Regular Open Surface with Physical Applications", *TWMS Journal of Applied and Engineering Mathematics* Vol.1 No.2 pp.203-222 (2011).
- [49] H. Grassmann and L. Kannenberg, *A New Branch of Mathematics: The "Ausdehnungslehre" of 1844 and Other Works*, Open Court Publishing, Chicago, 1995.
- [50] Cartan, E., *Les Systemes Différentielles Extérieurs*, Hermann, Paris, 1945.
- [51] H. Flanders, *Differential Forms with Applications to the Physical Sciences*, Dover, New York, 1963.

- [52] Warnick, K.F., R.H. Selfridge, and D.V. Arnold, “Teaching Electromagnetic Field Theory Using Differential Forms”, IEEE Trans. Educ., Vol. 40, No. 1, 53{68, 1997.
- [53] K.F. Warnick and P. Russer, “Differential Forms and Electromagnetic Field Theory”, Progress in Electromagnetics Research, Vol.148, 83-112, 2014.
- [54] F. Cajori, A History of Mathematical Notations, Dover Publ. 1993.
- [55] M.J. Crowe, A History of Vector Analysis, Dover Publ. 1994.