Another Matter of History

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

1 A Question and Some Quotations

My last communication to the AP Magazine dealt with the history of the AP Transactions. This time the subject is a more serious matter in the history of vector analysis.

Recently, I asked the following question to a new graduate student at The University of Michigan: "Can you tell me how you learned the derivation of the differential expression for the divergence of a vector function in the Cartesian system, namely,

\[ \text{div } \mathbf{f} = \nabla \cdot \mathbf{f} = \sum_i \frac{\partial f_i}{\partial x_i} \]  

(1)

where \( x_i \) denote the coordinate variables, and \( f_i \) the components of \( \mathbf{f} \) in that system with \( i = (1, 2, 3) \)?" His answer was that a convenient method of deriving that expression is to take the scalar product between \( \nabla \) and \( \mathbf{f} \) where the del operator, \( \nabla \), is defined by

\[ \nabla = \sum_i a_i \frac{\partial}{\partial x_i} \]  

(2)

and \( a_i \) denote the unit vectors in the Cartesian system, then,

\[ \nabla \cdot \mathbf{f} = \left( \sum_i a_i \frac{\partial}{\partial x_i} \right) \cdot \left( \sum_j f_j a_j \right) = \sum_i \frac{\partial f_i}{\partial x_i} \]  

(3)

I asked the same question to an older graduate student. His answer was practically the same, but he assured me of the validity of that approach by quoting the names of several popular books in applied mathematics and electromagnetics. He then added that the differential expression can be obtained by evaluating the flux per
unit volume in the Cartesian system by starting with the general
definition of the divergence in the form:

\[
\text{div } f = \lim_{\Delta v \to 0} \frac{\sum \Delta S_i}{\Delta v} \sum (n_i \cdot f) \Delta S_i
\]  
(4)

For convenience, we shall designate the first approach as the
'scalar product' model and the derivation based on (4) as the flux
model or the standard method. These two answers represent two
samples which I collected after conducting a survey of over fifty people
including some faculty members. Although some of them may
not remember the details of the flux model, most of them remember
the 'scalar product' model, in fact, vividly.

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, \( \nabla \cdot f \), who also introduced the notation
for the curl as \( \nabla \times f \). The fact that both the 'scalar product' and
the 'vector product' of \( \nabla \) and \( f \) do not exist can be illustrated by a
simple arithmetical 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 \( 6+ \). Neither of these two expressions is equivalent to the
original assembly.

Now if one considers Gibbs' notation for the divergence in Car-
tesian system as the 'scalar product' between \( \nabla \) and \( f \), then

\[
\nabla \cdot f = \left( \sum a_i \frac{\partial}{\partial x_i} \right) \cdot \left( \sum f_j a_j \right)
\]  
(5)

The right member of (5) 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 (5) then
one member of the assembly has the form

\[
a_1 \frac{\partial}{\partial x_1} \cdot f_1 a_1
\]  
(6)

Analogous to the arithmetical example, (6) 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.,

$$a_1 \cdot \frac{\partial}{\partial x_1} f_1 a_1 = \frac{\partial f_1}{\partial x_1} \tag{7}$$

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 (7), viz.,

$$\frac{\partial}{\partial x_1} (a_1 \cdot f_1 a_1) = \frac{\partial f_1}{\partial x_1} \tag{8}$$

Neither (7) nor (8) is equivalent to the original assembly (6). This is not a matter of interpretation; it is a manipulation which is not allowed in mathematics.

The importance of keeping the proper order in an operational method was emphasized by Feynman who stated [1]:

"With operators we must always keep the sequence right, so that the operations make proper sense. You will have no difficulty if you just remember that the operator $\nabla$ obeys the same convention as the derivative notation. What is to be differentiated must be placed on the right of the $\nabla$. The order is important.

Keeping in mind this problem or order, we understand that $T\nabla$ is an operator, but the product $\nabla T$ is no longer a hungry operator; the operator is completely satisfied."

In the present case, we are faced with the presence of the dot symbol after the differential sign in (5) and (6) so the differentiation cannot be applied to $f$ in (5) and $F_1 a_1$ in (6); it is blocked by the dot sign in the assembly. As in the arithmetical example, '2' cannot pass by the plus sign to multiply '3'.

The fact that the expression so arbitrarily created from (5) does represent the correct expression for the divergence in Cartesian system has fooled many people about the true nature of the 'scalar product' model. When the same model is applied to a curvilinear orthogonal system people found that it does not work [2] but they never questioned the meaning of the model itself. The amazing story
is that mathematicians, physicists, and engineers who used Gibbs' notations have practiced this manipulation for generations and it has reached every corner of the world. Let us quote some passages from several books. For uniformity, we have changed some of their notations to the ones used in this paper. The curvilinear orthogonal system is not involved in the following quotations:

1) From a book on Advanced Vector Analysis published in the twenties:

"To justify the notation \( \nabla \cdot \mathbf{f} \) we have only to expand the formal products according to the distributive law, then

\[
\nabla \cdot \mathbf{f} = \left( \sum_i a_i \frac{\partial}{\partial x_i} \right) \cdot \mathbf{f} = \sum_i \frac{\partial f_i}{\partial x_i} = \text{div} \ \mathbf{f}
\]

and similarly for \( \nabla \times \mathbf{f} \)."

We should remark here that any distributive law in mathematics should be proved. In this case, there is no distributive law to speak of because we are dealing with an assembly and not a meaningful mathematical expression.

2) Translated by this author from a German book on vector analysis published in the twenties:

The scalar product between \( \nabla \) with the field function \( \mathbf{f} \) is called the divergence of \( \mathbf{f} \)

\[
\text{div} \ \mathbf{f} = \nabla \cdot \mathbf{f} = \left( \sum_i a_i \frac{\partial}{\partial x_i} \right) \cdot \left( \sum_j a_j f_j \right) = \sum_i \frac{\partial f_i}{\partial x_i}
\]

The rotation of \( \mathbf{f} \) is denoted by the vector product between \( \nabla \) with the field function \( \mathbf{f} \). . . . and
\[
\text{div grad } f = \nabla \cdot \nabla f = \left( \sum_i a_i \frac{\partial}{\partial x_i} \right) \cdot \left( \sum_j a_j \frac{\partial f}{\partial x_j} \right) = \nabla^2 f
\]

It is seen that a term like \( a_1 \frac{\partial}{\partial x_1} \cdot a_1 \frac{\partial f}{\partial x_1} \) is again an assembly, not a meaningful expression.

3) From a book on electrodynamics published in the early forties:

"If we form the scalar product

\[
\nabla \cdot f = \sum_i \frac{\partial f_i}{\partial x_i}
\]

we obtain a proper scalar function called the divergence of \( f \). . . We may form the vector product of \( \nabla \) and \( f \) so as to obtain another proper vector function known as the curl or rotation of \( f \)."

4) From a book on Advanced Calculus published in the fifties:

"The formula \( \text{div } f = \sum_i \frac{\partial f_i}{\partial x_i} \) can be written

\[
\nabla \cdot f = \left( \sum_i \frac{\partial}{\partial x_i} a_i \right) \cdot \left( \sum_j f_j a_j \right) = \sum_i \frac{\partial f_i}{\partial x_i} = \text{div } f
\]

In this case the author not only did the manipulation but also wrote the del operator in the form of

\[
\nabla = \sum_i \frac{\partial}{\partial x_i} a_i
\]

instead of \( \nabla = \sum a_i \frac{\partial}{\partial x_i} \). That is very misleading because as it stands the expression in (9) is a differentiated function, which happens to be equal to zero because \( a_i \) is a constant vector; it is not an operator as Feyman has emphasized.

5) From a college textbook on calculus published in the sixties:
"The 'curl' of a vector is defined to be del cross \( \mathbf{f} \), that is

\[
\text{curl } \mathbf{f} = \nabla \times \mathbf{f} = \begin{vmatrix}
    a_1 & a_2 & a_3 \\
    \frac{\partial}{\partial z_1} & \frac{\partial}{\partial z_2} & \frac{\partial}{\partial z_3} \\
    f_1 & f_2 & f_3
\end{vmatrix}
\]

and the 'divergence' of a vector is defined to be del dot \( \mathbf{f} \), that is,

\[
\text{div } \mathbf{f} = \nabla \cdot \mathbf{f} = \sum_i \frac{\partial f_i}{\partial x_i}
\]

6) From a book on college physics published in the sixties:

"Let us try the dot product of \( \nabla \) with a vector field we know, say \( \mathbf{f} \). We write

\[
\nabla \cdot \mathbf{f} = \nabla_x f_x + \nabla_y f_y + \nabla_z f_z
\]

or

\[
\nabla \cdot \mathbf{f} = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}
\]

7) From a book on Vector Analysis for physicists published in England in the late seventies:

"... and find that \( \text{div } \mathbf{f} = \nabla \cdot \mathbf{f} \) in terms of the Cartesian operator \( \nabla \), and to be quite explicit

\[
\text{div } \mathbf{f} = \nabla \cdot \mathbf{f} = \sum_i \frac{\partial f_i}{\partial x_i}
\]

... the operator enters an equation just like a vector - to produce a scalar multiplication with another vector."

8) From an English translation of a Czechoslovak book on Applied Mathematics published in the sixties:
"The formula stated in ( ) can often be formally deduced if we note that the operator \( \nabla \) is given in vector form by ( ), for example,

\[
\nabla^2 = \nabla \cdot \nabla = \left( \sum_i a_i \frac{\partial}{\partial x_i} \right) \cdot \left( \sum_j a_j \frac{\partial}{\partial x_j} \right)
\]

\[
= \sum_i \frac{\partial^2}{\partial x_i^2} = \Delta.
\]

This is only a small sample from over one hundred books which all did the same erroneous manipulation.

We have also found books in Chinese, Japanese, and Russian adopting the same practice. In this country, many introductory textbooks in electromagnetics and some in fluid mechanics have followed this incorrect approach. Of course, in books using the old notations (grad \( f \), div \( f \), and curl \( f \) or rot \( f \)) exclusively, this problem does not occur. In particular, we would like to mention the books by King [3], Hallén [4], and Van Bladel [5].

2 The Introduction of Gibbs’ Notations to an Old German Treatise by Richard Gans

A classical book on Vector Analysis written in German by Richard Gans was published in 1905 [6], four years after the first book on Vector Analysis was published in the U.S. The American book was authored by Edwin B. Wilson [7], founded upon the lectures of J. Willard Gibbs, one of the pioneers in vector analysis, and the originator of the modern notations of the three key functions, viz. \( \nabla f \), \( \nabla \cdot f \), and \( \nabla \times f \). We shall review Gibbs’ and Wilson’s works in the final section of this paper.

It appears that Gans was the first author who gave a general definition of the three key functions in vector analysis in the form

\[
\text{grad } f = \lim_{V \to 0} \frac{\int f \, d\sigma}{V} \tag{10}
\]

\[
\text{div } f = \lim_{V \to 0} \frac{\int f \cdot d\sigma}{V} \tag{11}
\]
\[
\text{curl } \mathbf{f} = \lim_{V \to 0} \frac{\int \mathbf{d} \sigma \times \mathbf{f}}{V}
\]

(12)

Based on these definitions, the differential expressions of these functions in the Cartesian system were derived. Gans’ book was so successful that several revised editions followed. The sixth edition was translated into English in 1931. In this edition Gibbs’ notations were introduced for the first time, in addition to the old notations, (grad \( f \), div \( f \), and curl \( f \)). It was remarked on p. 49 of the English translation of the Sixth Edition [6]:

"... Thus, the operator \( \nabla \) denotes a differentiation. Seeing that \( \nabla U \equiv \text{grad } U \) has the components \( \frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z} \), that \( (\nabla \cdot A) \equiv \text{div } A = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}, \) and that \( [\nabla, A] \equiv \text{curl } A \) has the components \( \frac{\partial}{\partial y} A_x - \frac{\partial}{\partial z} A_y, \text{ etc. } \) We may formally regard the operator as a vector with components \( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \)."

It should be observed that the definition sign "\( \equiv \)" was used by Gans to identify \( \nabla U \) with grad \( U \), (\( \nabla \cdot A \)) with div \( A \), and \( [\nabla, A] \) with curl \( A \). The author used \( [\nabla, A] \) for Gibbs’ \( \nabla \times A \). There was no mention of 'scalar product' and 'vector product' between \( \nabla \) and \( A \). His interpretation of \( \nabla \) was based on the appearance of these differential expressions which were derived based on (10-12). There was no manipulation involved in his interpretation.

3 The Original Work of Gibbs and the Book by Wilson

Before we identify the person(s) who seems to be the first one to manipulate Gibbs’ notations for the divergence and the curl we should review Gibbs’ original work first. Although Gibbs, together with Heaviside, was recognized as one of the founders of Vector Analysis as a branch of applied mathematics and his notations are now almost universally accepted as the standard, his original work was never officially published. According to Crowe [8], a renowned historian of Vector Analysis, Gibbs’ notes on Vector Analysis which he prepared
for his students at Yale University in 1881 and 1884 [9] were sent to 130 scientists and mathematicians including Michelson, J.J. Thomson, Rayleigh, Stokes, Kelvin, Tait, Heaviside, Helmholtz, Kirchhoff, L.A. Lorentz, Weber, etc. Tait was then the proponent of quaternion mathematics, a forerunner of vector analysis. His comment on Gibbs' notes was [10]:

"Even Professor Willard Gibbs must be ranked as one of the retarders of Quaternion progress, in virtue of his pamphlet on Vector Analysis, a sort of the hermaphrodite monster, compounded of the notations of Hamilton and of Grassman."

It is surprising that such a rude comment could originate from a chaired professor in an institute of higher learning (University of Edinburgh). Fortunately, America's first Ph.D. in engineering, was truly a scholar of the first rank and a gentleman by nature. According to Heaviside [11]:

"Professor Gibbs' pamphlet (not published, New Haven, 1881-1884, pp. 83) is not a quaternionic treatise, but an able and in some respects original little treatise on vector analysis, through too condensed and also too advanced for learner's use; and that Professor Gibbs, being no doubt a little touched by Professor Tait's condemnation, has recently (in the pages of Nature) made a powerful defense of his position. . . . As regards his notation, however, I do not like it."

Incidently, Heaviside used some quaternionic notations in his treatment of vector analysis, and it was Gibbs' notations which finally prevailed. It is hoped that the future generations of American students will always remember the name of J. Willard Gibbs (1839-1901) as a great scientist and humanitarian [12, 13].

With this much historical background, let us review the original work of Gibbs now compiled in his Collected Works [14]

"54. Def. If $\omega$ is a vector having continuously varying values in space,
\[\nabla \omega = i \frac{dw}{dx} + j \frac{dw}{dy} + k \frac{dw}{dz}\]

\[\nabla \times \omega = i \frac{dw}{dx} + j \frac{dw}{dy} + k \frac{dw}{dz}\]

\(\nabla \omega\) is called the **divergence** of \(\omega\) and \(\nabla \times \omega\) its **curl**.

If we set

\[\omega = X i + Y j + Z k\]

we obtain by substitution the equations

\[\nabla \cdot \omega = \frac{dX}{dx} + \frac{dY}{dy} + \frac{dZ}{dz}\]

and

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

which may also be regarded as defining \(\nabla \omega\) and \(\nabla \times \omega\).

The key message here is that Gibbs **defines** the divergence and the curl as,

\[\nabla \cdot f = \sum_i a_i \frac{\partial f}{\partial x_i}\]

\[\nabla \times f = \sum_i a_i \times \frac{\partial f}{\partial x_i}\]

and he uses \(\nabla \cdot f\) and \(\nabla \times f\) as their notations.

The title for Secs. 66-72 is labelled "Combinations of the Operators \(\nabla, \nabla, \nabla \times\)." This is the first time the word 'operators' appeared in his notes. Although the scientist did not elaborate the meaning of these notations, there is no doubt that in view of his references to the definition of \(\nabla \omega\) in Sec. 52 and \(\nabla \omega\) and \(\nabla \times \omega\) in Sec. 54 they are meant to be

\[\nabla = \sum_i a_i \frac{\partial}{\partial x_i}\]
\[ \nabla = \sum_i a_i \cdot \frac{\partial}{\partial z_i} \]
\[ \nabla \times = \sum_i a_i \times \frac{\partial}{\partial z_i} \]

The important point is that he never spoke of the 'scalar product' and the 'vector product' between \( \nabla \) and \( f \), his \( \omega \). Greek letters were used to denote vectors in those days.

In 1901, Wilson published his book on Vector Analysis [7] based on Gibbs' lectures. There are two prefaces in that book. The preface by Professor Gibbs contains the following paragraph:

"I have not desired that Dr. Wilson should aim simply at the reproduction of my lectures, but rather that he should use his own judgement in all respects for the production of a text-book in which the subject shall 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 wrote:

"When I undertook to adopt the lectures of Professor Gibbs on Vector Analysis for publication in the Yale Bicentennial Series, Professor Gibbs himself was already as fully engaged upon 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 discretion like in the selection of the topics to be treated and in the mode of treatment."

In regard to the use of the operator \( \nabla \), Wilson's preface contains the following paragraph:

"It has been the aim here to give also an exposition of scalar and vector products, of the operator \( \nabla \), of divergence and curl which have gained such universal recognition since the publication of Maxwell's Treatise on Electricity and Magnetism, of slope, potential, linear vector function, etc., such as shall be adequate for the needs of students"
Actually Maxwell's treatise never used Gibbs' notations for the divergence and the curl. The last sentence of his preface concludes with:

"Finally, I wish to express my deep indebtedness to Professor Gibbs. For although he has been so pre-occupied as to be unable to read either manuscript or proof, he has always been ready to talk matters over with me, and it is he who has furnished me with inspiration sufficient to carry through the work."

Having reviewed Gibbs' original work and the history behind Wilson's book, we wish to call attention to the material in Wilson's book in regard to the misinterpretation of Gibbs' notations for the divergence and the curl. Sec. 70 of that book tells the whole story:

"70.] Although the operation $\nabla \mathbf{V}$ has not been defined and cannot be at present, two formal combinations of the vector operator $\nabla$ and a vector function $\mathbf{V}$ may be treated. These are the formal scalar product and the formal vector product of $\nabla$ into $\mathbf{V}$. They are

$$\nabla \cdot \mathbf{V} = \left( i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} \right) \cdot \mathbf{V} \quad (32)$$

$$\nabla \times \mathbf{V} = \left( i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} \right) \times \mathbf{V} \quad (33)$$

$\nabla \cdot \mathbf{V}$ reads del dot $\mathbf{V}$, and $\nabla \times \mathbf{V}$, del cross $\mathbf{V}$.

The differentiations $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ being scalar operators, pass by (underlining, this author's emphasis) the dot and the cross, that is

$$\nabla \cdot \mathbf{V} = i \cdot \frac{\partial \mathbf{V}}{\partial x} + j \cdot \frac{\partial \mathbf{V}}{\partial y} + k \cdot \frac{\partial \mathbf{V}}{\partial z} \quad (32)'$$

$$\nabla \cdot \mathbf{V} = i \times \frac{\partial \mathbf{V}}{\partial x} + j \times \frac{\partial \mathbf{V}}{\partial y} + k \times \frac{\partial \mathbf{V}}{\partial z} \quad (33)'$$

They may be expressed in terms of the components $V_1, V_2, V_3$ of $\mathbf{V}$."

The footnote '1' on that page reads:
"A definition of \(\nabla V\) will be given in Chap. VII."

We now have traced out the first written document on that incorrect manipulation. It is up to the researchers in the history of science to find out whether or not Gibbs' original lecture contained the two words "pass by" to change (32) and (33) to (32)' and (33)'. In view of the evidence which I have gathered here I personally doubt this possibility. In any event, a puzzle of eighty-nine years appears to have been resolved.

A technical report discussing in more detail some of the misunderstandings in vector analysis, particularly, formulations in the curvilinear orthogonal system, has now been completed [15]. It is hoped that the work will be published soon.

I wish to thank Professor Nenheng Fang, a Visiting Scholar at the University of Michigan, Professor Hao Ling of the University of Texas, and Dr. Jian Ming Jin of the University of Michigan for their valuable assistance. The comments which I have received from my colleagues Prof. Thomas B.A. Senior and Dr. John H. Bryant on this work are very much appreciated. I am also indebted to Ms. Bonnie Kidd for typing this manuscript in a very professional and skillful manner. Arthur and Bing, my sons, helped to edit this paper, a gratifying service.
REFERENCES


