Neither Div nor Curl nor Both Constitute the Derivative

 Charles S. Barnett

 Adjunct Mathematics Instructor

 Las Positas College

Presented at the 41st annual fall conference

California Mathematics Council, Community Colleges

Monterey, California

December 13 and 14, 2013

**Abstract**

The Div and Curl operators are derivative-like, but are only aspects of the derivative of a map from 3-space to 3-space. The derivative of such a given map is a linear map from 3-space to 3-space. By employing the Frechet difference quotient at the outset, we can construct that derivative without use of the methods of advanced analysis. The process, results, and properties parallel those of the one-dimensional case.

**Introduction**

In the usual calculus sequence we begin by defining the derivative of a real-valued function of a real variable via the limit of a difference quotient. Next come the techniques for finding the derivative of familiar functions without resort to the defining relation. As we work our way to the study of functions from *n*-space to *m*-space where and , we encounter “partial derivatives,” “gradient,” “divergence,” and “curl.” But “derivative” standing alone seems to disappear, and it reappears, if at all, in analysis courses above, or to the side of, Advanced Calculus. That traditional approach has the advantage of quickly getting to methods that have applications in engineering and scientific disciplines, the source of most students who study the calculus sequence. But these derivative-like concepts, especially Divergence and Curl, seem to appear out of a vacuum, not as a continuation of the development of the derivative concept. An alternative approach is available.

There exists a relatively smooth route from the definition of the derivative for the one-variable case to the definition of the derivative for the case in which both domain and range can have any (not necessarily equal) dimensions. The approach involves constructing the limit for the Frechet difference quotient, essentially a directional derivative. If this limit satisfies certain smoothness conditions, then a derivative emerges, and it is a linear map; the linearity is demonstrable directly from the defining relation.

When both the domain and range of a function are real 3-space, then the derivative is a linear map from 3-space to 3-space, and the Divergence and Curl are aspects of that derivative, but neither one nor both constitute the derivative. Further properties of the derivative emerge, without the use of coordinates, along lines that parallel those used to derive comparable properties of the derivative for the one-variable case. We do return to coordinates when performing calculations.

This approach to the multi-dimensional derivative illustrates the beauty and utility of generalization without having to resort to the more advanced methods of higher analysis.

What follows is expository in nature, hence the level of justification for the assertions given varies from a minimum of none to, in a few cases, a maximum of relatively acceptable plausibility arguments. I assert that solid proofs of the claims displayed appear in the references cited in the annotated bibliography.

The usual calculus course proceeds through the material somewhat as follows: (1) scalar-valued functions of a scalar, (2) vector-valued functions of a scalar, (3) scalar-valued functions of a vector, and (4) vector-valued functions of a vector. I will follow that sequence to the main goal of (4), except that I will skip (2).

I begin by defining the derivative of a real-valued function of a real variable in a way that may seem strange and unnecessarily convoluted. However, the definition I use sets the stage for the definitions that apply to the other two cases. This first section, and only this section, is new; at least I have not seen it in print. All else is, well, derivative.

**Scalar-valued Functions of a Scalar**

Situation: represents the real numbers. .

Goal: Define the derivative of at , .

For consider the following limit:

. (1)

If this limit exists, then is called “the derivative of at with respect to .”

Definition: is “differentiable” in if exists for each and each . If further, for each is continuous in , then is said to be “continuously differentiable in .”

Claim: is linear at variable

Pick and fix in and consider . Show first that = . Well,

.

Therefore, (2)

Now let and denote two non-zero elements of and consider .

for some

 . (3)

Justifications for the numbered steps:

 in (2)

 in (2)

So, is linear in the second variable.

Definition: The derivative of at , denoted by , is the linear map from defined by .

Examples:

1. is a constant function, , where is a constant. Then , i.e. , sends all to zero for every

Argument: .

1. is a linear map, i.e., , where is an arbitrary but fixed number.

 Then is a constant map, i.e.,

 for every .

 Argument: .

1. implies that , i.e.,

 all

Argument:

Now, divide by and pass to the limit to see that

 .

**Best affine approximation to**

Pick and fix . Then the mean value theorem says that

. (4)

You will see direct parallels to (4) as we work our way up to functions from .

**Scalar-valued Functions of a Vector**

Situation:

Goal: Define the derivative of at , . For , consider the following limit that defines function .

If this limit exists, then is called “the derivative of at with respect to ”.

Definition: The function is said to be *differentiable* in if exists for each in . It is said to be *continuously* *differentiable* if, for each , exists for each and is continuous in .

Claim: is linear in the variable . First, show that for .

Pick and fix in . If , the difference quotient is . If , then

Now, consider . We must show that ,

where

.

There exists a version of the mean value theorem that we will need:

If exists for , then there exists such that

Now consider

Justifications:

 Subtract and add

 Clear

 Apply Eqn. (6) with and to obtain the first term.

 Apply definition of to yield the second term.

 Apply Eqn. (5) with .

 Execute the limit operation.

Definition: If is continuously differentiable, then the derivative of at is the linear map

 ,

defined by

**The Gradient**

Thus far I have not mentioned the gradient; classical textbooks introduce and discuss the gradient early in developing the theory of functions from . I turn now to connecting our derivative to the gradient.

Much of what follows could be developed without reference to a specific basis or a specific scalar product, but henceforth I will employ the usual dot product and the standard basis. I use the following notation for base vectors.

, , .

Let

Then

Justifications:

 Clear

 is linear in the second variable.

 is linear in the second variable.

 Definition of

 Components of are equal to those of except for the th, which is .

 Therefore, the difference quotient used to define is identical to that used to define the partial derivative.

So,

Now, there exists a theorem from linear algebra which says that if is linear, then there exists a unique vector, , say, in such that

Invoke the fact that is linear in the second variable, compare Eqns. (9) and (10), and you see that the unique vector is in the case at hand. That is,

So, knowing the gradient, , is equivalent to knowing the derivative but is not the derivative. is a member of is a linear map from to.

**Best affine approximation to**

Given , consider the best affine approximation to in the neighborhood of a fixed point . We see from Eqn. (11) that the best such approximation is

or, in terms of the gradient,

Equation (12’) represents the classical version.

**Vector-valued Functions of a Vector**

Situation:

Goal: Define the derivative of at , , and discuss some of its properties. In particular, study the relationship of to the Divergence and Curl.

For , consider the following limit that defines :

if the limit exists.

Definition: Function is said to be differentiable if exists for every and . is continuously differentiable if it is differentiable and is continuous in .

In what follows I assume that all functions possess the smoothness required for whatever operation is indicated. The development is formal and sketchy. The references cited in the annotated bibliography contain extensive, careful, and broader-based arguments that lead to results much more general than those displayed here.

Examples:

1. is a constant function.

 for every . is a fixed vector.

 Then .

 So, .

1. is a linear map.

 , where is linear.

 Then .

 .

Now, divide by and pass to the limit to see that

 for all .

Let represent the usual basis in . Then

 ,

where the represent the component functions. Let represent the derivative with respect to of the component functions. In the previous section we sketched an argument that showed that the are linear in the variable . It can be shown that inherits linearity in from the . So,

is linear in the variable .

Definition: The derivative of at is the linear transformation

defined by

Examples (restatement in terms of rather than ):

1. The derivative of a constant function is the zero vector. So transforms all of into the zero vector.
2. The derivative of a linear function is a constant.

 linear, implies that

 for every .

Now we can get back to more familiar, classical ground. is a linear map. Given a basis, a linear map can be represented by a matrix relative to that basis. So, as above, let represent the standard basis. Then

 ,

where the scalar-valued are the component functions of .

Then

I have invoked Eqn. (9) of the previous section and the linearity of to arrive at Eqn. (15).

Observe from Eqn. (15) that, relative to the standard basis, is represented by

where here and below I use to indicate the matrix of .

**Best affine approximation to *F***

As in the previous and cases, we have a best approximation via Taylor’s formula:

We now have available the definition of the derivative of a function from . That derivative is a linear map from. To define Div and Curl and to show their connection to the derivative, I need some results from Linear Algebra. They appear in Appendix A as assertions without proof. In Appendix A I restrict attention to results about real -space where , which suffices for present purposes. More general versions of most exist.

**Divergence and Curl of a Vector Field**

Given as above, then is a linear map from . Linear maps may be broken into the sum of their symmetric and skew-symmetric parts.

Let represent the adjoint of . Then

 and

 *skew-symmetric part*. (19)

Observe that sum of the two parts.

Formula (16) displays the matrix representation of , and I wish to display some additional matrix representations. So, to simplify typography, until further notice let

Then Formula (16) becomes

and the matrix of is

The matrix representations and (21) combined with Eqns. (18) and (19) and followed by some matrix algebra lead to the matrix representations of and

They are

Definition: The Divergence of is a map , defined by

The *Trace* of is equal to the sum of the roots of the characteristic polynomial of , a property of the linear map that is invariant to matrix representation. The trace of any linear map from is equal to the sum of the diagonal elements of any matrix representation. Hence, the familiar version of the Divergence:

Observe also that

Definition: The Curl of at is defined by

How do we know that a vector that makes the right side of Eqn. (24) equal to the left-hand side exists? Well, is a skew-symmetric transformation and item A5 of the Appendix asserts that such a vector exists for such maps; we name it . Also, according to A5, if the matrix of a skew-symmetric map looks like

then the vector at issue ( here) is

Compare Eqns. (23), (25), and (26) and you see that

or, in the usual notation,

**Divergence and Curl of a Steady Flow (an application)**

Consider a vector field and the flow induced by the differential equation

Note that time does not appear on the right hand side of Eqn. (28), hence “steady flow.”

Pick and fix . Expand around via Taylor’s formula; substitute the result into Eqn. (28) to yield

Split into the sum of its symmetric and skew-symmetric components

and substitute into Eqn. (29) to obtain

Now, consider the motion of the points within a small sphere centered at at time zero, say. How has that spherical test ball changed in a small time ? The terms on the right-hand side yield an approximate answer.

Consider first the motion induced by

Equation (32) says that the test ball rotates about an axis through parallel to with angular velocity . That motion is volume-preserving.

Consider next the motion described by

 is a symmetric map centered at . Let represent its eigenvalues and denote the corresponding orthonormal set of eigenvectors. Then the corresponding eigenvalues of the flow described by Eqn. 33 are . That flow distorts the spherical ball slightly into an ellipsoid. The volume of the ball changes by a factor that is the determinant of the transformation. That determinant is and its time rate of change when is .

Finally, consider the motion induced by

Equation (34) says that the ellipsoid experiences a small translation with velocity vector .

**The General Case and the Chain Rule**

The chain rule is so important in differential calculus that its multidimensional version deserves a brief discussion. Before doing so, I point out that our restriction to the case above was not necessary; I invoked the restriction because of the emphasis on the undergraduate calculus sequence.

Let and represent natural numbers and consider . If is differentiable at , then its derivative can be defined via the Frechet difference quotient just as we did in the earlier cases. The process is identical to the one used for the case. The derivative is a linear transformation as before, and all goes through except the discussion of the Curl, which is unique to the three-dimensional case. With this generalization all cases are subsumed, again except for the Curl, by this general case.

Consider now the chain rule in the general case. Let and represent natural numbers. Situation: is differentiable at , and is differentiable at , then is differentiable at and

where represents function composition. Observe that in Eqn. (35) is a linear map and is a linear map; hence the composition of the two maps is a linear map, and all is well. Use brackets to denote matrix representation, and the matrix version of Eqn. (35) becomes

Recall the one-variable chain rule.

where the juxtaposition on the right-hand side of Eqn. (37) denotes multiplication, but could just as well have been written

because, in the one-dimensional case, composition of linear functions is multiplication. Compare Eqn. (35) and Eqn. (37’) and you see that Eqn. (37’) is just a special case of Eqn. (35).

**Appendix**

This appendix contains results from Linear Algebra that I need to show how Divergence and Curl are related to the derivative as defined and discussed in this note. The assertions appear without proofs; proofs appear in the references.

Consider vector space over the field of real numbers endowed with the usual basis and the corresponding scalar product, the dot product.

A1. If is linear, then there exists unique vector such that

 .

A2. Given that is linear, pick and fix . Construct linear map by

 .

 By A1 above, there exists unique vector such that

 .

 Rename vector where . Then

 .

 Fact: , called the *adjoint* *of* , is linear.

 Fact: Relative to an orthonormal basis, the matrix of is the transpose of the

 matrix of

A3. Definition: is said to be symmetric (or self-adjoint) if . is skew-symmetric if .

A4. For any linear ,

 is symmetric.

 is skew-symmetric.

 and are called the symmetric and skew-symmetric parts of because .

A5. If is skew-symmetric, then there exists unique vector such that

 for all

 Relative to the standard basis , the matrix of a skew symmetric map looks like the following matrix.

 and the of the theorem is

 .

A6. If is symmetric, then there exists an orthonormal basis, , and real numbers (not necessarily distinct) such that for .

A7. Situation: is linear. is the identity map. For each define linear map by . Then det , the determinant of , is a polynomial of degree 3 in .

 Let denote the roots of det . Then

 .

det is the *characteristic polynomial of* .

det is an invariant of (independent of the matrix representation of ).

The , and since det , we can write the characteristic polynomial of as

 **Annotated Bibliography**

1. H. K. Nickerson, D. C. Spencer, and N. E. Steenrod, *Advanced* *Calculus*, Dover Publications, 2011. Unabridged republication of the work originally published in 1959 by the D. Van Nostrand Company.

Most of what appears in this presentation comes from this book. To give you some idea of the nature of its content, I offer the following three items: (1) a note that appears on the outside of its back cover, (2) an excerpt from the authors’ preface, and (3) a partial list of chapter titles.

First, the note on the cover:

“This book is a radical departure from all previous concepts of advanced calculus,” declared the *Bulletin of the American Mathematics Society*, “and the nature of this departure merits serious study of the book by everyone interested in undergraduate education in mathematics.” Classroom-tested in a Princeton University honors course, it offers students a unified introduction to advanced calculus.

Next, the preface material:

 These notes were prepared for the honors course in Advanced Calculus, Mathematics 303-304, Princeton University.

 The standard treatises on this subject, at any rate those available in English, tend to be omnibus collections of seemingly unrelated topics. The presentation of vector analysis often degenerates into a list of formulas and manipulative exercises, and the student is not brought to grips with the underlying mathematical ideas.

 In these notes a unity is achieved by beginning with an abstract treatment of vector spaces and linear transformations. This enables us to introduce a single basic derivative (the Frechet derivative) in an invariant form. All other derivatives (gradient, divergence, curl and exterior derivative) are obtained from it by specialization. The corresponding theory of integration is likewise unified, and the various multiple integral theorems of advanced calculus appear as special cases of a general Stokes’ formula concerning the integration of exterior forms. In a final chapter these concepts are applied to analytic functions of complex variables.

Finally, the partial list of chapter titles:

 I. The Algebra of Vector Spaces

 II. Linear Transformations of Vector Spaces.

 III. The Scalar Product

 IV. Vector Products in

 V. Endomorphisms

 VI. Vector-valued Functions of a Scalar

 VII. Scalar-valued Functions of a Vector

 VIII. Vector-valued Functions of a Vector

So, this reference is not your usual advanced calculus book. It was published in an 8” by 11” typescript form and that first edition was the last. The Dover edition is a photo-reduced form of the original publication.

Clearly, this book has had little influence on modern-day teaching of calculus and advanced calculus. They are taught today much as they were in 1959 when the authors wrote this book; perhaps that persistence is appropriate for all but a few exceptional students. We emphasize mechanical how-to-do-it methods rather than “underlying mathematical ideas.” Since most of our students will be consumers of and users of mathematics rather than producers of and teachers of mathematics, the present emphasis on manipulative skills is probably appropriate.

However, if you are interested in the deeper aspects of the subject, I recommend Nickerson, Spencer, and Steenrod. It is thorough, carefully written, includes proofs (some off the beaten path) of almost all content, and will reward effort, especially if precision and clarity interest you. It may be “advanced calculus on steroids,” but it is well worth the moderate price.

1. Spivak, Michael, *Calculus on Manifolds*: *A Modern Approach to Classical Theorems of Advanced Calculus*,” W. A. Benjamin, Inc., New York, Amsterdam, 1965.

Excerpt from the preface:

 This little book is especially concerned with those portions of “advanced calculus” in which the subtlety of the concepts and methods makes rigor difficult to attain at an elementary level. The approach taken here uses elementary versions of modern methods found in sophisticated mathematics. The formal prerequisites include only a term of linear algebra, a nodding acquaintance with the notation of set theory, and a respectable first-year calculus course (one which at least mentions the upper bound (sup) and the greatest lower bound (inf) of a set of real numbers). Beyond this a certain (perhaps latent) rapport with abstract mathematics will be found almost essential.

This book too treats foundational material carefully and rigorously. Its scope is narrower and deeper than that of Nickerson, Spencer, and Steenrod. I include it here because it contains a nice proof of the chain rule for functions from vector spaces to vector spaces.

1. H. M. Schey, *Div, Grad, Curl, and All That: An Informal Text on Vector Calculus*, W.W. Norton & Company, Inc., New York, 1973.

As a counterpoint to the two previous references, I call attention to this delightful paperback book. It was designed for undergraduate physics students at M.I.T. It emphasizes the physical significance of the vector concepts of Div, Grad, and Curl and makes no claim to mathematical rigor. Here is an excerpt from the preface:

I undertook to write this short text on vector analysis as a result of my experience several years ago in teaching electricity and magnetism to M.I.T. undergraduates. The course began with elementary concepts and culminated in a brief discussion of Maxwell’s equations. It follows that en route we encountered a fair amount of vector calculus, and I was chagrined to find my students unequal to the challenge. (That puts it rather mildly; they took to hissing every time I said divergence, gradient, or curl.) I tried to provide some of the mathematics in lecture, but couldn’t do as much of this as necessary since my job was to teach physics, not mathematics. I sorely needed a short (and preferably inexpensive) auxiliary text that my students could use to learn the vector calculus they needed in their study of electromagnetic theory, and it was to fill such a need that I wrote this book. However, I think it can also be used by students who, without any applications to physics in mind, would like to learn vector calculus just for its own sweet self.

Needless to say, what I have written is certainly not a mathematician’s account of vector calculus, for I have paid scant attention to mathematical rigor. This treatment is very “relaxed,” and physical and geometrical arguments are used throughout wherever I felt they helped to make a point.

This book contains a nice demonstration of the connection between the curl and the angular velocity of a swirling liquid. Schey uses as model the image of a liquid rotating in the (x,y) plane as in water draining from a bathtub. If is the velocity of the liquid as a function of the distance from the origin, and is the angular velocity of the liquid as a function of that same distance, then Schey shows that

 (B1)

Recall, from Eqn. (24), that

 (B2)

Perhaps Eqn. (B1) helps explain the “2” in Eqn. (B2)