3

AN INTRODUCTION TO VECTOR CALCULUS

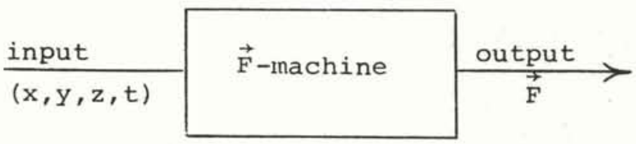
A

Introduction

In the same way that we studied numerical calculus after we learned numerical arithmetic, we can now study vector calculus since we have already studied vector arithmetic. Quite simply (and this will be explored in the remaining sections of this chapter), we might have a vector quantity that varies with respect to another variable, either a scalar or a vector. In this chapter we shall be most interested in the case where we have a vector which varies with respect to a scalar.

A rather simple physical situation of this case would be the problem where we measure the force on an object at different times. That is, force (a vector) is then a function of time (a scalar). In this context we can ask whether our force is a continuous function of time, a differentiable function of time, and so forth.

At the same time, this type of exploration opens up new avenues of investigation and supplies us with excellent practical reasons for exploring the rather abstract notion of higher dimensional spaces. More specifically, let us keep in mind that it is a rather overly-simple situation in real life when the quantity under consideration depends on only one other variable. For example, with respect to our above illustration of force depending upon time, let us note that in a real-life situation the object we are studying has non-negligible size and for this reason a force applied at one point on the object has different effects at other points. Thus, we might find that our vector force depends on the four variables x, y, z, and t. That is, the force may be a function of both position (three dimensions) and time. In terms of a function, our input is a 4-tuple (that is, an ordered array [sequence] of four real numbers x, y, z, and t). Pictorially,



Now we have already used 2-tuples to abbreviate vectors in the plane and we have used 3-tuples to abbreviate vectors in space. That is, we have let (a,b) denote $a\vec{i} + b\vec{j}$ and (a,b,c) denote $a\vec{i} + b\vec{j} + c\vec{k}$. In this context, we may think of a 4-tuple as denoting a 4-dimensional "vector," even though we may not be able to visualize it in the usual sense. We have earlier come to grips with this problem in the form of exponents. Namely, if we look at b^1, b^2, and b^3 we can interpret each of these pictorially very nicely. That is, b^1 is a length of b units, b^2 is the area of a square whose side has length b, and b^3 is the volume of a cube whose side has length b. But while b^4 does not have such a simple geometric interpretation, it is just as meaningful to multiply four (or more) factors of b as it is to multiply one, two, or three. In other words, we are back to a fundamental topic of Part 1 of our course - a picture is worth a thousand words <u>if you can draw it</u>.

In any event, what we are saying is that when we study functions of n variables we are in effect looking at an n-dimensional space. In fact this offers us a rather straight-forward, non-mystic interpretation as to why time is often referred to as the fourth dimension. Namely, in most physical situations the variable we are measuring (as we discussed earlier) depends on position and time. Position in general involves three dimensions and time is then the fourth dimension (variable).

Notice that in this respect we can have a scalar (as well as a vector) function of n-dimensional vectors (meaning the input consists of n variables while the output is a number). For example, we might be studying temperature (a scalar) as a function of position and time, in the sense that temperature in general does vary from point to point and at the same point it usually varies from time to time.

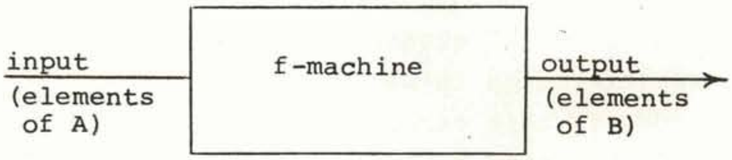
In order not to introduce too many new ideas at once, we shall devote the remainder of this chapter only to the calculus of 2 and 3-dimensional vectors. Once we gain some familiarity with this basic concept, we shall extend our frontiers to include the more general study of n-dimensional vectors. This topic will be discussed in a later chapter of supplementary notes which will be covered during our study of Block 3.

B

Functions Revisited

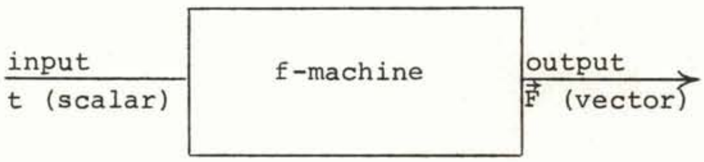
When we first introduced the concept of functions, we mentioned that a function was a rule which assigned to members of one set members of another set. We pointed out that the two sets involved could be arbitrary but that in the study of real variables our attention, <u>by definition</u>, is confined to the case in which both sets are subsets of the real numbers.

We also developed the visual aid of the <u>function machine</u>, wherein if $f:A \to B$, we had



The point of this chapter is that, in the real world, we are often confronted with situations in which we are dealing with vectors rather than scalars, whereupon we have the situation that either the input or the output of our f-machine could be vectors.

In our earlier example, we spoke of force (a vector) as a function of time (a scalar). In this case our f-machine looks like



More concretely, we might have a force $\vec{F}$ defined in the xy-plane by

$$\vec{F} = t\vec{i} + t^2\vec{j} \qquad (1)$$

From (1) it is clear that different values of the <u>scalar</u> t yield different values of the <u>vector</u> $\vec{F}$. For instance, if $t = 3$ then $\vec{F} = 3\vec{i} + 9\vec{j}$. Moreover, it should seem fairly obvious that we might like to abbreviate the above information in the form

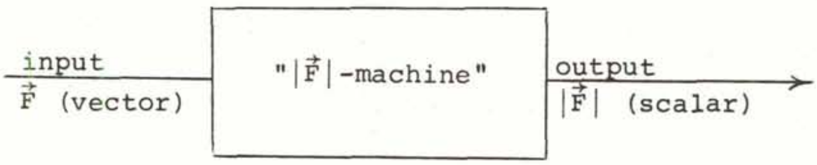
$$\vec{F}(3) = 3\vec{i} + 9\vec{j} \qquad (2)$$

Notice how the language in (2) closely parallels the notation that we have already used for "regular" functions. The only difference is that before we were talking about scalar functions (i.e. the "outputs" were scalars) of scalar variables (i.e. the "inputs" were also scalars), while in this case we are dealing with vector functions (i.e. the outputs are vectors) of scalar variables.

In this same vein, we might write

$f(\vec{x})$ (3)

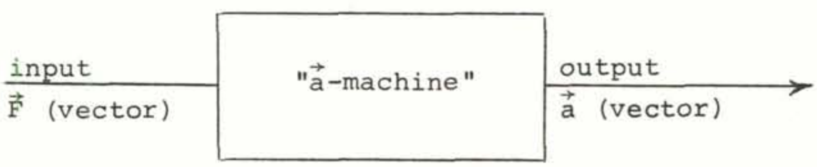
to indicate that we have a scalar function (i.e. the output of f is a scalar) of a vector variable (i.e. the input of the f-machine is a vector). As an example in "real life" of equation (3), consider a situation in which all we were interested in was the magnitude of various forces being considered. In this case, given a force $\vec{F}$ we would compute its magnitude $|\vec{F}|$. In this way our procedure is to take a vector (force) and convert it to a scalar (magnitude of the force). Pictorially,



The final possibility is that we have

$\vec{f}(\vec{x})$ (4)

Equation (4) denotes a vector function of a vector variable. An example of this situation might be that at any given instant we know the force (a vector) being exerted on a particle from which we want to determine the acceleration of the particle (acceleration is also a vector - in fact in Newtonian physics the direction of the acceleration is the same as that of the force, which means that the "law" f = ma remains correct in the vector form $\vec{f} = m\vec{a}$). Again pictorially,



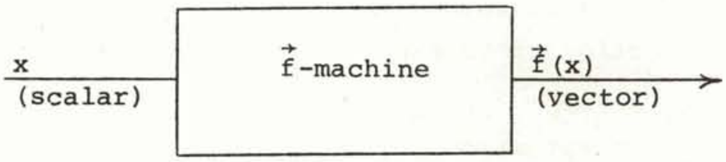
In summary then, if we allow vectors as well as scalars to be our variables, then the study of functions of a single variable reduces to one of four types

$f(x)$, $\vec{f}(x)$, $f(\vec{x})$, or $\vec{f}(\vec{x})$

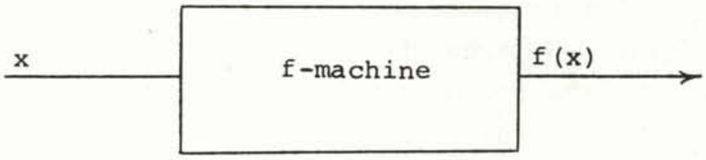
In this block we shall, when dealing with vector functions, assume that we have a vector function of a scalar variable. That is, we shall deal with functions of the form

$\vec{f}(x)$

Pictorially,



The key observation to be made from the picture above is that if the arrows are omitted our diagram is identical with our previous diagram of a function machine in the study of real variables. Namely,



This observation is the backbone of the next section.

C

The "Game" Applied to Vector Calculus

In defining $\lim_{x\to a} f(x) = L$, we used our intuition to make sure that we knew what we wanted the expression to mean, and we then proceeded to make the definition more rigorous using ε's and δ's. The rigorous way seemed quite frightening at first, but, after a while, we began to notice that it only said in precise mathematical terms what we already believed to be true intuitively.

What may have gone unnoticed was that once we had the rigorous definition, every proof we gave concerning limits consisted of applying the rules of mathematics to the mathematical definition of a limit. In terms of our game idea, we defined what a limit was, used the accepted rules of numerical arithmetic, and then proved theorems about limits.

The key point here is that if it should happen that our definition of limit in our discussion of vector functions of a scalar variable has the same structure as our definition of limit when we were dealing with scalar functions of scalar variables, and if it happens that every accepted rule of numerical arithmetic that was used in the earlier proofs of our limit theorems also happens to be an accepted rule for vector arithmetic, <u>then</u>, in terms of the philosophy of our game concept, every theorem about limits that was valid in the study of "numerical limits" will remain valid in the study of "vectorial limits." The purpose of this section is to explore this idea in more computational detail.

How can we test whether the vector situation parallels the numerical case? There is a very elegant yet simple technique that we shall employ here. We shall begin by writing the definition of limit in the scalar case. We shall then go through the definition and insert arrows over those symbols that will now denote vectors rather than scalars (and, admittedly, we must take care in determining what is still a scalar and what has become a vector). In this way, we are sure that the basic structure cannot be different since the only changes we made were in the <u>names</u> of the variables. We then check to see if the resulting definition obtained from the scalar definition by this "arrowizing" technique is still meaningful. For example, if we tried to vectorize the numerical statement that $a/b = c$, we would obtain $\vec{a}/\vec{b} = \vec{c}$ which makes no sense, since vector division is undefined. (Of course, had we vectorized all but b, the resulting statement would make sense but this is not the point we are trying to make here.)

In any event, carrying out our instructions so far, first we write the original definition of limit in the scalar case, namely

$$\lim_{x \to a} f(x) = L$$

means given any $\varepsilon > 0$, we can find $\delta > 0$ such that whenever

$$0 < |x - a| < \delta$$

then

$$|f(x) - L| < \varepsilon.$$

Next, we rewrite this definition putting in arrows wherever appropriate. Recall that we are dealing with vector functions of scalar variables. Hence, x is a scalar. But f and L should be represented as vectors since we have already agreed that we are dealing with vector functions and this implies that our output, in this case f(x) and L, are also vectors. Notice, however, that even though we have now introduced vectors, ε and δ are still scalars. The reason for this is that ε and δ denote absolute values (magnitudes) of quantities, and the magnitude of both scalars and vectors are non-negative real numbers. In any event, if we now rewrite our old definition with the appropriate "arrowization" we obtain

$$\lim_{x \to a} \vec{f}(x) = \vec{L}$$

means given any $\varepsilon>0$ we can find $\delta>0$ such that whenever

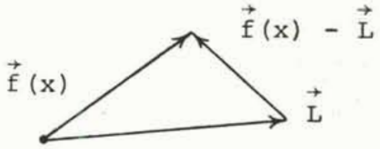
$$0 < |x - a| < \delta$$

then

$$|\vec{f}(x) - \vec{L}| < \varepsilon.$$

Our first observation of this definition should show that it is <u>meaningful</u>. That is, every part of the expression is defined (unlike $\vec{a}/\vec{b}$).

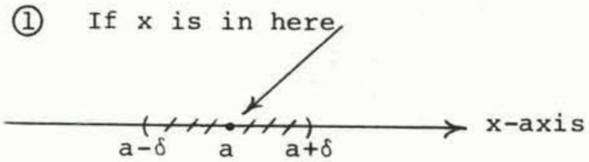
Our next task is to see whether the definition captures what we would like $\lim_{x \to a} \vec{f}(x) = \vec{L}$ to mean. After all, what good is it if our definition is meaningful but it doesn't capture the meaning we have in mind? To this end, it is probably fair to assume that we would like $\lim_{x \to a} \vec{f}(x) = \vec{L}$ to mean that if x is "sufficiently close to" <u>a</u> then $\vec{f}(x)$ is "sufficiently close to" $\vec{L}$. Notice that for two vectors to be "nearly" equal their difference must be small. In other words, for any pair of vectors A and B, if we assume that they originate at a common point then their difference in magnitude is the length of the vector that goes from the head of one to the head of the other. Clearly, the smaller this vector is in magnitude the more "together" are the heads of the two vectors. In terms of our above definition, notice that $\vec{f}(x) - \vec{L}$ means pictorially that

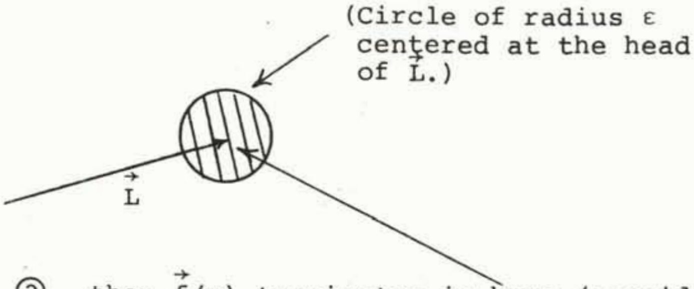


The smaller the difference between $\vec{f}(x)$ and $\vec{L}$ the closer is the head of $\vec{f}$ to the head of $\vec{L}$ which means that $\vec{f}$ is more nearly equal to $\vec{L}$.

In other words, then, in terms of what the analytic expressions mean geometrically, we should be convinced that the formal definition agrees with our intuition. Again, in terms of a picture

① If x is in here





② then $\vec{f}(x)$ terminates in here (provided it originates at the tail of $\vec{L}$).

So far, so good, but if we want to be able to capitalize on our game idea, the crucial test now lies ahead. Recall that all our formulas for derivatives stemmed from certain basic properties of limits (such as the limit of a sum is the sum of the limits, and so on). These properties of limits followed mathematically from certain properties of real numbers. The key now is to show that the proofs apply <u>verbatim</u> when we replace the scalars by appropriate vectors.

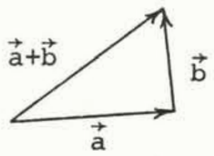
For example, what were some of the properties we used about absolute values in dealing with our proofs about limits? Well, for one thing, we used the property that for any numbers a and b

$$|a + b| \leq |a| + |b| \qquad (1)$$

The important thing is that if we now vectorize (1) to obtain

$$|\vec{a} + \vec{b}| \leq |\vec{a}| + |\vec{b}| \qquad (2)$$

we see that (2) is also true. Pictorially,



The length of one side of a triangle cannot exceed the sum of the lengths of the other other two sides.

We also used such facts as

$$|a| \geq 0 \text{ and } |a| = 0 \qquad (3)$$

if and only if a = 0. If we now vectorize (3) appropriately, we obtain

$$|\vec{a}| \geq 0 \text{ and } |\vec{a}| = 0 \qquad (4)$$

if and only if $\vec{a} = \vec{0}$. Recalling that $|\vec{a}|$ means the magnitude of $\vec{a}$, it is trivial to see that (4) is also a true statement.

<u>We must be careful not to become too glib in our present approach</u>. That is, we must not feel that it is a truism that we can just vectorize everything in sight and obtain true vector statements from true scalar statements. One place where we must be very careful, for example, is when we vectorize any numerical statement that involves products. Consider the property of absolute values that for any numbers a and b

$$|ab| = |a|\ |b| \qquad (5)$$

If we try to vectorize (5) by brute force we find for one thing that the resulting equation may well be meaningless. For instance, the statement

$$|\vec{a}\vec{b}| = |\vec{a}|\ |\vec{b}| \qquad (6)$$

is meaningless (i.e., undefined) since we have not yet defined an "ordinary" product of two vectors $\vec{a}\vec{b}$. That is, the left side of Equation (6) is undefined. If we interpret the multiplication as scalar multiplication and vectorize Equation (5) accordingly, we obtain

$$|\vec{a}\cdot\vec{b}| = |\vec{a}|\ |\vec{b}|$$

or

$$|\vec{a} \times \vec{b}| = |\vec{a}|\ |\vec{b}| \tag{8}$$

(where on the right side we mean the usual multiplication since both $|\vec{a}|$ and $|\vec{b}|$ are <u>numbers</u>).

The crucial point is that both forms of equation (8) are meaningful but they both happen to be <u>false</u>. Namely,

$$\vec{a}\cdot\vec{b} = |\vec{a}|\ |\vec{b}|\ \cos\angle^{\vec{b}}_{\vec{a}}\ .$$

Hence,

$$|\vec{a}\cdot\vec{b}| = |\vec{a}|\ |\vec{b}| \left|\cos\angle^{\vec{b}}_{\vec{a}}\right| \tag{9}$$

since $\left|\cos\angle^{\vec{b}}_{\vec{a}}\right| \leqslant 1$, (9) yields the interesting, but perhaps shattering result that

$$|\vec{a}\cdot\vec{b}| \leqslant |\vec{a}|\ |\vec{b}|\ .$$

In a similar way,

$$|a \times b| = |\vec{a}|\ |\vec{b}| \left|\sin\angle^{\vec{b}}_{\vec{a}}\right| \leqslant |\vec{a}|\ |\vec{b}| \quad \left(\text{since } \left|\sin\angle^{\vec{b}}_{\vec{a}}\right| \leqslant 1\right).$$

In other words, any limit proof that involved the result that the absolute value of a product is the product of the absolute values might well be false in absolute value or limit problems involving either the dot or the cross product of vectors.

Rather than belabor this point, let us illustrate our ideas by means of a few examples; and to emphasize their importance, let us utilize a new section to discuss them.

D

Some Limit Theorems for Vectors

Theorem 1

Suppose $\lim_{x \to a} \vec{f}(x) = \vec{L}_1$ and $\lim_{x \to a} \vec{g}(x) = \vec{L}_2$.

Then $\lim_{x \to a} (\vec{f}(x) + \vec{g}(x)) = \vec{L}_1 + \vec{L}_2$.

To prove this theorem, we write down the analogous proof in the scalar case. (We will not motivate this proof since this was already done in Part 1 of our course. However, it will be an interesting check for you to see how "natural" the proof now seems to you.)

We had:

Suppose $\lim_{x \to a} f(x) = L_1$ and $\lim_{x \to a} f(x) = L_2$.

Then $\lim_{x \to a} [f(x) + g(x)] = L_1 + L_2$.

The proof went as follows

(1) Let $h(x) = f(x) + g(x)$. We must show that $\lim_{x \to a} h(x) = L_1 + L_2$.

(2) Let $\varepsilon > 0$ be arbitrarily given. Then there exist numbers δ_1 and δ_2 such that

$$0 < |x - a| < \delta_1 \to |f(x) - L_1| < \frac{\varepsilon}{2}$$

and

$$0 < |x - a| < \delta_2 \to |g(x) - L_2| < \frac{\varepsilon}{2} .$$

(3) Let $\delta = \min \{\delta_1, \delta_2\}$.
Therefore, $0 < |x - a| < \delta \to |f(x) - L_1| + |g(x) - L_2| < \varepsilon \quad (= \frac{\varepsilon}{2} + \frac{\varepsilon}{2})$.

(4) But $|f(x) - L_1| + |g(x) - L_2| > |[f(x) - L_1] + [g(x) - L_2]|$.

(5) Therefore, $0 < |x - a| < \delta \to |[f(x) - L_1] + [g(x) - L_2]| < \varepsilon$.

(6) $|[f(x) - L_1] + [g(x) - L_2]| = |[f(x) + g(x)] - [L_1 + L_2]|$

$$= |h(x) - (L_1 + L_2)| .$$

(7) Therefore, $0 < |x - a| < \delta \rightarrow |h(x) - (L_1 + L_2)| < \varepsilon$
Therefore, $\lim_{x \to a} h(x) = L_1 + L_2$.

The key now is to observe that every step in our proof remains valid when vectors are introduced. Why? Well, for example, the validity of step (1) requires that the sum of two numbers be a number; otherwise, $h(x) = f(x) + g(x)$ would be meaningless. To convert this into a vector statement we only require that the sum of two vectors be a vector, and this we know is true. The validity of step (2) depends only on the definition of limit and we have taken care to mimic the scalar definition in forming the vector definition. Step (3) remains valid by "default" since only scalar properties are being used. That is, both $|f(x) - L_1|$ and $|\vec{f}(x) - \vec{L}_1|$ are scalars. The validity of step (4) hinges on the "triangle in equality," $|a + b| \leq |a| + |b|$, but this, as we have seen, is also valid for vectors. Step (5) is merely substitution of scalar quantities. As for step (6), this follows from the commutative and associative properties of numerical addition, and, as we have seen, these properties are also obeyed in vector addition.

In other words, if we "vectorize" the scalar proof, the resulting proof is a valid vector proof. When we do this (just for practice) we obtain:

(1) Let $\vec{h}(x) = \vec{f}(x) + \vec{g}(x)$. We must show that $\lim_{x \to a} \vec{h}(x) = \vec{L}_1 + \vec{L}_2$.

(2) Let $\varepsilon > 0$ be arbitrarily given. Then there exist numbers δ_1 and δ_2 such that

$$0 < |x - a| < \delta_1 \rightarrow |\vec{f}(x) - \vec{L}_1| < \frac{\varepsilon}{2}$$

and

$$0 < |x - a| < \delta_2 \rightarrow |\vec{g}(x) - \vec{L}_2| < \frac{\varepsilon}{2} .$$

(3) Let $\delta = \min\{\delta_1, \delta_2\}$.
Therefore, $0 < |x - a| < \delta \rightarrow |\vec{f}(x) - \vec{L}_1| + |\vec{g}(x) - \vec{L}_2| < \varepsilon$.

(4) But $|\vec{f}(x) - \vec{L}_1| + |\vec{g}(x) - \vec{L}_2| \geq |[\vec{f}(x) - \vec{L}_1] + [\vec{g}(x) - \vec{L}_2]|$.

(5) Therefore, $0 < |x - a| < \delta \rightarrow |[\vec{f}(x) - \vec{L}_1] + [\vec{g}(x) - \vec{L}_2]| < \varepsilon$.

(6) $|[\vec{f}(x) - \vec{L}_1] + [\vec{g}(x) - \vec{L}_2]| = |[\vec{f}(x) + \vec{g}(x)] - (\vec{L}_1 + \vec{L}_2)|$

$$= |\vec{h}(x) - (\vec{L}_1 + \vec{L}_2)|.$$

(7) Therefore, $0 < |x - a| < \delta \rightarrow |\vec{h}(x) - (\vec{L}_1 + \vec{L}_2)| < \varepsilon$.

Therefore, $\lim_{x\to a} \vec{h}(x) = \vec{L}_1 + \vec{L}_2$.

It is, of course, important to notice that our vector proof as it stands is self-contained. That is, it exists in its own right even had we not obtained it from an earlier scalar proof. It's just that by observing the similarity in structure, we may deduce the properties of certain vector functions from the corresponding properties of the scalar functions, and in this way, we again benefit from the parallel structure.

Theorem 2

If $\lim_{x\to a} \vec{f}(x) = L_1$ and if $\lim_{x\to a} \vec{g}(x) = \vec{L}_2$, then if $h(x) = \vec{f}(x) \cdot \vec{g}(x)$,

$\lim_{x\to a} h(x) = \vec{L}_1 \cdot \vec{L}_2$.

Before proceeding with the proof of Theorem 2, we should note that there is more than one way to form a product of two vectors. We have arbitrarily elected to talk about the dot product. A similar proof holds for the cross product and will be left as an exercise. It is worth noting that the fact that we can form a product of two vectors in more than one way will cause certain complications in vector calculus that did not exist in scalar calculus, since there, a product of two numbers could be interpreted in only one way.

Proof

Had this been written in terms of scalars rather than vectors, our proof would have taken the form

$$f(x) = [f(x) - L_1] + L_1$$

$$g(x) = [g(x) - L_2] + L_2$$

therefore, $|f(x)\ g(x) - L_1\ L_2| = |L_2[f(x) - L_1] + L_1[g(x) - L_2] + [f(x) - L_1][g(x) - L_2]|$

$$\leq |L_2||f(x) - L_1| + |L_1||g(x) - L_2| + |f(x) - L_1||g(x) - L_2|^* \qquad (1)$$

Since $|L_1|$ and $|L_2|$ are bounded and both $|f(x) - L_1|$ and $|g(x) - L_2|$ can be made arbitrarily small by choosing x sufficiently near a, the right side of (1) gets arbitrarily small as $x \to a$. This guarantees that

$$\lim_{x \to a} (f(x)\ g(x) - L_1\ L_2) = 0$$

or

$$\lim_{x \to a} f(x)\ g(x) = L_1\ L_2\ .$$

If we now repeat this argument with the appropriate "vectorization," we have

$$\vec{f}(x) = [\vec{f}(x) - \vec{L_1}] + \vec{L_1}$$

$$\vec{g}(x) = [\vec{g}(x) - \vec{L_2}] + \vec{L_2}\ .$$

Therefore, $\vec{f}(x) \cdot \vec{g}(x) = \{[\vec{f}(x) - \vec{L}_1] + \vec{L}_1\} \cdot \{[\vec{g}(x) - \vec{L}_2] + \vec{L}_2\}$

$$= [\vec{f}(x) - \vec{L}_1] \cdot [\vec{g}(x) - \vec{L}_2] + \vec{L}_1 \cdot [\vec{g}(x) - \vec{L}_2] + \vec{L}_2 \cdot [\vec{f}(x) - \vec{L}_1] + \vec{L}_1 \cdot \vec{L}_2\ ^{**}$$

---

*Notice that while, for example, $|L_1[g(x) - L_2]| = |L_1||g(x) - L_2|$ our string of inequalities requires only the weaker result that $|L_1[g(x) - L_2]| \leq |L_1||g(x) - L_2|$. This, as we shall soon see, is very important because of the vector properties that $|\vec{a} \cdot \vec{b}| \leq |\vec{a}||\vec{b}|$ and $|\vec{a} \times \vec{b}| \leq |\vec{a}||\vec{b}|$.

**We put this step in to emphasize the fact that the rules of ordinary algebra carry over very nicely here to dot products.

therefore, $|\vec{f}(x) \cdot \vec{g}(x) - \vec{L}_1 \cdot \vec{L}_2| = |\vec{L}_2 \cdot [\vec{f}(x) - \vec{L}_1] + \vec{L}_1 \cdot [\vec{g}(x) - \vec{L}_2] + [\vec{f}(x) - \vec{L}_1] \cdot [\vec{g}(x) - \vec{L}_2]|$

$$\leq |\vec{L}_2 \cdot [\vec{f}(x) - \vec{L}_1]| + |\vec{L}_1 \cdot [\vec{g}(x) - \vec{L}_2]| + |[\vec{f}(x) - \vec{L}_1] \cdot [\vec{g}(x) - \vec{L}_2]| .$$

(This last step is independent of any vector arithmetic since all dot products are numbers not vectors.)

Now since $|\vec{a} \cdot \vec{b}| \leq |\vec{a}||\vec{b}|$ (observe that we don't need $|\vec{a} \cdot \vec{b}| = |\vec{a}||\vec{b}|$), our last inequality yields

$$|\vec{f}(x) \cdot \vec{g}(x) - \vec{L}_1 \cdot \vec{L}_2| \leq |\vec{L}_2||\vec{f}(x) - \vec{L}_1| + |\vec{L}_1||\vec{g}(x) - \vec{L}_2| + |\vec{f}(x) - \vec{L}_1| |\vec{g}(x) - \vec{L}_2|$$

(Notice that there are no dot products here since $|\vec{L}_2|$, $|\vec{f}(x) - \vec{L}_1|$ etc. are numbers.)

Our last equation has the same properties as (1) and so the desired result follows.

Additional examples are left for the exercises. Our aim in this section was simply to illustrate that our limit theorems for scalars do indeed carry over, and in a rather natural way, to vectors.

E

Derivatives of Vector Functions of Scalar Variables

In the study of ordinary calculus, we mentioned that the study of differentiation could be viewed as an application of the limit theorems. That is, when we defined f' by

$$f'(x_1) = \lim_{\Delta x \to 0} \left[\frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} \right] \tag{1}$$

every property of f' was then computed by the use of the limit theorems applied to (1).

The point is that we now do the same thing for vectors. We observe that the expression

$$\frac{\vec{f}(x_1 + \Delta x) - \vec{f}(x_1)}{\Delta x}$$

is well-defined since it involves the quotient of a vector and a scalar and this may ve viewed as scalar multiplication. That is,

$$\frac{\vec{f}(x_1 + \Delta x) - \vec{f}(x_1)}{\Delta x} = \underset{\substack{\uparrow \\ \text{scalar}}}{\frac{1}{\Delta x}} \left[\vec{f}(x_1 + \underset{\substack{\uparrow \\ \text{vector}}}{\Delta x}) - \vec{f}(x_1)\right] .$$

Without belaboring this point, this expression can indeed be thought of as representing an average rate of change, and, if we then proceed to the limit, we may think of this as an instantaneous rate of change.

With this as motivation, all we are saying is that it is meaningful to define $\vec{f}'$ by

$$\vec{f}'(x_1) = \lim_{\Delta x \to 0} \left[\frac{\vec{f}(x_1 + \Delta x) - \vec{f}(x_1)}{\Delta x}\right] .$$

The key point now is that since the usual recipes for derivatives follow from the basic properties of limits and since the same limit theorems are true for vectors as for scalars, it follows that there will be recipes for vector derivatives similar to those for scalar derivatives.

For example, it is provable that the derivative of a sum is the sum of the derivatives. The proof is again verbatim from the proof in the scalar case.

Without bothering to copy the scalar proof, we write down the vector proof with the hope that you can recognize that it is the scalar proof with appropriate vectorization. We have

Let $\vec{h}(x) = \vec{f}(x) + \vec{g}(x)$.

Then $\vec{h}'(x_1) = \lim\limits_{\Delta x \to 0} \left[\dfrac{\vec{h}(x_1 + \Delta x) - \vec{h}(x_1)}{\Delta x}\right]$

$$= \lim_{\Delta x \to 0} \left[\frac{\vec{f}(x_1 + \Delta x) + \vec{g}(x_1 + \Delta x) - [\vec{f}(x_1) + \vec{g}(x_1)]}{\Delta x}\right]$$

$$= \lim_{\Delta x \to 0} \left[\left\{\frac{\vec{f}(x_1 + \Delta x) - \vec{f}(x_1)}{\Delta x}\right\} + \left\{\frac{\vec{g}(x_1 + \Delta x) - \vec{g}(x_1)}{\Delta x}\right\}\right] \quad \text{by vector arithmetic}$$

$$= \lim_{\Delta x \to 0} \left[\frac{\vec{f}(x_1 + \Delta x) - \vec{f}(x_1)}{\Delta x}\right] + \lim_{\Delta x \to 0} \left[\frac{\vec{g}(x_1 + \Delta x) - \vec{g}(x_1)}{\Delta x}\right]$$

(since the limit of a sum is the sum of the limits, as shown in the previous section).

$$= \vec{f}'(x_1) + \vec{g}'(x_1)$$

In a similar way (and the details are left as exercises), the same product rule applies to vectors. The only difference is that we now have several types of products. For example, we may have the product of a scalar function with a vector function, or we may have the dot product of two vector functions, or we may have the cross product of two vector functions. In any case, we obtain the results

$$\frac{d}{dx}[f(x)\,\vec{g}(x)] = f(x)\,\vec{g}'(x) + f'(x)\,\vec{g}(x)$$

$$\frac{d}{dx}[\vec{f}(x) \cdot \vec{g}(x)] = \vec{f}(x) \cdot \vec{g}'(x) + \vec{f}'(x) \cdot \vec{g}(x)$$

$$\frac{d}{dx}[\vec{f}(x) \times \vec{g}(x)] = [\vec{f}(x) \times \vec{g}'(x)] + [\vec{f}'(x) \times \vec{g}(x)]$$

(In this last case beware of changing the roles of the factors $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$.)

We even have a form of the quotient rule provided that our quotient is a vector function divided by a scalar function (for, otherwise, the quotient is not defined). The formula is

$$\frac{d}{dx}\left[\frac{\vec{f}(x)}{g(x)}\right] = \frac{g(x)\,\vec{f}'(x) - g'(x)\,\vec{f}(x)}{[g(x)]^2}\,.$$

These formulas have a nice application to Cartesian coordinates. Suppose $\vec{F}$ is a vector function of the single real variable t. For example, we may be considering force as a function of time. Suppose $\vec{F}$ has the form

$$\vec{F} = t\vec{i} + (3t^2 + 1)\,\vec{j} + e^t\,\vec{k}$$

and we desire to compute $d\vec{F}/dt$.

Since the derivative of a sum is the sum of the derivatives, we have

$$\frac{d\vec{F}}{dt} = \frac{d(t\vec{i})}{dt} + \frac{d([3t^2 + 1]\,\vec{j})}{dt} + \frac{d(e^t\,\vec{k})}{dt}\,.$$

Then since a constant times a variable has as its derivative the constant times the derivative of the variable, and since $\vec{i}$, $\vec{j}$, and $\vec{k}$ are constants, we obtain:

$$\frac{d\vec{F}}{dt} = (1)\vec{i} + (6t)\vec{j} + (e^t)\vec{k}$$

$$= \vec{i} + 6t\vec{j} + e^t\,\vec{k}\ .$$

While this was a specific example, it should be fairly easy to see that, in general, the derivative is obtained by differentiating the components. Further details are left to the text and the exercises, but we want to emphasize again the fact that our definition of limit for vectors, being so closely modeled after the scalar situation, guarantees that the differentiation formulas with which we are already familiar in the scalar case are also valid in the vector case.

MIT OpenCourseWare
http://ocw.mit.edu

Resource: Calculus Revisited: Multivariable Calculus
Prof. Herbert Gross

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.