Leibniz notation for vector fields

Canonical name	LeibnizNotationForVectorFields
Date of creation	2013-03-22 15:26:54
Last modified on	2013-03-22 15:26:54
Owner	stevecheng (10074)
Last modified by	stevecheng (10074)
Numerical id	13
Author	stevecheng (10074)
Entry type	Topic
Classification	msc 58A10
Classification	msc 34-00
Classification	msc 53-00
Related topic	VectorField
Related topic	DerivativeNotation
Related topic	DifferentialForms
Related topic	FirstOrderOperatorsInRiemannianGeometry
Related topic	LieDerivative

Consider a vector field V, such as

$$V(x, y) = (y, -x) \qquad \text{for } (x, y) \in \mathbb{R}^2.$$

Oftentimes V will be expressed in the following Leibnizian notation

$$y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y},$$

which is extremely strange the first time one encounters it. After all, why not just write

$$V(x, y) = y e_1 - x e_2$$

where $e_1 = (1, 0)$, $e_2 = (0, 1)$? What do partial derivatives have to do with anything? (That was the question I asked.)

However, if viewed in the right way, it is not so strange. For if a vector field V is given, then one natural thing that can be done with it is to differentiate a scalar-valued function f in the direction of V. In functional notation, this would be

$$\mathrm{D} f(p) \cdot V(p)$$

where p is a point, $V(p)$ is gives the particular vector in the vector field at the point p, and $\mathrm{D} f(p) \cdot v$ denotes the directional derivative of f at p with respect to the direction v.

For our example, $\mathrm{D} f(p) \cdot V(p)$ (where $p = (x, y)$) equals

$$\begin{aligned}
\mathrm{D}_1 f(x, y) \cdot y + \mathrm{D}_2 f(x, y) \cdot (-x) &= \frac{\partial f}{\partial x} y - \frac{\partial f}{\partial y} x \\
&= y \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial y} \\
&= \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) [f] = V[f].
\end{aligned}$$

We have written out the steps explicitly to show where the partial derivative notation for V comes from. At the second last step we considered $\partial/\partial x$ and $\partial/\partial y$ as operators acting on the function f.

The Leibniz notation on manifolds

But there is more. We can consider a more general situation, where V is a vector field on a manifold. In this case, because the tangent space to a

manifold varies with each point p, we cannot fix certain basis vectors e_i to describe our vector field anymore. Loosely speaking, the basis vectors now have to vary smoothly.

Suppose we have a coordinate system $\{x^i\}_{i=1,\ldots,n}$ on the manifold. Then the *tangent vector* on the manifold corresponding to an infinitesimal change in x^i is often written

$$\frac{\partial}{\partial x^i} .$$

This makes total sense, because one of the favorite ways to *define* tangent vectors on abstract manifolds is to identify them with directional derivatives. The basis $\partial/\partial x^i$ (which varies smoothly with p) then becomes the replacement for the fixed basis e_i in Euclidean space. Note that this is consistent with the calculus notation $\partial f/\partial x^i$ for the partial derivative of f with respect to the x^i variable, because partial derivatives are merely directional derivatives with respect to the direction e_i.

If $\{y^j\}$ is another coordinate system for the manifold, then we have the formula (for the derivation, see the entry on `http://planetmath.org/VectorField`vector fields)

$$\frac{\partial}{\partial y^j} = \frac{\partial x^i}{\partial y^j} \frac{\partial}{\partial x^i} , \tag{1}$$

We have used the Einstein summation convention above to emphasize the mnemonic cancelling of fractions.

The quantity $\partial x^i/\partial y^j$ is the directional derivative in the direction $\partial/\partial y^j$ of the ith coordinate *function*, mapping a point p on the manifold to the coordinate x^i. Notice the subtle subtle change of viewpoint here: we are not considering x^i as mere "variables", but as *functions* of the point p.

Formula (**??**), which is a *linear combination* of *vectors* in a tangent space, of course looks like the chain rule learned in elementary multivariate calculus, but it is much more than that: we are saying that the formula holds for general curvilinear coordinate systems on manifolds! This is definitely one of the virtues of the Leibnizian notation — making advanced concepts look simple.

As a simple example to get used to this notation, consider the function $f \colon \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x, y, z) = x^2 + y^2 + z^2$. The Euclidean space $\mathbb{R}^3$ can also be thought of as a manifold, and suppose we use a spherical coordinate system (r, θ, ϕ) on it.

Let us compute $\partial f/\partial r$. If f is "viewed as a function of (r, θ, ϕ)", then $f(r, \theta, \phi) = r^2$, so we certainly hope that $\partial f/\partial r = 2r$ with our definition of

2

directional derivatives. This easily follows from the formula for differential forms on manifolds:

$$df = \frac{\partial f}{\partial r}\,dr + \frac{\partial f}{\partial \theta}\,d\theta + \frac{\partial f}{\partial \phi}\,d\phi\,.$$

But let us see this by calculating from (**??**) too. We have

$$x = r\sin\theta\cos\phi\,,\quad y = r\sin\theta\sin\phi\,,\quad z = r\cos\theta\,.$$

So

$$\frac{\partial x}{\partial r} = \sin\theta\cos\phi\,,\quad \frac{\partial y}{\partial r} = \sin\theta\sin\phi\,,\quad \frac{\partial z}{\partial r} = \cos\theta\,,$$

and substituting in (**??**),

$$
\begin{aligned}
\frac{\partial f}{\partial r} &= \frac{\partial x}{\partial r}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial f}{\partial y} + \frac{\partial z}{\partial r}\frac{\partial f}{\partial z} \\
&= (\sin\theta\cos\phi)(2x) + (\sin\theta\sin\phi)(2y) + (\cos\theta)(2z) \\
&= 2r^{-1}(xr\sin\theta\cos\phi + yr\sin\theta\sin\phi + zr\cos\theta) \\
&= 2r^{-1}(x^2 + y^2 + z^2) \\
&= 2r\,.
\end{aligned}
$$

Needless to say, this calculation can be done with the usual functional notation, but it will be somewhat clumsy. We have to say: let α be the spherical coordinate chart; and then $D_1(f \circ \alpha)(r, \theta, \phi) = D\,f(\alpha(r, \theta, \phi)) \cdot D_1\,\alpha(r, \theta, \phi)$ would be the quantity $\partial f/\partial r$. That is not to say Leibnizian notation has no disadvantages. For example, one of the typical objections to the Leibnizian formula

$$\frac{df}{dx} = \frac{df}{dy}\frac{dy}{dx}$$

is that the function f means something different on the two sides of the equation. However, formula (**??**) partly escapes this objection: we can consider f to be a function on a *manifold*, ignoring the vector space structure of $\mathbb{R}^3$. Cartesian coordinates (x, y, z) simply become another coordinate chart. Spherical coordinates constitute another. Then $\partial f/\partial r$ (i.e. the directional derivative $\partial/\partial r$ applied to f) is a natural quantity to consider, rather than "the derivative of $f \circ \alpha$' with respect to the first variable" (which is what the functional notation says).

Physicists seem to grasp the Leibnizian formalism very readily, and it is a shame that many calculus textbooks do not fully explain the logic behind the formalism — probably because it looks to be unrigorous — but the point we are trying to drive here is that when differentials are suitably interpreted, they *are* rigorous.

The dual to the tangent vectors $\partial/\partial x^i$

There is a subtle ambiguity in the Leibniz notation that we should also discuss here. Suppose we are on a two-dimensional manifold with coordinates u and v. The notation $\partial f/\partial u$ is ambiguous, because it implicitly depends on the v coordinate as well. What we really mean when we refer to $\partial f/\partial u$ is a displacement where u changes at a uniform rate of 1, and where v does not change at all.

Say, for some bizarre reason, I decided to use a coordinate system on the Euclidean plane made up of the Euclidean x coordinate and the radial r coordinate. Now when I write $\partial f/\partial x$, I mean something quite different than when I write $\partial f/\partial x$ relative to the Euclidean coordinates. In the first instance the derivative is with respect to the vector field

$$e_1 - (y/x)e_2 \,.$$

In the second instance, the derivative is with respect to the vector field

$$e_1 + 0e_2 \,.$$

On closer thought, we can see that $\partial/\partial x$ in the elementary calculus interpretation has the same ambiguity, but the problem is so trivial that we often forget that it exists. For instance, if we have a function $f = xyz$, and we stipulate that also $y = x^2$, then obviously $\partial f/\partial x \neq yz$, because y is changing at the same time as x. The definition of a partial derivative with respect to x is the derivative when x changes and all the other variables are held *fixed*. So this rule should be applied when working with the tangent vectors $\partial/\partial x^i$ on a manifold too.

If we agree to use different letters for each coordinate system, and not mix them up (always a reasonable thing to do), then we will not make any mistakes arising from this ambiguity with the Leibniz notation.

Another way to understand the ambiguity is as follows. In a vector space, there is no natural isomorphism between it and its dual space (unless

we involve the inner product or something like that). On manifolds, the role of the dual space is taken by the space of differential one-forms. (See `http://planetmath.org/DifferentialForms`differential forms for the rigorous details.) A basis for this dual space is dx^j for $j = 1, \ldots, n$. There is a basis in the tangent space that is dual to dx^i: namely, these are the $\partial/\partial x^i$:

$$dx^j \left(\frac{\partial}{\partial x^i} \right) = \delta_i^j \quad \text{(Kronecker delta)}. \tag{2}$$

But if we are given a lone element dx^j, we cannot produce a unique vector $\partial/\partial x^j$ from it (i.e. there is no isomorphism) — we need to be given the *entire* basis $\{dx^j\}_{j=1,\ldots,n}$. So it is not surprising why the vectors $\partial/\partial x^i$ should depend on each other.

Motivation for the notation of differential forms

Incidentally, the formula (**??**) explains the following identity involving differential forms (often seen in calculus textbooks with hardly any explanation of what it means):

$$df = \frac{\partial f}{\partial x^j} \, dx^j \,. \tag{3}$$

The various d's floating around obscures the essential idea somewhat, but the derivation of this formula is basic linear algebra. Since df is a linear functional on the tangent space, *defined* by $df(V) = V[f]$, it can be written as a linear combination of the dual basis dx^j. That is, for some a_j, we have

$$df = a_j \, dx^j \,.$$

And these a_j are solved for by evaluating at the tangent vectors $V = \partial/\partial x^i$:

$$\frac{\partial f}{\partial x^i} = \frac{\partial}{\partial x^i}[f] = df \left(\frac{\partial}{\partial x^i} \right) = a_j dx^j \left(\frac{\partial}{\partial x^i} \right) = a_j \delta_i^j = a_i \,,$$

giving formula (**??**).

For those who are not familiar with the language of differential forms, the definition $df(V) = V[f]$ just used might seem to be somewhat artificial, designed solely to make the classical formula (**??**) work out. The following comment by Spivak[**?**] might help clarify matters:

Classical differential geometers (and classical analysts) did not hesitate to talk about "infinitely small" changes dx^i of the coordinates x^i, just as Leibniz had. No one wanted to admit that this was nonsense, because true results were obtained when these infinitely small quantities were divided into each other (provided one did it in the right way). Eventually it was realized that the closest one can come to describing an infinitely small change is to describe a direction in which this change is supposed to occur, i.e. a tangent vector. Since df is supposed to be the infinitesimal change of f under an infinitesimal change of the point, df must be a function of this change, which means that df should be a function on tangent vectors. The dx^i themselves then metamorphosed into functions, and it became clear that they must be distinguished from the tangent vectors $\partial/\partial x^i$. Once this realization came, it was only a matter of making new definitions, which preserved the *old* notation, and waiting for everybody to catch up. In short, all classical notions involving infinitely small quantities became functions on tangent vectors, like df, except for quotients of infinitely small quantities, which became tangent vectors, like dc/dt.

We can also give an analogy as follows (also from [**?**]). Suppose f is a function on a manifold. Let "$x^i = x^i(t)$" be a curve on this manifold (this is the classical notation). Then by the chain rule,

$$\frac{df}{dt} = \frac{\partial f}{\partial x^i}\frac{dx^i}{dt}\,,$$

where f on the left side really means $f(x^i(t))$. The formal identity obtained by multiplying both sides by dt,

$$df = \frac{\partial f}{\partial x^i}\,dx^i\,,$$

means that "true results are obtained by dividing by dt again, *no matter what the functions $x^i(t)$ are.*" Also, the left- and right-hand sides individually do not depend on any particular curve $x^i(t)$ at all, but really just the tangent vectors dx^i/dt to that curve. Again, this leads us to realization that df and dx^i should be treated as functions of a tangent vector.

References

[1] Vladimir I. Arnol'd (trans. Roger Cooke). *Ordinary Differential Equations*. Springer-Verlag, 1992.

[2] Michael Spivak. *A Comprehensive Introduction to Differential Geometry, Volume I*. Publish or Perish, 1979.

[3] Michael Spivak. *Calculus on Manifolds*. Perseus, 1965.