Infinitesimals Reloaded

There is no vacation

Michael P. Windham

April 28, 2005

Also available at

http://www.southalabama.edu/mathstat/personal_pages/windham/

1. Introduction

In *Where Have You Gone Infinitesimals* I described how intuitive notions of infinitesimals, the basis for calculus in the seventeenth century, were related to tensors, the ultimate calculus of the twentieth century. I bemoaned the fact that abstraction and rigor had replaced intuition and informality. I also mentioned at the end of the paper that I used infinitesimals all the time suggesting that they were actually alive and well. Alive and well may be stretching it, but useful in all their intuitive glory is not. Let me tell you how they can be used. Along the way it will get a little weird, but that will be a temporary state, I hope.

This paper is a sequel to *Where Have You Gone Infinitesimals*. It may be worth your while to read the first two and a half pages of that paper,[1] through the paragraph ending "I rest my case." If I need to refer to the paper again, I will call it *WHYGI*.

The way I use infinitesimals in teaching calculus is in integration. In one variable calculus, you have a function f and an interval $[a, b]$ and you define the Riemann integral

$$\int_a^b f(x)\,dx$$

I tell people, you can think of $f(x)$ as the height of an infinitesimally thin rectangle and dx as its width, so that $f(x)\,dx$ is its area. The integral just adds up these infinitesimal areas to give the total area under the graph of f over the interval $[a, b]$.

Of course, then you calculate $\int_{-1}^{1} x\,dx$ and get zero! So, what is going on? The fact of the matter is that the integral is not built to calculate area, it is built to build antiderivatives, that is, to solve the problem: given f, find F, so that $F' = f$. That the integral does these things is the content of the Fundamental Theorem of Calculus, which can be described in two ways.

- Recovering total change from instantaneous rates of change: $\displaystyle\int_a^b \frac{dy}{dx}\,dx = y(b) - y(a) = \Delta y$ for $\Delta x = b - a$, and
- Building antiderivatives: $F(x) = \displaystyle\int_a^x f(t)\,dt$ is an antiderivative of f.

These things are what I tell people about integration now, but if I lived in the seventeenth century I would do it differently. The Riemann integral as an adding up of infinitesimal quantities, I would not need to change. Recovering

[1] www.southalabama.edu/mathstat/personal_pages/windham/

change requires only a modification of the phraseology, but little else. In *WHYGI* I observed that differentials are a primary source for infinitesimals, that is, if y is a variable, the differential of y is the infinitesimal change in y per infinitesimal change in x, given by

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

so that, recovering the total change in y is simply $\int_a^b dy = \int_a^b \frac{dy}{dx}\,dx = y(b) - y(a) = \Delta y$. The antiderivative problem looks a good deal different from my modern formulation. It becomes given $f(x)\,dx$, find y, so that $dy = f(x)\,dx$.

What I really want to tell you is how to do these things in R^n, but not the usual way people are told in Calculus III, where n is never bigger than three, ten different integrals are defined, five different theorems that are really the Fundamental Theorem of Calculus are given, and three different antiderivative problems are discussed. Or, the way people are told in more advanced mathematics where one finds exterior algebra, differential forms, integration on chains, and something called Stokes' Theorem. I will tell you how to do both of these, from a reasonably elementary approach using infinitesimals. We begin with the Riemann integral.

2. Riemann again

Consider, if you will, the geometry of a point. That won't take long you say. Ah, but it is not a simple as you think. In R^3, for instance, a point can be a three dimensional object, if it is part of a solid, a two dimensional object, as part of a surface, a one dimensional object, as part of a curve, and a zero dimensional object in its own right. When you integrate a function over a set, you evaluate the function at a point, multiply by the infinitesimal size of the point and add up the results for each point in the set. The size of the point depends on the dimension of the set you are integrating over. So, how big is a point?

The good thing about a point is that its size is infinitesimal, so that you can throw away higher order terms when you calculate the size. What that really means is that you can think of a point as a box with *straight* sides, not necessarily perpendicular to each other, but at least parallel.

So, let me leave the infinitesimal world for a minute and look at the volume of a p dimensional box in R^n. I should call it a parallelepiped, but I will just call it a box. What you call the size depends on its dimension, but I will call it volume for the moment. So, let me get it over with. If $b_1, \ldots, b_p$

are vectors in R^n, letting $B = [b_1 \dots\ b_p]$, the $n \times p$ matrix with the vectors as columns, the volume of the p dimensional box obtained by using the vectors as adjacent edges, vol(B), is

$$\mathrm{vol}(B) = \sqrt{\det(B^T B)} \tag{1}$$

In R^3 for $p = 1$ the box is a single vector. The equation says that the volume of the box is the length of the vector, $|b_1|$. For $p = 2$, the equation gives $(|b_1|^2|b_2|^2 - (b_1 \cdot b_2)^2)^{\frac{1}{2}} = |b_1||b_2| \sin\theta$ where θ is the angle between the two vectors, that is, the length of the base times the height. Finally, for $p = 3$, both sides are just $|\det(B)|$. If you have not seen this calculation before, it really is what it should be. The equation says that the volume of a p dimensional box is the volume of its $p-1$ dimensional base times its height. I leave that as an exercise in linear algebra to verify.

Now back to the point - in R^3. A p dimensional set is made up of p dimensional points, at least, that is how we are going to think about it. So, a curve is a one dimensional set and the size of a point would be its length. The most natural way to quantitatively describe a curve is with a parameterization $(x(t), y(t), z(t))$ - a description of a one dimensional object by one parameter, t, that tells you how to draw it. Drawing one point on the curve is done with an infinitesimal displacement dt that produces an infinitesimal displacement vector along the curve

$$\langle dx, dy, dz\rangle = \left\langle \frac{dx}{dt}\,dt, \frac{dy}{dt}\,dt, \frac{dz}{dt}\,dt \right\rangle = \left\langle \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right\rangle dt$$

so that the length of the point according to (1) is

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\,dt$$

A surface can be dealt with in an entirely analogous manner. Draw the surface using a parameterization $(x(u, v), y(u, v), z(u, v))$, then infinitesimal displacements in each of the parameters produce an infinitesimal displacement along the point on the surface

$$\begin{aligned}\langle dx, dy, dz\rangle &= \left\langle \frac{\partial x}{\partial u}\,du + \frac{\partial x}{\partial v}\,dv, \frac{\partial y}{\partial u}\,du + \frac{\partial y}{\partial v}\,dv, \frac{\partial z}{\partial u}\,du + \frac{\partial z}{\partial v}\,dv \right\rangle \\ &= \left\langle \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right\rangle du + \left\langle \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right\rangle dv \end{aligned} \tag{2}$$

suggesting that the point is a two dimensional box whose sides are the two infinitesimal displacement vectors in (2), and that the area of the point dA according to (1) is the length of the cross product of the displacements.

Finally, a three dimensional point would have volume $dV = dxdydz$ the product of the lengths of its sides, $\langle dx, 0, 0\rangle$, $\langle 0, dy, 0\rangle$, and $\langle 0, 0, dz\rangle$. Note that (1) still works. In fact, it works if you draw the point with a three dimensional parameterization $(x(u, v, w), y(u, v, w), z(u, v, w))$, which in context of integration we call a change of variables (sound familiar?). Using displacements in the parameters again, a displacement across the point is the sum of three vectors, the sides of a three dimensional box, with volume[2]

$$dV = \left| \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix} \right| dudvdw$$

That, dear reader, is the geometry of a point in R^3, at least with regard to its size.

A general formula for the volume of a p dimensional point in R^n, is now within our grasp. You draw the point with a parameterization, decompose the vector $\langle dx_1, \ldots, dx_n\rangle$ into a linear combination of p vectors each of which is multiplied by a single du_i and apply (1) to those vectors to get an infinitesimal amount of volume $dV = \mathrm{vol}(J)\, du_1 \ldots du_p$, the p dimensional volume of the point, where

$$J = \left[\frac{\partial x_i}{\partial u_j}\right]$$

is the $n \times p$ Jacobian matrix of the parameterization.

As I have already said, the geometry of size is used in integration of functions. In general, it is simply the following.

For an integral over an n dimensional set in R^n, use the Riemann integral.

For a p dimensional set, S, parameterize the set using p variables. A p dimensional point is the box obtained from displacements in each of the parameters and its volume dV is computed by (1). The integral of a function f over S becomes the function in terms of the parameters, times the volume

[2] I have used the standard but unfortunate notation ds, dA, and dV for infinitesimal *amounts* of size. The notation is unfortunate because it suggests a differential, an infinitesimal change in a variable and that is definitely not the case here. If A is a variable that represents area, say $A = xy$, then $dA = ydx + xdy$ is the change in area due to a change in the sides. On the other hand, $dA = dxdy$ is not a change in area but an infinitesimal amount of area. We have a name for this, it is called an abuse of the notation. But, this one has been going on so long, most people don't even know they are being abused.

of a point in R^n in terms of the parameters, added up over the set in the parameter space that draws S in R^n, $\int_S f(x)\,dV$.

I can even take it further. A point in its own right is a zero dimensional set. I define the integral of a function f over a point x to be $f(x)$. BAM! If you are having fun now, wait till you see what comes next.

3. Taking it up a notch or two

Let's take a look at some happenings in R^2. The first happening is recovering change. If w is a variable then $dw = \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy$ is the infinitesimal change in w due to infinitesimal changes in x and y. If you want to recover the total change in w as you move from point A to point B, then, first of all you would need to describe how you get from A to B. That's easy enough you draw a curve using a parameterization, so that for the parameter going through the interval $[a, b]$, the curve is drawn from A to B. The variable w along the curve is now a function of t, so that the logical way to recover the total change in w would be

$$\int_a^b \frac{dw}{dt}\,dt = w(B) - w(A)$$

which using the chain rule can be written

$$\int_a^b \left(\frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \right) dt$$

We get two things from these observations beyond the recovery of change. First, the integral suggests a way of integrating any infinitesimal $\omega = P\,dx + Q\,dy$ along a curve C from A to B, namely, parameterize the curve and define

$$\int_C \omega = \int_a^b \left(P\frac{dx}{dt} + Q\frac{dy}{dt} \right) dt$$

It might occur to you that drawing a curve from A to B is somehow different from drawing the same curve from B to A. In fact, it is different and the effect on the integral is to change the sign. Other than that, the parameterization does not matter, as a change of variables in the integral would show. When you say which end of the curve is the beginning point and which is the end point, you have chosen what has called an *orientation* for the curve. Actually you can specify an orientation by simply choosing the parameterization.

The other thing we get is just a reiteration. I can say that the Fundamental Theorem of Calculus can be written

$$\int_C dw = w(B) - w(A) = \int_{\{B\}} w - \int_{\{A\}} w$$

Big deal, but humor me. Let me read it this way: The value an integral of the differential of a variable on the set is determined by the values of the variable on the boundary of the set (the boundary of the curve being its end points). One amazing consequence of this statement is that is says the total change in w in going from A to B does not matter how you get from A to B. That is amazing, so the question is are there other versions of the philosophy that "the values of some kind integral of a differential over a set are determined by the values of what it is the differential of on the boundary of the set." Well, watch this.

Before you do let me simplify the notation a bit. We are about to be bombarded by partial derivatives, so, let me use the other notation, namely $w_x = \frac{\partial w}{\partial x}$ and save a lot of paper and typing.

Now watch. Suppose I have an infinitesimal $\omega = P\,dx + Q\,dy$ consider the set $C = \{(x,y) : 0 \le x \le 1, 0 \le y \le 1\}$, the unit square in R^2, then the boundary of the square I can parameterize easily as curves using u and v. In particular, if I call the boundary ∂C, then

$$\int_{\partial C} \omega = \int_0^1 P(x,0)\,dx + \int_0^1 Q(1,y)\,dy - \int_0^1 P(x,1)\,dx - \int_0^1 Q(0,y)\,dy$$

which parameterizes the boundary of the square counterclockwise. The minus signs have the effect of reversing the integral so that direction of travel around the boundary is continuous. Continuing

$$\begin{aligned}
\int_{\partial C} \omega &= \int_0^1 (Q(1,y) - Q(0,y))\,dy - \int_0^1 (P(x,1) - P(x,0))\,dx \\
&= \int_0^1 \left(\int_0^1 Q_x(x,y)dx \right) dy - \int_0^1 \left(\int_0^1 P_y(x,y)dy \right) dx \\
&= \int_C (Q_x - P_y)\,dA
\end{aligned}$$

where I have used the one-variable Fundamental Theorem of Calculus. This result fits the basic philosophy of integrating some derivatives of something over a set being determined by the values of the something on the boundary of the set. Of course the unit square is a pretty simple set, but that is not a problem.

If D is the image of the unit square using a function $(u, v) \to (x, y)$, then ∂D would be the image of ∂C, usually, anyway. Suppose $\omega = P\,dx + Q\,dy$ is defined on D. Since, $dx = x_u\,du + x_v\,dv$ and $dy = y_u\,du + y_v\,dv$, we have $\omega = P\,dx + Q\,dy = (Px_u + Qy_u)\,du + (Px_v + Qy_v)\,dv$, so that,

$$
\begin{aligned}
\int_{\partial D} \omega &= \int_{\partial C} (Px_u + Qy_u)\,du + (Px_v + Qy_v)\,dv \\
&= \int_0^1 \int_0^1 ((Px_v + Qy_v)_u - (Px_u + Qy_u)_v)\,dudv \\
&= \int_0^1 \int_0^1 (Q_x - P_y)(x_u y_v - x_v y_u)\,dudv \qquad (3)
\end{aligned}
$$

The last line looks very much like $\int_D (Q_x - P_y)\,dA$ after a change of variables from (x, y) to (u, v), except that $dA = |x_u y_v - x_v y_u|\,dudv$. Some absolute value bars seem to be missing. To get out of this gracefully I need to make a long-winded, but fascinating digression.

In R^2, take two infinitesimals, $\alpha = a_1\,dx + a_2\,dy$ and $\beta = b_1\,dx + b_2\,dy$, at the point (x, y). An algebraically natural way of multiplying them together could begin something like this.

$$
\begin{aligned}
\alpha\beta &= (a_1\,dx + a_2\,dy)(b_1\,dx + b_2\,dy) \\
&= a_1 b_1\,dxdx + a_1 b_2\,dxdy + b_1 a_2\,dydx + a_2 b_2\,dydy
\end{aligned}
$$

The second line in the calculation assumes that the multiplication is distributive and commutes with the coefficients, or put another way, it is linear in each of the pieces, $\alpha(b_1\,dx + b_2\,dy) = b_1\,\alpha\,dx + b_2\,\alpha\,dy$. The next step would be to simplify. One perfectly reasonable way to simplify would be

$$\alpha\beta = a_1 b_1\,dx^2 + (a_1 b_2 + b_1 a_2)\,dxdy + a_2 b_2\,dy^2$$

but in doing so we have assumed that $dydx = dxdy$, that is, the multiplication is commutative. I tacitly made that assumption when I said things like $dA = dxdy$ when talking about the Riemann integral. I want to explore exactly the opposite assumption. I want to assume that the multiplication is *anticommutative*, that is, $dydx = -dxdy$. Anticommutativity implies that $dxdx = -dxdx$, so that $dxdx = dydy = 0$. Now simplify using the anticommutativity and you get

$$\alpha\beta = (a_1 b_2 - a_2 b_1)\,dxdy$$

That's interesting. It will come as no surprise that this is not my idea. This multiplication has a name, the *exterior* product, and even a symbol $\alpha \wedge \beta$ to distinguish it from assuming commutativity. So, we write

$$\alpha \wedge \beta = (a_1 b_2 - b_1 a_2)\, dx \wedge dy$$

It also looks like we have created a new kind of infinitesimal, a "second order" infinitesimal, $\gamma = c\, dx \wedge dy$, and we have. There would be no third order infinitesimal in R^2. If we follow the rules, a third order infinitesimal built from dx and dy, would always have at least one of them repeated, so would be zero.

Now back to (3). We could define for a second order differential $\gamma = g\, dx \wedge dy$ an integral,

$$\int_D \gamma = \int_D g\, dx \wedge dy = \int_C g(x_u y_v - x_v y_u) du dv$$

and all we would have to do to remember how to calculate it is to substitute $x_u\, du + x_v\, dv$ for dx, and $y_u\, du + y_v\, dv$ for dy and integrate the coefficient of the $du \wedge dv$ over C using the ordinary Riemann integral. I gets even better. Since $dP = P_x\, dx + P_y\, dy$ and $dQ = Q_x\, dx + Q_y\, dy$, we have using exterior arithmetic

$$\begin{aligned} dP \wedge dx + dQ \wedge dy &= (P_x\, dx + P_y\, dy) \wedge dx + (Q_x\, dx + Q_y\, dy) \wedge dy \\ &= (Q_y - P_x) dx \wedge dy \end{aligned}$$

So, for a first order infinitesimal $\omega = P\, dx + Q\, dy$ define the *differential* of ω to be $d\omega = dP \wedge dx + dQ \wedge dy$ and (3) becomes the following Fundamental Theorem of Calculus

$$\int_D d\omega = \int_{\partial D} \omega$$

This is so much fun, let's do it again, except this time instead of drawing a set in R^2 with the unit square, draw a 2-dimensional surface D in R^3. We could call the boundary of ∂D the image of ∂C. The orientation ∂D is determined by the counter-clockwise orientation of ∂C.

If you have an infinitesimal $\omega = P\, dx + Q\, dy + R\, dz$ then the technique for integrating along an oriented curve in R^2 generalizes to integrating over an oriented curves in R^3 in the obvious way.[3] All that changes in first line of (3)

[3] In fact, in R^n we can define for $\omega = \sum w_i\, dx_i$ and an oriented curve c the integral of *omega* along c to be $\int_c \omega = \int_a^b \sum w_i \frac{dx_i}{dt}\, dt$ where t is the parameter of a parameterization that draws the curve in the direction defined by the orientation.

is that we have $Px_u + Qy_u + Rz_u$ instead of $Px_u + Qy_u$ and $Px_v + Qy_v + Rz_v$ instead of $Px_v + Qy_v$. The bottom line becomes

$$\int_{\partial D} \omega = \int_0^1 \int_0^1 (Q_x - P_y)(x_u y_v - x_v y_u) + (R_x - P_z)(x_u z_v - x_v z_u) + (R_y - Q_z)(y_u z_v - y_v z_u)\, du dv$$

which is not nearly as bad as it looks if we do exterior arithmetic in R^3 exactly as we did in in R^2.

In particular, a second order infinitesimal in R^3 is $\beta = L\, dx \wedge dy + M\, dx \wedge dz + N\, dy \wedge dz$, the the integral of β over D is defined to be

$$\int_D \beta = \int_0^1 \int_0^1 L(x_u y_v - x_v y_u) + M(x_u z_v - x_v z_u) + N(y_u z_v - y_v z_u)\, du dv \quad (4)$$

which is easy to remember by just substituting for dx, dy, and dz, the expressions in terms of du and dv and simplifying with exterior arithmetic. Moreover, defining $d\omega = dP \wedge dx + dQ \wedge dz + dR \wedge dz$ as usual, (4) becomes

$$\int_D d\omega = \int_{\partial D} \omega$$

yet another Fundamental Theorem of Calculus.

We have a pattern here. Before I spell it out in all its glory let me do one more example, except that I will use a famous mathematical technique. Write down the answer and see if there is a question to go with it.

We have first and second order infinitesimals in R^3, what about third order infinitesimals? Why not. They are simple, $\omega = w\, dx \wedge dy \wedge dz$. The simplicity follows from the anticommutativity rule, which leads to the fact that if you multiply three of dx, dy, and dz together, then the result would be zero if any two are the same and $\pm\, dx \wedge dy \wedge dz$ otherwise. But, that is it. A fourth or higher order would be zero since at least one of dx, dy, or dz would appear more than once.

We have the differential of a variable and the differential of a first order infinitesimal, what about differentials for the higher order infinitesimals? Why not. For $\omega = L\, dx \wedge dy + M\, dx \wedge dz + N\, dy \wedge dz$ let

$$\begin{aligned} d\omega &= dL \wedge dx \wedge dy + dM \wedge dx \wedge dz + dN \wedge dy \wedge dz \\ &= (L_z - M_y + N_x)\, dx \wedge dy \wedge dz \end{aligned}$$

For higher order infinitesimals the differential would be zero.

Integrating third order differentials over three dimensional sets should go this way. Suppose D is a three dimensional set in R^3, that is the image of a function $(u, v, w) \to (x, y, z)$ defined on the unit cube $C = \{(u, v, w) : 0 \leq u, v, w \leq 1\}$, then ∂D is the image of ∂C. Since, $dx = x_u\, du + x_v\, dv + x_w\, dw$, and similarly for dy and dz,

$$dx \wedge dy \wedge dz = \det \begin{bmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{bmatrix} du \wedge dv \wedge dw$$

so that

$$\int_D \omega = \int_D w\, dx \wedge dy \wedge dz = \int_0^1 \int_0^1 \int_0^1 w \det \begin{bmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{bmatrix} dudvdw$$

So, what is the Fundamental Theorem of Calculus? That's easy.

$$\int_D d\omega = \int_{\partial D} \omega$$

but what of that calculation do I not know how to do? Not much, but something. Integrating over ∂D is done by integrating over the six faces of C, but these six integrals must be combined so that the everything fits when I use the one variable Fundamental Theorem of Calculus. It boils down to picking which ones should be positive and which negative and I am not going to tell you how to do it - yet. I am not being sadistic, I am sparing you a great mess that is much easier to describe in general than in this special case and a general description is the next step. The fact of the matter is that there is a theorem.

4. The big picture

Suppose R^n has a coordinate system with points described by $(x_1, \ldots, x_n)$ and dx_i is an infinitesimal change in the coordinate x_i, then a *first order infinitesimal* is $\omega = \sum w_i dx_i$ where w_i is a variable measuring an amount per unit change in x_i.

For $p > 1$ a *p-th order infinitesimal* in R^n is a formal sum

$$\sum w\ \omega_1 \wedge \ldots \wedge \omega_p$$

where the coefficients w are variables and $\omega_1, \ldots, \omega_p$ are any first order infinitesimals. The sums are manipulated with the following basic rules of *exterior* arithmetic.

- $(a\,\omega + b\,\sigma) \wedge \omega_2 \wedge \ldots \wedge \omega_p = a\,\omega \wedge \omega_2 \wedge \ldots \wedge \omega_p + b\,\sigma \wedge \omega_2 \wedge \ldots \wedge \omega_p$
- $\omega_1 \wedge \ldots \wedge \omega_p = 0$, if any two of the ω_i's are the same
- $\omega_1 \wedge \ldots \wedge \omega_p$ changes sign if any two of the ω_i's are interchanged

A little thought will show you that these rules say $\omega_1 \wedge \ldots \wedge \omega_p$ is linear in each factor and that any p-th order infinitesimal is a linear combination of the infinitesimals $dx_{i_1} \wedge \ldots \wedge dx_{i_p}$ where $1 \le i_1 < \ldots < i_p \le n$. In fact this observation is so useful I will streamline the notation for it. Denote by I a p-tuple of indices $(i_1, i_2, \ldots, i_p)$ where $1 \le i_1 < i_2 < \ldots < i_p \le n$. There are $\binom{n}{p} = n!/(p!(n-p)!)$ possibilities. Also, let $dx_I = dx_{i_1} \wedge \ldots \wedge x_{i_p}$. With this notation a p-th order infinitesimal can be written

$$\omega = \sum w_I \, dx_I$$

All we have really done here is concoct a slick way of calculating and manipulating determinants. In fact for $\omega_j = \sum w_{ij} dx_i$ we have $\omega_1 \wedge \ldots \wedge \omega_p = \sum w_I dx_I$ where the coefficient w_I is the determinant of the $p \times p$ sub-matrix obtained from rows $i_1, \ldots, i_p$ of the $n \times p$ matrix $W = [w_{ij}]$. This fact is not hard to believe if you think about some of the famous properties of the determinant. It is linear in each of the columns or rows, zero if two columns or rows are the same, and changes sign if you interchange two columns or rows - which are exactly the basic rules of exterior arithmetic.

Finally, we finish off the arithmetic by defining a multiplication of a p-th order infinitesimal $\omega = \sum w_I dx_I$ and a q-th order infinitesimal $\sigma = \sum s_J dx_J$ to be

$$\omega \wedge \sigma = \sum \sum w_I s_J dx_I \wedge \, dx_J$$

The multiplication acts exactly like a multiplication should except that it is not commutative. In fact, $\sigma \wedge \omega = (-1)^{pq} \omega \wedge \sigma$.

Now for the calculus of infinitesimals.

Differentiation is easy enough. The *differential* of a variable w is what it always has been, $dw = \sum \frac{\partial w}{\partial x_i} dx_i$ and describes the infinitesimal change in w due to infinitesimal change in the x_i's. The *differential* of a p-th order infinitesimal $\omega = \sum w_I \, dx_I$ is

$$d\omega = \sum dw_I \wedge dx_I$$

So, one could say it is an infinitesimal change in ω due to infinitesimal changes in its coefficients. For convenience I will make statements for p-th order infinitesimals that will make sense for variables if you use $p = 0$ in the statement. In particular, a 0-th order infinitesimal is a variable.

Naturally, you can obtain the rules of the game for manipulating the differential operation. If α and β are p-th order infinitesimals, then $d(\alpha+\beta) = d\alpha+d\beta$ and if γ is a q-th order infinitesimal, then $d(\alpha\wedge\gamma) = d\alpha\wedge\gamma+(-1)^p\alpha\wedge d\gamma$.

Integration will come as no surprise. If D in R^n is a p-dimensional surface parameterized by a function of parameters $u_1, \ldots, u_p$ defined on the unit cube $C_p = \{(u_1, \ldots, u_p) : 0 \leq u_1, \ldots, u_p \leq 1\}$ and ω is a p-th order infinitesimal, then substitute $\sum_i \frac{\partial x_j}{\partial u_i} du_i$ for dx_j in ω and simplify to get $w\, du_1 \wedge \ldots \wedge u_p$ the resulting p-th order infinitesimal on C_p, then

$$\int_D \omega = \int_0^1 \ldots \int_0^1 w(u_1, \ldots, u_p) du_1 \ldots du_p$$

If D is the union of a finite number of such sets, $D_1, \ldots, D_k$, that overlap only on sets of dimension less than p, then $\int_D \omega = \sum \int_{D_i} \omega$. Finally, if w is a 0-th order infinitesimal, that is a variable, and $D = \{x\}$ is a 0-dimensional set, that is a point, the $\int_D w$ is defined to be $w(x)$ as I have already done.

A Fundamental Theorem of Calculus? Of course, it is

$$\int_D d\omega = \int_{\partial D} \omega$$

where all I have to tell you is how to integrate over ∂C_{p+1} since it draws ∂D. Just to keep it simple let me look a C_p instead. The p dimensional cube C_p has $2p$ faces that are copies of C_{p-1}. They can be parameterized using $p-1$ of the u_i's at a time as follows. For $i = 1, \ldots, p$ and $\alpha = 0$ or 1, let

$$C_{i,\alpha} = \{(u_1, \ldots, u_{i-1}, \alpha, u_{i+1}, \ldots, u_p) \text{ in } C_p\}$$

then

$$\int_{\partial C} \omega = \sum (-1)^{i+\alpha} \int_{C_{i,\alpha}} \omega$$

Vóila!

For what it's worth we can pose an antiderivative problem as well. As a matter of fact it is worth a great deal, but I'll explain the deal later. The antiderivative problem is:

> Given a p-th order infinitesimal ω, find a $(p-1)$-st order infinitesimal σ so that $d\sigma = \omega$.

In R the problem always has a solution, that is what the Fundamental Theorem of Calculus says. But for R^n, in general it does not. In particular for a p-th order infinitesimal $d(d\omega) = 0$. Just look at $p = 0$ and you will see that the result is just a statement of the equality of mixed partial derivatives. The result follows for $p > 0$, since the differential of an infinitesimal is determined by the differentials of its coefficients. So, if $d\omega \neq 0$, then ω is not somebody's differential.

On the other hand, if $d\omega = 0$ on a suitably nice set like an open ball in R^n or R^n itself, then it is somebody's differential. In fact you can even write a formula for the solution, but it is so disgusting that I refer you to Michael Spivak's *Calculus on Manifolds* for the details.

Well, that is the big picture.

There are a couple of things I need to confess. All of this has been around for many years. People do not use the term p-th order infinitesimal, they say p-th order differential form. I do not like the term because it suggests that you have the differential of something, but you may not. Those people call what I have called the differential the exterior derivative, so they are not confused, at least among themselves.

The purists among you are no doubt disgusted with me completely ignoring hypotheses and would claim that much of what I said is not true. So, the hypotheses are whatever it takes to make true what should be true.

There is a problem with what are called non-orientable sets. An example is the Möbius strip, a two dimensional surface in R^3 drawn with the unit square by $x = \cos(2\pi u)(1+(v-\frac{1}{2})\cos(\pi u))$, $y = \sin(2\pi u)(1+(v-\frac{1}{2})\cos(\pi v))$, and $z = (v-\frac{1}{2})\sin(\pi u)$. Essentially you glue opposite sides of the unit square together after you give it a half-twist. The problem is that if you want to integrate a second order infinitesimal, say $dx \wedge dy$ then you are faced with the embarrassing fact that $dx \wedge dy = \pi du \wedge dv$ and $dx \wedge dy = 3\pi du \wedge dv$ at the same point, namely (1,0,0). You get the first answer when you draw the point with $u = 0$ and $v = \frac{1}{2}$ and the second when you draw it again with $u = 1$ and $v = \frac{1}{2}$. So, you should only integrate over orientable surfaces, which essentially means you do not get different values for the same infinitesimal at the same point on the surface. The details are a bit technical but the description of an orientable set can be made precise, some other time.

I have one more thing to do here - for the record. I need to bring these infinitesimals into the modern world. Differential forms fit into the modern world mathematically, if not intuitively. A p-th order differential form ω is nothing more than a type $(0, p)$ tensor[4] with one additional property. It is

[4]To get a clue about what tensors are, go to *WHYGI*.

Infinitesimals	Vector fields	Infinitesimals	Vector fields
Linear combinations	Linear combinations	dw	∇w (gradient)
$\alpha \wedge \beta$	$a \times b$	$d\alpha$	curl a
$\alpha \wedge *\beta$	$a \cdot b$	$d*\alpha$	div a
$\int_D \omega$	Iterated integrals	$\int_D d\omega = \int_{\partial D} \omega$	Green's Theorem
	Line integrals		Stokes' Theorem
	Surface integrals		Gauss' Theorem

Table 1: Multivariable calculus in a nutshell

completely antisymmetric, that is, for tangent vectors $V_1, \ldots, V_p$,

$$\omega(V_1, \ldots, V_i, \ldots, V_j, \ldots, V_p) = -\omega(V_1, \ldots, V_j, \ldots, V_i, \ldots, V_p)$$

for all i and j. The antisymmetry gives you the anticommutativity and that's all it takes to get the ball rolling.

The previous section was meant to show you how one might come upon this big picture. The next section tells you how it relates to the picture everyone sees.

5. Calculus III

So, how does this big picture relate to what we do in multivariable calculus courses? The Riemann integral of a function over a p-dimensional is defined the same way in both.

For the rest, look in R^3. In the calculus course you talk about variables (functions) and vector fields. Suppose w is a variable, and $a = \langle a_1, a_2, a_3 \rangle$ and $b = \langle b_1, b_2, b_3 \rangle$ are vector fields. From a vector field you can build two infinitesimals, that is, from a you get $\alpha = a_1\, dx + a_2\, dy + a_3\, dz$ and $*\alpha = a_1\, dy \wedge dz + a_2\, dz \wedge dx + a_3\, dx \wedge dy$. With these associations you have Table 1.

By the way, the formula $d(d\omega) = 0$, includes $\text{curl}(\nabla w) = 0$ and $\text{div}(\text{curl } a) = 0$, as well.

In fact, I have the impression that things like the curl and divergence and surface integrals were built to have a Fundamental Theorem of Calculus. For example, how do you explain what the curl means? You look at a limit something like this. A fluid is flowing through space and you integrate its velocity vector around a circle of radius r to see how the fluid is circulating. Then apply Stokes' Theorem using the disc that the circle encloses, divide the integral over the disc by its area, let r go to zero and out pops the curl as a measure of the circulation per unit area. It is a vector, so that it has some direction associated to it that allows you to determine the axis about which the circulation is the greatest. It would be interesting to know which came first, Stokes' Theorem or the curl. Let me know what you discover.

6. Infinitesimals reloaded

Vectors fields, gradients and curls, and so on, have well-established interpretations in the real world and students need to know about them. What real world interpretations are there for higher order infinitesimals and their differentials. To put it another way, if the people in the seventeenth century had had them, would they have been able to use them in a natural way to understand the world around them. When I started writing this paper, I think I had in mind that I wanted to revive the intuitive appeal of infinitesimal quantities. So, here is my attempt to do so.

We have two slightly different notions of infinitesimal. The first is the function times the volume element and the second is the differential form.

The first is very easy to use intuitively, and I already have. The values of the function measure some quantity per unit volume, the "density" of the quantity. The volume element is an infinitesimal amount of volume. So, multiplying the two together gives an infinitesimal amount of the quantity. Integration just adds them up to get a finite amount. This type of infinitesimal is what you might call "static", it is just an amount.

Differential forms, my p-th order infinitesimals, should also represent infinitesimal amounts, but they have an aura of direction about them. Consider one variable calculus. The equation $dy = f(x)\,dx$ measures the infinitesimal change in y due to an infinitesimal change in x, and whether $f(x)$ is positive or negative is significant. If $f(x) > 0$, then y does the same thing that x does, that is, increasing x increases y. If $f(x) < 0$, then x and y move in opposite directions.

There is a sense of direction for infinitesimals in any R^n. An infinitesimal change in a variable $dw = \sum \frac{\partial w}{\partial x_i}\,dx_i$ is easy enough to imagine, but what if the coefficients are not partial derivatives? I have been saying all along that you have an infinitesimal amount of something, but what? It is, in fact, hard to say. In truth you have something that has a magnitude and a direction and you want to measure its effect relative to other directions, which you get by looking at curves. So, you get something concrete by "dividing through by dt". In other words, for $\omega = \sum w_i\, dx_i$, the meaningful quantity is

$$\frac{\omega}{dt} = \sum w_i \frac{dx_i}{dt} \tag{5}$$

the component of the vector $w = \langle w_1, \ldots, w_n \rangle$ in the direction of the tangent vector to a curve at a point on the curve. So, you could say that a first order infinitesimal is just another way of writing a vector field. You don't take the infinitesimal too seriously until you divide through by dt. Vector fields are a reasonably natural thing. After all, they represent things with magnitude and direction in a natural, geometrical way. The velocity field of a fluid flowing through space, the force due to gravity, an electric field, friction $\ldots$, all are comfortably described by vector fields. And, vector fields are real, not some imaginary "infinitesimal."

You could say a p-th order infinitesimal is some "quantity" with magnitude and a "direction" that you measure the effect of simultaneously in perhaps several directions, determined by a surface. The real meaning of the quantity is in the coefficients of the infinitesimal and the form just organizes the coefficients in a convenient way so that you use them to your advantage. Well, I am not giving up on the infinitesimal idea yet, but there is definitely something to this vector field thing, so, let me go on with it for a bit.

The meaningful quantity you get from the infinitesimal is the coefficient you get when you parameterize a surface and substitute $\sum \frac{\partial x_i}{\partial u_j}\, du_j$ for dx_i. In particular, for p-th order infinitesimal $\omega = \sum w_I\, dx_I$, and a p-dimensional surface parameterized by $u_1, \ldots, u_p$, the relevant quantity is

$$\sum w_I J_I$$

where J_I is the determinant of the $p \times p$ sub-matrix of the Jacobian matrix $J = \left[\frac{\partial x_i}{\partial u_j}\right]$ whose rows are specified by I. You then integrate the quantity over the surface to get something wonderful. That is,

$$\int_D \omega = \int_C \sum w_I J_I\, du_1 \ldots du_p = \int_C \sum w_I \frac{J_I}{\mathrm{vol}(J)}\, dV \tag{6}$$

The infinitesimal amount relative to the surface is $\sum w_I J_I / \mathrm{vol}(J)\, dV$.

The second part of the equation is very interesting. To show you why, I need to call upon the first of a few (obscure) facts from the world of determinants.[5] If B and C are $n \times p$ matrices with $p \leq n$, then

$$\det(C^T B) = \sum C_I B_I$$

[5] I am beginning to feel a little guilty about all of these obscure facts about determinants. Maybe I will put a document on my home page that goes into them. OK, I did it.

where I is the usual ordered p-tuple of indices and for an $n \times p$ matrix A, A_I is the determinant of $p \times p$ sub-matrix of A whose rows are given by I. In particular, $\mathrm{vol}(J) = \sqrt{\sum J_I^2}$. You could put the coefficients J_I into a vector $\langle \ldots, J_I, \ldots \rangle$ and $\mathrm{vol}(J)$ would be its length. You could put the coefficients of ω into a vector, then $\sum w_I J_I / \mathrm{vol}(J)$ would be the dot product of $\langle \ldots, w_I, \ldots \rangle$ with a unit vector $\langle \ldots, J_I, \ldots \rangle / \mathrm{vol}(J)$ obtained from tangent vectors to the surface, the columns of J. Another thing to notice is that if you change the parameterization, then $\langle \ldots, J_I, \ldots \rangle / \mathrm{vol}(J)$, at most only changes sign, so that it only depends on the orientation of the parameterization, not the details. Let me spell that out a little more. If $(x_1, \ldots, x_n)$ is drawn by $u_1, \ldots, u_p$ with one parameterization and $\bar{u}_1, \ldots, \bar{u}_p$ with another parameterization of the surface, then $\langle \ldots, J_I, \ldots \rangle / \mathrm{vol}(J) = \pm \langle \ldots, \bar{J}_I, \ldots \rangle / \mathrm{vol}(\bar{J})$ when evaluated at the point. It would seem prudent, then, to see what this vector $\langle \ldots, J_I, \ldots \rangle / \mathrm{vol}(V)$ might have to do with life.

OK, another adventure in determinants is required. If B and C are $n \times p$ matrices and the columns of B are linearly independent, then $P = B(B^T B)^{-1} B^T C$ is a matrix whose columns are the projections of the columns of C onto the plane spanned by the columns of B, a not so obscure fact from linear algebra. The box built from columns of P would be the projection of the box built from the columns of C onto the plane built from the columns of B and its volume is

$$\mathrm{vol}(P) = \frac{|\det C^T B|}{\mathrm{vol}(B)} = \left| \sum C_I \frac{B_I}{\mathrm{vol}(B)} \right|$$

This takes a little caculation, but not much. If you put J in for B and throw away the absolute value, it begins to look like what you are integrating in (6). What does throwing away the absolute value bars have to do with anything? It maintains a sense of direction. You are calculating an oriented volume, in the sense that the directions of the vectors in B and C relative to each other effect the sign of the result.

So, a p-th order infinitesimal is a quantity that has an effect *along* p-dimensional surfaces that depends on the orientation of the surface. The effect is realized at a point on the surface by projecting the quantity onto tangent vectors to the surface at the point, so to speak the tangential component of the quantity along the surface.

The infinitesimal can also be represented as a "vector field" defined at the points in R^n. Why, you ask, do I keep putting "vector field" in quotes? These guys are not the vector fields we refer to when we use the phrase in multivariable calculus. For one thing the are $\binom{n}{p}$ dimensional objects. Now the rules of exterior arithmetic say that the p-th order infinitesimals are a vector space over the reals when evaluated at a point and over the variables in general.[6] These "vector fields" technically represent an isomorphism between the p-th order infinitesimals and the usual vector fields on $R^{\binom{n}{p}}$. Of course, if $p = 1$, then the dimension of the p-th order infinitesimals is n and we have been using them to accomplish some of the things the usual vector fields on R^n are used for, but I want, even in that case, to maintain a distinction. The reason is the next part of the adventure.

You may have noticed that I also said that you are measuring the effect of the p-th order infinitesimal *along* the p-dimensional surface. You are projecting the quantity onto the p-dimensional tangent plane at each point on the surface. But, the impact of the quantity at a point on a p-dimensional surface would also be the impact of the quantity in the direction orthogonal to an $(n-p)$-dimensional surface orthogonal to the p-dimensional surface at the point. The impact is *along* the p-dimensional surface and *across* the $(n-p)$-dimensional surface. However, if you are to use this with an $(n-p)$-dimensional surface, you would begin by parameterizing the surface, so that you have the tangent vectors to the surface. You would like to calculate a projection on the p-dimensional plane orthogonal to the surface using the tangent vectors you have. To show you how to do this, I need to take you on one, final excursion into the world of obscure facts about determinants.

This time we have B is an $n \times (n-p)$ matrix with independent columns, and C is an $n \times p$ matrix. Now, for each ordered p-tuple I, there is an ordered $(n-p)$-tuple $I' = (i'_1, \ldots, i'_{n-p})$ containing the indices not in I. Let ε_I be determined as follows: if $i_1, \ldots, i_p, i'_1, \ldots, i'_{n-p}$ can be rearranged into $(1, \ldots, n)$ with an even number of interchanges of adjacent indices, then $\varepsilon_I = 1$, otherwise $\varepsilon_I = -1$. For example, if $I = (3)$ and $I' = (1, 2)$, then $(3, 1, 2) \to (1, 3, 2) \to (1, 2, 3)$, so that $\varepsilon_I = 1$, but if $I = (2)$ and $I' = (1, 3)$, then $(2, 1, 3) \to (1, 2, 3)$, so that $\varepsilon_I = -1$. Now for an obscure fact, let $[C\,B]$ be the $n \times n$ matrix built from the columns of C and B, then

$$\det[C\,B] = \sum \varepsilon_I C_I B_{I'}$$

Actually, this fact is believable enough. What it says is the determinant of a matrix is the sum of the determinants of pairs of matrices one of which is obtained from p rows and columns of the matrix and the other from what

[6] OK, a module over the variables for you purists, but I'm playing this fast and loose. I am after the ideas and do not want to be confused by stating all of the hypotheses. You can get them by reading a real book on the subject. In fact, if you read *WHYGI* you may recognize that these guys are really tensor fields on R^n.

you have left by eliminating those rows and columns, together with -1's thrown in appropriately. In fact, for $p = 1$, the calculation is the one used to define the determinant recursively from the entries in its first column and the corresponding $(n-1) \times (n-1)$ minors. Not so bad, but the next one...

Build an $n \times p$ matrix $B_\perp$ whose columns linearly independent and orthogonal to the columns of B so that $\det[B_\perp\, B] > 0$, then

$$\frac{\det C^T B_\perp}{\mathrm{vol}(B_\perp)} = \frac{\det[C\, B]}{\mathrm{vol}(B)}$$

the last of the obscure facts. What it says is that $\sum \varepsilon_I C_I B_{I'}$ is the volume of the box built from the columns of C projected onto the p-dimensional plane orthogonal to the plane spanned by the columns of B. Bingo. In fact, there are no absolute values so it is an oriented volume determined by B. What does this have to do with infinitesimals?

If the p-th order infinitesimal $\omega = \sum w_I\, dx_I$ projects the quantity tangentially along p-dimensional surfaces, then the $(n-p)$-th order infinitesimal

$$*\omega = \sum \varepsilon_I w_I\, dx_{I'}$$

projects the quantity orthogonally across $(n-p)$-dimensional surfaces. For example, a vector field $v = \langle v_1, v_2, v_3 \rangle$ in R^3 is projected tangentially along curves with $\omega = v_1\, dx + v_2\, dy + v_3\, dz$ and orthogonally across 2-dimensional surfaces by $*\omega = v_3\, dx \wedge dy - v_2\, dx \wedge dz + v_1\, dy \wedge dz$, which is usually written $*\omega = v_1\, dy \wedge dz + v_2\, dz \wedge dx + v_3\, dx \wedge dy$, so that the coefficients appear in the order they appear in the vector field. Going from ω to $*\omega$ is called applying the Hodge * operator[7] to ω. If you look at where I used the $*$ notation earlier, I was sneaking in the Hodge * operator at the time.

So far I have talked about what a p-th order infinitesimal looks like, but how would you actually come upon them in the real world, the seventeenth century world to be precise, which is, after all, what I wanted to do all along. So, try this on for size.

You have a fluid flowing through space.[8] The density of the fluid is given by $\rho(x, y, z, t)$ and the velocity through space is $v(x, y, z, t) = \langle v_1, v_2, v_3 \rangle$.

[7] The Hodge * depends on what you mean by orthogonal. I have used the usual euclidean dot product. If you were in spacetime and used the Minkowski metric the * operator would look different.

[8] I have to thank Merrill Jenkins in Physics for loaning me a book that told me about fluid flow. Harley Flanders' book *Differential Forms with Applications to the Physical Sciences* also gave me some ideas.

Mathematically we are in R^4, but physically we are in R^3, so visualize what is happening in the three space dimensions. When I speak of a point it will be a point in space with coordinates (x, y, z) and a rectangular box whose sides have infinitesimal lengths dx, dy, and dz.

The volume of the fluid in a point is obtained as follows. Since fluid is flowing through the point in each of the coordinate directions the net amount along the x direction would be $dx - v_1\,dt$, similarly of the other two directions, so that the infinitesimal volume of fluid in the point would be

$$\begin{aligned}
\nu &= (dx - v_1\,dt) \wedge (dy - v_2\,dt) \wedge (dz - v_3\,dt) \\
&= dx \wedge dy \wedge dz - (v_1\,dy \wedge dz + v_2\,dz \wedge dx + v_3\,dx \wedge dy) \wedge dt \\
&= *(dt + v_1\,dx + v_2\,dy + v_3\,dz)
\end{aligned}$$

where the first line says that because of the nature of the situation, namely "flow", it makes sense to use an oriented volume to incorporate the relation between the direction the fluid is flowing and the coordinate system we chose to locate points in space. The other two lines are the result of exterior arithmetic, including the Hodge * operator, but provide other ways of interpreting the infinitesimal volume. For example, the infinitesimal $v_1\,dx + v_2\,dy + v_3\,dz$ is the spacial velocity field of the fluid, and its * is $v_1\,dy \wedge dz + v_2\,dz \wedge dx + v_3\,dx \wedge dy$, the flow per unit time of the fluid across the faces of the point, that is, out of the point. That interpretation is so nice I will give the infinitesimal a name, $\phi = v_1\,dy \wedge dz + v_2\,dz \wedge dx + v_3\,dx \wedge dy$. So, $\nu = dx \wedge dy \wedge dz - \phi \wedge dt$.

If you have a region D in R^3, then at a fixed point in time you would have $dt = 0$, so that $\int_D \nu = \int_D dx \wedge dy \wedge dz$ is just the volume of D, hence the volume of fluid in D at that point in time.[9] If you have an interval of time T, then $\int_{\partial D \times T} \nu = -\int_{\partial D \times T} \phi \wedge dt$ is the volume of the fluid that has left the region across the boundary during the time interval T.

The change in volume is

$$\begin{aligned}
d\nu &= -d(\phi \wedge dt) \\
&= -d\phi \wedge dt \\
&= -\mathrm{div}(v)\,dx \wedge dy \wedge dz \wedge dt
\end{aligned}$$

[9]Actually you would be integrating over a three dimensional surface in R^4 where the parameterization has t constant so that the dt terms really do disappear. In what you are about to see the $dx \wedge dy \wedge dz$ really goes away because the x, y, and z variables are dependent so that the part of the Jacobian corresponding to $dx \wedge dy \wedge dz$ vanishes. So, I am using an abuse of the notation, at least, but the result is correct.

where $\mathrm{div}(v) = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}$ the usual spacial divergence of a vector field. The formula $d\nu = -d\phi \wedge dt$ suggests that we can define the "rate of change of volume per unit time" by

$$\frac{d\nu}{dt} = -d\phi = -\mathrm{div}(v)\, dx \wedge dy \wedge dz$$

So, if D is a region in R^3

$$\int_D \frac{d\nu}{dt} = -\int_D d\phi = -\int_{\partial D} \phi$$

which says that the volume of fluid in a region decreases at a rate equal to the amount crossing the surface of the region per unit time. Sounds good!

The mass of the fluid in the point would be $\mu = \rho\nu$. The infinitesimal change in mass would be

$$\begin{aligned} d\mu &= d\rho \wedge \nu + \rho d\nu \\ &= -\left(\frac{\partial \rho}{\partial t}\, dx \wedge dy \wedge dz + d(\rho\phi)\right) \wedge dt \\ &= -\left(\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho v)\right)\, dx \wedge dy \wedge dz \wedge dt \end{aligned}$$

Now, mass is conserved in nature, so that $d\mu = 0$, which says nothing more than what is called in fluid dynamics the *continuity equation* of fluid flow

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho v) = 0$$

We also have that the "rate of change in mass per unit" time is zero, so that

$$\frac{d\mu}{dt} = -\left(\frac{\partial \rho}{\partial t} dx \wedge dy \wedge dz + d(\rho\phi)\right) = 0$$

and for a region D in space

$$\int_D \frac{\partial \rho}{\partial t}\, dx \wedge dy \wedge dz = \frac{d}{dt}\int_D \rho\, dx \wedge dy \wedge dz = -\int_{\partial D} \rho\phi$$

which says the mass in D is decreasing at the rate that it is flowing out across the boundary of ∂D.

All this infinitesimal stuff leads to sensible facts!

But wait, $d\mu = 0$ means that $\mu = d\omega$ for some second order infinitesimal ω. You can think about that one.

7. Fini

The fluid mechanics example does it. It develops infinitesimals intuitively, gets more with exterior arithmetic and exterior derivatives and uses the Fundamental Theorem of Calculus to understand things. Infinitesimals, I love 'em.