Proof of Stokes' Theorem PhL 8.11.09

This is the second of two theorems which underlie all the other integral theorems in my "integral theorems.doc". The other theorem is the Divergence Theorem which I proved in a separate document.

CONTENTS

1. Statement of Stokes' Theorem 1

2. Proof of Stokes' Theorem 2

3. Comments on the curl. 5

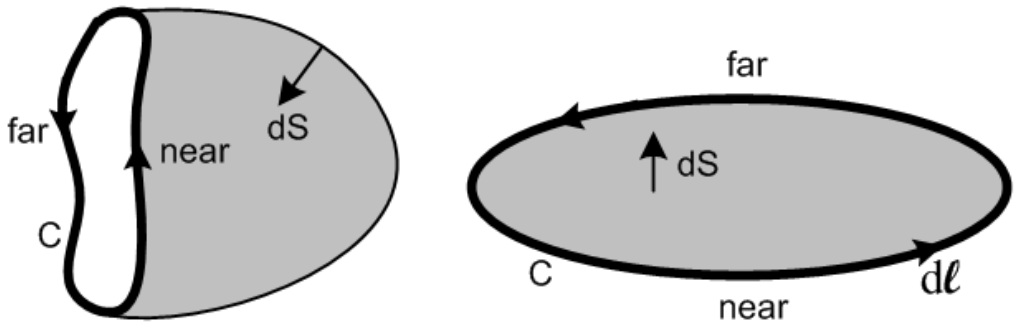
4. History of the two basic integral theorems. 5

1. Statement of Stokes' Theorem

My interest in this theorem is limited to its application in 3-dimensional space, call it 3D. It can be generalized to higher dimensions, but I am not going to do that here because I have had no use (yet) for such a generalized theorem. So here is the 3D Stokes' Theorem:

Stokes' theorem (for me) $$\int d\mathbf{S} \bullet \nabla x \mathbf{A} \;=\; \oint_C d\boldsymbol{\ell} \bullet \mathbf{A}$$

In this theorem, **A**(**x**) is some arbitrary vector field. We have a closed contour C which exists in 3D space and is in general not contained in a plane (though it could be). The surface integral is over any smooth surface that terminates on the contour C in what I would call "the obvious manner". Think of the contour C as a child's plastic (possibly non-planar) bubble ring, and think of the surface as a "half bubble" that terminates on this bubble ring. Here on the left is a crude picture of such a situation,



We show on the right a simpler situation where the contour C is just a circle which we are viewing from above at some angle of elevation, and where the surface is the disk bounded by the circle. Obviously if the contour and the surface were stretchy wire and rubber, you could gradually deform from the case on the right to the case on the left, something that is probably important in the general topological study of this theorem.

The labels near and far attempt to cue the viewer how to see the picture. The arrows on the contour indicate the sense (direction) of the contour integration appearing in the theorem, while the arrow marking

the vector differential area d**S** shows its direction. The relationship is given by the venerable "right hand rule": put fingers in the sense of the contour, then the thumb points in the direction of d**S .** Thus, in our situation on the left, d**S** is pointing in, not out. Of course we could redraw the picture reversing all three arrows.

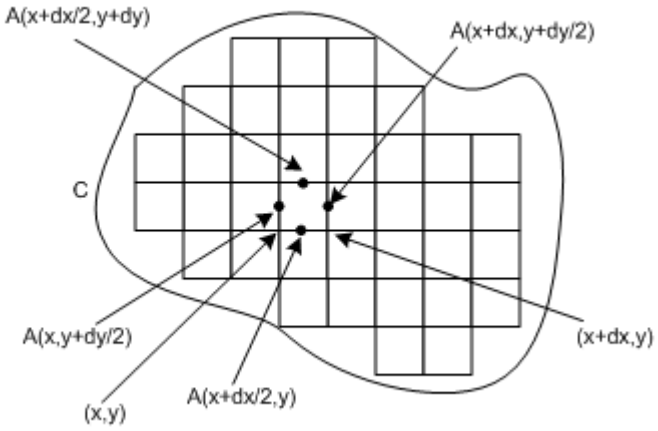
There is no reason for the bound surface to lie entirely on one side of the contour as suggested in the first figure. It could wave back and forth between the two sides.

It is important to note that Stokes' theorem involves what one would call an "open surface", whereas the d**S** integral in the divergence theorem $\int dV \nabla \bullet \mathbf{A} = \int d\mathbf{S} \bullet \mathbf{A}$ is over a "closed surface" that completely encloses a volume over which $\nabla \bullet \mathbf{A}$ is integrated.

2. Proof of Stokes' Theorem

In our proof, we are going to "tile" the entire open surface with little squares, and we are going to make the tiles get very small. When this is done, any small piece of the surface can be regarded as being flat. That is to say, at some point on the surface, we model the surface by its tangent plane at that point. At that point, we make a little local coordinate system x,y,z with x to the right, y up, and z out to the viewer. Obviously the surface must be "smooth" for this to be possible, it cannot contain weird delta function spikes for example (at least not for our purposes). As we move around on the bound surface, this x,y,z coordinate system rotates as needed to align with the local tangent plane.

So, all we really need consider is a small patch of surface that we regard as planar, and we draw our little squares. In fact, we can even "steal" or divergence theorem picture to this end.



However, we need to think of this now as a schematic picture because surface and the contour are both curved and in general non-planar. However, for the set of squares we focus in on at the center of the picture, we regard the picture as being accurate and flat. For these tiny squares, the paper is the tangent plane. The x axis is to the right, the y axis is up, and the z axis points to the viewer (right handed coordinate system). The dimensions of the squares are dx and dy, and the vector surface area points out to the viewer, so that $dS_z = dxdy\,\hat{\mathbf{z}}$ **.** Due to this fact, we have $d\mathbf{S} \bullet \nabla \times \mathbf{A} = dS_z\,[\nabla \times \mathbf{A}]_z$, so we only care about the z component of the curl (of course in our little "local" x,y,z coordinates). We can write this as:

$$[\nabla x \mathbf{A}]_z(x,y) = \partial_x A_y(x,y) - \partial_y A_x(x,y) = [A_y(x+dx/2,y) - A_y(x-dx/2,y)]/dx + [A_x(x,y+dy/2) - A_x(x,y-dy/2)]/dy$$

We can then express the curl at the center of our square as (eg, replace $x \rightarrow x + dx/s$, etc)

$$[\nabla x \mathbf{A}]_z(x+dx/2,y+dy/2) = [A_y(x+dx,y+dy/2) - A_y(x,y+dy/2)]/dx + [A_x(x+dx/2,y+dy) - A_x(x+dx/2,y)]/dy$$

Looking at the labels on the above figure, we see that we have a similar situation that we had in our proof of the 2D divergence theorem. We use this shorthand,

$$[\nabla x \mathbf{A}]_z(\text{center}) = [A_y(\text{right}) - A_y(\text{left})]/dx + [A_x(\text{top}) - A_x(\text{bottom})]/dy$$

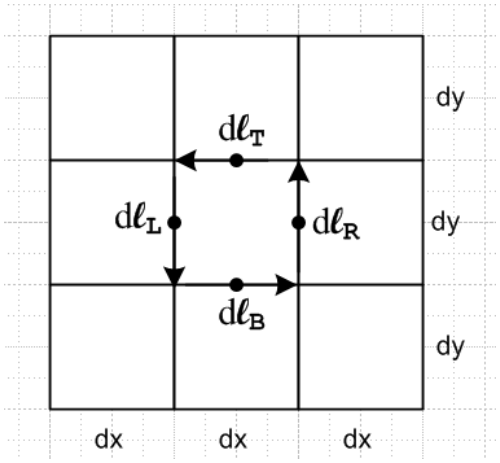
Now multiply through by dxdy to get

$$dxdy\, [\nabla x \mathbf{A}]_z(\text{center}) = [A_y(\text{right}) - A_y(\text{left})]\, dy + [A_x(\text{top}) - A_x(\text{bottom})]\, dx \quad (*)$$

At this point, we introduce four little directed vectors for the edges of our central square:

$d\boldsymbol{\ell}_R$ $d\boldsymbol{\ell}_L$ $d\boldsymbol{\ell}_T$ $d\boldsymbol{\ell}_B$ for example, $d\boldsymbol{\ell}_R = dy\, \hat{\mathbf{y}}$ and $d\boldsymbol{\ell}_L = -dy\, \hat{\mathbf{y}}$

and here is a blow-up of the central region showing how these are defined:



Using these four vector differential distances, we can rewrite our equation (*) above as

$$dxdy\, [\nabla x \mathbf{A}]_z(\text{center}) = A_y(\text{right})dy - A_y(\text{left})dy + A_x(\text{top})dx - A_x(\text{bottom})dx$$

$$= \mathbf{A}(\text{right}) \bullet d\boldsymbol{\ell}_R + \mathbf{A}(\text{left}) \bullet d\boldsymbol{\ell}_L + \mathbf{A}(\text{top}) \bullet d\boldsymbol{\ell}_T + \mathbf{A}(\text{bottom}) \bullet d\boldsymbol{\ell}_B$$

Notice that all the signs are now plus. We can of course rewrite this in more compact notation:

$$dxdy\ [\nabla x \mathbf{A}]_z(\text{center}) = \oint_{\text{center square}} d\boldsymbol{\ell} \bullet \mathbf{A}$$

where we use the traditional CCW sense of the contour. This then is the contribution to the LHS of Stokes' theorem from this one little square. We write it again as follows:

$$d\mathbf{S} \bullet [\nabla x \mathbf{A}](\text{center}) = \oint_{\text{center square}} d\boldsymbol{\ell} \bullet \mathbf{A} \qquad \text{since } dS_z = dxdy\ \hat{\mathbf{z}}$$

Now suppose we ask about the contribution of all 9 squares in the picture above. We add up 9 copies of the above algebra, and of course the main thing we notice is that "things cancel" on all internal edges of the squares. For example, the contribution along $d\boldsymbol{\ell}_R$ of the right side of the center square (which contribution is $\mathbf{A}(\text{right}) \bullet d\boldsymbol{\ell}_R$) will be exactly cancelled by the contribution $d\boldsymbol{\ell}_L$ from the square to the right of the central square. The value of **A** at the dot is the same, but the distance is "up" for the central square's contribution, and "down" for the contribution of the square to the right.

The upshot is that the contribution from all 9 squares is this:

$$\Sigma_{\text{nine squares}}\ d\mathbf{S} \bullet [\nabla x \mathbf{A}] = \oint_{\text{boundary around the nine squares}} d\boldsymbol{\ell} \bullet \mathbf{A}$$

So we see that we can mark off the contribution of each small group of squares by a CCW line integral around the boundary of that group. We can then do this for the entire surface. Yes, the surface is curved, but for any group of squares of interest, we use the a local coordinate system which we call x,y,z . It should be fairly clear that the main idea is that we get cancellation along internal square edges, and this is true whether or not the surface is completely flat, or is "gently curved". We can even imagine our little squares being deformed to exactly match the surface if we like, so the each "square" edge is now slightly curved. The squares are then four-edged surface patches. Fine. The upshot is the same. As we make the squares very very small, we end up with the total LHS being a jagged contour which approximates the boundary curve C.

So, we have now arrived at our discrete result which is

$$\Sigma_{\text{all inside squares}}\ d\mathbf{S} \bullet [\nabla x \mathbf{A}] = \oint_{\text{jagged boundary around the all squares}} d\boldsymbol{\ell} \bullet \mathbf{A}$$

and then we take the limit as the squares get very small and we have

$$\int d\mathbf{S} \bullet [\nabla x \mathbf{A}] = \oint_C d\boldsymbol{\ell} \bullet \mathbf{A} \qquad \text{Stokes' Theorem has been proved!}$$

Now we can address the same question that came up in our derivation of the 2D divergence theorem, which theorem also produces $\oint_C d\boldsymbol{\ell} \bullet \mathbf{A}$ as its RHS. Strictly speaking, our outer contour is a jagged contour with only "vertical and horizontal" pieces, and this is true no matter how small we make the squares, and there is always some "unfilled" space around the edges. At the end of the 2D divergence section we show why it is that, despite the presence of this small unfilled space, we can replace the jagged

contour integral with a smooth one which goes exactly around the boundary. Since we made the argument there, there is no reason to repeat it here.

3. Comments on the curl.

As in the divergence case, I think I have done all of the above for general curvilinear coordinates based on M&M stuff, but let's not worry about that here where we are in Cartesian coordinates.

Looking above at the contribution of the central square, it seems clear that the "curl" is a property of a vector field $\mathbf{A}(x)$ at the point $\mathbf{x}$ such that if you take a tiny contour integral of $\mathbf{A}\bullet d\boldsymbol{\ell}$ in some plane containing that point, you might get a non-zero result. The curl is a vector, $\nabla x\mathbf{A}$, and if you want to know the size of the curl in the $\hat{\mathbf{n}}$ direction, you run your little contour integral in a circle perpendicular to that direction and centered at the point $\mathbf{x}$ (and related to it by the right hand rule).

The quantity $d\mathbf{S} \bullet [\nabla x\mathbf{A}]$ will be a maximum when $d\mathbf{S}$ points in the direction of $[\nabla x\mathbf{A}]$. Thus, the curl points in the direction in which the little perpendicular-plane contour integral $\oint_C d\boldsymbol{\ell} \bullet \mathbf{A}$ is a maximum. This is certainly a very strange property of a vector field. The curl was originally called "the rotation" and was written Rot($\mathbf{A}$). One often sees pictures showing the vector mapping with little arrows of a vector field $\mathbf{A}$ which has "curl" in some direction. And these pictures are compared to other similar pictures showing a field that has "divergence", which, however, is a scalar quantity.

4. History of the two basic integral theorems.

Wiki claims this theorem was "exposed" in 1850 by a certain Bill Thomson, but Sir George Stokes (1819-1903) got his name associated with the theorem in 1854. Both knew Maxwell and it was in this era the Maxwell's Equations appeared, perhaps gradually in the range 1860-1870, and of course both the curl and divergence appear in these equations. Stokes was a big shot at Cambridge and is the first person that was the famous Lucasian chair of physics, president of the Royal Society, and member of Parliament, all at the same time. He worked a lot with other hot-shots Maxwell and Lord Kelvin. Probably his most famous work relates to fluid dynamics. I am not quite sure who in the present day one would think of as being a similar big shot in math and physics. There are some particle people names that come to mind, but it is not so clear what these people have actually succeeded in doing.

As for the divergence theorem, here is some wiki on that:

"It was first discovered by Joseph Louis **Lagrange** in 1762, then later independently rediscovered by Carl Friedrich **Gauss** in 1813, by **George Green** in 1825 and in 1831 by Mikhail Vasilievich **Ostrogradsky**, who also gave the first proof of the theorem. Subsequently, variations on the Divergence theorem are called Gauss's Theorem, Green's theorem, and Ostrogradsky's theorem."

So this theorem predates Stokes' theorem by anywhere from 90 to 20 years, depending on how you want to measure things.