

Vector Curl

by Patrick Rutkowski, Columbia University

www.patrick-rutkowski.com

[Discovery (2D)]

Let us consider a vector function $\mathbf{F}(x, y)$ of two-dimensional space. Place a square at the origin with width Δx and height Δy . The lower left and upper right corners will be at (x_0, y_0) and (x, y) . Let the square's center be at the origin. The midpoints of the square's sides will be $x_m = 0$ and $y_m = 0$. We could write these midpoints using the literal number 0, but we will keep the notation x_m and y_m to keep the meaning explicit.

Our goal is to find an approximate answer for the line integral of \mathbf{F} over this square. We can do this by sampling the vector function at the midpoints of the sides:

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx [\mathbf{F}(x_m, y_0) \cdot (\Delta x, 0)] + [\mathbf{F}(x, y_m) \cdot (0, \Delta y)] + [\mathbf{F}(x_m, y) \cdot (-\Delta x, 0)] + [\mathbf{F}(x_0, y_m) \cdot (0, -\Delta y)]$$

We can rearrange this approximation to

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx - [\mathbf{F}(x_m, y) - \mathbf{F}(x_m, y_0)] \cdot (\Delta x, 0) + [\mathbf{F}(x, y_m) - \mathbf{F}(x_0, y_m)] \cdot (0, \Delta y)$$

Next we can take a first order approximation for the difference terms

$$\begin{aligned} \mathbf{F}(x, y_m) - \mathbf{F}(x_0, y_m) &\approx \frac{\partial \mathbf{F}}{\partial x} (x_0, y_m) \cdot \Delta x \\ \mathbf{F}(x_m, y) - \mathbf{F}(x_m, y_0) &\approx \frac{\partial \mathbf{F}}{\partial y} (x_m, y_0) \cdot \Delta y \end{aligned}$$

Given the above we can now write

$$\begin{aligned} \oint \mathbf{F} \cdot d\mathbf{c} &\approx - \left[\frac{\partial \mathbf{F}}{\partial y} \Delta y \right] \cdot (\Delta x, 0) + \left[\frac{\partial \mathbf{F}}{\partial x} \Delta x \right] \cdot (0, \Delta y) \\ &= \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \Delta x \Delta y \end{aligned}$$

This combination of partial derivatives is called the curl. For the curl to be well defined at (x_m, y_m) the functions $\partial F_y / \partial x$ and $\partial F_x / \partial y$ should be well defined at the point (x_m, y_m) , since this is the point where we end up evaluating the partials as the differential square shrinks. At the very least, the partials should approach some sensible value in the limit as (x, y) goes to (x_m, y_m) .

[Discovery (3D)]

We now wish to generalize the previous analysis to three dimensions. We need to formulate an expression for vector rotation. We want to take a vector \mathbf{r} and rotate it by angle θ about some unit axis $\hat{\boldsymbol{\omega}}$. You can verify for yourself that the following expression serves exactly this purpose:

$$\mathbf{R}(\mathbf{r}) = -\hat{\boldsymbol{\omega}} \times (\hat{\boldsymbol{\omega}} \times \mathbf{r}) \cdot \cos \theta + (\hat{\boldsymbol{\omega}} \times \mathbf{r}) \cdot \sin \theta + \hat{\boldsymbol{\omega}} \cdot \mathbf{r}$$

This function is complicated, but fortunately we will not need to manipulate it much in the work to come. We will only need to refer to its general properties.

We can now use this function \mathbf{R} to calculate the path integral over a square that is rotated in some arbitrary way (and not just in the xy -plane). What's more, \mathbf{F} can now be a function of three inputs $\mathbf{F}(x, y, z)$:

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx \left[\mathbf{F}(\mathbf{R}(x_m, y_0, 0)) \cdot \mathbf{R}(\Delta x, 0, 0) \right] + \left[\mathbf{F}(\mathbf{R}(x, y_m, 0)) \cdot \mathbf{R}(0, \Delta y, 0) \right] + \left[\mathbf{F}(\mathbf{R}(x_m, y, 0)) \cdot \mathbf{R}(-\Delta x, 0, 0) \right] + \left[\mathbf{F}(\mathbf{R}(x_0, y_m, 0)) \cdot \mathbf{R}(0, -\Delta y, 0) \right]$$

Note that \mathbf{R} obeys the property

$$\mathbf{R}(-\mathbf{r}) = -\mathbf{R}(\mathbf{r})$$

We can therefore rearrange the approximation just like before:

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx - \left[\mathbf{F}(\mathbf{R}(x_m, y, 0)) - \mathbf{F}(\mathbf{R}(x_m, y_0, 0)) \right] \cdot \mathbf{R}(\Delta x, 0, 0) + \left[\mathbf{F}(\mathbf{R}(x, y_m, 0)) - \mathbf{F}(\mathbf{R}(x_0, y_m, 0)) \right] \cdot \mathbf{R}(0, \Delta y, 0)$$

The first order approximation of these differences are not as simple as before. The function \mathbf{F} is now a function of three inputs as opposed to just two, and it is now being evaluated over the function \mathbf{R} instead of directly over Cartesian space. However, the aggregate $\mathbf{F} \circ \mathbf{R}$ is evaluated with only x or y varying, while its other inputs remain fixed. We thus have functions of just one variable in each case. Perhaps the first order approximations are not so difficult after all. We can now rewrite our approximation as

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx - \left[\frac{\partial}{\partial y} \left[\mathbf{F} \circ \mathbf{R} \right] \circ \left(x_m, y_0, 0 \right) \right] \Delta y \cdot \mathbf{R}(\Delta x, 0, 0) + \left[\frac{\partial}{\partial x} \left[\mathbf{F} \circ \mathbf{R} \right] \circ \left(x_0, y_m, 0 \right) \right] \Delta x \cdot \mathbf{R}(0, \Delta y, 0)$$

The partial derivatives in this expression come out to

$$\begin{aligned} \frac{\partial}{\partial x} \left[\mathbf{F} \circ \mathbf{R} \right] &= \left(\frac{\partial \mathbf{F}}{\partial x} \circ \mathbf{R} \right) \frac{\partial R_x}{\partial x} + \left(\frac{\partial \mathbf{F}}{\partial y} \circ \mathbf{R} \right) \frac{\partial R_y}{\partial x} + \left(\frac{\partial \mathbf{F}}{\partial z} \circ \mathbf{R} \right) \frac{\partial R_z}{\partial x} \\ \frac{\partial}{\partial y} \left[\mathbf{F} \circ \mathbf{R} \right] &= \left(\frac{\partial \mathbf{F}}{\partial x} \circ \mathbf{R} \right) \frac{\partial R_x}{\partial y} + \left(\frac{\partial \mathbf{F}}{\partial y} \circ \mathbf{R} \right) \frac{\partial R_y}{\partial y} + \left(\frac{\partial \mathbf{F}}{\partial z} \circ \mathbf{R} \right) \frac{\partial R_z}{\partial y} \end{aligned}$$

Let us now look at the partial derivatives of \mathbf{R} that appear in this result. For example, consider the derivative $\partial R_y / \partial x$:

$$\frac{\partial R_y}{\partial x} = \frac{\partial}{\partial x} \left[R_y(\mathbf{r}) \right] = \frac{\partial}{\partial x} \left[\mathbf{R}(\mathbf{r}) \cdot \hat{\mathbf{j}} \right] = \mathbf{R} \left(\frac{\partial \mathbf{r}}{\partial x} \right) \cdot \hat{\mathbf{j}} = \mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{j}}$$

This is just $\hat{\mathbf{j}}$ dotted with the rotation of $\hat{\mathbf{i}}$ about $\hat{\boldsymbol{\omega}}$. In summary we have:

$$\frac{\partial R_y}{\partial x} = \mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{j}}$$

We can work out the other partial derivatives on \mathbf{R} in a similar way to get

$$\begin{aligned} \frac{\partial R_x}{\partial x} &= \mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{i}} & \frac{\partial R_y}{\partial x} &= \mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{j}} & \frac{\partial R_z}{\partial x} &= \mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{k}} \\ \frac{\partial R_x}{\partial y} &= \mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{i}} & \frac{\partial R_y}{\partial y} &= \mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{j}} & \frac{\partial R_z}{\partial y} &= \mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{k}} \end{aligned}$$

The partial derivatives thus reduce out

$$\begin{aligned} \frac{\partial}{\partial x} \left[\mathbf{F} \circ \mathbf{R} \right] &= \left(\frac{\partial \mathbf{F}}{\partial x} \circ \mathbf{R} \right) \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{i}} \right) + \left(\frac{\partial \mathbf{F}}{\partial y} \circ \mathbf{R} \right) \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{j}} \right) + \left(\frac{\partial \mathbf{F}}{\partial z} \circ \mathbf{R} \right) \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{k}} \right) \\ \frac{\partial}{\partial y} \left[\mathbf{F} \circ \mathbf{R} \right] &= \left(\frac{\partial \mathbf{F}}{\partial x} \circ \mathbf{R} \right) \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{i}} \right) + \left(\frac{\partial \mathbf{F}}{\partial y} \circ \mathbf{R} \right) \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{j}} \right) + \left(\frac{\partial \mathbf{F}}{\partial z} \circ \mathbf{R} \right) \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{k}} \right) \end{aligned}$$

For brevity let us omit writing out the nested calls to \mathbf{R} . If we are being strict about our notation we should leave it in, but doing that would coming algebra would quickly become intractable. From here onward you should simply remember that each partial derivative of \mathbf{F} is to be evaluated over \mathbf{R} . Our notation will now be:

$$\begin{aligned} \frac{\partial}{\partial x} \left[\mathbf{F} \circ \mathbf{R} \right] &= \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{i}} \right) \frac{\partial \mathbf{F}}{\partial x} + \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{j}} \right) \frac{\partial \mathbf{F}}{\partial y} + \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{k}} \right) \frac{\partial \mathbf{F}}{\partial z} \\ \frac{\partial}{\partial y} \left[\mathbf{F} \circ \mathbf{R} \right] &= \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{i}} \right) \frac{\partial \mathbf{F}}{\partial x} + \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{j}} \right) \frac{\partial \mathbf{F}}{\partial y} + \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{k}} \right) \frac{\partial \mathbf{F}}{\partial z} \end{aligned}$$

We can now write our approximation as

$$\begin{aligned} \oint \mathbf{F} \cdot d\mathbf{c} &\approx - \left[\left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{i}} \right) \frac{\partial \mathbf{F}}{\partial x} + \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{j}} \right) \frac{\partial \mathbf{F}}{\partial y} + \left(\mathbf{R}(\hat{\mathbf{j}}) \cdot \hat{\mathbf{k}} \right) \frac{\partial \mathbf{F}}{\partial z} \right] \Delta y \cdot \mathbf{R}(\Delta x, 0, 0) \\ &\quad + \left[\left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{i}} \right) \frac{\partial \mathbf{F}}{\partial x} + \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{j}} \right) \frac{\partial \mathbf{F}}{\partial y} + \left(\mathbf{R}(\hat{\mathbf{i}}) \cdot \hat{\mathbf{k}} \right) \frac{\partial \mathbf{F}}{\partial z} \right] \Delta x \cdot \mathbf{R}(0, \Delta y, 0) \end{aligned}$$

Next notice that \mathbf{R} obeys

$$\mathbf{R}(c\mathbf{r}) = c\mathbf{R}(\mathbf{r})$$

and thus that

$$\begin{aligned}\mathbf{R}(0, \Delta y, 0) &= \mathbf{R}(0, 1, 0) \Delta y = \mathbf{R}(\hat{j}) \Delta y \\ \mathbf{R}(\Delta x, 0) &= \mathbf{R}(1, 0, 0) \Delta x = \mathbf{R}(\hat{i}) \Delta x\end{aligned}$$

The next logical iteration of our approximation is therefore

$$\begin{aligned}\oint \mathbf{F} \cdot d\mathbf{c} \approx & - \left[\left(\mathbf{R}(\hat{j}) \cdot \hat{i} \right) \frac{\partial \mathbf{F}}{\partial x} + \left(\mathbf{R}(\hat{j}) \cdot \hat{j} \right) \frac{\partial \mathbf{F}}{\partial y} + \left(\mathbf{R}(\hat{j}) \cdot \hat{k} \right) \frac{\partial \mathbf{F}}{\partial z} \right] \cdot \mathbf{R}(\hat{i}) \Delta y \Delta x \\ & + \left[\left(\mathbf{R}(\hat{i}) \cdot \hat{i} \right) \frac{\partial \mathbf{F}}{\partial x} + \left(\mathbf{R}(\hat{i}) \cdot \hat{j} \right) \frac{\partial \mathbf{F}}{\partial y} + \left(\mathbf{R}(\hat{i}) \cdot \hat{k} \right) \frac{\partial \mathbf{F}}{\partial z} \right] \cdot \mathbf{R}(\hat{j}) \Delta x \Delta y\end{aligned}$$

We can now do some trivial vector algebra to pull out the partials of \mathbf{F} :

$$\begin{aligned}\oint \mathbf{F} \cdot d\mathbf{c} \approx & + \frac{\partial \mathbf{F}}{\partial x} \cdot \left[\left[\mathbf{R}(\hat{i}) \cdot \hat{i} \right] \mathbf{R}(\hat{j}) - \left[\mathbf{R}(\hat{j}) \cdot \hat{i} \right] \mathbf{R}(\hat{i}) \right] \Delta x \Delta y \\ & + \frac{\partial \mathbf{F}}{\partial y} \cdot \left[\left[\mathbf{R}(\hat{i}) \cdot \hat{j} \right] \mathbf{R}(\hat{j}) - \left[\mathbf{R}(\hat{j}) \cdot \hat{j} \right] \mathbf{R}(\hat{i}) \right] \Delta x \Delta y \\ & + \frac{\partial \mathbf{F}}{\partial z} \cdot \left[\left[\mathbf{R}(\hat{i}) \cdot \hat{k} \right] \mathbf{R}(\hat{j}) - \left[\mathbf{R}(\hat{j}) \cdot \hat{k} \right] \mathbf{R}(\hat{i}) \right] \Delta x \Delta y\end{aligned}$$

The inner terms involving the \mathbf{R} 's and the basis vectors are really just triple cross products. According to $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ we can reduce things down to

$$\begin{aligned}\oint \mathbf{F} \cdot d\mathbf{c} \approx & + \frac{\partial \mathbf{F}}{\partial x} \cdot \left[\hat{i} \times (\mathbf{R}(\hat{j}) \times \mathbf{R}(\hat{i})) \right] \Delta x \Delta y \\ & + \frac{\partial \mathbf{F}}{\partial y} \cdot \left[\hat{j} \times (\mathbf{R}(\hat{j}) \times \mathbf{R}(\hat{i})) \right] \Delta x \Delta y \\ & + \frac{\partial \mathbf{F}}{\partial z} \cdot \left[\hat{k} \times (\mathbf{R}(\hat{j}) \times \mathbf{R}(\hat{i})) \right] \Delta x \Delta y\end{aligned}$$

The cross product $\mathbf{R}(\hat{j}) \times \mathbf{R}(\hat{i})$ is just a cross of the second and first basis the vectors of our rotated coordinate system, namely $\mathbf{R}(\hat{j}) \times \mathbf{R}(\hat{i}) = \mathbf{R}(\hat{k})$. It must therefore be the third basis vector of that rotated system, but negated. When we take this into account we get

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx \frac{\partial \mathbf{F}}{\partial x} \cdot \left[\hat{i} \times \mathbf{R}(\hat{k}) \right] \Delta x \Delta y + \frac{\partial \mathbf{F}}{\partial y} \cdot \left[\hat{j} \times \mathbf{R}(\hat{k}) \right] \Delta x \Delta y + \frac{\partial \mathbf{F}}{\partial z} \cdot \left[\hat{k} \times \mathbf{R}(\hat{k}) \right] \Delta x \Delta y$$

We can do more with this if we rearrange how the triple products are organized, such that $\mathbf{R}(\hat{k})$ appears as the outermost term. We can also pull out the factors of $\Delta x \Delta y$:

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx \left(\mathbf{R}(\hat{k}) \cdot \left[\frac{\partial \mathbf{F}}{\partial x} \times \hat{i} \right] + \mathbf{R}(\hat{k}) \cdot \left[\frac{\partial \mathbf{F}}{\partial y} \times \hat{j} \right] + \mathbf{R}(\hat{k}) \cdot \left[\frac{\partial \mathbf{F}}{\partial z} \times \hat{k} \right] \right) \Delta x \Delta y$$

In fact, we can now factor out the $\mathbf{R}(\hat{k})$ term altogether:

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx \left(\left[\frac{\partial \mathbf{F}}{\partial x} \times \hat{i} \right] + \left[\frac{\partial \mathbf{F}}{\partial y} \times \hat{j} \right] + \left[\frac{\partial \mathbf{F}}{\partial z} \times \hat{k} \right] \right) \cdot \mathbf{R}(\hat{k}) \Delta x \Delta y$$

When we expand out the cross products we get

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx \left[\frac{\partial F_y}{\partial x} \hat{\mathbf{k}} - \frac{\partial F_z}{\partial x} \hat{\mathbf{j}} + \frac{\partial F_z}{\partial y} \hat{\mathbf{i}} - \frac{\partial F_x}{\partial y} \hat{\mathbf{k}} + \frac{\partial F_x}{\partial z} \hat{\mathbf{j}} - \frac{\partial F_y}{\partial z} \hat{\mathbf{i}} \right] \cdot \mathbf{R}(\hat{\mathbf{k}}) \Delta x \Delta y$$

Next we can group the terms around their basis vectors:

$$\oint \mathbf{F} \cdot d\mathbf{c} \approx \left[\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{\mathbf{k}} \right] \cdot \mathbf{R}(\hat{\mathbf{k}}) \Delta x \Delta y$$

When we take the square to be differentially small the result goes from approximate to exact and we write $\Delta x \Delta y$ as the differential area $dx dy$:

$$\oint \mathbf{F} \cdot d\mathbf{c} = \left[\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{\mathbf{k}} \right] \cdot \mathbf{R}(\hat{\mathbf{k}}) dx dy$$

For brevity let us define a notation for this new vector quantity

$$\nabla \times \mathbf{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{\mathbf{k}}$$

The vector $\mathbf{R}(\hat{\mathbf{k}}) dx dy$ is normal to the surface of the square, and it has a magnitude equal to the square's area. We can denote it more concisely as $d^2\mathbf{A}$:

$$d^2\mathbf{A} = \mathbf{R}(\hat{\mathbf{k}}) \Delta x \Delta y$$

Our path integral can now be written as

$$\oint \mathbf{F} \cdot d\mathbf{c} = (\nabla \times \mathbf{F}) \cdot d^2\mathbf{A}$$

After much work we finally have our answer. It appears that taking a path integral of \mathbf{F} around a very small square is equivalent to taking the dot product of that square's area vector with this new vector $\nabla \times \mathbf{F}$, which is composed of various partials of \mathbf{F} . Since the path integral depends on $\mathbf{R}(\hat{\mathbf{k}})$ the result will differ when taken over two squares which are centered at the same point but oriented differently. However, the vector $\nabla \times \mathbf{F}$ is clearly independent of the orientation of the square around which the path integral is being taken. The path integral's value will be maximized when $d^2\mathbf{A}$ points in the direction of $\nabla \times \mathbf{F}$.