Divergence and Curl

E. L. Lady

(Last revised October 7, 2004)

A **vector field** $\mathbf{F}$ defined on a certain region in n-dimensional Euclidian space consists of an n-dimensional vector defined at every point in this region. (I. e. $\mathbf{F}$ is a function which assigns a vector in $\mathbb{R}^n$ to every point in the given region.)

In the case of three-dimensional space, a vector field would have the form

$$\mathbf{F} = P(x,y,z)\,\mathbf{i} + Q(x,y,z)\,\mathbf{j} + R(x,y,z)\,\mathbf{k}\,,$$

where P, Q, and R are scalar functions.

The usual way to think of differentiation in this situation is to think of the derivative of $\mathbf{F}$ as being given by the **Jacobean matrix**

$$\mathcal{J} = \begin{bmatrix} \dfrac{\partial P}{\partial x} & \dfrac{\partial P}{\partial y} & \dfrac{\partial P}{\partial z} \\ \dfrac{\partial Q}{\partial x} & \dfrac{\partial Q}{\partial y} & \dfrac{\partial Q}{\partial z} \\ \dfrac{\partial R}{\partial x} & \dfrac{\partial R}{\partial y} & \dfrac{\partial R}{\partial z} \end{bmatrix}.$$

If one writes $\mathbf{x} = (x,y,z)$ and $\Delta\mathbf{x} = (x,y,z) - (x_0,y_0,z_0)$, and $\Delta\mathbf{F} = \mathbf{F}(x,y,z) - \mathbf{F}(x_0,y_0,z_0)$, and if one thinks of $\Delta\mathbf{x}$ and $\Delta\mathbf{F}$ as being column vectors rather than row vectors, then one has

$$\Delta\mathbf{F} = \begin{bmatrix} \Delta P \\ \Delta Q \\ \Delta R \end{bmatrix} \approx \begin{bmatrix} \dfrac{\partial P}{\partial x} & \dfrac{\partial P}{\partial y} & \dfrac{\partial P}{\partial z} \\ \dfrac{\partial Q}{\partial x} & \dfrac{\partial Q}{\partial y} & \dfrac{\partial Q}{\partial z} \\ \dfrac{\partial R}{\partial x} & \dfrac{\partial R}{\partial y} & \dfrac{\partial R}{\partial z} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \mathcal{J}\,\Delta\mathbf{x}$$

in the same way that, with a function $f(x)$ of one variable, one has $\Delta f \approx f'(x_0)\Delta x$.

One can also think of the Jacobean matrix (or, better still, the linear transformation corresponding to the matrix $\mathcal{J}$) as a way of getting the directional derivative of $\mathbf{F}$ in the direction of a unit vector $\mathbf{u}$:

$$\lim_{h\to 0} \frac{\mathbf{F}(\mathbf{x}+h\mathbf{u}) - \mathbf{F}(\mathbf{x})}{h} = \mathcal{J}\mathbf{u}\,,$$

where in computing the matrix product here we think of $\mathbf{u}$ as a column vector, i. e. a 3×1 matrix.

It is useful to notice that the first column of $\mathcal{J}$ gives the directional derivative of $\mathbf{F}$ in the direction $\mathbf{i}$, the second column is the directional derivative in the direction $\mathbf{j}$, and the third column is the directional derivative in the $\mathbf{k}$ direction.

Divergence

The **divergence** of the vector field $\mathbf{F}$, often denoted by $\nabla \bullet \mathbf{F}$, is the **trace** of the Jacobean matrix for $\mathbf{F}$, i. e. the sum of the diagonal elements of $\mathcal{J}$. Thus, in three dimensions,

$$\nabla \bullet \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\,.$$

Now the concept of the trace is surprisingly useful in matrix theory, but it in general is also a very difficult concept to interpret in any meaningful intuitive way. Thus it is not surprising that the divergence seems a rather mysterious concept in vector analysis.

Example. Let $\mathbf{F}(x,y) = 3e^{5x}\cos 3y\,\mathbf{i} - 5e^{5x}\sin 3y\,\mathbf{j} = e^{5x}(3\cos 3y\,\mathbf{i} - 5\sin 3y\,\mathbf{j})$. Then $\nabla \bullet \mathbf{F} = 15e^{5x}\cos 3y - 15e^{5x}\cos 3y = 0$.

More generally, if $g(x,y)$ is any function with continuous second partial derivatives and $\mathbf{F}(x,y) = \dfrac{\partial g}{\partial y}\,\mathbf{i} - \dfrac{\partial g}{\partial x}\,\mathbf{j}$, then $\nabla \bullet \mathbf{F} = \dfrac{\partial^2 g}{\partial x\,\partial y} - \dfrac{\partial^2 g}{\partial y\,\partial x} = 0$.

In trying to understand any relationship involving differentiation, it is usually most enlightening to start with the case where the derivative is constant. For instance in studying the relationship between distance, time, and velocity, one begins in high school with the case where velocity is constant.

The principle of Taylor Series expansions shows that in a small neighborhood of a point (x_0, y_0) in two-dimensional space, it will usually be true for a vector field $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j}$ that

$$\mathbf{F}(x,y) = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j} \approx \mathbf{c} + (a_1 x + a_2 y)\,\mathbf{i} + (b_1 x + b_2 y)\,\mathbf{j},$$

where $\mathbf{c}$ is a constant vector and

$$a_1 = \frac{\partial P}{\partial x}(x_0,y_0), \quad a_2 = \frac{\partial P}{\partial y}(x_0,y_0), \qquad b_1 = \frac{\partial Q}{\partial x}(x_0,y_0), \quad b_2 = \frac{\partial Q}{\partial y}(x_0,y_0)\,.$$

(More precisely,

$$P(x,y) \approx P(x_0,y_0) + \frac{\partial P}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial P}{\partial y}(x_0,y_0)(y-y_0)$$

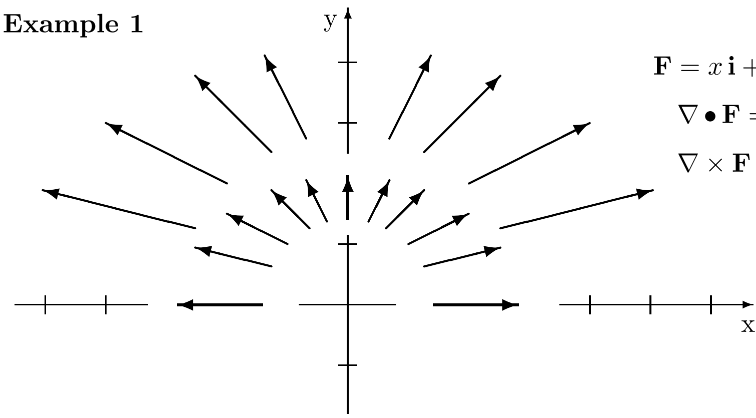
$$Q(x,y) \approx Q(x_0,y_0) + \frac{\partial Q}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial Q}{\partial y}(x_0,y_0)(y-y_0)$$

in a sufficiently small neighborhood of (x_0, y_0). Actually, this will be true provided that P and Q are differentiable at (x_0, y_0), even if they don't have Taylor Series expansions.)

Now a_2 and b_1 don't affect $\nabla \bullet F$, so to get an intuitive sense of how divergence works in the plane, it makes sense to look at some examples of vector fields $\mathbf{F} = ax\,\mathbf{i} + by\,\mathbf{j}$, where a and b are constants. Then $\nabla \bullet \mathbf{F} = a + b$.

Here are some pictures. (For future reference, the values for $\nabla \times \mathbf{F}$ are also given.)

Example 1

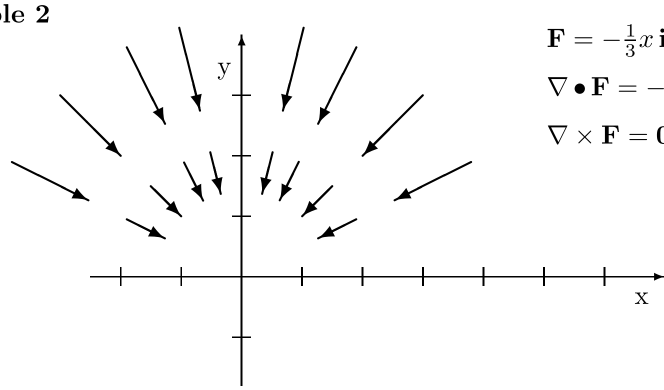


$\mathbf{F} = x\,\mathbf{i} + .5y\,\mathbf{j} = \nabla(.5x^2 + .25y^2)$

$\nabla \bullet \mathbf{F} = 1.5$

$\nabla \times \mathbf{F} = \mathbf{0}$

Example 2

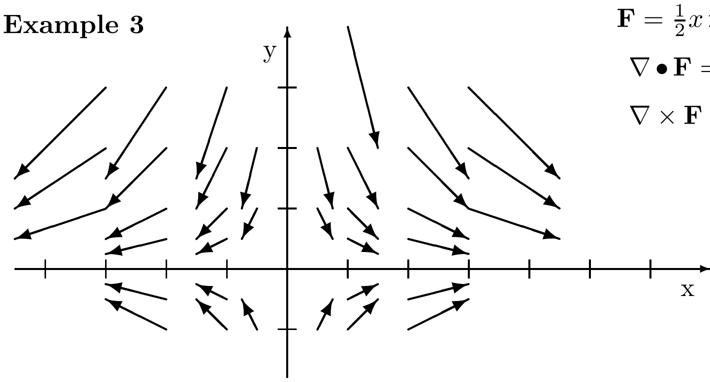


$\mathbf{F} = -\frac{1}{3}x\,\mathbf{i} - \frac{1}{3}y\,\mathbf{j} = \nabla(-\frac{1}{6}x^2 - \frac{1}{6}y^2)$

$\nabla \bullet \mathbf{F} = -2/3$

$\nabla \times \mathbf{F} = \mathbf{0}$

Example 3



$\mathbf{F} = \frac{1}{2}x\,\mathbf{i} - \frac{1}{2}y\,\mathbf{j} = \nabla(\frac{1}{4}x^2 - \frac{1}{4}y^2)$

$\nabla \bullet \mathbf{F} = 0$

$\nabla \times \mathbf{F} = \mathbf{0}$

Consider now the special case of a vector field $\mathbf{F}$ with constant direction. Then we can write $\mathbf{F}(x, y) = f(x, y)\,\mathbf{u}$, where $\mathbf{u}$ is a constant vector with magnitude 1. A calculation then shows that

$$\nabla \bullet \mathbf{F} = \nabla f \bullet \mathbf{u} = D_{\mathbf{u}}(f)\,,$$

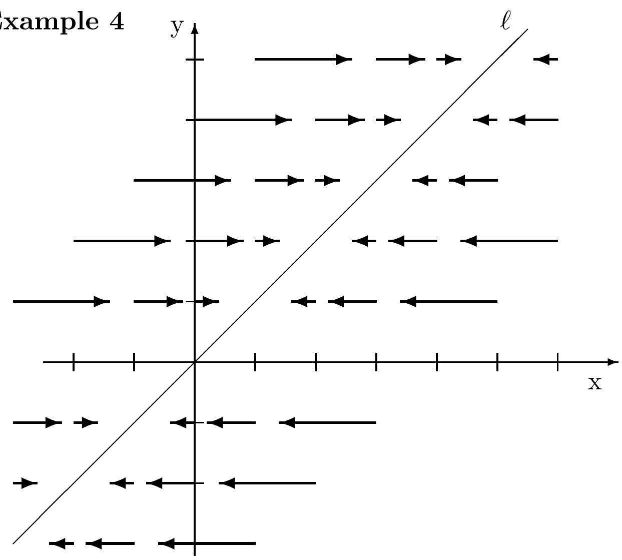
where $D_{\mathbf{u}}(f)$ denotes the directional derivative of $f(x, y)$ in the direction $\mathbf{u}$. Assume that $f(x, y) \geq 0$ at the particular point of interest. Then $f(x, y) = ||\mathbf{F}(x, y)||$. Since $\mathbf{F}$ is parallel to $\mathbf{u}$, this shows that

Proposition. If $\mathbf{F}(x,y)$ is a vector field with constant direction $\mathbf{u}$, then $\nabla \bullet \mathbf{F}$ is the rate of increase in $||\mathbf{F}||$ in the direction $\mathbf{u}$.

(Note that if $f(x,y) < 0$ then the direction of $\mathbf{F}$ is $-\mathbf{u}$ and $||\mathbf{F}|| = -f(x,y)$, so that $\nabla \bullet \mathbf{F} = D_{\mathbf{u}}(f) = D_{-\mathbf{u}}(-f)$ is still the rate at which $||\mathbf{F}||$ is changing as one moves in the direction of $\mathbf{F}$.)

The principle here is equally valid in three dimensions or even in higher dimensional spaces.

Example 4



A field with constant direction and negative divergence.

$\mathbf{F} = a(y-x)\,\mathbf{i}$

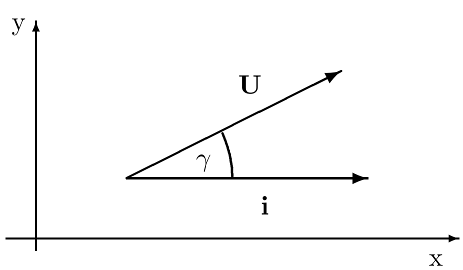
$\nabla \bullet \mathbf{F} = -a$

$\nabla \times \mathbf{F} = -a\,\mathbf{k}$

$\mathbf{F} = \mathbf{0}$ along the line ℓ.

On the other hand, consider a vector field $\mathbf{U}$ in two dimentions with constant magnitude. Suppose, say, that $||\mathbf{U}(x,y)|| = 1$ everywhere. Then basic trigonometry shows that if $\gamma(x,y)$ is the angle that $\mathbf{U}(x,y)$ makes to the horizontal, i. e. γ is the angle between $\mathbf{U}$ and $\mathbf{i}$, then

$$\mathbf{U}(x,y) = \cos\gamma(x,y)\,\mathbf{i} + \sin\gamma(x,y)\,\mathbf{j}\,.$$



$||\mathbf{U}|| = 1$

We then see that

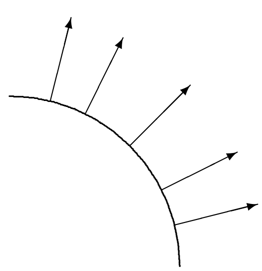
$$\begin{aligned}\nabla \bullet \mathbf{U} &= \frac{\partial}{\partial x}(\cos\gamma) + \frac{\partial}{\partial y}(\sin\gamma)\\ &= -\sin\gamma\,\frac{\partial\gamma}{\partial x} + \cos\gamma\,\frac{\partial\gamma}{\partial y}\\ &= (-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}) \bullet (\frac{\partial\gamma}{\partial x}\,\mathbf{i} + \frac{\partial\gamma}{\partial y}\,\mathbf{j})\\ &= (-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}) \bullet \nabla\gamma\,.\end{aligned}$$

Thus $\nabla \bullet \mathbf{U}$ is the directional derivative of $\gamma(x,y)$ in the direction $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$. This shows that

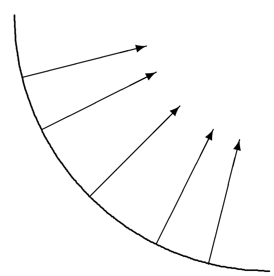
Proposition. If $\mathbf{U}(x,y)$ is a vector field in the plane with constant magnitude 1, then $\nabla \bullet \mathbf{U}(x,y)$ equals the rate at which $\mathbf{U}$ turns as (x,y) moves at unit speed in a direction which is 90° counterclockwise to the direction of $\mathbf{U}(x,y)$.

To get a clearer idea of what this means, look at a curve $\mathcal{C}$ orthogonal to the vector field, i. e. the tangent to $\mathcal{C}$ is perpendicular to $\mathbf{U}(x,y)$ at every point (x,y) on $\mathcal{C}$. Then what we see for a vector field $\mathbf{U}$ of constant magnitude is that as we move along the orthogonal curve in a direction 90° counterclockwise from the direction of $\mathbf{U}$, in case of positive divergence the vector field $\mathbf{U}$ will be turning counter-clockwise, and consequently $\mathcal{C}$ will be curving in the direction opposite to the direction of $\mathbf{U}$. For a field with constant magnitude and negative divergence, on the other hand, as one moves along the curve in the direction 90° counterclockwise from the direction of $\mathbf{U}$, $\mathbf{U}$ will be turning clockwise, so $\mathcal{C}$ will be curving in the in the same direction as that in which $\mathbf{U}$ is pointing.

Vector fields with constant magnitude.



Positive divergence



Negative divergence

Vector fields having constant magnitude are unusual. For instance, linear vector fields (see below) never do. Certainly we can contrive such a field simply by using any differentiable function $\gamma(x,y)$ and the formula $\mathbf{U}(x,y) = \cos\gamma\,\mathbf{i} + \sin\gamma\,\mathbf{j}$ or by setting $\mathbf{U}(x,y,z) = \mathbf{F}(x,y,z)/\,||\mathbf{F}||$, for any vector field which is never zero in the region of interest. One of the more natural examples of this sort with constant magnitude 1 is $\mathbf{F}(x,y,z) = \dfrac{x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}{r}$ where $r = \sqrt{x^2+y^2+z^2}$.

It is easily seen that $\mathbf{F} = \nabla g$, where $g(x,y,z) = \sqrt{x^2+y^2+z^2}$.

One can show that $\nabla \bullet \mathbf{F} = \dfrac{2}{r}$. In fact, $\mathbf{F} = r^{-1}\mathbf{r}$, where $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, and so

$$\nabla \bullet \mathbf{F} = \nabla(r^{-1}) \bullet \mathbf{r} + r^{-1}\,\nabla \bullet \mathbf{r} = -r^{-3}\mathbf{r} \bullet \mathbf{r} + 3r^{-1} = 2r^{-1}.$$

Since we can write any vector field $\mathbf{F}$ as $f(x,y)\mathbf{U}(x,y)$ where $f(x,y) = ||\mathbf{F}(x,y)||$ and $\mathbf{U}(x,y) = \dfrac{\mathbf{F}(x,y)}{||\mathbf{F}||}$, we can apply the product rule to get

$$\nabla \bullet \mathbf{F} = (\nabla f) \bullet \mathbf{U} + f\ \nabla \bullet \mathbf{U}.$$

In words, then,

Proposition. The divergence of a two-dimensional vector field $\mathbf{F}(x,y)$ is the sum of the following two terms:

(1) The rate at which $||\mathbf{F}||$ is increasing when (x,y) moves at unit speed in the direction of $\mathbf{F}$ (this is negative, of course, if $||\mathbf{F}||$ is decreasing);

(2) the product of $||\mathbf{F}||$ and the rate at which $\mathbf{F}$ is turning as (x,y) moves in the direction 90° counterclockwise to the direction of $\mathbf{F}$ at unit speed.

There is a more sophisticated and more conceptual way of deriving this result, namely by making a change of coordinates. Of course one has to know the fact that a change of coordinates will not change $\nabla \bullet \mathbf{U}$. One should also remember that this change of coordinates has its limitations as a practical tool, since except for linear fields (see below), it will only simplify the Jacobean matrix at one particular point.

The key to the change in coordinates comes from a fundamental fact about the derivatives of vectors. Namely, the derivative $\dfrac{d\mathbf{v}(t)}{dt}$ of a vector $\mathbf{v}(t)$ at a given point can be decomposed into two components, one in the direction of the vector, which shows the rate at which the magnitude of $\mathbf{v}$ is changing, and the other orthogonal to $\mathbf{v}(t)$, which points in the direction towards which $\mathbf{v}$ is turning and whose magnitude is the product of $||\mathbf{v}||$ and the rate at which $\mathbf{v}$ is turning.

Thus the first component of the directional derivative of a vector field $\mathbf{F}$ in a given direction at a given point gives the rate at which the magnitude of $\mathbf{F}$ is changing and the second component is determined by the rate at which $\mathbf{F}$ turns as one moves in the given direction at unit speed. Recall that if $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j}$ and we denote the directional derivatives of $\mathbf{F}$ in the $\mathbf{i}$ and $\mathbf{j}$ directions by $D_{\mathbf{i}}(\mathbf{F})$ and $D_{\mathbf{j}}(\mathbf{F})$, then

$$D_{\mathbf{i}}(\mathbf{F}) = \frac{\partial P}{\partial x}\,\mathbf{i} + \frac{\partial Q}{\partial x}\,\mathbf{j}$$

$$D_{\mathbf{j}}(\mathbf{F}) = \frac{\partial P}{\partial y}\,\mathbf{i} + \frac{\partial Q}{\partial y}\,\mathbf{j}.$$

This makes it useful to choose an orthogonal coordinate system so that at a given point (x_0, y_0), $\mathbf{i}$ is in the direction of $\mathbf{F}$. As usual, we make $\mathbf{j}$ in the direction $90°$ counterclockwise from the direction of $\mathbf{i}$. Then $\dfrac{\partial P}{\partial x}$ is the first component of the directional derivative of $\mathbf{F}$ in the direction of $\mathbf{F}$, and thus gives the rate at which $||\mathbf{F}||$ changes as we move in the direction of $\mathbf{F}$. And $\dfrac{\partial Q}{\partial y}$ is the second component of the directional derivative of $\mathbf{F}$ in the direction $90°$ perpendicular to $\mathbf{F}$, and thus equals the product of $||\mathbf{F}||$ and the rate at which $\mathbf{F}$ turns as we move in the direction $90°$ counterclockwise to $\mathbf{F}$. Since $\nabla \bullet \mathbf{F} = \dfrac{\partial P}{\partial x} + \dfrac{\partial Q}{\partial y}$, we thus recover the characterization of the divergence of a two-dimensional field given in the preceding Proposition.

In general, the Proposition should not be looked at as a practical method for computing $\nabla \bullet \mathbf{F}$, since in most cases the original definition is easily used for calculation. Instead, it is a way of attempting to see the intuitive conceptual meaning of $\nabla \bullet \mathbf{F}$. However in certain cases, the differentiation required by the original formula is a slight nuisance. For instance, consider the vector field $\mathbf{F}(x,y) = \dfrac{y\,\mathbf{i} - x\,\mathbf{j}}{r^n}$, where $r = \sqrt{x^2+y^2}$. Here we can easily see from the Proposition above that $\nabla \bullet \mathbf{F} = 0$ since $\mathbf{F}$ is tangent to the circles $x^2 + y^2 = \text{const}$ around the origin and $||\mathbf{F}||$ is constant as one moves around such circles, so that the directional derivative of $||\mathbf{F}||$ as (x,y) moves in the direction of $\mathbf{F}$ is zero. Furthermore, the direction of $\mathbf{F}$ does not change when (x,y) moves in a direction perpendicular to $\mathbf{F}$, i. e. in the direction of the radial lines from the origin. As a computational check, with $P = y/r^n$, $Q = -x/r^n$, then, using the fact that $\partial r/\partial x = x/r$, $\partial r/\partial y = y/r$, one finds that

$$\frac{\partial P}{\partial x} = \frac{\partial}{\partial x}\left(\frac{y}{r^n}\right) = \frac{-nxy}{r^{n+2}}$$
$$\frac{\partial Q}{\partial y} = \frac{\partial}{\partial y}\left(\frac{-x}{r^n}\right) = \frac{nxy}{r^{n+2}}$$

so that indeed $\nabla \bullet \mathbf{F} = \dfrac{\partial P}{\partial x} + \dfrac{\partial Q}{\partial y} = 0$.

One should also note that the Proposition breaks down at points (x_0, y_0) where $\mathbf{F}(x_0, y_0) = \mathbf{0}$, since at those points the direction of $\mathbf{F}$ is not well defined. In some cases, this is only an apparent difficulty. If one writes $\mathbf{F}(x,y) = f(x,y)\mathbf{u}(x,y)$, where $\mathbf{u}(x,y)$ is a unit vector, and if $\mathbf{F}(\mathbf{x}_0) = \mathbf{F}(x_0, y_0) = \mathbf{0}$, then one may be able to get away by setting $\mathbf{u}(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{u}(\mathbf{x})$. But in some important cases this limit will not exist. Two examples are the fields $\mathbf{F}(x,y) = x\,\mathbf{i} + y\,\mathbf{j}$ and $\mathbf{F}(x,y) = -y\,\mathbf{i} + x\,\mathbf{j}$, with $\mathbf{x}_0 = (0,0)$ for both examples. In cases like this, however, one ought to at least get away with using the formula $\mathbf{F}(\mathbf{x}_0) = \lim_{\mathbf{x}\to\mathbf{x}_0} \nabla \bullet \mathbf{F}(\mathbf{x})$.

Since the characterization of divergence for two-dimensional vector fields given by the preceding Proposition is so nicely conceptual and coordinate-free, it seems only natural to hope that if one can find the appropriate way to state it, then it will continue to be true for three-dimensional fields as well. Or at least there should be some analogous result for three dimensions. This is pretty much true, as one will see at the end of the article.

The Jacobean matrix for a vector field does not tell us the magnitude or direction of the field, since what it describes is the way the field changes. However recall our earlier observation that the directional derivative of a field $\mathbf{F}$ in any direction $\mathbf{u}$ consists of two components, one (in the direction of the field) telling us the rate at which $||\mathbf{F}||$ changes as one moves at unit speed in the direction $\mathbf{u}$, and the other (orthogonal to the first) telling us the rate at which $\mathbf{F}$ turns.

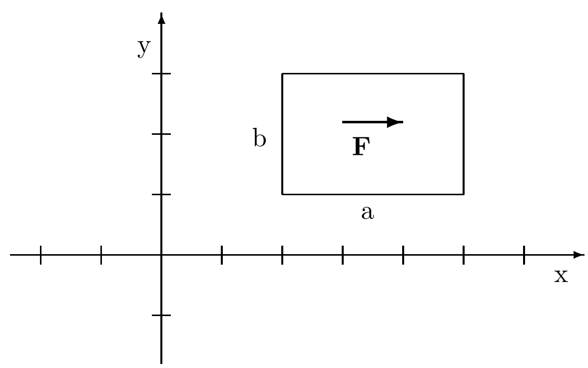
If $\mathbf{F}$ has constant magnitude, then the first of these components (the one in the direction of $\mathbf{F}$) of the directional derivative in any direction will be $\mathbf{0}$. And if $\mathbf{F}$ has constant direction, then the second component (the one orthogonal to $\mathbf{F}$) wil be $\mathbf{0}$.

Applying this to the directional derivatives $D_{\mathbf{i}}(\mathbf{F})$ and $D_{\mathbf{j}}(\mathbf{F})$, one sees that

A vector field $\mathbf{F}$ has constant magnitude if and only if at every point the columns of the Jacobean matrix are orthogonal to $\mathbf{F}$.
A vector field $\mathbf{F}$ has constant direction if and only if at every point the columns of the Jacobean matrix are parallel to $\mathbf{F}$.

Why Is It Called Divergence?

Let's go back to the simple example of a horizontal vector field $\mathbf{F} = P(x,y)\,\mathbf{i}$ and suppose that this vector field $\mathbf{F}$ represents the velocity of a moving fluid. Now construct a small box in the xy-plane with horizontal and vertical sides.



Suppose in particular that $\nabla \bullet \mathbf{F} = \dfrac{\partial P}{\partial x} = 5$. Let a denote the horizontal width of the box, as indicated. Since we are assuming that $\mathbf{F}$ is directed toward the right (i. e. that P is positive) then the rate at which the fluid is flowing out of the right end of the box is $5a$ more than the rate at which it is entering the left end. If, on the other hand, P is negative, then fluid flows out of the left end of the box at a rate $5a$ more than it flows into the right end. In either case, if the box has height b, this means that there is $5ab$ more fluid flowing out of the box than there is fluid flowing in.

On the other hand, if we were to have a vertical vector field $Q\,\mathbf{j}$ representing the flow of a fluid, and if $\dfrac{\partial Q}{\partial y}$ were constant, say $\dfrac{\partial Q}{\partial y} = 8$, this would mean that the velocity at which fluid is leaving through the top of the box (since we have assumed that Q is positive) would exceed the velocity at which it enters through the bottom by $8b$, and there would be $8ab$ more fluid flowing out of the box than fluid flowing in.

In either case, the difference between the amount of fluid leaving the box and the amount entering is equal to the divergence of the vector field times the area of a box.

But we can write a general vector field $\mathbf{F}$ in the two-dimensional case as the sum of a horizontal field $P\,\mathbf{i}$ and a vertical one $Q\,\mathbf{j}$. And the divergence of $\mathbf{F}$ is the sum of the divergence of the horizontal and vertical components:

$$\nabla \bullet F = \nabla \bullet (P\,\mathbf{i}) + \nabla \bullet (Q\,\mathbf{j})\,.$$

We conclude that

Lemma. If the velocity of a fluid in the plane is given by a vector field $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j}$ such that $\dfrac{\partial P}{\partial x}$ and $\dfrac{\partial Q}{\partial y}$ are constant, then for any rectangular box with horizontal and vertical sides, the amount of fluid inside the box will decrease per unit time at a rate equal to the divergence of $\mathbf{F}$ times the area of the box. (It will increase if the divergence is negative.)

Now if one looks at an arbitrary differentiable two-dimensional vector field, then in a sufficiently small neighborhood of a given point (x_0, y_0), the partial derivatives $\dfrac{\partial P}{\partial x}$ and $\dfrac{\partial Q}{\partial y}$ will be very nearly constant. Thus the conclusion of the Lemma above will be true to a very close degree of approximation for any sufficiently small box constructed about (x_0, y_0). And one can in fact see that the divergence of $\mathbf{F}$ will equal the limit of the ratio between the rate at which fluid flows out of the box to the area of the box, as the size of the box shrinks to zero.

But since any reasonbly-shaped region can be closely approximated by dividing it up into tiny boxes, one gets the following standard characterization of divergence.

Theorem. Suppose that a two-dimensional vector field $\mathbf{F}$ represents the velocity of a fluid at a particular moment in time. Consider small regions surrounding a given point (x_0, y_0) and take the ratio of the rate at which fluid flows out of such a region (taken as negative if fluid is flowing inward) to the area of the region. Then the divergence of $\mathbf{F}$ at the point (x_0, y_0) is the limit of this ratio as the region around (x_0, y_0) shrinks to the single point (x_0, y_0).

An alternate way of stating this may be easier to grasp intuitively.

Theorem. Suppose that a two-dimensional vector field $\mathbf{F}$ represents the velocity of a fluid at a particular moment in time. Suppose that a region Ω in the plane is small enough so that the divergence $\nabla \bullet \mathbf{F}$ does not change very much over the region Ω. Then the rate at which fluid is flowing out of Ω is approximately equal to the product of the area of Ω and the divergence of $\mathbf{F}$ at any point within Ω.

It may be enlightening to compare the relationship between divergence and the rate of dissipation of fluid out of a region to the relationship between density and mass. If we are looking at the surface of a solid (perhaps metal) plate with varying thickness or density of meterial, and if $M(\Omega)$ represents the mass of that portion of the plate within a region Ω and $\mu(\mathbf{x})$ is the density of the plate (in grams per square centimeter, for instance) at a point $\mathbf{x}$, then we know that if Ω is small enough so that the density does not change a whole lot within Ω, then $M(\Omega) \approx A(\Omega)\mu$, where $A(\Omega)$ is the area and μ is the density at any point within Ω.

We see then that the relationship between divergence and dissipation of fluid from a region is exactly the same as the relationship beween desnity and mass.

This suggests the following theorem, which in fact is easy to prove using the preceding results:

Theorem. Suppose that a two-dimensional vector field $\mathbf{F}$ represents the velocity of a fluid at a particular moment in time. Let Ω be a finite region in the plane. Then the rate at which fluid flows out of Ω is given by

$$\iint_\Omega \nabla \cdot \mathbf{F}\, dx\, dy\,.$$

(If the integral is negative, then of course fluid is accumulating in the region rather than dissipating.)

How does one mathematically describe "the rate at which fluid is flowing out of a region Ω?"

If Ω is bounded by a simple closed curve $\mathcal{C}$ (i. e. one which is connected and does not intersect itself), let $\mathbf{n}$ denote the unit outward normal to this curve. I. e. at a point $\mathbf{x}$ on $\mathcal{C}$, $\mathbf{n}(\mathbf{x})$ is a unit vector perpendicular to $\mathcal{C}$ (i. e. to the tangent vector to $\mathcal{C}$) and pointing away from Ω. Then the rate at which fluid is flowing out of Ω at the point $\mathbf{x}$ is given by the product of $||\mathbf{F}||$ and the cosine of the angle between $\mathbf{F}(\mathbf{x})$ and $\mathbf{n}(\mathbf{x})$. (Thus the rate of flow is negative if the angle is obtuse, i. e. if $\mathbf{F}$ is directed toward the interior of Ω.) Since $\mathbf{n}$ is a unit vector, this product is given by $\mathbf{F}(\mathbf{x}) \bullet \mathbf{n}(\mathbf{x})$.

The total flow outward from Ω will then be given by the integral of $\mathbf{F} \bullet \mathbf{n}$ over the curve $\mathcal{C}$. Thus the preceding Theorem can be restated as follows:

Theorem. Suppose that a two-dimensional vector field $\mathbf{F}$ represents the velocity of a fluid at a particular moment in time. Let Ω be a region bounded by the simple closed curve $\mathcal{C}$ and let $\mathbf{n}$ denote the unit outward normal to the curve $\mathcal{C}$. Then

$$\oint_\mathcal{C} \mathbf{F}(\mathbf{x}) \bullet \mathbf{n}(\mathbf{x})\, ds = \iint_\Omega \nabla \bullet \mathbf{F}\, dx\, dy\,.$$

(Here ds is the differential corresponding to arc length on $\mathcal{C}$.)

This is the two-dimensional case of the **Divergence Theorem**.

This two-dimensional case can also be proved by means of Green's Theorem. Parametrize the curve by functions $x(t)$ and $y(t)$ such that $\sqrt{(x'(t)^2 + y'(t)^2} = 1$. (This is easy to do in principle, although often difficult in practice because the calculation can be nasty.) Then (possibly after a reversal of sign) $x'(t)\,\mathbf{i} + y'(t)\,\mathbf{j}$ is the unit tangent vector to $\mathcal{C}$ and so $y'(t)\,\mathbf{i} - x'(t)\,\mathbf{j}$ is the unit outward normal (rotated clockwise from the unit tangent). Let $\mathbf{F}(x,y) = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j}$. Then

$$\oint_{\mathcal{C}} \mathbf{F}(\mathbf{x}) \bullet \mathbf{n}(\mathbf{x})\, ds = \int_{t_0}^{t_1} P(x,y)\, y'(t) - Q(x,y)\, x'(t)\ dt$$
$$= \oint_{\mathcal{C}} -Q\, dx + P\, dy\,,$$

and by Green's Theorem this equals

$$\iint_{\Omega} \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\, dx\, dy = \iint_{\Omega} \nabla \bullet \mathbf{F}\, dx\, dy\,.$$

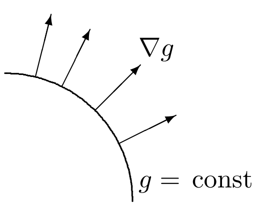
All this reasoning (except for the use of Green's Theorem) works just as well in three-dimensional space. One needs to consider a three-dimensional region T, whose boundary will be a surface S rather than a curve. One gets the three-dimensional theorem which is generally referred to as the Divergence Theorem.

Divergence Theorem. Suppose that a three-dimensional vector field $\mathbf{F}$ represents the velocity of a fluid at a particular moment in time. Let T be a finite region inf three-space bounded by a simple closed surface S. Then

$$\iint_S \mathbf{F} \bullet \mathbf{n}\, d\sigma = \iiint_T \nabla \bullet \mathbf{F}\, dx\, dy\, dz\,.$$

The Laplacean and the Heat Equation.

The divergence of the gradient field for a function g, i. e. $\nabla \bullet \nabla g$, is called the **Laplacean** of the function. This is relevant to the study of the flow of heat in a solid (usually metal).



It is worth noting that for a gradient field $\mathbf{F} = \nabla g$ in the plane, $\mathbf{F}$ is perpendicular to the level curves $g(x,y) = \text{const}$, and thus the rate at which $\mathbf{F}$ is turning as (x,y) moves in a direction perpendicular to $\mathbf{F}$ is the same as the rate at which the tangent vector to the level curve turns as one moves at unit speed along the level curve. This is, up to sign, the curvature of the level curve. (By definition, curvature is always positive.)

The Laplacean of a planar function $g(x,y)$ is given as follows:

(1) If ∇g points away from the direction in which the level curve for g curves at the given point, then $\nabla \bullet g$ is the sum of the rate at which $||\nabla g||$ changes as one moves in the direction of ∇g plus the product of $||\nabla g||$ times the curvature of the level curve for g at the given point.

(2) If ∇g points in the same direction towards which the level curve is curving, then $\nabla \bullet g$ is the sum of the rate at which $||\nabla g||$ changes as one moves in the direction of ∇g *minus* the product of $||g||$ and the curvature of the level curve.

For reasons which I'm not completely sure about, functions g with $\nabla \bullet \nabla g = 0$ are often called **harmonic**.

Note that any linear function $f(x,y) = ax + by$ is certainly harmonic, as are the functions $x^2 - y^2$ and xy. And by a calculation done above, the function $f(x,y) = \tan^{-1}(y/x)$ is harmonic, since $\nabla f = (y\,\mathbf{i} - x\,\mathbf{j})/r^2$, and we have seen that the vector fields $(y\,\mathbf{i} - x\,\mathbf{j})/r^n$ all have divergence 0. A very important set of harmonic functions consists of those of the form $f(x,y) = e^{ax}\sin ay$ and $f(x,y) = e^{ax}\cos ay$.

Furthermore, if $u_i(x,y)$ are harmonic functions, then any linear combination $\sum c_i u_i(x,y)$ is also harmonic. In fact, this is true (at least under reasonable restrictions) even when $\sum c_i u_i$ is an infinite series.

Based on the reasoning given above, we can see what it means for a planar function to be harmonic by looking at its gradient field and level curves. One can distinguish two cases: Case 1 is when ∇g increases when one moves away from the point (x,y) in the direction of increasing g (i. e. the direction of ∇g). And Case 2 is the case when ∇g decreases when one moves in the direction of ∇g. In Case 1, in order that $\nabla \bullet \nabla g(x,y) = 0$, the level curve through (x,y) must curve towards the direction of ∇g and furthermore the product of $||\nabla g||$ times the curvature of the level curve must equal the rate of increase of $||\nabla g||$ when (x,y) moves away in the direction of ∇g. And in Case 2, the level curve at (x,y) must curve away from the direction of ∇g and the product of $||g||$ and the curvature of the level curve must be the negative of the rate of increase of ∇g when (x,y) moves in the direction of ∇g.

The Heat Equation. Although heat is energy rather than mass, it flows in a way that makes it very much like a fluid in terms of mathematical structure. Heat flows from one point to another when there is a temperature difference between the two points, and the rate of flow is proportional to the temperature difference, but in the opposite direction (from a point with high temperature to one with lower). Thus if we look at a metal plate of homogeneous material and constant thickness, and let $u(x,y)$ be the temperature at a point (x,y), then heat flow is proportional to $-\nabla u$. (It is traditional in texts on partial differential equations to use the variable u for the function being considered.)

As we have mentioned, for a vector field $\mathbf{F}$ representing the velocity of a fluid in the plane, $-\nabla \bullet \mathbf{F}(x,y)$ gives the rate at which the fluid accumulates at (x,y). More precisely, we look at a small region Ω surrounding (x,y), and look at the amount of fluid accumulating in Ω in a unit time interval, and take the ratio of this to the area of Ω. If we then take the limit as we shrink the region Ω down to the single point (x,y), this value will be $-\nabla \bullet \mathbf{F}$.

Now when the “fluid” in question is heat, and the plane corresponds to a metal plate of constant thickness and homogeneous material, then the rate of accumulation of heat is proportional to the rate of temperature increase, i. e. $\dfrac{\partial u}{\partial t}$. Thus the changing temperature in a metal plate will be governed by the **heat equation**

$$a\frac{\partial u}{\partial t} = \nabla \bullet \nabla u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2},$$

where a is a positive constant of proportionality.

Now if one leaves the plate alone, then eventually the temperature will stabilize, so that $\partial u/\partial t = 0$. (In principle, it takes infinitely long before the temperature completely stabilizes, but for practical purposes it usually happens fairly quickly.) Now one might think that if the temperature stabilizes, this would mean that the plate would have a constant temperature all over. But this will not be the case if heat is being continually applied (or removed) at points around the boundary of the plate. For instance, one edge of the plate might be submerged in a bucket of ice water, thus giving it a temperature of exactly $0°$ Celsius, and the other edge in a flame or furnace.

Thus we have the *steady-state heat equation*, also known as **Laplace's Equation**,

$$\nabla \bullet \nabla u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0\,,$$

which describes the temperature distribution in a metal plate where points on the boundary are held at prescribed temperatures and the temperature distribution over the plate has stabilized.

This is a *partial differential equation* for the unknown function $u(x,y)$. The problem with this equation is not that it's difficult to find solutions. In fact, there are infinitely many solutions to the equation and many of them are quite well known. (These are the functions that we've called harmonic.) What is challenging is to find a solution that will take the prescribed values on the boundary of the plate. This sort of problem is known as a *boundary value problem.*

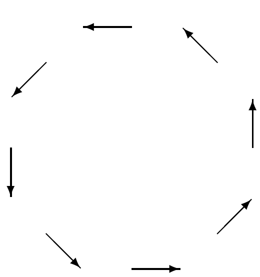
Curl

Unlike divergence, curl is something that only exists in three-dimensional space. It is usually defined by way of the cross product, and the cross product does not exist in the plane or in four-dimensional space.

In fact, defining curl as a vector is, in the language of computer programming, a "clever hack." Concepts such as curl, angular momentum, and torque really should be second order tensors (i. e. 3×3 matrices). But these particular matrices skew-symmetric, i. e. they are 0 on the diagonal and the half below the diagonal is the mirror of the half aobve, except with change of sign. Thus the matrix has only three distinct entries, and these can be used as the components of a vector.

Although defining the curl as a cross product works well on the symbolic level in several respects, it can also be misleading if taken too seriously. For instance, we know that the cross product of two vectors is perpendicular to each of them. Therefore it is plausible to conclude that $\nabla \times \mathbf{F}$ will be perpendicular to $\mathbf{F}$. However this is often not true. For instance if $\mathbf{F} = z\,\mathbf{i} + x\,\mathbf{j} + y\,\mathbf{k}$, then $\nabla \times \mathbf{F} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, and this is perpendicular to $\mathbf{F}$ only at points on the plane $z + x + y = 0$.

If we think of a vector field as representing the velocity of a fluid, then the curl corresponds roughly to the extent to which the fluid is swirling at a particular point. One tends to



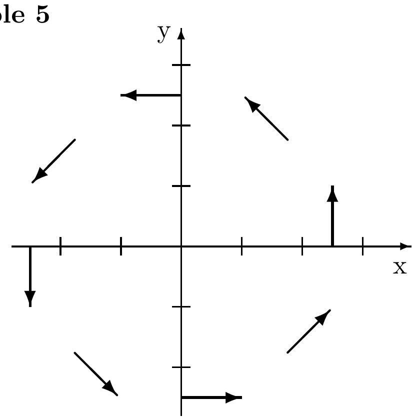
think of a vector field with a non-zero curl at a particular point as looking in a neighborhood of the point somewhat like the picture to the left. In fact, as one looks at this planar vector field as the reference point moves around the circle, one notices that what happens is that as one moves in the direction of the field and the direction of increasing x (look at the the bottom of the circle), as the vector rotates counter-clockwise, the $\mathbf{j}$-component increases (or becomes less negative), i. e. the direction of the vector moves upwards. And as one moves in the direction of increasing y, the

i-ccmponent decreases (or becomes more negative). Thus we have $\dfrac{\partial Q}{\partial x}$ and $-\dfrac{\partial P}{\partial y}$ both positive, thus giving a positive value to $\dfrac{\partial Q}{\partial x} - \dfrac{\partial P}{\partial y}$, which corresponds in this planar example to $\nabla \times \mathbf{F}$.

This seems to give a bit of intuitive significance to the formula for curl, however it is a bit simplistic in at least two ways. First of all, it is not only the direction of $\mathbf{F}$ that contributes to the curl. Surely if the magnitude of $\mathbf{F}$ is decreasing rapidly enough, for instance, then the magnitude of the x-coordinate and y-ccordinate of $\mathbf{F}$ will also decrease, regardless of which direction the vector is turning. Secondly, in this example $\mathbf{F}$ turns as we move in the same direction as $\mathbf{F}$. The analysis breaks down if the turning of $\mathbf{F}$ happens when we move in a direction roughly perpendicular to $\mathbf{F}$, as in Examples 1 and 2. In Example 1 we see that what happens is that as we move around the cirle and x increases, the vector turns in a negative direction (clockwise), but y at first increases and then decreases. And in fact, in Examples 1 and 2, $\nabla \times \mathbf{F} = \mathbf{0}$.

Thus it is incorrect to think that a non-zero curl corresponds to a twisting of the field.

Example 5



$\mathbf{F} = -ay\,\mathbf{i} + ax\,\mathbf{j}$

$\nabla \bullet \mathbf{F} = 0$

$\nabla \times \mathbf{F} = 2a\,\mathbf{k}$

However we will see that a turning of the vector field *when one moves in the direction of the field* is indeed one of the things that contributes to curl. In fact, the fundamental and archetypical example of a planar vector field with non-zero curl is the field

$$\mathbf{F} = -ay\,\mathbf{i} + ax\,\mathbf{j}$$

(Example 5). This is a linear field and its Jacobean matrix is $\begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}$. This represents the velocity of a point in the plane if the entire plane is rotated with an angular speed a. (One can think of a wooden disk being rotated, or an old-fashioned phonograph record on a turntable.) We have $\nabla \times \mathbf{F} = 2a\,\mathbf{k}$. Since the entire plane is being rotated, and the axis of rotation is $\mathbf{k}$, this makes sense. It's important to remember, though, that curl is something that happens at an individual point, not on the plane as a whole or merely at the origin. So one should consider that if one is on a rotating disk, and is walking away from the origin, one will experience a twisting, since the foot which is further away from the origin will be moving slightly faster than the other one. And in fact, if one stands on the rotating disk during the time interval in which the disk makes a complete revolution, in addition to traveling around the disk, one's body will also be rotated through $360°$ with respect to an external frame of reference. By the time the disk has made a complete revolution, one will have faced all four compass points. (Or, if one prefers to think of $\mathbf{F}$ as the velocity of a fluid, one can think of a person trapped in a whirlpool.)

A three-dimensional analog for the field in Example 5 is the field

$$\mathbf{F}(x,y,z) = (-cy+bz)\,\mathbf{i} + (cx-az)\,\mathbf{j} + (-bx+ay)\,\mathbf{j}\,.$$

This turns out to be the velocity vector for a rotation around the axis $a\,\mathbf{i}+b\,\mathbf{j}+c\,\mathbf{k}$. And $\nabla\times\mathbf{F} = 2(a\,\mathbf{i}+b\,\mathbf{j}+c\,\mathbf{k})$. I will come back to this example later, but it is easy to verify that for all $\mathbf{x}$, $\mathbf{F}(\mathbf{x})$ is perpendicular to $a\,\mathbf{i}+b\,\mathbf{j}+c\,\mathbf{k}$ and perpendicular to $\mathbf{x}$. (Except that $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ if the position vector $\mathbf{x}$ is parallel to $a\,\mathbf{i}+b\,\mathbf{j}+c\,\mathbf{k}$.)

These rotational examples are fundamental and yet also somewhat misleading. The non-zero curl here is not merely a consequence of the fact that the vector field rotates around the origin. To see this, change Example 5 a little, letting

$$\mathbf{G}(x,y) = r^n\mathbf{F}\,,$$

where $\mathbf{F}$ is the vector field $\mathbf{F} = -ay\,\mathbf{i} + ax\,\mathbf{j}$ of Example 5, $r = \sqrt{x^2+y^2}$, and n is an integer. The differentiation here may seem awkward because of the square root, but we can note that r^n increases (or decreases, if $n<0$) most rapidly in the direction radially away from the origin, and the rate of increase is $\dfrac{d}{dr}(r^n) = nr^{n-1}$. Thus $\nabla(r^n) = nr^{n-1}\mathbf{u}_r$, where $\mathbf{u}_r = (x\,\mathbf{i}+y\,\mathbf{j})/r$. Note that since $\mathbf{F}$ and $\mathbf{u}_r$ are perpendicular and $||\mathbf{F}|| = |a|r$, we have $\mathbf{u}_r\times\mathbf{F} = ar\,\mathbf{k}$. Thus

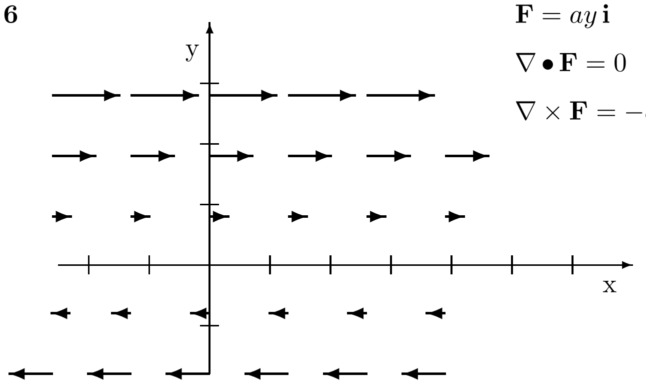
$$\begin{aligned}\nabla\times\mathbf{G} = \nabla(r^n)\times\mathbf{F} + r^n\,\nabla\times\mathbf{F} &= nr^{n-1}\mathbf{u}_r\times\mathbf{F} + 2ar^n\,\mathbf{k}\,.\\ &= anr^n\,\mathbf{k} + 2ar^n\,\mathbf{k} = (n+2)ar^n\,\mathbf{k}\,.\end{aligned}$$

In particular, if $n=-2$ then $\nabla\times\mathbf{G} = \mathbf{0}$. (Note that if n is negative, then $\mathbf{F}$ is discontinuous at $(0,0)$. In particular, for this case one should not try to apply Stoke's Theorem for any region containing the origin.)

Example 6. Let $\mathbf{F}(x,y,z) = ay\,\mathbf{i}$, where a is a constant. Then $\nabla\times\mathbf{F} = -a\,\mathbf{k}$.

This is a horizontal vector field in the plane, and $\nabla\times\mathbf{F}$ gives us the rate at which $\mathbf{F}$ changes as we move vertically. Although this is not a rotation, the non-zero curl coresponds to a twisting (or shearing) effect. If we were to walk in the y direction through this force field, we would feel a twisting effect, since, assuming that $a>0$, the force on our forward foot would be a little greater than that on the other foot.

Example 6



$\mathbf{F} = ay\,\mathbf{i}$

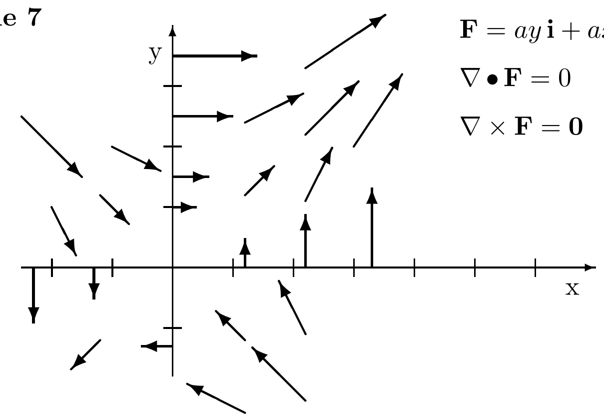
$\nabla \bullet \mathbf{F} = 0$

$\nabla \times \mathbf{F} = -a\,\mathbf{k}$

Or if we imagine $\mathbf{F}$ as representing the flow of water in a river, and if we were to move a boat in the y direction, we can see that the river would constantly be trying to turn the boat towards the x direction, since the force on the bow of the boat would be greater than that on the stern.

However one should not consequently make the simplistic assumption that somehow curl can be equated to torque. Unfortunately, the calculations just don't work out. Furthermore, consider the planar vector field $\mathbf{F} = ay\,\mathbf{i} + ax\,\mathbf{j}$ (Example 7). We see that $\nabla \times \mathbf{F} = \mathbf{0}$. However if this field represents the velocity of a current and if we move through this current with a boat (presumably longer than it is wide) in the $\mathbf{j}$ direction then the current will be pushing the boat more strongly at the bow (front) than at the stern (rear), and consequently will be exerting a torque, trying to turn the boat clockwise towards the $\mathbf{i}$ direction. On the other hand, if we the boat in the $\mathbf{i}$ direction, then the current will attempt to turn it counterclockwise towards the $\mathbf{j}$ direction.

Example 7



$\mathbf{F} = ay\,\mathbf{i} + ax\,\mathbf{j} = \nabla(axy)$

$\nabla \bullet \mathbf{F} = 0$

$\nabla \times \mathbf{F} = \mathbf{0}$

This shows that for a vector field $\mathbf{F}$ there is not a simple relationship between $||\nabla \times \mathbf{F}||$ and the torque exerted on an object placed in the field $\mathbf{F}$ in terms of the area (or volume) of that object.

To understand curl more systematically, start by considering a vector field with constant direction:

$$\mathbf{F}(x, y, z) = f(x, y, z)\mathbf{u}\,,$$

where $\mathbf{u}$ is a constant unit vector. Then

$$\nabla \times \mathbf{F} = \nabla f \times \mathbf{u}\,.$$

We see then that the curl of $\mathbf{F}$ is perpendicular to both the direction of $\mathbf{F}$ and to the gradient of f (i. e. the gradient of $||\mathbf{F}||$, assuming that $f > 0$). The curl is zero if the gradient of f is parallel to $\mathbf{u}$, i. e. to $\mathbf{F}$. And

$$||\nabla \times \mathbf{F}|| = ||\nabla f \times \mathbf{u}|| = ||\nabla f|| \sin\varphi = ||\nabla f|| \cos\psi\,,$$

where φ is the angle between ∇f and $\mathbf{F}$ and ψ is the complementary angle (in the plane of ∇f and $\mathbf{F}$). In other words, in the case of a field of constant direction, $\nabla \times \mathbf{F}$ measures the rate of change of $||\mathbf{F}||$ in a direction perpendicular to $\mathbf{F}$.

There are, of course, only two directions perpendicular to both ∇f and $\mathbf{F}$ (assuming that these two are not parallel), and if $\mathbf{w}$ is a vector representing one of these directions, then $-\mathbf{w}$ represents the other. The correct choice for the direction of $\nabla \times \mathbf{F}$ (or, in the planar case, for the sign of the curl, if one thinks of the curl as always being a multiple of $\mathbf{k}$) will be determined by the right-hand rule. For a field in the plane, $\nabla \times \mathbf{F}$ will be a positive multiple of $\mathbf{k}$ when ∇f is in a clockwise direction from $\mathbf{u}$, i. e. f increases in a direction clockwise from $\mathbf{u}$ (but not necessarily perpendicular to it).

On the other hand, consider a vector field in the plane with constant magnitude 1:

$$\mathbf{U}(x,y,z) = \cos\gamma\;\mathbf{i} + \sin\gamma\;\mathbf{j}\,,$$

where $\gamma(x,y)$ is the angle between $\mathbf{U}$ and $\mathbf{i}$. Then

$$\begin{aligned}
\nabla \times \mathbf{U} &= (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\,\mathbf{k} = \left(\frac{\partial}{\partial x}(\sin\gamma) - \frac{\partial}{\partial y}(\cos\gamma)\right)\,\mathbf{k} \\
&= (\cos\gamma\,\frac{\partial\gamma}{\partial x} + \sin\gamma\,\frac{\partial\gamma}{\partial y})\,\mathbf{k} \\
&= (\cos\gamma\,\mathbf{i} + \sin\gamma\,\mathbf{j}) \bullet (\frac{\partial\gamma}{\partial x}\,\mathbf{i} + \frac{\partial\gamma}{\partial y}\,\mathbf{j})\,\mathbf{k} \\
&= (\mathbf{U} \bullet \nabla\gamma)\,\mathbf{k}\,.
\end{aligned}$$

Since $\mathbf{U}$ is by assumption a unit vector, the scalar in parentheses here is the directional derivative of γ in the direction of $\mathbf{U}$.

I.e. for a vector field $\mathbf{U}$ in the plane with constant magnitude 1, $\nabla \times \mathbf{u}$ shows the rate at which $\mathbf{U}$ is turning as (x,y) moves in the direction of $\mathbf{U}$.

Since this theorem is so simple, it is very tempting to believe that it would also hold for vector fields in three dimensions. However this is not the case. Consider the the following example.

Example 8. $\quad \mathbf{F}(x,y,z) = \dfrac{-y\,\mathbf{i} + x\,\mathbf{j} + \sqrt{3}r\,\mathbf{k}}{2r} = \dfrac{-y\,\mathbf{i} + x\,\mathbf{j}}{2r} + \dfrac{\sqrt{3}}{2}\,\mathbf{k},$

where, as usual, $r = \sqrt{x^2 + y^2 + z^2}$. We have $||\mathbf{F}|| = 1$. Furthermore, the Jacobian matrix for $\mathbf{F}$, and hence also the curl, are the same as for the planar field $\frac{1}{2}\mathbf{G}$, where $\mathbf{G} = (-y\,\mathbf{i} + x\,\mathbf{j})/r$, which was considered immediately after Example 5. We found that $\nabla \times \mathbf{G} = \mathbf{k}/r$. Thus

$$\nabla \times \mathbf{F} = \frac{\mathbf{k}}{2r}\,.$$

Since $\mathbf{G}$ is a planar vector field with constant magnitude 1, the Theorem then tells us that the directional derivative of $\mathbf{G}$ in the direction of $-y\,\mathbf{i}+x\,\mathbf{j}$ is $1/r$, so the directional derivative of $\mathbf{F}$ in this direction is $1/2r$. Since $\mathbf{F}$ has constant magnitude 1, this is the same as the rate at which $\mathbf{F}$ is turning when (x,y) moves in this direction. Now $\mathbf{F}$ is at an angle of $30°$ to $\mathbf{k}$ and $\mathbf{F}$ does not change when z increases. From this we can see that the directional derivative of $\mathbf{F}$ in the direction of $\mathbf{F}$ is $1/2$ the directional derivative in the direction of $-y\,\mathbf{i}+x\,\mathbf{j}$. Thus for this three-dimensional case of a vector field with constant magnitude 1, the curl is not the same as the rate at which $\mathbf{F}$ turns as (x,y) moves in the direction of $\mathbf{F}$.

In fact, a close look at this example reveals that it would in fact impossible for a theorem to hold stating that $\nabla\times\mathbf{F}$ is the rate at which $\mathbf{F}$ turns when (x,y,z) moves in the direction of $\mathbf{F}$. Because $\nabla\times\mathbf{F}$ is completely determined by the Jacobean matrix $\mathcal{J}$ for the field $\mathbf{F}$ at the particular point (x,y,z), but the suggested theorem is stated in terms of the direction of $\mathbf{F}$. And $\mathcal{J}$ does not tell us either the direction or the magnitude of $\mathbf{F}$; it only tells us about how $\mathbf{F}$ changes.

In general, a planar vector field can be written as a product $\mathbf{F}(x,y)=f(x,y)\mathbf{U}(x,y)$, where $f(x,y)$ is a scalar function and $\mathbf{U}(x,y)$ is a vector field with constant magnitude 1. Then

$$(\star)\qquad \nabla\times\mathbf{F}=(\nabla f)\times\mathbf{U}+(f)(\nabla\times\mathbf{U})\,.$$

Theorem A. For a general vector field $\mathbf{F}$ in two dimensional space, $\nabla\times\mathbf{F}$ will be the product of $\mathbf{k}$ with sum of the following two terms:

(1) The rate at which $||\mathbf{F}||$ changes as one movess at unit speed in a direction $90°$ clockwise to $\mathbf{F}$;

(2) The product of $||\mathbf{F}||$ and the rate at which $\mathbf{F}$ turns as one moves in the direction $\mathbf{F}$ at unit speed.

One can also easily derive this result by using a change of coordinates. Fix a point (x_0,y_0) and introduce a new coordinate system so that $\mathbf{i}$ is in the direction of $\mathbf{F}$ at (x_0,y_0). Then if $\mathbf{F}(x,y)=P\,\mathbf{i}+Q\,\mathbf{j}$, the directional derivative of $\mathbf{F}$ in the $\mathbf{i}$ direction is $\dfrac{\partial P}{\partial x}\,\mathbf{i}+\dfrac{\partial Q}{\partial x}\,\mathbf{j}$. Recall that the first component here, which is in the direction of $\mathbf{F}$, gives the rate at which $||\mathbf{F}||$ is changing as we move in the direction of $\mathbf{F}$, and the second component, viz. $\dfrac{\partial Q}{\partial x}(x_0,y_0)$, equals the product of $||\mathbf{F}(x_0,y_0)||$ and the rate at which $\mathbf{F}$ is turning as one moves in the the direction of $\mathbf{F}(x_0,y_0)$. On the other hand, the directional derivative of $\mathbf{F}$ in the $-\mathbf{j}$ direction, is given by $-\dfrac{\partial P}{\partial y}\,\mathbf{i}-\dfrac{\partial Q}{\partial y}\,\mathbf{j}$. The component $-\dfrac{\partial P}{\partial y}\,\mathbf{i}$ is in the direction of $\mathbf{F}$ and thus $\dfrac{\partial P}{\partial y}$ is the rate of change of $||\mathbf{F}||$ as one moves away from (x_0,y_0) in the $-\mathbf{j}$ direction, i.e. the direction $90°$ clockwise to $\mathbf{F}(x_0,y_0)$, But $\nabla\times\mathbf{F}=(\dfrac{\partial Q}{\partial x}-\dfrac{\partial P}{\partial y})\,\mathbf{k}$. Thus the theorem follows.

The Theorem is not meant as a practical method for computing $\nabla\times\mathbf{F}$, since the basic formula is already quite simple to use in most cases. Instead, it is a way of trying to find the conceptual intuitive meaning of $\nabla\times\mathbf{F}$. However there are a few cases where this approach is slightly simpler than doing the differentiations. For instance, consider the planar vector field

$$\mathbf{F} = \frac{x\,\mathbf{i} + y\,\mathbf{j}}{r^n}$$

with $r = \sqrt{x^2 + y^2}$. Since this field is directed radially away from the center, the direction of $\mathbf{F}(x,y)$ does not change as (x,y) moves in the direction of $\mathbf{F}$. Furthermore, $||\mathbf{F}||$ is constant on the circles $x^2 + y^2 = \text{const}$, hence the directional derivative of $||\mathbf{F}||$ is zero in the direction perpendicular to the direction of $\mathbf{F}$. Thus we see that $\nabla \times \mathbf{F} = \mathbf{0}$.

Also look again at the field

$$\mathbf{G} = r^n(-ay\,\mathbf{i} + ax\,\mathbf{j})$$

which we considered immediately after Example 5. This field is tangent to the circles $x^2 + y^2 = \text{const}$, and thus the rate at which $\mathbf{G}$ turns as (x,y) moves in the direction of $\mathbf{G}$ equals the curvature of the circle around the origin through (x,y), namely $1/r$. Since $||\mathbf{G}|| = ar^{n+1}$, the second summand indicated in the Theorem equals $ar^{n+1}\,\dfrac{1}{r}\,\mathbf{k} = ar^n\,\mathbf{k}$. Furthermore, the directional derivative of $||\mathbf{G}||$ as one moves in a direction $90°$ clockwise to $\mathbf{G}$, i. e. along a radial line away from the center, is $a(n+1)r^n$, so the first summand indicated in the Theorem equals $a(n+1)r^n\,\mathbf{k}$. Thus one gets $\nabla \times \mathbf{G} = (n+2)ar^n\,\mathbf{k}$, as previously calculated.

This is not the end of the discussion of curl, but the rest will have to be postponed until I talk about some topics in linear algebra.

Eigenvectors and Eigenvalues.

As mentioned, in a small neighborhood of a point (x_0, y_0), a vector field in the plane can be closely approximated by one of the form

$$\mathbf{F}(x,y) = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j} = \mathbf{C} + (a_1x + a_2y)\,\mathbf{i} + (b_1x + b_2y)\,\mathbf{j},$$

where $\mathbf{C}$ is a constant vector. The Jacobean matrix corresponding to this field is

$$\mathcal{J} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}.$$

The nature of this matrix, and consequently the behavior of the original vector field within a small neighborhood, can be understood in terms of the eigenvalues and eigenvectors.

To say that a non-zero vector $\mathbf{v}$ is an **eigenvector** for a matrix A with corresponding **eigenvalue** c is to say that $A\mathbf{v} = c\mathbf{v}$. Any non-zero multiple of an eigenvector is also an eigenvector with the same eigenvalue, so that an eigenvector really corresponds more to a direction than to a particular vector.

As examples of eigenvectors, we can notice that the matrix $\begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$ has the eigenvectors $\mathbf{i} + 2\,\mathbf{j}$

and $2\,\mathbf{i}-\mathbf{j}$ with corresponding eigenvalues 5 and and 10. In fact,

$$\begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 20 \\ -10 \end{bmatrix} = 10 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Likewise the matrix $\begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}$ has the eigenvectors $\mathbf{i}+\mathbf{j}$ and $\mathbf{i}-\mathbf{j}$ with corresponding eigenvalues 6 and 4.

It is shown in linear algebra that the eigenvalues for an $n \times n$ matrix are the roots of a certain polynomial of degree n. It is also a fact that every polynomial of odd degree must have at least one real root, since its graph must cross the x-axis.

> It is a theorem in linear algebra that every 3×3 matrix, or for that matter, every $n \times n$ matrix for *odd* n, has at least one real eigenvector and corresponding eigenvalue.

Counting Eigenvectors. We have already mentioned that any multiple of an eigenvector is also an eigenvector, for the same eigenvalue. When we are counting eigenvectors, we don't want multiples to count as separate. On the other hand, it is possible that for certain matrices that two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ which are not multiples of each other both be eigenvectors corresponding to the same eigenvalue. This happens, so instance, with the matrix $\begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{bmatrix}$, where both $\mathbf{i}$ and $\mathbf{j}$ are eigenvectors corresponding to the same eigenvalue a. In this case, it is easy to see that any combination $r\mathbf{v}_1 + s\mathbf{v}_2$ is also an eigenvector corresponding to the same eigenvalue. We don't want to count these infinitely many eigenvectors separately, so in counting we indicate that this matrix has two linearly independent eigenvectors corresponding to the given eigenvalue, but no more.

An $n \times n$ matrix can have at most n linearly independent eigenvectors, since more than n vectors in $\mathbb{R}^n$ cannot be linearly independent. Also, it can be proven (fairly easily) that any set of eigenvectors corresponding to *distinct* eigenvalues is always linearly independent. So it is only when two or more eigenvectors correspond to the same eigenvalue that the issue of linear independence arises.

It is known from linear algebra that a *symmetric* $n \times n$ matrix (see below) always has n linearly independent eigenvectors. (It is possible that this can also happen for matrices which are not symmetric.)

> If an $n \times n$ matrix has n linearly independent eigenvectors, then the trace of the matrix is the sum of the corresponding eigenvalues.

Thus in a lot of cases we will be able to interpret the divergence at a particular point of a vector field in terms of the eigenvalues of the Jacobean matrix.

In terms of vector fields, consider a vector planar field $\mathbf{F}$ with corresponding Jacobean matrix $\mathcal{J}$ at a point (x_0, y_0) and suppose $\mathbf{v}$ is an eigenvector for $\mathcal{J}$ with corresponding eigenvalue c. Let us suppose that $\mathbf{v}$ is fairly small (as we may, since only the direction is crucial). Now if (x, y) is another point such that $(x, y) - (x_0, y_0) = \Delta\mathbf{x} = \mathbf{v}$, then $\mathcal{J}\Delta\mathbf{x} = c\,\Delta\mathbf{x}$, so that

$$\mathbf{F}(x, y) - \mathbf{F}(x_0, y_0) = \Delta\mathbf{F}(x_0, y_0) \approx \mathcal{J}\Delta\mathbf{x} = c\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}.$$

The corresponding equation in three dimensions, or for that matter in space of any dimensionality, is equally valid.

We want to define the concept of an eigenvector for a vector field $\mathbf{F}$ in such a way that an eigenvector for $\mathbf{F}$ is the same as an eigenvector for the Jacobean matrix. This can be accomplished by use of the directional derivative. Unfortunately, there's a slight technicality in that directional derivatives are normally only defined in terms of a *unit* vector in a given direction, but it's inconvenient to require that eigenvectors be unit vectors.

Definition. A vector $\mathbf{v}$ is an eigenvector for a vector field $\mathbf{F}$ at a point $\mathbf{x}$ if the directional derivative for $\mathbf{F}$ in the direction of $\mathbf{v}$ at the point $\mathbf{x}$ is a multiple of $\mathbf{v}$. If $\mathbf{u} = \mathbf{v}/\,||\mathbf{v}||$ and $D_{\mathbf{u}}\mathbf{F} = c\,\mathbf{u}$, then c is called the eigenvalue corresponding to $\mathbf{v}$.

Note that for a *linear* vector field

$$\mathbf{F}(x, y) = P(x, y)\,\mathbf{i} + Q(x, y)\,\mathbf{j} = \mathbf{F}(x_0, y_0) + (a_1 x + a_2 y)\,\mathbf{i} + (b_1 x + b_2 y)\,\mathbf{j},$$

the Jacobean matrix $\mathcal{J}$ is constant. From the above, this means that the vector field $\mathbf{F}$ looks the same, no matter what point we look at, except for the summand $\mathbf{F}(x_0, y_0)$:

$$\mathbf{F}(x, y) = \mathbf{F}(x_0, y_0) + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} = \mathbf{F}(x_0, y_0) + \begin{bmatrix} a_1(x - x_0) + a_2(y - y_0) \\ b_1(x - x_0) + b_2(y - y_0) \end{bmatrix}.$$

Thus for a linear field we can get the general idea by looking at $\mathbf{F}(x_0, y_0)$ where (x_0, y_0) is the origin. Furthermore, it makes it a lot easier to think about it if we replace $\mathbf{F}$ by the vector field $\mathbf{F} - \mathbf{F}(x_0, y_0)$, which has the same Jacobian. In other words, often we might as well consider the case where $\mathbf{F}(x_0, y_0) = \mathbf{F}(0, 0) = \mathbf{0}$.

Note that if $\mathbf{F}$ is a linear vector field with $\mathbf{F}(0, 0) = \mathbf{0}$, then all the vectors making up $\mathbf{F}$ are linear combinations of the columns of the Jacobian matrix $\mathcal{J}$:

$$\mathbf{F}(x, y) = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + y\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}.$$

If every two-dimensional vector is an eigenvector for $\mathcal{J}$, this says that $\mathcal{J} = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$, which is to say that for all points (x, y), $\mathbf{F}(x, y) = cx\,\mathbf{i} + cy\,\mathbf{j} = c\mathbf{x}$, in other words, the vector field $\mathbf{F}$ is radially directed away from the origin (or towards it if c is negative).

Otherwise, for a linear vector field $\mathbf{F}$ with $\mathbf{F}(0,0)=\mathbf{0}$, we see that $\mathbf{x}=x\,\mathbf{i}+y\,\mathbf{j}$ is an eigenvector of $\mathcal{J}$, with corresponding eigenvalue c, if

$$\mathbf{F}(x,y)=\mathbf{F}(0,0)+\mathcal{J}\begin{bmatrix}x\\y\end{bmatrix}=\mathbf{0}+\begin{bmatrix}cx\\cy\end{bmatrix}=\begin{bmatrix}cx\\cy\end{bmatrix}.$$

In other words, for a linear vector field with $\mathbf{F}(0,0)=\mathbf{0}$, the eigenvectors for $\mathcal{J}$ correspond to those directions such that when (x,y) lies in that direction from the origin, $\mathbf{F}(x,y)$ is directed radially away from (or towards) the origin.

> For a linear vector field $\mathbf{F}$ with $\mathbf{F}(0,0)=\mathbf{0}$, the eigenvectors can be recognized as those vectors in $\mathbf{F}$ which lie on straight lines through the origin, and also those position vectors $\mathbf{x}$ (aside from the origin) such that $\mathbf{F}(\mathbf{x})=\mathbf{0}$.

Thus if $\mathcal{J}$ has two linearly independent eigenvectors, then when we look at the vector field $\mathbf{F}$, we will see that some of the vectors in the field form two lines emanating from the origin (or directed toward the origin), and all the other vectors making up $\mathbf{F}$ will be pointing in at least slightly skewed directions. In fact, since for a linear vector field the Jacobean matrix is the same at all points, we see that corresponding to a planar vector field $\mathbf{F}$ with two linearly indepedent eigenvectors $\mathbf{v}$ and $\mathbf{w}$, there will be two key directions, the directions of $\mathbf{v}$ and $\mathbf{w}$. The difference between any two vectors $\mathbf{F}(x,y)$ along a straight line with the direction of $\mathbf{v}$ will be a multiple of $\mathbf{v}$ and likewise for $\mathbf{w}$. In particular, if a straight line in the direction $\mathbf{v}$ contains $(0,0)$, or any point (x_0,y_0) with $\mathbf{F}(x_0,y_0)$ in the direction of $\mathbf{v}$, then $\mathbf{F}(x,y)$ will be in the direction of $\mathbf{v}$ for all the points on that line. (And likewise, of course, for lines in the direction $\mathbf{w}$.)

In Example 3, $\mathbf{i}$ is an eigenvector corresponding to the eigenvalue $1/2$ and $\mathbf{j}$ is an eigenvector corresponding to $-1/2$. In Example 7, $\mathbf{i}+\mathbf{j}$ is an eigenvector corresponding to the eigenvalue 1 and $\mathbf{i}-\mathbf{j}$ is an eigenvector corresponding to the eigenvalue -1.

It is shown in linear algebra that every $n\times n$ matrix has at least one eigenvalue and corresponding eigenvector, provided that we allow complex numbers as eigenvalues and entries in eigenvectors. However since we are concerned here only with real numbers, it is possible that there may be no (real) eigenvector for a matrix $\mathcal{J}$. The most standard example of this is a matrix corresponding to a linear transformation which rotates vectors through an angle α, combined with scaling by a factor m. We have

$$\mathcal{J}=\begin{bmatrix}m\cos\alpha & -m\sin\alpha\\ m\sin\alpha & m\cos\alpha\end{bmatrix}$$

and

$$\mathbf{F}=(xm\cos\alpha-ym\sin\alpha)\,\mathbf{i}+(xm\sin\alpha+ym\cos\alpha)\,\mathbf{j}+\mathbf{F}(0,0)\,.$$

If $\mathbf{F}$ is a vector field which has $\mathcal{J}$ as its Jacobean matrix and if $\mathbf{F}(0,0)=\mathbf{0}$, then at every point (x,y), $\mathbf{F}(x,y)$ is turned counter-clockwise at an angle α from the radius vector $x\,\mathbf{i}+y\,\mathbf{j}$ and, if $m>0$, then $||\mathbf{F}||=mr$, where $r=\sqrt{x^2+y^2}$. In this case, we have $\nabla\bullet\mathbf{F}=2m\cos\alpha$ and $\nabla\times\mathbf{F}=2m\sin\alpha\;\mathbf{k}$.

The matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, in particular, represents a rotation of $90°$ counter-clockwise, and so has no (real) eigenvectors and eigenvalues. This matrix is the Jacobean matrix for the vector field

$$\mathbf{G} = y\,\mathbf{i} - x\,\mathbf{j} + \mathbf{G}(0,0)\,.$$

This may be a good moment to point out that when using the term "rotate," it is easy to get confused between two, or perhaps three, different things. A linear vector field is completely determined by its Jacobean matrix $\mathcal{J}$, but the vector field is not the same thing as the matrix, as one sees with the example $\mathbf{G} = ay\,\mathbf{i} - ax\,\mathbf{j}$, where $\mathcal{J} = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$. This vector field $\mathbf{G}$ here gives the velocity vector for a rotation of the plane with an angular velocity of a. On the other hand, $\mathcal{J}$ is the matrix of a linear transformation that rotates vectors through $90°$, combined with a expansion by a factor of a. This represents the fact that if we move away from a point $\mathbf{x}$ by a displacement $\Delta\mathbf{x}$, then $\Delta\mathbf{G}$ will be obtained by rotating $\Delta\mathbf{x}$ through an angle of $90°$ and multiplying its magnitude by a. Since $\mathbf{G}$ is linear and $\mathbf{G}(0,0) = \mathbf{0}$, this also represents the fact that we can obtain $\mathbf{G}(\mathbf{x})$ by rotating the position vector $\mathbf{x}$ by $90°$ and scaling by a. It is the scalar a, not the angle $90°$, that gives the angular speed of the rotation of the plane for which $\mathbf{G}$ is the velocity vector. (It will also be true that the vector $\mathbf{G}(x,y)$ will be turning, or one might sometimes say rotating, when the point (x,y) moves.)

Although eigen*vectors* are by definition non-zero, 0 is an allowable eigen*value*. If $\mathbf{v}$ is a eigenvector corresponding to the eigenvalue 0, then $\mathcal{J}\mathbf{v} = 0$. Thus an $n \times n$ matrix $\mathcal{J}$ has 0 as one of its eigenvalues if and only if $\mathcal{J}$ is a singular matrix. From linear algebra, we know that this is the case when the determinant of $\mathcal{J}$ is zero. There will then be a line through the origin (in the direction $\mathbf{v}$) along which $\mathbf{F}$ is constant (thus $\mathbf{F} = \mathbf{0}$ on this line in the special case $\mathbf{F}(0,0) = \mathbf{0}$).

Two by Two Matrices. In the case of a planar linear field for which there is an eigenvector $\mathbf{v}$ corresponding to the eigenvalue 0, there are now two possibilities. If there are any points (x,y) with $\mathbf{F}(x,y)$ is not parallel to $\mathbf{v}$, let $\mathbf{w} = \mathbf{F}(x,y)$ for such a point. Then $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, so that every vector in $\mathbb{R}^2$, in particular the position vector for any point, is a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Since $\mathbf{F}(x,y)$ remains constant when moving in the direction $\mathbf{v}$, we can see that $\mathbf{F}$ must be parallel to $\mathbf{w}$ in the whole plane (except along the line where it is $\mathbf{0}$). In particular, $\mathbf{F}(\,\mathbf{i})$ and $\mathbf{F}(\,\mathbf{j})$ will be multiples of $\mathbf{w}$, which is to say that the two columns of $\mathcal{J}$ are both multiples of $\mathbf{w}$. Furthermore, $\mathbf{F}(\mathbf{w})$ will be a multiple of $\mathbf{w}$. ($\mathbf{F}(\mathbf{w}) \neq \mathbf{0}$, otherwise $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ for every $\mathbf{x} \in \mathbb{R}^2$.) Thus $\mathbf{w}$ will be a second eigenvector for $\mathcal{J}$, with the corresponding eigenvalue being non-zero.

If, on the other hand, $\mathbf{F} \neq \mathbf{0}$ and $\mathbf{F}(\mathbf{v}) = \mathbf{0}$ and the entire vector field $\mathbf{F}$ is parallel to $\mathbf{v}$, then $\mathbf{F}(\mathbf{F}(\mathbf{x})) = \mathbf{0}$ for every $\mathbf{x} \in \mathbb{R}^2$, so that $\mathcal{J}^2$ is the zero matrix. $\mathbf{F}$ can have no second eigenvector (or eigenvalue) since if $\mathbf{w}$ were such an eigenvector, then $\mathbf{F}(\mathbf{w}) = c\mathbf{w}$ where $c \neq 0$ (otherwise it would follow that $\mathbf{F} = \mathbf{0}$, since $\mathbf{F}(\mathbf{v}) = 0$), but

$$\mathbf{0} = \mathbf{F}(\mathbf{F}(\mathbf{w})) = \mathbf{F}(c\mathbf{w}) = c\mathbf{F}(\mathbf{w}) = c^2\mathbf{w}\,.$$

a contradiction.

In either case, if $\mathbf{F}$ is a linear planar vector field and $\mathbf{F}(\mathbf{0}) = \mathbf{0}$ and $\mathbf{F}$ has an eigenvector corresponding to the eigenvalue 0, then $\mathbf{F}$ has constant direction.

Restated in terms of matrices rather than vector fields, we have the following:

Two by Two Theorem. For a non-zero 2×2 matrix $\mathcal{J}$, there are three possibilities.

(1) Neither of the two columns of $\mathcal{J}$ are multiples of the other.

(2) The two columns are multiples of each other or one column is zero. Furthermore, if $\mathbf{w}$ is a non-zero column, then $\mathcal{J}\mathbf{w}\neq\mathbf{0}$. In this case, $\mathbf{w}$ is an eigenvector for $\mathcal{J}$ corresponding to a non-zero eigenvalue. And $\mathcal{J}$ has a second eigenvector corresponding to the eigenvalue 0.

(3) The two columns of $\mathcal{J}$ are multiples of each other (or one of them is $\mathbf{0}$) and if $\mathbf{w}$ is either of these columns then $\mathcal{J}\mathbf{w}=\mathbf{0}$. In this case, $\mathbf{w}$ (if not zero) is the only eigenvector for $\mathcal{J}$ and corresponds to the eigenvalue 0. Furthermore, in this case $\mathcal{J}^2$ is the zero matrix.

Case (1) (the case where $\mathcal{J}$ is a non-singular matrix) could be divided into still further subcases, but it is Cases (2) and (3) that we're really interested in at the moment.

The Jacobean matrix for a planar vector field with either constant direction or constant magnitude will fall under Case (2) or Case (3). (However fields with constant direction are never linear.)

Examples 1, 2, and 3 are all examples of Case 1 vector fields.

Example 4, with $\mathcal{J}=\begin{bmatrix}-a & a\\ 0 & 0\end{bmatrix}$ is an example of Case (2). The eigenvector $\mathbf{i}$ corresponds to the non-zero eigenvalue $-a$ and $\mathbf{i}+\mathbf{j}$ is an eigenvector corresponding to the eigenvalue 0. On the line ℓ through the origin with equation $y=x$, $\mathbf{F}$ is 0. More generally, any time we move away from a given point in the direction $\mathbf{i}+\mathbf{j}$, $\mathbf{F}$ does not change. We have $\nabla\bullet\mathbf{F}=-a$.

Example 6, with Jacobean matrix $\begin{bmatrix}0 & a\\ 0 & 0\end{bmatrix}$ (for $a\neq 0$), is an example of Case (3). The only eigenvector is $a\,\mathbf{i}$ (or any multiple of it, in particular $\mathbf{i}$), corresponding to the eigenvalue 0.

Trajectories. The same ideas we have been using for linear vector fields can be applied to look at any vector field $\mathbf{F}$ at any point $\mathbf{x}_0$, if we remember that we are getting information about a very good approximation to $\mathbf{F}$ in a small neighborhood of $\mathbf{x}$. For eigenvectors to be visually apparent, though, one really needs to look at the field $\mathbf{F}-\mathbf{F}(\mathbf{x}_0)$.

The best way to see the field visually may be to look at the **trajectories** or **integral curves** for the field through a given point. These are the curves $\mathbf{x}(t)$ whose tangent vectors belong to the field $\mathbf{F}$, i. e. $\dfrac{d}{dt}\mathbf{x}(t)=\mathbf{F}(\mathbf{x}(t))$. (This is the curve that will be followed by a cork placed in the field, if the field is two-dimensional and represents the surface of a moving stream.)

We've seen that for a linear vector field $\mathbf{F}$ with $\mathbf{F}(\mathbf{0})=\mathbf{0}$, the eigenvectors can be recognized as those vectors $\mathbf{F}$ which lie along straight lines through the origin, as well as those position vectors $\mathbf{x}$ such that $\mathbf{F}(\mathbf{x})=\mathbf{0}$. For a non-linear field, the difference is that one should look for trajectories instead of straight lines. Furthermore, since the Jacobean matrix $\mathcal{J}$ is not constant, one cannot assume that the point $\mathbf{x}_0$ of interest is the origin.

The eigenvectors for $\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{x}_0)$, at the point $\mathbf{x}_0$ will correspond to the trajectories for $\mathbf{F}-\mathbf{F}(\mathbf{x}_0)$ that go through $\mathbf{x}_0$. (One also needs to include here curves through $\mathbf{x}_0$ on which $\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{x}_0)=\mathbf{0}$.)

(Strictly speaking, according to the definition given, one could object that a curve $\mathbf{x}'(t)$ going through $\mathbf{x}_0$ cannot be considered a trajectory at $\mathbf{x}_0=\mathbf{x}(t_0)$ since it doesn't make sense to say that $\mathbf{x}'(t_0)=\mathbf{F}(x,y)-\mathbf{F}(x_0)$ because $\mathbf{F}(\mathbf{x})-\mathbf{F}(x_0)$ is zero there, hence no direction is specified. But visually, a curve $\mathbf{x}(t)$ will be seen to be a trajectory for $\mathbf{F}$ at $\mathbf{x}_0$ if for points on the curve very close to $\mathbf{x}_0$, the tangent vector to the curve is headed directly towards, or directly away from, $\mathbf{x}_0$. To express this more formally, the condition required is that if $\mathbf{x}=\mathbf{x}(t)$ is very close to $\mathbf{x}_0=\mathbf{x}(t_0)$, so that $\Delta\mathbf{x}=\mathbf{x}-\mathbf{x}_0$ is small, then the tangent vector to the curve at $\mathbf{x}$, should be in the same direction as $\Delta\mathbf{x}$. But the tangent vector to the curve at $\mathbf{x}$ which is $\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{x}_0)$ since the curve is a trajectory for $\mathbf{F}-\mathbf{F}(\mathbf{x}_0)$. So this is the same as saying that $\Delta\mathbf{x}$ is an eigenvector for $\mathbf{F}-\mathbf{F}(\mathbf{x}_0)$, and also that for this specific $\mathbf{x}$, $\mathbf{F}(\mathbf{x})-\mathbf{F}(x_0)$ is an eigenvector for $\mathbf{F}-\mathbf{F}(\mathbf{x}_0)$. This little technicality also explains that it is possible for more than one trajectory for $\mathbf{F}-\mathbf{F}(\mathbf{x}_0)$ to go through the point $\mathbf{x}_0$.)

It is fairly easy to see these in the examples shown in the preceding graphics, since it is easy to visualize the trajectories. For instance, in Example 7, the only trajectories which pass through the origin are the straight lines with slopes of ± 1, corresponding to the eigenvectors $\mathbf{i}+\mathbf{j}$ and $\mathbf{i}-\mathbf{j}$ at the origin. (The Jacobean matrix is $\begin{bmatrix} 0 & .4 \\ .4 & 0 \end{bmatrix}$.)

If we look at a vector field as representing the velocity of a moving fluid, then what this says is that the eigenvectors at a point $\mathbf{x}_0$ correspond, roughly speaking, to little streams within the current that flow either either directly towards or directly away from $\mathbf{x}_0$, and the corresponding eigenvalues correspond to the speeds of these streams, except that, contrary to the usual usage, we here use the word "speed" with the understanding that it take have negative as well as positive values. An eigenvector corresponding to the eigenvalue 0 would correspond to a line (or curve, unless the scope of our vision is totally microsopic) going through the point in question where the fluid absolutely motionless. (Think of a vertical curve climbing through the eye of a hurricane, for example.)

If there are three linearly independent eigenvectors for the field (or two, in the planar case), then the divergence of the field is the sum of the corresponding eigenvalues. With the interpretation we have given, it is thus easy to see the fact that the divergence equals the rate at which fluid disappears (or accumulates, if the divergence is negative) at the given point.

We mentioned earlier that the canonical example of a matrix with no real eigenvectors is

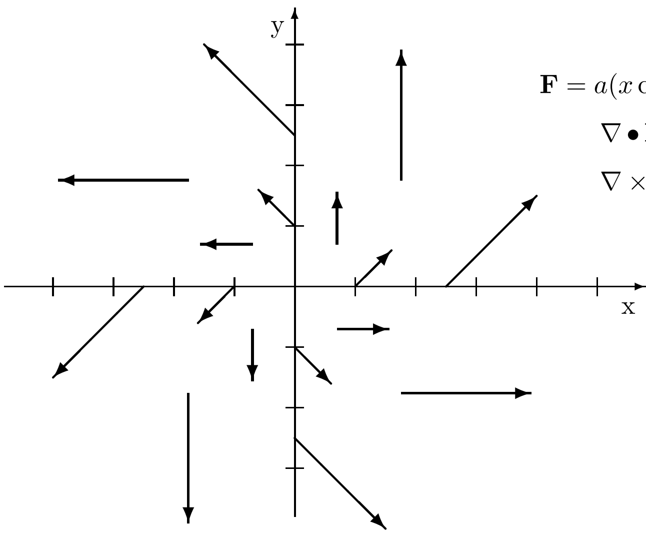
$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

This is the matrix for the linear transformation which rotates each vector counterclockwise through an angle of α. It is the Jacobean matrix for the planar linear vector field

$$\mathbf{F}=(x\cos\alpha-y\sin\alpha)\,\mathbf{i}+(x\sin\alpha+y\cos\alpha)\,\mathbf{j}$$

(Example 9). If $0<\alpha<\pi/2$, as in the picture, then we see that $\mathbf{F}$ is the velocity vector for the motion of a fluid which is swirling around the origin and spiraling outward. We have $\nabla\bullet\mathbf{F}=2\cos\alpha$, corresponding to the fact that fluid is moving outward away from the origin, even through there are no trajectories showing movement directly away from the origin.

Example 9



$$\mathbf{F} = a(x\cos\alpha - y\sin\alpha)\,\mathbf{i} + a(x\sin\alpha + y\cos\alpha)\,\mathbf{j}$$

$$\nabla \bullet \mathbf{F} = 2a\cos\alpha$$

$$\nabla \times \mathbf{F} = 2a\sin\alpha\,\mathbf{k}$$

Curl and Skew-symmetric Matrices

As already mentioned, curl is a phenomenon that occurs in three-dimensional space. The two-dimensional fields we have looked at so far are enlightening up to a point, but they really only partially address the concept. In particular, so far we have no insight whatsoever into the significance of the direction of the vector $\nabla \times \mathbf{F}$.

To penetrate this mystery we will need a few simple concepts from matrix theory. The **(main) diagonal** of a matrix consists of the entries running diagonally from the upper left corner to the upper right corner:

$$\begin{bmatrix} D & * & * \\ * & D & * \\ * & * & D \end{bmatrix}.$$

The **transpose** A^{tr} of a matrix A is its mirror image if one lays a mirror along the main diagonal. (Another way of saying this is that the columns of the transpose are the same as the rows of the original matrix.) A matrix is **symmetric** if it is the same as its transpose, i. e. if the entries above the diagonal are the mirror image of the ones below. And a matrix is **skew-symmatric** if it is the negative of its transpose, i. e. the entries above the main diagonal are the mirror image of the ones below except for a change in sign. (The diagonal entries of a skew-symmetric matrix must be 0.)

For example, if $A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 10 & 11 & 12 \end{bmatrix}$, then $A^{\text{tr}} = \begin{bmatrix} 1 & 5 & 10 \\ 2 & 6 & 11 \\ 3 & 7 & 12 \end{bmatrix}$. The matrix $\begin{bmatrix} 7 & 4 & 5 \\ 4 & 8 & 6 \\ 5 & 6 & 10 \end{bmatrix}$ is symmetric and $\begin{bmatrix} 0 & 6 & 7 \\ -6 & 0 & 8 \\ -7 & -8 & 0 \end{bmatrix}$ is skew-symmetric.

It is easy to see that the sum of a matrix and its transpose will be symmetric and the difference of a matrix and its transpose will be skew-symmetric. Any $n \times n$ matrix can be written as the sum of a symmetric matrix and a skew symmetric one, since we have

$$A = \frac{1}{2}(A + A^{\text{tr}}) + \frac{1}{2}(A - A^{\text{tr}}).$$

Suppose that a particular point the Jacobean matrix for a vector field $\mathbf{F}$ is

$$\mathcal{J} = \begin{bmatrix} a_{11}(\mathbf{x}) & a_{12}(\mathbf{x}) & a_{13}(\mathbf{x}) \\ a_{21}(\mathbf{x}) & a_{22}(\mathbf{x}) & a_{23}(\mathbf{x}) \\ a_{31}(\mathbf{x}) & a_{32}(\mathbf{x}) & a_{33}(\mathbf{x}) \end{bmatrix}.$$

Then

$$\nabla \times \mathbf{F} = (a_{32}(\mathbf{x}) - a_{23}(\mathbf{x}))\,\mathbf{i} - (a_{31}(\mathbf{x}) - a_{13}(\mathbf{x}))\,\mathbf{j} + (a_{21}(\mathbf{x}) - a_{12}(\mathbf{x}))\,\mathbf{k}.$$

Then what we notice is that $\nabla \times \mathbf{F} = \mathbf{0}$ if and only if $\mathcal{J}$ is a symmetric matrix.

We have seen that a 2×2 skew-symmatrix matrix $\begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$ (for $a \neq 0$) is the Jacobean matrix corresponding to a rotation of the plane with angular velocity $a\,\mathbf{k}$ and has no real eigenvectors. We will see that a 3×3 skew-symmetric matrix has exactly one real eigenvector, and this corresponds to the eigenvalue 0.

Now an arbitrary 3×3 skew-symmetric matrix looks like

$$\mathcal{J} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}.$$

This is the Jacobean matrix for the linear vector field

$$\mathbf{F} = \mathbf{F}(0,0,0) + (-cy + bz)\,\mathbf{i} + (cx - az)\,\mathbf{j} + (-bx + ay)\,\mathbf{k}.$$

We will look at the case where the constant term $\mathbf{F}(0,0,0)$ is $\mathbf{0}$. We see that $\nabla \times \mathbf{F} = 2(a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k})$. Now look at $\mathbf{F}(\mathbf{x})$ for an arbitrary point $\mathbf{x} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$. Calculation shows

$$\mathbf{F}(x,y,z) = (-cy + bz)\,\mathbf{i} + (cx - az)\,\mathbf{j} + (-bx + ay)\,\mathbf{k} = \tfrac{1}{2}(\nabla \times \mathbf{F}) \times \mathbf{x}.$$

From this we see two things: (1) $\nabla \times \mathbf{F}$ is an eigenvector for $\mathcal{J}$ corresponding to the eigenvalue 0; (2) $\mathbf{F}$ consists of the vectors in the plane perpendicular to $\nabla \times \mathbf{F}$ and also perpendicular to the radius vector $\mathbf{x}$. Thus under the asuumption that $\mathbf{F}(0,0,0) = \mathbf{0}$, the vector field $\mathbf{F}$ is the velocity of a rotation of three-space around the axis through the origin in the direction $a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k}$ with an angular velocity of $||a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k}||$.

It is very tempting to believe that if $\mathbf{F}$ is a vector field with Jacobean matrix $\mathcal{J}$, and we decompose $\mathcal{J}$ into the sum of a symmetric and a skew-symmetric matrix, then there would exist a corresponding way of writing $\mathbf{F}$ as a sum of two fields, one of which is the field of velocity vectors for a rotation and the other is a field with three mutually orthogonal eigenvectors at every point. This is certainly easily done in the case that $\mathbf{F}$ is linear, but in general it is not feasible, because it is usually not possible to find a vector field with pre-assigned Jacobean matrix. This is the problem of finding three functions P, Q, and R whose gradients are the rows of the given matrix, and is not possible unless certain compatibility conditions are satisfied. (See my article on integrating vector fields.)

The best one can say then is that in a sufficiently small neighborhood of a point of interest a vector field $\mathbf{F}$ can be reasonably well approximated by a linear field, and therefore it will look pretty much as though it is the sum of two fields, one of which has three mutually orthogonal eigenvectors and has zero curl, and the other of which represents the velocity vectors (within the given neighborhood) of a rotation of all of three space around the axis $\nabla\times\mathbf{F}$ with an angular speed of $\frac{1}{2}\,||\nabla\times\mathbf{F}||$.

It seems to me that this is the best intuitive interpretation that one can give for the concept of curl. However, such a decomposition of even a linear vector field with $\mathbf{F}(\mathbf{0})=\mathbf{0}$ into the sum of a field with zero curl and one which corresponds to a rotation of 3-space is often not at all apparent visually.

Consider again Example 9:

$$\mathbf{F}(x,y)=(x\cos\alpha-y\sin\alpha)\,\mathbf{i}+(x\sin\alpha+y\cos\alpha)\,\mathbf{j}\,,$$

where α is a constant. We saw that this linear field is the velocity vector for the motion of a fluid which is swirling around the origin and spiralling outward. We can break break $\mathbf{F}$ up into the sum of two terms, one having a symmetric Jacobean matrix and the other a skew-symmetric one. Namely, $\mathbf{F}_1=(x\,\mathbf{i}+y\,\mathbf{j})\cos\alpha$, and $\mathbf{F}_2=(-y\,\mathbf{i}+x\,\mathbf{j})\sin\alpha$. The first field represents the velocity vectors of a fluid streaming radially away from the origin, and the second the velocity vector for a rotation of the plane around the origin with an angular velocity of $\sin\alpha\,\mathbf{k}$. In terms of this decomposition, it makes perfectly good sense that that $\nabla\bullet\mathbf{F}=2\cos\alpha$ and $\nabla\times\mathbf{F}=2\sin\alpha\,\mathbf{k}$.

Now look again at Example 6, $\mathbf{F}=ay\,\mathbf{i}$. When we first looked at it, it seemed a little surprising that it had non-zero curl: $\nabla\times\mathbf{F}=-a\,\mathbf{k}$. Now write $\mathbf{F}=\mathbf{F}_1+\mathbf{F}_2$, where $\mathbf{F}_1(x,y)=\frac{1}{2}(ay\,\mathbf{i}+ax\,\mathbf{j})$, and $\mathbf{F}_2=\frac{1}{2}(ay\,\mathbf{i}-ax\,\mathbf{j})$. The Jacobean matrix for $\mathbf{F}_1$ is symmetric and the one for $\mathbf{F}_2$ is skew-symmetric. $\mathbf{F}_1$ is actually Example 7, multiplied by $\frac{1}{2}$, the velocity vector for a current that contains one substream streaming directly toward the origin at a 45° angle and a speed of a, and a perpendicular one streaming directly way from the origin at the same speed. $\mathbf{F}_2$ on the other hand is the velocity vector for a clockwise rotation of the plane with an angular velocity of $a\,\mathbf{k}$. But even after one knows this, it seems rather hard to look at $\mathbf{F}$ and see $\mathbf{F}_1$ and $\mathbf{F}_2$.

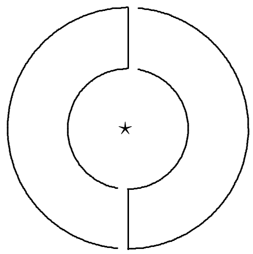
A Vector Field with No Curl is a Gradient.

(1) For a vector field $\mathbf{F}$ defined in a given region (in two or three dimensional space), $\nabla\times\mathbf{F}=\mathbf{0}$ if and only if each point $\mathbf{x}$ in the region is surrounded by some subregion in which $\mathbf{F}$ is a gradient field, i. e. there exists a function g in that subregion such that $\mathbf{F}=\nabla g$ in that subregion.

(2) If $\nabla\times\mathbf{F}=\mathbf{0}$ for a vector field $\mathbf{F}$ at a point $\mathbf{x}_0$ in $\mathbb{R}^n$, then $\mathbf{F}$ has n mutually orthogonal eigenvectors at $\mathbf{x}_0$.

(3) If $\mathbf{x}_0$ is a critical point for a twice-differentiable function g in $\mathbb{R}^n$, then g has a maximum at $\mathbf{x}_0$ if all the eigenvalues of ∇g at $\mathbf{x}_0$ are strictly negative, and a minimum if all the eigenvalues are strictly positive. If some of the eigenvalues are strictly positive and some are strictly negative (including the possibility that others are zero), then g has a saddle point at $\mathbf{x}_0$.

Statement (1) is discussed and proved in my article on integrating vector fields.



At first glance, the way this first statement is stated, in terms of subregions, seems a bit odd. One would think that one could take all these subregions, and the corresponding functions defined in them, and paste them together to get one function g defined on the whole original region such that $\mathbf{F} = \nabla g$. In fact, this is most often the case. But a problem occasionally occurs when the original region winds around a discontinuity for $\mathbf{F}$. The gradient ∇g determines g only up to a constant summand, and there may be a difficulty in consistently choosing this constant in a way that makes g continuous throughout the whole original region. One can get a situation similar to the picture on the left. One may be able to make the functions defined in the left-hand and right-hand regions agree where they meet along the top boundary line, but they may then be inconsistent along the bottom boundary. (The $\star$ in the middle indicates a point of discontinuity of the vector field.)

The archetypical example of this is the vector field

$$\mathbf{F} = \frac{y\,\mathbf{i} - x\,\mathbf{j}}{x^2 + y^2}\,.$$

In either the left half-plane or right half-plane, one sees that $\mathbf{F}(x, y)$ is the gradient of the function $\tan^{-1}(y/x)$. But $\tan^{-1}(y/x)$ is discontinuous along the y-axis. One can easily adjust g to get a function continuous in the upper half-plane or lower half-plane with $\mathbf{F} = \nabla g$, or in fact in any region which does not wind around the origin. What we need is a function $g(x, y)$ so that if we set $\theta = g(x, y)$, then the point (x, y) has polar coordinates $r = \sqrt{x^2 + y^2}$ and θ. But one cannot define θ in a continuous manner in a region which winds around the origin without having a discontinuous break in θ along at least one line, because as the point (x, y) circles the origin to return to an original starting point, θ will have increased by 2π, thus producing an inconsistency. (Well, if one wants to be really weird in the way one defines g, then one could make the discontinuities occur along some curve through the origin other than a straight line. But the point is, there have to be discontinuities.)

Statement (2) is a direct consequence of a theorem in linear algebra that an $n \times n$ matrix A is symmetric if and only if it has n mutually orthogonal eigenvectors. Part of this is not difficult to prove. It's easy to see that if an $n \times n$ matrix A is symmetric, then for any two n-dimensional column vectors $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v} \bullet A\mathbf{w} = \mathbf{w} \bullet A\mathbf{v}$. If now $\mathbf{v}$ and $\mathbf{w}$ are eigenvectors for A with corresponding eigenvalues m and n, then we get

$$n\mathbf{v} \bullet \mathbf{w} = \mathbf{v} \bullet A\mathbf{w} = \mathbf{w} \bullet A\mathbf{v} = m\mathbf{v} \bullet \mathbf{w}\,.$$

If $m \neq n$, it then follows that $\mathbf{v} \bullet \mathbf{w} = 0$, i.e. $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

However it is unfortunately not very easy to prove in general that an $n \times n$ symmetric matrix has n linearly independent eigenvectors. But in the 2×2 case, it's easy to see that a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ will have two real eigenvectors provided that b and c have the same sign (for instance if $b = c$, making A symmetric). In fact, in linear algebra it is known that the eigenvalues for A are the solutions to the equation $(x - a)(x - d) - bc = 0$. Now the graph of $y = (x - a)(x - d)$ is a parabola directed upwards, and this intersects the x-axis at $x = a$ and $x = d$. If b and c have the same sign then bc is positive, and so the parabola $y = (x - a)(x - d)$ intersects the horizontal line $y = bc$, and intersects it at two points, except in the case $a = d$ and $b = c = 0$. Therefore A will have two distinct real eigenvalues and therefore two real eigenvectors. (In the exceptional case $a = c$ and $b = d = 0$, we have $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, so A has only the single eigenvalue a, but every vector in $\mathbb{R}^2$ is an eigenvector.)

Now let's look at statement (3). Let $g(x, y)$ be differentiable function of two variables. We will assume that g is continuously twice differentiable, meaning that the second partials $\dfrac{\partial^2 g}{\partial x^2}$, $\dfrac{\partial^2 g}{\partial x\, \partial y}$ and $\dfrac{\partial^2 g}{\partial y^2}$ exist and are continuous. Now $\nabla g = \dfrac{\partial g}{\partial x}\, \mathbf{i} + \dfrac{\partial g}{\partial y}\, \mathbf{j}$, and the Jacobean matrix for ∇g is

$$\mathcal{J} = \begin{bmatrix} \dfrac{\partial^2 g}{\partial x^2} & \dfrac{\partial^2 g}{\partial x\, \partial y} \\ \dfrac{\partial^2 g}{\partial x\, \partial y} & \dfrac{\partial^2 g}{\partial y^2} \end{bmatrix} .$$

Because this matrix is is symmetric, it has two linearly independent eigenvectors.

Now let $\mathbf{x}_0$ be a critical point for g, i. e. a point where $\nabla g = \mathbf{0}$. We have seen that if a vector field $\mathbf{F}$ in n-dimensional space such that $\mathbf{F}(\mathbf{x}_0) = \mathbf{0}$ has n linearly independent eigenvectors and all the corresponding eigenvalues at $\mathbf{x}_0$ are strictly positive, then $\mathbf{F}$ will be directed away from $\mathbf{x}_0$ throughout some neighborhood of $\mathbf{x}_0$. But if ∇g is directed away from $\mathbf{x}_0$ for all points near $\mathbf{x}_0$, this says that g is increasing when we move in any direction away from $\mathbf{x}_0$, which shows that g has a minimum at $\mathbf{x}_0$.

And if on the other hand all the eigenvalues are strictly negative, then throughout some neighborhood of $\mathbf{x}_0$, ∇g will be directed toward $\mathbf{x}_0$. This says that g is decreasing as (x, y) moves away from $\mathbf{x}_0$ from any direction, so g has a maximum at $\mathbf{x}_0$.

However if there are two non-zero eigenvalues with opposite signs, then g is sometimes increasing and sometimes decreasing as (x, y) moves away from $\mathbf{x}_0$, so that $\mathbf{x}_0$ is a saddle point for g.

Going back to the 2-dimensional case, write $A = \dfrac{\partial^2 g}{\partial x^2}$, $C = \dfrac{\partial^2 g}{\partial y^2}$, and $B = \dfrac{\partial^2 g}{\partial x\, \partial y}$. Then $AC - B^2$ is the determinant of the Jacobean matrix for ∇g. But it is known from linear algebra that the determinant of a matrix equals the product of the eigenvalues. Therefor $AC - B^2$ will be strictly positive if both eigenvalues for the Jacobean matrix are non-zero and have the same sign. If $\mathbf{x}_0$ is a critical point for g, this then indicates that g has either a maximum or a minimum at $\mathbf{x}_0$. We can determine the sign of the eigenvalues, and thus determine whether g has a maximum or a minimum at $\mathbf{x}_0$, by looking at the sum of the two eigenvalues, i. e. at $\nabla \bullet \nabla g = A + C$. Furthermore, if $AC - B^2 > 0$, then necessarily A and C have the same sign, so we can say that g has a maximum at $\mathbf{x}_0$ if $A < 0$ and a minimum if $A > 0$.

And $AC - B^2$ will be strictly negative if the eigenvalues are non-zero with opposite signs, indicating that $\mathbf{x}_0$ is a saddle point for g.

Finally, if $AC - B^2 = 0$, this indicates that $\mathcal{J}$ has an eigenvector corresponding to the eigenvalue 0. In this case, the behavior of the function is too delicate to figure out on the basis of the information in the Jacobean. For instance, consider the two functions $g_1(x, y) = x^2 + y^3$ and $g_2(x, y) = x^2 + y^4$. The gradients are $\nabla g_2 = 2x\, \mathbf{i} + 3y^2\, \mathbf{j}$ and $\nabla g_2 = 2x\, \mathbf{i} + 4y^3\, \mathbf{j}$. For both functions, the only critical point is at $(0,0)$. The Jacobeans are

$$\mathcal{J}_1 = \begin{bmatrix} 2 & 0 \\ 0 & 6y \end{bmatrix}$$

$$\mathcal{J}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 12y^2 \end{bmatrix}.$$

At the critical point $(0,0)$, $\mathcal{J}_1$ and $\mathcal{J}_2$ are equal and $\mathbf{i}$ is an eigenvector corresponding to the eigenvalue 2 and $\mathbf{j}$ is an eigenvector corresponding to the eigenvalue 0. But it is evident that $(0,0)$ is a minimum for x^2+y^4 and is neither a maximum nor a minimum for x^2+y^3 (and thus by definition a saddle point, although the graph looks nothing like a saddle).

What sort of vector field has both zero curl and zero divergence? Answer: The gradient field for a harmonic function. (See, for instance, Examples 1, 2, 3, 7, and 11.)

A vector field in n-space with zero curl and zero divergence at a particular point can be characterized by the fact that at the given point it has n orthogonal eigenvectors, and the sum of the corresponding eigenvalues is 0.

A planar vector field which has zero curl and zero divergence at a certain point will have a Jacobean matrix at that point of the form $\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$. The eigenvalues for this matrix are the solutions to the equation $(x^2-a^2)-b^2=0$, i.e. $\pm\sqrt{a^2+b^2}$. Corresponding eigenvectors are $\begin{bmatrix} -b \\ a-\sqrt{a^2+b^2} \end{bmatrix}$ and $\begin{bmatrix} -b \\ a+\sqrt{a^2+b^2} \end{bmatrix}$, which are orthogonal to each other. (Here, of course, $a = \dfrac{\partial P}{\partial x} = -\dfrac{\partial Q}{\partial y}$ and $b = \dfrac{\partial Q}{\partial x} = \dfrac{\partial P}{\partial y}$.)

The Maximum Principle for Harmonic Functions. If $g(x,y)$ is a harmonic function, then $\nabla \bullet \nabla g = \mathbf{0}$. This shows that the two eigenvalues for ∇g have opposite signs, and this in turn indicates that any critical point for g is a saddle point.

Now consider a set Ω which is bounded and closed (i. e. includes all its boundary points). It is known from a theorem in topology that, as a continuous function, g must have a maximum (and also a minimum) somewhere in Ω. But we have seen that this maximum or minimum cannot occur at a critical point for g. Therefore it must occur at a boundary point. This give us the Maximum Principle, which turns out to be extremely useful: **The points within a closed bounded set Ω where a harmonic function g takes its maximum and minimum must lie on the boundary of that set.**

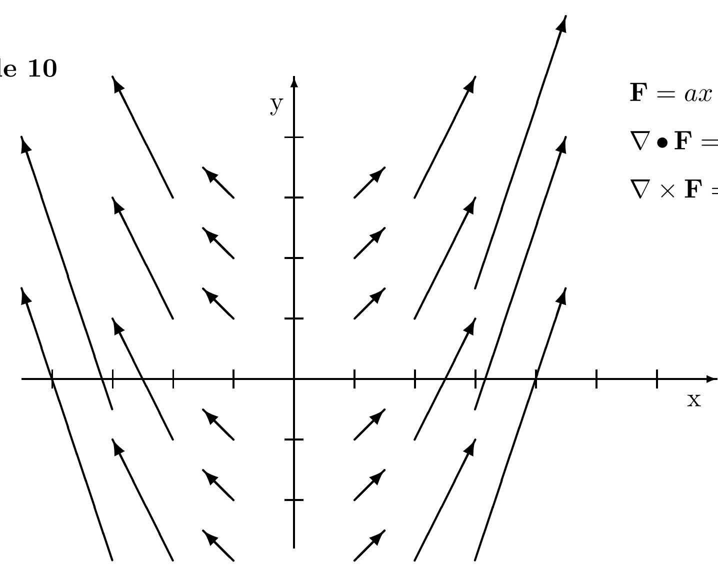
Finally, note that if, say, a 3×3 matrix $\mathcal{J}$ has three mutually orthogonal eigenvectors, this in turn means that by means of a rotation of coordinates, $\mathcal{J}$ can be put the form $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$. (In considering studying curl, it is not good to make any change of coordinates other than combinations of translations and rotations, i. e. rigid motions.) Conceptually, this is very enlightening in understanding the nature of vector fields with zero curl. However one should remember that a specific change of coordinates will usually only simplify the form of the Jacobean matrix at a single point.

Some Non-linear Vector Fields

We have seen that a linear planar vector field $\mathbf{F}$, has constant direction if and only if the Jacobean matrix $\mathcal{J}$ has 0 as an eigenvalue, or equivalently, $\mathcal{J}$ is a singular matrix. (Except for the trivial case where $\mathbf{F}=\mathbf{0}$ everywhere. In this case, $\mathcal{J}$ is certainly singular, since $\mathcal{J}=0$, but $\mathbf{F}$ has no direction at all.) More generally, in the case of a field in 3-space or higher dimensional space, a linear vector field $\mathbf{F}$ will have constant direction if and only if the Jacobean matrix $\mathcal{J}$ has rank one.

However the same need not be true if the vector field $\mathbf{F}$ is not linear. For instance, the Jacobean for the planar vector field $ax\,\mathbf{i}+ax^2\,\mathbf{j}$ in Example 10 is $\begin{bmatrix} a & 0 \\ 2ax & 0 \end{bmatrix}$, which is singular at every point. The vector $\mathbf{j}$ is always an eigenvector, and corresponds to the eigenvalue 0. (Notice that if (x,y) moves in the direction of $\mathbf{j}$, i. e. the direction of increasing or decreasing y, then $\mathbf{F}$ does not change.) But $\mathbf{F}$ is in the direction $\mathbf{i}+x\,\mathbf{j}$, which is not constant. At a point (x,y), the second eigenvector for $\mathcal{J}$ is the vector $\mathbf{i}+2x\,\mathbf{j}$, with corresponding eigenvalue a.

Example 10



$\mathbf{F}=ax\,\mathbf{i}+ax^2\,\mathbf{j}$

$\nabla\bullet\mathbf{F}=a$

$\nabla\times\mathbf{F}=2ax\,\mathbf{k}$

As another example, consider a planar vector field $\mathbf{U}$ with constant magnitude 1. As $\mathbf{U}$ moves it will change direction but not length, so we expect its directional derivatives to all be perpendicular to it. Thus if $\mathbf{w}$ is the unit vector perpendicular to $\mathbf{U}$ turned in, say, the positive (counterclockwise) direction from $\mathbf{U}$, then the directional derivative $D_{\mathbf{w}}(\mathbf{U})$ will be a multiple (possibly negative) of $\mathbf{w}$. Thus $\mathbf{w}$ will be an eigenvector for $\mathbf{U}$ and the corresponding eigenvalue will indicate how fast $\mathbf{U}$ turns as (x,y) moves in the direction $\mathbf{w}$.

Furthermore, the two columns of $\mathbf{U}$, which are the directional derivatives of $\mathbf{U}$ in the $\mathbf{i}$ and $\mathbf{j}$ directions, will also be multiples of $\mathbf{v}$, and hence are multiples of each other, unless one of them is zero. So this puts us in Case (2) or Case (3) of the Two by Two Theorem, so that $\mathbf{U}$ also has an eigenvector who corresponding eigenvalue is 0. Since $\nabla \bullet \mathbf{U}$ equals the sum of the two eigenvalues, this gives a second (although certainly not easier!) proof of the fact that $\nabla \bullet \mathbf{U}$ equals the rate at which $\mathbf{U}$ turns as one moves in the direction 90° counterclockwise from $\mathbf{U}$.

We can also see this computationally. We have seen that $\mathbf{U}$ has the form

$$\mathbf{U}(x,y) = \cos\gamma(x,y)\,\mathbf{i} + \sin\gamma(x,y)\,\mathbf{j}\,.$$

The Jacobean matrix for this non-linear vector field is

$$(\star\star)\qquad \mathcal{J} = \begin{bmatrix} -\dfrac{\partial\gamma}{\partial x}\sin\gamma & -\dfrac{\partial\gamma}{\partial y}\sin\gamma \\ \dfrac{\partial\gamma}{\partial x}\cos\gamma & \dfrac{\partial\gamma}{\partial y}\cos\gamma \end{bmatrix}.$$

The vector field $\mathbf{U}$, and consequently the matrix $\mathcal{J}$ are thus functions of a single parameter $\gamma(x,y)$. Thus the directional derivative of $\mathbf{U}$ will be zero when (x,y) moves in a direction perpendicular to $\nabla\gamma$, i.e. in the direction $\dfrac{\partial\gamma}{\partial y}\mathbf{i} - \dfrac{\partial\gamma}{\partial x}\mathbf{j}$. Thus $\dfrac{\partial\gamma}{\partial y}\mathbf{i} - \dfrac{\partial\gamma}{\partial x}\mathbf{j}$ is an eigenvector for $\mathcal{J}$ corresponding to the eigenvalue 0.

But despite having 0 as an eigenvalue, $\mathbf{U}$ is very definitely not a field with constant direction, except in the uninteresting case when $\gamma(x,y)$ is constant.

We see that the two columns of $\mathcal{J}$ are multiples of $-\sin\gamma\;\mathbf{i} + \cos\gamma\;\mathbf{j}$, which is the direction perpendicular to $\mathbf{U}$. Thus for any vector $\mathbf{v}$, $\mathcal{J}\mathbf{v}$ is always a multiple of $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$, and this tells us that this is the only possible eigenvector for $\mathcal{J}$ corresponding to a non-zero eigenvalue. (See the Case (2) of the Two by Two Theorem above.) Checking, we see that

$$\mathcal{J}\begin{bmatrix} -\sin\gamma \\ \cos\gamma \end{bmatrix} = \begin{bmatrix} -\dfrac{\partial\gamma}{\partial x}\sin\gamma & -\dfrac{\partial\gamma}{\partial y}\sin\gamma \\ \dfrac{\partial\gamma}{\partial x}\cos\gamma & \dfrac{\partial\gamma}{\partial y}\cos\gamma \end{bmatrix}\begin{bmatrix} -\sin\gamma \\ \cos\gamma \end{bmatrix}$$

$$= (\,(-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}) \bullet \nabla\gamma\,)\begin{bmatrix} -\sin\gamma \\ \cos\gamma \end{bmatrix}$$

so that $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$ is an eigenvector for $\mathcal{J}$ with corresponding eigenvalue $(-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}) \bullet \nabla\gamma$, which is the rate at which $\mathbf{U}$ is turning as (x,y) moves in the direction $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$. This confirms what we figured out above with the more conceptual approach.

One can note that there is an exceptional case here if $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$ is perpendicular to $\nabla\gamma$ (and therefore $\nabla\gamma$ is parallel to $\mathbf{U}$), which says that the directional derivative of γ in the direction perpendicular to $\mathbf{U}$ is 0. In this case, $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$ is a multiple of $\dfrac{\partial\gamma}{\partial y}\mathbf{i} - \dfrac{\partial\gamma}{\partial x}\mathbf{j}$, so one sees that $\mathcal{J}(-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}) = \mathbf{0}$. So this is Case (3) of the Two by Two Theorem. As previously seen, $-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j}$ is then the only eigenvector for $\mathcal{J}$ (since $\dfrac{\partial\gamma}{\partial y}\mathbf{i} - \dfrac{\partial\gamma}{\partial x}\mathbf{j}$ is a multiple of it).

This tells us not much of anything new about $\mathbf{U}$, but it's interesting to see how everything fits together.

Example 8 Redux. It's interesting now to go back and look at Example 8 again:

$$\mathbf{F} = \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{2\sqrt{x^2+y^2}} + \frac{\sqrt{3}\,\mathbf{k}}{2}\,.$$

At the time, we observed that we should not expect to be able to prove a theorem showing that for a vector field like $\mathbf{F}$ in three dimensions with constant magnitude 1, $\nabla\times\mathbf{F}$ equals the rate that $\mathbf{F}$ turns when (x,y) moves in the direction of $\mathbf{F}$, because $\mathcal{J}$, which determines $\nabla\times\mathbf{F}$, does not tell us the direction of $\mathbf{F}$.

However Example 8 is far from a typical example of a three-dimensional field with constant magnitude. We have observed that the Jacobean matrix for $\mathbf{F}$ is the same as for the planar field $\mathbf{G}_1 = \dfrac{-y\,\mathbf{i} + x\,\mathbf{j}}{2\sqrt{x^2+y^2}}$, except for an added row and added column of zeros. We can write

$\mathbf{G}_1 = \frac{1}{2}(\cos\gamma\,\mathbf{i} + \sin\gamma\,\mathbf{j})$, where $\gamma = \theta + \pi/2$ and θ is the usual polar coordinate. Since $\pi/2$ is a constant, $\nabla\gamma = \nabla\theta$. In fact, $\nabla\gamma(x,y) = \nabla\theta(x,y) = 2\mathbf{G}_1(x,y)$, so we are in the exceptional case for vector fields of the form $\cos\gamma\,\mathbf{i} + \sin\gamma\,\mathbf{j}$, as discussed above. We can use formula $(\star\star)$ to get

$$\mathcal{J} = \frac{1}{2r^3}\begin{bmatrix} xy & -x^2 & 0 \\ y^2 & -xy & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The two non-zero columns of $\mathcal{J}$ are both multiples of $-x\,\mathbf{i} - y\,\mathbf{j}$, which is the direction in which $\mathbf{G}_1$ turns when (x,y) moves and which is also an eigenvector for $\mathcal{J}$ corresponding to the eigenvalue 0.

We know that for vector fields of constant magnitude the columns of the Jacobean matrix are orthogonal to the field. And $\mathbf{G}_1 = (-y\,\mathbf{i} + x\,\mathbf{j})/2r$ is in fact perpendicular to the columns of its Jacobean matrix $\mathcal{J}$. But what makes Example 8 possible is that the vector $(-y\,\mathbf{i} + x\,\mathbf{j})/2r + a\,\mathbf{k}$ is also perpendicular to these columns, for any scalar a.

In the typical case of a vector field in three dimensions with constant magnitude 1, on the other hand, where two of the three columns of the Jacobean matrix are linearly independent, the direction of the field is in fact determined up to a plus-and-minus sign by the Jacobean matrix. In fact, if $\mathbf{v}_1$ and $\mathbf{v}_2$ are two columns of the Jacobean matrix which are not multiples of each other, then the vector field, since it has constant magnitude, is perpendicular to $\mathbf{v}_1$ and $\mathbf{v}_2$ and thus is parallel to $\mathbf{v}_1\times\mathbf{v}_2$. This means that one can't rule out the possibility that there is a nice theorem relating $\nabla\times\mathbf{F}$ to the rate at which $\mathbf{F}$ turns, except in those exceptional cases where the Jacobean matrix has rank one. But I doubt that such a theorem exists.

It may be interesting to compare the field $2\mathbf{G}_1 = (y\,\mathbf{i} - x\,\mathbf{j})/r$ above, and the linear vector field $\mathbf{G}_0 = -y\,\mathbf{i} + x\,\mathbf{j}$ (Example 5), which we have seen gives the velocity vectors for a rotation of the entire plane with an angular velocity of one radian per unit time, with the non-linear vector field

$$\mathbf{G}_2 = \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{x^2+y^2} = \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{r^2} = \nabla\left(\tan^{-1}(\frac{y}{x})\right).$$

All three of these vector fields represent the velocity vectors of circular motion around the origin, but the their behavior with respect to curl and especially eigenvectors are very different. The angular velocity for the vector field $(-y\,\mathbf{i} + x\,\mathbf{j})/r^n$ is $r^{-n}\,\mathbf{k}$. We have seen earlier that the curl is

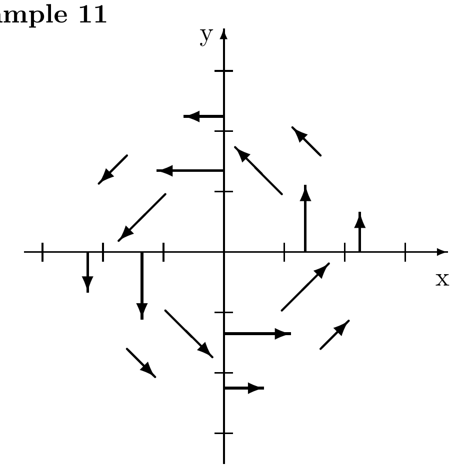
$(-n+2)r^{-n}\,\mathbf{k}$, so that

$$\begin{aligned}
\nabla\times\mathbf{G}_0 &= \nabla\times r^0(-y\,\mathbf{i}+x\,\mathbf{j}) = 2\,\mathbf{k}\\
\nabla\times 2\mathbf{G}_1 &= \nabla\times r^{-1}(-y\,\mathbf{i}+x\,\mathbf{j}) = \frac{\mathbf{k}}{r}\\
\nabla\times\mathbf{G}_2 &= \nabla\times r^{-2}(-x\,\mathbf{i}+y\,\mathbf{j}) = \mathbf{0}\,.
\end{aligned}$$

Above, we have computed the Jacobean matrices for the vector fields $\mathbf{G}_0$ and $2\mathbf{G}_1$ and seen that the first has no eigenvectors and the second has exactly one.

If $\mathcal{J}$ is the Jacobean matrix for $\mathbf{G}_2$ at a point $\mathbf{x}$, $\mathcal{J}\mathbf{x}$ is not perpendicular to the position vector $\mathbf{x}$, although the field vector $\mathbf{G}_2(\mathbf{x})$ is of course perpendicular to $\mathbf{x}$. (Since $\mathbf{G}_2$ is not linear, the vector $\mathcal{J}\mathbf{x}$ is actually irrelevant to the situation.) Since $\nabla\times\mathbf{G}_2=\mathbf{0}$, we know that the Jacobean matrix $\mathcal{J}$ at every point will be symmetric. Linear algebra then tells us that $\mathcal{J}$ will have two orthogonal eigenvectors.

Example 11



$$\mathbf{G}_2 = \frac{-ay\,\mathbf{i}+ax\,\mathbf{j}}{x^2+y^2} = a\,\nabla\left(\tan^{-1}(\frac{y}{x})\right)$$

$$\nabla\bullet\mathbf{G}_2 = 0$$

$$\nabla\times\mathbf{G}_2 = \mathbf{0}$$

Routine calculation produces

$$\mathcal{J} = \begin{bmatrix} \dfrac{\partial P}{\partial x} & \dfrac{\partial P}{\partial y}\\[2mm] \dfrac{\partial Q}{\partial x} & \dfrac{\partial Q}{\partial y}\end{bmatrix} = \frac{1}{r^4}\begin{bmatrix} -2xy & x^2-y^2\\ x^2-y^2 & 2xy\end{bmatrix} = \frac{1}{r^4}\begin{bmatrix} -2xy & (x+y)(x-y)\\ (x+y)(x-y) & 2xy\end{bmatrix}.$$

The eigenvectors for $\mathcal{J}$ turn out to be $(x-y)\,\mathbf{i}+(x+y)\,\mathbf{j}$ and $(x+y)\,\mathbf{i}+(y-x)\,\mathbf{j}$. (The fact that these are perpendicular to each other is predictable by linear algebra because the matrix $\mathcal{J}$ is symmetric.) The corresponding eigenvalues are $\pm 1/r^2$. As verification, we'll look at the calculation for the first eigenvector:

$$\frac{1}{r^4}\begin{bmatrix} -2xy & (x+y)(x-y) \\ (x+y)(x-y) & 2xy \end{bmatrix}\begin{bmatrix} x-y \\ x+y \end{bmatrix} = \frac{1}{r^4}\begin{bmatrix} (x-y)(\,-2xy+(x+y)^2\,) \\ (x+y)(\,(x-y)^2+2xy\,) \end{bmatrix}$$

$$= \frac{x^2+y^2}{r^4}\begin{bmatrix} x-y \\ x+y \end{bmatrix} = \frac{1}{r^2}\begin{bmatrix} x-y \\ x+y \end{bmatrix}.$$

It is enlightening to write the eigenvectors here in terms of the radial and tangential vectors for the circle. We have

$$(x-y)\,\mathbf{i}+(x+y)\,\mathbf{j} = (x\,\mathbf{i}+y\,\mathbf{j}) + (-y\,\mathbf{i}+x\,\mathbf{j}\,)$$
$$(x+y)\,\mathbf{i}+(y-x)\,\mathbf{j} = (x\,\mathbf{i}+y\,\mathbf{j}) - (-y\,\mathbf{i}+x\,\mathbf{j}\,).$$

To interpret this intuitively, imagine moving away from a point (x,y) in the direction of the eigenvector $(x-y)\,\mathbf{i}+(x+y)\,\mathbf{j}$, so that

$$\Delta\mathbf{x} = a(\,(x-y)\,\mathbf{i}+(x+y)\,\mathbf{j}\,) = a(x\,\mathbf{i}+y\,\mathbf{j}\,) + a(-y\,\mathbf{i}+x\,\mathbf{j}\,),$$

where a is a small real number. Then one is moving farther from the origin (assuming that a is positive) by a distance ra and then a distance of ra in a direction tangential to the circle, which is roughly equivalent, if a is small enough, to moving along the circle itself.

Now for the movement away from the origin, the vector $\mathbf{G}_2(x,y)$ keeps the same direction, but gets a little shorter: for this move, $\Delta\,||\mathbf{G}_2|| = -a/r$ and $\Delta\mathbf{G}_2 = -a(-y\,\mathbf{i}+x\,\mathbf{j})/r^2$. (The vector $-y\,\mathbf{i}+x\,\mathbf{j}$ is in the direction of $\mathbf{G}_2$ but is not a unit vector, hence the extra factor of r in the denominator here.) And for the move in the tangential direction, $||\mathbf{G}_2||$ stays the same, but $\mathbf{G}_2$ turns through an angle of a/r. (One can see this if one remembers that $\mathbf{G}_2$ is perpendicular to the radius vector.) Since $\mathbf{G}_2$ is turning but not changing magnitude during this tangential relocation, the derivative for the motion of $\mathbf{G}_2$ is directed toward the center of the circle, and one sees that for this move, one has $\Delta\mathbf{G}_2 \approx -\dfrac{a}{r^2}(x\,\mathbf{i}+y\,\mathbf{j})$.

Putting all this together, one gets that

$$\Delta\mathbf{G}_2 \approx \frac{-a(-y\,\mathbf{i}+x\,\mathbf{j}\,)}{r^2} + \frac{-a(x\,\mathbf{i}+y\,\mathbf{j}\,)}{r^2} = \frac{-1}{r^2}\Delta\mathbf{x},$$

again confirming that Δx is one of the eigenvectors for $\mathbf{G}_2$.

On a purely qualitative level, what we see here is that the crucial difference between $\mathbf{G}_2$ and $\mathbf{G}_0$ is that for $\mathbf{G}_2$, when (x,y) moves radially away from the origin, the corresponding $\Delta\mathbf{G}_2$ is a vector tangential to the circle and in the clockwise direction, so that in this case $\Delta\mathbf{G}_2$ is at a clockwise angle to $\Delta\mathbf{x}$, whereas when (x,y) moves in a direction tangential to the circle in a counter-clockwise direction, $\mathbf{G}_2$ turns and the corresponding $\Delta\mathbf{G}_2$ is a vector directed radially towards the origin, so that in this case $\Delta\mathbf{G}_2$ is at a counterclockwise angle to $\Delta\mathbf{x}$. From this, it seems reasonable it should be possible to construct a vector which is an appropriate linear combination of a motion radially away from the center and a motion tangential to the circle in the counter-clockwise direction so that if $\Delta\mathbf{x}$ is given by this motion, then the corresponding $\Delta\mathbf{G}_2$ should be exactly opposite to $\Delta\mathbf{x}$. This vector will be an eigenvector for $\mathbf{G}_2$, and by an analogous process, one can construct a second eigenvector, which will be perpendicular to it. (Of course since $\mathbf{G}_2$ is a non-linear vector field, one ought to also choose $\Delta\mathbf{x}$ to be small.)

On the other hand, in the case of $\mathbf{G}_0 = -y\,\mathbf{i} + x\,\mathbf{j}$ one finds that whether $\Delta\mathbf{x}$ is radially directed away from the origin or tangential to the circle, multiplication of $\Delta\mathbf{x}$ by the Jacobean matrix corresponding to $\mathbf{G}_0$ will produce a vector $\Delta\mathbf{G}_0$ which is rotated 90° counter-clockwise from $\Delta\mathbf{x}$. Thus there is no possibility of combining vectors in these two directions to obtain a vector $\Delta\mathbf{x}$ which is not rotated when multiplied by the Jacobean matrix for $\mathbf{G}_0$, which is precisely what would be required to have an eigenvector.

In light of the theorem that a vector field with zero curl is a gradient (at least within some neighborhood), it is tempting to reason that it is obvious that the field $\mathbf{F} = a\,\mathbf{i}$ of Example 6 cannot have zero curl (as in fact it doesn't). Because imagine that $\mathbf{F} = \nabla g$ for some function $g(x,y)$. Since $\mathbf{F} = \nabla g$ is a field in the $\mathbf{i}$ direction, this would mean that $g(x,y)$ does not increase when (x,y) moves in the $\mathbf{j}$ direction. But then it seems impossible that ∇g could increase when (x,y) moves in the $\mathbf{j}$ direction, as is the case for the field $\mathbf{F}$.

Although the intuitive basis for this reasoning is sound, the reasoning itself is simplistic. For consider the previously discussed vector field

$$\mathbf{F} = \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{x^2 + y^2} = \nabla g$$

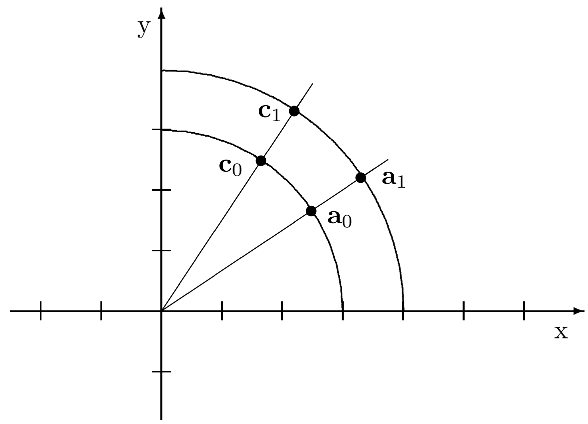
where

$$g(x,y) = \tan^{-1}\left(\frac{y}{x}\right)$$

(see Example 11). Since ∇g is tangent to the circles $x^2 + y^2 = \text{const}$, the function $g(x,y)$ does not change when (x,y) moves in a direction radially towards or away from the origin. (This is obvious in any case, since movement in such a direction does not change y/x.) On the other hand, $||\,\nabla g\,|| = 1/r$, so that ∇g decreases in magnitude when one moves in a direction radially away from the origin, despite the fact that $g(x,y)$ does not change.

In fact, from Theorem A above about the curl of planar vector fields, the fact that $||\mathbf{F}|| = 1/r$ and $\nabla \times \mathbf{F} = \mathbf{0}$ and that $\mathbf{F}$ turns at a rate of $1/r$ as (x,y) moves in the direction of $\mathbf{F}$ (i. e. tangent to the circle through (x,y) around the origin) means that necessarily $\mathbf{F}$ must decrease at a rate of $1/r^2$ when (x,y) moves in a direction perpendicular to $\mathbf{F}$, i. e. radially away from the origin, and in fact, $\dfrac{d}{dr}\,||\mathbf{F}|| = \dfrac{d}{dr}\left(\dfrac{1}{r}\right) = -\dfrac{1}{r^2}$.

Because it does seem strange that ∇g can decrease when (x,y) moves in a direction that does not change g, its worth looking at this situation more closely to see how this can happen. Here's a picture.



Consider points $\mathbf{a}_0 = (a_0, b_0)$ and $\mathbf{a}_1 = (a_1, b_1)$ on the same radial line and with distances r_0 and r_1 from the origin. And likewise points $\mathbf{c}_0 = (c_0, d_0)$ and $\mathbf{c}_1 = (c_1, d_1)$ on a different radial line. Now $g(x, y) = \tan^{-1}(y/x)$ does not change between $\mathbf{a}_0$ and $\mathbf{a}_1$ or between $\mathbf{c}_0$ and $\mathbf{c}_1$. Furthermore, a basic theorem for line integrals (the analogue of the Fundamental Theorem of Calculus) says that

$$\int_{\mathbf{a}_0}^{\mathbf{c}_0} \nabla g \bullet d\mathbf{r} = g(\mathbf{c}_0) - g(\mathbf{a}_0)$$

$$= g(\mathbf{c}_1) - g(\mathbf{a}_1) = \int_{\mathbf{a}_1}^{\mathbf{c}_1} \nabla g \bullet d\mathbf{r}\,.$$

So the integral over the circular path from $\mathbf{c}_0$ to $\mathbf{a}_0$ is the same as the integral over the path from $\mathbf{c}_1$ to $\mathbf{a}_1$. But the path further away from the origin is longer. This certainly seems to *suggest* that ∇g is smaller on the second path (as, in fact, in this example it certainly is).

To make this more convincing, notice also that the line integrals are taken over paths where ∇g is in the direction of the tangent vector to the path. Therefore the line integrals reduce to ordinary integrals

$$\int_{s_0}^{s_1} ||\nabla g(x(s), y(s))||\; ds$$

where we have parametrized the two curves with respect to the arc-length variable s.

But what we see in this example is that the two trajectory curves (the two circular arcs) are curved, the orthogonal curves (the level curves for g, which in this example are the radial lines from the origin) spread out as we move away from the first trajectory along the level curve in the direction opposite to the direction of curvature for this trajectory. And so the outside trajectory is longer, and it is a greater distance from $\mathbf{a}_1$ to $\mathbf{c}_1$ than from $\mathbf{a}_0$ to $\mathbf{c}_0$. But the function g changes by the same amount in both cases. So it makes sense that ∇g is smaller on the outside curve. (This inference would not make any sense were it not for the fact that the direction of ∇g is parallel to the curve. In this specific example, of course, $||\nabla g||$ is constant along the two trajectory curves, so the two integrals reduce to $(s_1 - s_0)\,||(g(x, y)||$, making it very easy to see that for the two integrals to be equal, $||\nabla g||$ must be smaller on the outside curve than on the inside.)

It's worth seeing how the reasoning we have looked at applies to the example $g(x, y) = x^2 + y^2$, which is actually described by the same picture that we have used for $\tan^{-1}(y/x)$. In this case, the trajectories (or integral curves) determined by $\nabla g(x,y)$, which is given by

$$\nabla g = 2x\,\mathbf{i} + 2y\,\mathbf{j}\,,$$

are the radial lines through the origin, and thus have curvature 0, and the level curves are the circles centered at the origin. So the roles of the points in the picture are now switched. We will want to consider the line integrals over the radial lines, $\int_{\mathbf{a}_0}^{\mathbf{a}_1} \nabla g \bullet \mathbf{dr}$ and $\int_{\mathbf{c}_0}^{\mathbf{c}_1} \nabla g \bullet \mathbf{dr}$. Although the trajectories here (the radial lines) have zero curvature, they are slanted, and the level curves, the circles, do spread out as we move along the radial lines away from the origin. Despite this fact, the distance from $\mathbf{a}_0$ to $\mathbf{a}_1$ along one radial line is that same as the distance from $\mathbf{c}_0$ to $\mathbf{c}_1$ along the other. We therefore surmise, as is readily apparent, that $||\nabla g||$ does not change when (x,y) moves along a circle centered at the origin.

Examples can be misleading, because they often have special properties that are not true in the general case. If we did not already know that a gradient field ∇g has to decrease in magnitude as we move away from a point in the direction opposite to the direction of curvature of the trajectory, the reasoning just given would be considerably less than totally convincing, since there are too many possibilities it doesn't take into account. In my own mathematical investigations, I almost always start by looking at examples. But when when I notice a certain phenomenon which holds for all the examples I have found, I then ask myself what it is about these examples that make this phenomenon occur, and then try and see whether I can construct another example were it does not occur. If I am repeatedly unsuccessful at constructing a counter-example, then I ask myself what the stumbling blocks are that I constantly run up against. If I can identify these stumbling blocks, then it is possible that I have found a proof that the phenomenon in question is true in general. (More often, though, I am successful in constructing the counter-example, and therefore have to kiss the theorem goodbye. But this may enable me to see how to change the theorem I was trying to prove in order to obtain one that is in fact valid.)

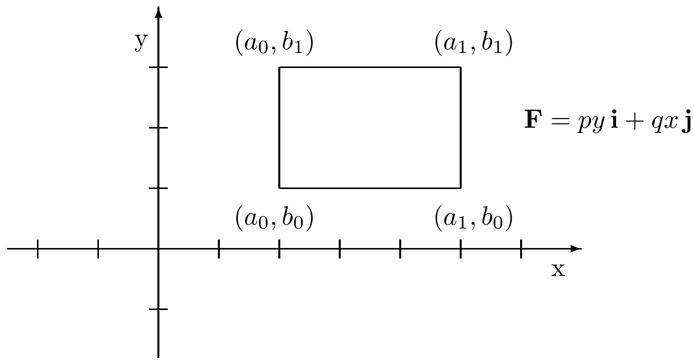
Green's Theorem and Stoke's Theorem.

Although we need 3-space to define curl, we can think of a vector field $\mathbf{F}(x,y) = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j}$ in the plane as having a curl $\nabla \times \mathbf{F}$ in the $\mathbf{k}$ direction. I. e.

$$\nabla \times (P\,\mathbf{i} + Q\,\mathbf{j}) = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\,\mathbf{k}\,.$$

As was the case for divergence, we can see the idea of curl most clearly if we start by looking at linear vector fields $\mathbf{F}(x,y,z) = (a_1 x + a_2 y)\,\mathbf{i} + (b_1 x + b_2 y)\,\mathbf{j}$. We may further simplify by supposing the a_1 and b_2 are 0, since these will not contribute to $\nabla \times \mathbf{F}$.

Consider then a planar vector field $\mathbf{F}(x,y) = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j} = py\,\mathbf{i} + qx\,\mathbf{j}$, where p and q are scalar constants. Now look at a rectangle with sides parallel to the axes as shown below.



Here $\nabla \times \mathbf{F} = (q-p)\,\mathbf{k}$. Then $(b_1 - b_0)\dfrac{\partial Q}{\partial x} = (b_1 - b_0)q$ is the amount that $\mathbf{F}$ increases (or decreases, in the negative case) between the bottom of the rectangle and the top. And $(a_1 - a_0)\dfrac{\partial P}{\partial y} = (a_1 - a_0)p$ is the amount that $\mathbf{F}$ increases between the left side and the right.

If we consider $\mathbf{F}$ as denoting the velocity of a fluid, and if we adopt the convention that positive flow is counter-clockwise, then since $py\,\mathbf{i} = pb_1\,\mathbf{i}$ is the horizontal component of $\mathbf{F}$ along the top of the rectangle, it is seems plausible to say that $-(a_1 - a_0)P(x, b_1) = -(a_1 - a_0)pb_1$ is a measure of the flow along the top of the rectangle. (I don't know of an easy conceptual definition of the word flow, but I hope it becomes clear in context.) Likewise $(a_1 - a_0)pb_0$ equals the flow along the bottom edge. Thus $-(a_1 - a_0)(b_1 - b_0)p$ equals the flow along the top and bottom edges of the rectangle. Likewise, since $qx\,\mathbf{j}$ is the vertical component of $\mathbf{F}$, on sees that $(a_1 - a_0)(b_1 - b_0)q$ equals the flow along the left-hand and right-hand sides. (Since $a_1 q(b_1 - b_0$ is the flow along the RH side and $-a_0(b_1 - b_0$ is the flow along the left.) Thus $(q-p)(a_1 - a_0)(b_1 - b_0)$ equals the total flow around the rectangle. This is generally called the **circulation** of the vector field $\mathbf{F}$.

More generally, if we assume merely that $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j}$, where $P(x, y)$ and $Q(x, y)$ are differentiable functions, then we see that the flow along the top edge of the rectangle considered above will be given by

$$-\int_{a_0}^{a_1} P(x, b_1)\,dx$$

and the flow along the bottom edge by

$$\int_{a_0}^{a_1} P(x, b_0)\,dx\,.$$

Furthermore, we see that for given x,

$$P(x, b_1) - P(x, b_0) = \int_{b_0}^{b_1} \frac{\partial P}{\partial y}\,dy\,.$$

Thus the total flow along the top and bottom edges will be given by

$$-\int_{a_0}^{a_1} P(x, b_1)\,dx + \int_{a_0}^{a_1} P(x, b_0)\,dx = -\int_{a_0}^{a_1}\int_{b_0}^{b_1} \frac{\partial P}{\partial y}\,dy\,dx\,.$$

The same reasoning shows that the flow along the left-hand and right-hand edges will be given by

$$\int_{a_0}^{a_1}\int_{b_0}^{b_1} \frac{\partial Q}{\partial x}\,dy\,dx\,.$$

Thus the total flow around the rectangle will be given by

$$\int_{a_0}^{a_1}\int_{b_0}^{b_1}\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\,dy\,dx\,.$$

This is **Green's Theorem**, as applied to the case of a rectangle.

It is tempting to say that for a planar vector field $\mathbf{F}=P\,\mathbf{i}+Q\,\mathbf{j}$, $\dfrac{\partial Q}{\partial x}-\dfrac{\partial P}{\partial y}=||\nabla\times\mathbf{F}||$. However this may not be quite correct because the left-hand side is possibly negative. What is true is that $\dfrac{\partial Q}{\partial x}-\dfrac{\partial P}{\partial y}$ is the **k**-coordinate of $\nabla\times\mathbf{F}$, and the correct way of expressing this is by the equation

$$\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}=\mathbf{k}\bullet\nabla\times\mathbf{F}\,.$$

Thus Green's Theorem, for the case of a rectangle, states that the total flow of a vector field $\mathbf{F}$ around a rectangle as described above is given by

$$\int_{a_0}^{a_1}\int_{b_0}^{b_1}\mathbf{k}\bullet\nabla\times\mathbf{F}\,dy\,dx\,.$$

The proof given can easily be generalized to prove the general statement of Green's Theorem. Namely, if $\mathbf{F}(x,y)$ is a planar vector field defined in a portion of the plane containing a region Ω bounded by a simple closed curve $\mathcal{C}$, then

$$\iint_{\Omega}\mathbf{k}\bullet\nabla\times\mathbf{F}\;dx\,dy=\oint_{\mathcal{C}}\mathbf{F}\bullet d\mathbf{r}\,.$$

I have given the proof (which is essentially the same as that found in many standard texts) in my article on Green's Theorem.

Vector Fields in Three-space with Constant Magnitude

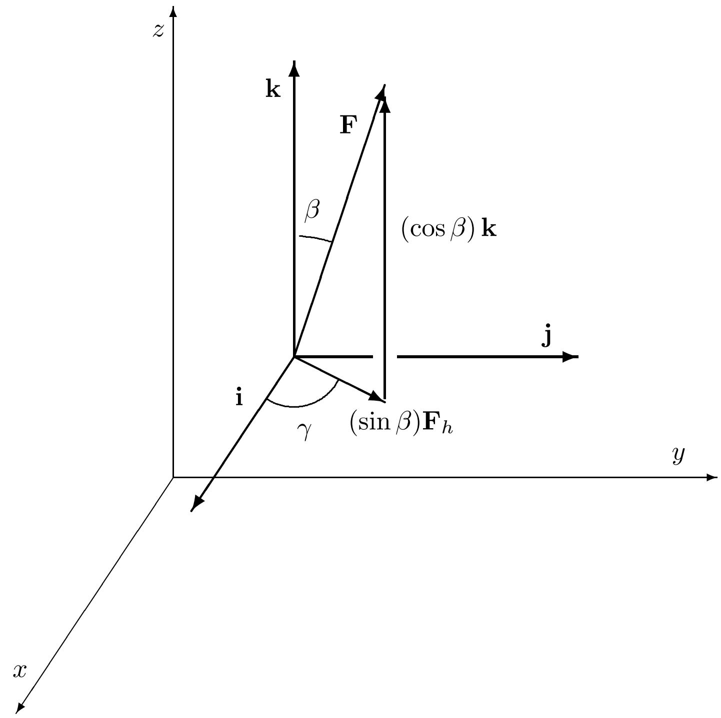
We can recall that in the two-dimensional case, we were able to understand a vector field of constant magnitude by using polar coordinates. In three-space, we can attempt to use spherical coordinates, although with less success.

Consider a vector field $\mathbf{F}$ with constant magnitude 1. Let β be the angle between $\mathbf{F}$ and $\mathbf{k}$. We can write $\mathbf{F}$ as the sum of a vertical component and a component in the xy-plane. Since $\mathbf{F}$ is a unit vector, clearly the vertical component equals $\cos\beta\,\mathbf{k}$. Let γ be the angle between $\mathbf{i}$ and the horizontal component of $\mathbf{F}$ and let $\mathbf{F}_h$ be a unit vector in the xy-plane in the same direction as this planar component. Thus

$$\mathbf{F}_h=\cos\gamma\,\mathbf{i}+\sin\gamma\,\mathbf{j}\,.$$

Since $\mathbf{F}$ and $\mathbf{F}_h$ are unit vectors, we see that the magnitude of the component of $\mathbf{F}$ in the xy-plane is $\sin\beta$. (By standard convention, we choose $0\le\beta\le\pi$ so that $\sin\beta\ge 0$.) Thus we have

$$\mathbf{F}=(\sin\beta)\,\mathbf{F}_h+(\cos\beta)\,\mathbf{k}=(\sin\beta)(\cos\gamma\;\mathbf{i}+\sin\gamma\;\mathbf{j})+\cos\beta\,\mathbf{k}\,.$$



Since the formulas now get rather complicated, we will resort to the standard notations $\beta_x = \dfrac{\partial \beta}{\partial x}$, $\gamma_y = \dfrac{\partial \gamma}{\partial y}$, etc. We get

$$\mathcal{J} = \begin{bmatrix} \beta_x \cos\beta\cos\gamma - \gamma_x \sin\beta\sin\gamma & \beta_y \cos\beta\cos\gamma - \gamma_y \sin\beta\sin\gamma & \beta_z \cos\beta\cos\gamma - \gamma_z \sin\beta\sin\gamma \\ \beta_x \cos\beta\sin\gamma + \gamma_x \sin\beta\cos\gamma & \beta_y \cos\beta\sin\gamma + \gamma_y \sin\beta\cos\gamma & \beta_z \cos\beta\sin\gamma + \gamma_z \sin\beta\cos\gamma \\ -\beta_x \sin\beta & -\beta_y \sin\beta & -\beta_z \sin\beta \end{bmatrix}.$$

Note that the directional derivative of $\mathbf{U}$ will be zero in a direction that is perpendicular to both $\nabla\beta$ and $\nabla\gamma$. Thus $\nabla\beta \times \nabla\gamma$ is an eigenvector for $\mathcal{J}$ corresponding to the eigenvalue 0. From the geometry it is apparent that if either β or γ change, then $\mathbf{U}$ will change, so that $\nabla\beta \times \nabla\gamma$ is in fact the only possible eigenvector for $\mathcal{J}$ corresponding to the eigenvalue 0. (There is a exceptional and somewhat tricky special case when β is 0 or π. We ignore this for the moment, as well as the exceptional cases when $\nabla\beta$ or $\nabla\gamma$ are $\mathbf{0}$ or the two are parallel to each other.)

To clarify the pattern, let us use block notation to write

$$\mathcal{J} = \begin{bmatrix} \beta_x \mathbf{w}_1 & \beta_y \mathbf{w}_1 & \beta_z \mathbf{w}_1 \end{bmatrix} + (\sin\beta) \begin{bmatrix} \gamma_x \mathbf{w}_2 & \gamma_y \mathbf{w}_2 & \gamma_z \mathbf{w}_2 \end{bmatrix},$$

where

$$\mathbf{w}_1 = \begin{bmatrix} \cos\beta\cos\gamma \\ \cos\beta\sin\gamma \\ -\sin\beta \end{bmatrix} \qquad \text{and} \qquad \mathbf{w}_2 = \begin{bmatrix} -\sin\gamma \\ \cos\gamma \\ 0 \end{bmatrix}.$$

We can now note that all three columns of $\mathcal{J}$ are linear combinations of the two unit vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ which are, predictably, orthogonal to $\mathbf{U}$, and are also orthogonal to each other. Thus any possible eigenvector with non-zero eigenvalue will be a linear combination of these two.

For an arbitrary vector $\mathbf{v} = a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k}$, we get

$$\begin{aligned}\mathcal{J}\mathbf{v} &= \Big[\beta_x\,\mathbf{w}_1 + \gamma_x\,\sin\beta\;\mathbf{w}_2 \quad \beta_y\,\mathbf{w}_1 + \gamma_y\,\sin\beta\;\mathbf{w}_2 \quad \beta_z\,\mathbf{w}_1 + \gamma_z\,\sin\beta\;\mathbf{w}_2\Big]\,\mathbf{v}\\ &= (\nabla\beta\bullet\mathbf{v})\,\mathbf{w}_1 \;+\; \sin\beta\;(\nabla\gamma\bullet\mathbf{v})\,\mathbf{w}_2\,.\end{aligned}$$

If we choose $\mathbf{v}$ as a unit vector, we can rewrite this as

$$\mathcal{J}\mathbf{v} = D_{\mathbf{v}}(\beta)\,(\cos\beta\,(\cos\gamma\,\mathbf{i} + \sin\gamma\,\mathbf{j}) - \sin\beta\,\mathbf{k}) + D_{\mathbf{v}}(\gamma)\;\sin\beta\,(-\sin\gamma\,\mathbf{i} + \cos\gamma\,\mathbf{j})\,.$$

I don't see any way of getting anything useful out of this.

However consider the special case of a vector field $\mathbf{U}(x,y,z)$ with constant magnitude 1 and such that $\nabla\times\mathbf{U} = \mathbf{0}$. Then we know that, at least within a sufficiently small neighborhood, $\mathbf{U}$ is the gradient field for some function $g(x,y,z)$. Thus $\mathbf{U}$ is orthogonal to the level surface $g(x,y,z) = \text{const}$. Since we are assuming that $\mathbf{U}$ has constant magnitude 1, we see that $\mathbf{U}$ is in fact the unit normal vector for such a level surface. And in this case, we can get some additional insight by means of differential geometry.

If $\mathbf{v}$ is any vector, then the fact that $\mathbf{U}$ has constant magnitude 1 implies that the directional derivative of $\mathbf{U}$ at a point (x_0,y_0,z_0) in the direction of $\mathbf{v}$, which will be the vector $\mathcal{J}\mathbf{v}$, will be perpendicular to $\mathbf{U}$, and thus will be tangent to the level surface. Furthermore, $\mathcal{J}\mathbf{v}$ will point in the direction in which $\mathbf{U}(x,y,z)$ turns as (x,y,z) moves away from (x_0,y_0,z_0) in the direction $\mathbf{v}$, and its magnitude will show the speed of this turning.

The fact that all the vectors $\mathcal{J}\mathbf{v}$ lie in a plane indicates that we are in the situation analogous to Case 2 or Case 3 of the Two-by-Two Theorem for 2×2 matrices. In other words, the 3×3 matrix $\mathcal{J}$ is a singular matrix and has at most two eigenvectors corresponding to non-zero eigenvalues, and these eigenvectors are all parallel to the tangent plane to the level surface at the given point.

But we have seen earlier that the assumption that $\nabla\times\mathbf{U} = \mathbf{0}$ implies that the 3×3 Jacobian matrix $\mathcal{J}$ corresponding to the derivative of $\mathbf{F}$ is symmetric and thus has three mutually orthogonal eigenvectors. Thus there must be a third eigenvector perpendicular to the tangent plane of the level surface and with 0 as the corresponding eigenvalue. Since all vectors perpendicular to the tangent plane are multiples of each other, this shows that $\mathbf{U}$ itself is an eigenvector for the Jacobean matrix $\mathcal{J}$ corresponding to the eigenvalue 0.

It is worth recording this conclusion as a proposition.

Proposition. If we vector field $\mathbf{U}(x,y,z)$ in three-space is the gradient field for a function $g(x,y,z)$ and has constant magnitude 1, then at any point (x_0,y_0,z_0), $\mathbf{U}(x_0,y_0,z_0)$ is an eigenvector for the Jacobean matrix $\mathcal{J}(x_0,y_0,z_0)$ corresponding to the eigenvalue 0. Furthermore, $\mathcal{J}$ has two more eigenvectors which are parallel to the tangent plane to the level surface $g(x,y,z) = \text{const}$ at the point (x_0,y_0,z_0).

Now consider a curve on the surface through a point (x_0, y_0, z_0) and let $\mathbf{v}$ be the unit tangent vector to this curve. Any curve on the level surface will have curvature arising from two different causes: first, the curvature that it may have as seen by someone actually on the surface itself and looking at it from that perspective, and second the curvature that it unavoidably has because it is on a curved surface. This second curvature, at a point (x_0, y_0, z_0) on the surface, is called the *normal curvature* of the surface at the point.

In other words, if one is building a road on the surface through the point (x_0, y_0, z_0), then one may do one's very best to make the road straight, but even so it will have some curvature, because it is on a curved surface. This unavoidable curvature when the road goes in a direction $\mathbf{v}$ is what we mean by the normal curvature of the surface in that direction.

To give the definition in more practical, although possibly a little less intuitive fashion, at the given point (x_0, y_0, z_0) on the surface, and for a given direction $\mathbf{v}$, one should construct the plane through the given point and going through both $\mathbf{v}$ and the normal vector $\mathbf{U}(x_0, y_0, z_0)$ to the surface at this point. This plane will then cut the surface in a curve, and the curvature of this curve is that is meant by the normal curvature of the surface at the given point in the given direction $\mathbf{v}$. (If one does this on a sphere, then the curve thereby constructed will be a great circle. This is in fact the most practical way of describing what is meant by a great circle.)

I am belaboring this issue, because the concept of the normal curvature of a surface at a given point in a given direction is difficult to master if one doesn't know differential geometry, but it is essential for gaining an intuitive understanding of the divergence of a vector field with constant magnitude and zero curl in three dimensions.

Now watch what happens to the normal vector $\mathbf{U}(x, y, z)$ as (x, y, z) moves away from a point (x_0, y_0, z_0) along a normal curve in the level surface with unit tangent vector $\mathbf{v}$. $\mathbf{U}$ turns in the direction of its directional derivative $D_{\mathbf{v}}(\mathbf{U})$, which is perpendicular to $\mathbf{U}$ and therefore is in the tangent plane. Now this directional derivative can be written as the sum of two components: one of which is in the direction of $\mathbf{v}$ and the second of which is orthogonal to $\mathbf{v}$.

Now the first component, which shows the rate at which $\mathbf{U}$ is turning *towards* $\mathbf{v}$, is the rate at which the projection of $\mathbf{U}$ onto the plane of the normal curve (i. e. the plane determined by $\mathbf{v}$ and $\mathbf{U}$) turns, and this will be the same as the rate at which the tangent vector to the normal curve is turning, since $\mathbf{U}$ and the tangent vector are perpendicular. Thus the magnitude of the first component of $D_{\mathbf{v}}(\mathbf{U})$ is the same as the curvature of the normal curve.

The second component of $D_{\mathbf{v}}$, orthogonal to the normal curve, shows the rate at which $\mathbf{U}$ is turning away from the plane of the normal curve, due to the fact that the surface is, we might say, tilting sideways as we move along the curve.

Intuitively, one can think of this in terms of driving a car on a mountain at a place which has not been smoothed out into a road. Generally, when one drives in a particular direction the car tips forward or backward in the direction one is driving, but also tilts from side to side in the perpendicular direction. The normal curvature of the road corresponds only to the forward or backward tipping. (If the car has the sort of antenna that sticks straight up from the roof, then this antenna would be in the direction of the normal vector $\mathbf{U}$.)

Theorem. Suppose that the gradient field $\mathbf{U} = \nabla g$ for a function $g(x, y, z)$ has constant magnitude 1. Let $\mathbf{v}$ and $\mathbf{w}$ be two orthogonal unit vectors tangent to the level surface $g(x, y, z) = \text{const}$ at a point (x_0, y_0, z_0) on surface. Let $\mathcal{J}_0$ be the Jacobian matrix $\mathcal{J}(x_0, y_0, z_0)$ for the derivative of $\mathbf{U}$ at the point (x_0, y_0, z_0). Then

$$\mathcal{J}_0\mathbf{v} = a_{11}\mathbf{v} + a_{12}\mathbf{w}$$

$$\mathcal{J}_0\mathbf{w} = a_{21}\mathbf{v} + a_{22}\mathbf{w}$$

where

(1) a_{11} is the normal curvature of the level surface $g = \text{const}$ at the point (x_0, y_0, z_0) in the $\mathbf{v}$ direction and a_{22} is the normal curvature in the $\mathbf{w}$ direction. Here, contrrary to the usual convention, one should interpret the curvature as a signed number which is positive if the normal curve curves in the direction away from $\mathbf{U}$ and negative if it curves in the direction towards $\mathbf{U}$.
(2) Furthermore, a_{12} is the rate at which the surface tilts toward the $\mathbf{w}$ direction as one moves on it away from (x_0, y_0, z_0) along the normal curve in the $\mathbf{v}$ direction at unit speed, and a_{21} is the rate at which the surface tilts toward the $\mathbf{v}$ direction as one moves along the normal curve in the $\mathbf{w}$ direction.
(3) Furthermore, $a_{12} = a_{21}$.
(4) The divergence of $\mathbf{U}$ equals $a_{11} + a_{22}$.

PROOF: The normal curve in the direction $\mathbf{v}$ to the level surface has $\mathbf{v}$ as its tangent vector.

$\mathcal{J}\mathbf{v}$ is the directional derivative $D_{\mathbf{v}}(\mathbf{U})$ of $\mathbf{U}$ as (x, y) moves away from (x_0, y_0) at unit speed in the direction $\mathbf{v}$, i. e. along the normal curve through (x_0, y_0) in the direction $\mathbf{v}$. Now since $\mathbf{U}$ has constant magnitude, this directional derivative is perpendicular to $\mathbf{U}$ and gives the rate at which $\mathbf{U}$ is turning as (x, y) moves along the normal curve at unit speed. Since $D_{\mathbf{v}}(\mathbf{U})$ is perpendicular to $\mathbf{U}$, it lies in the tangent plane to the level surface, and thus can be broken up into two orthogonal components, one in the direction of $\mathbf{v}$ and one perpendicular to it: $D_{\mathbf{v}}(\mathbf{U}) = a_{11}\mathbf{v} + a_{12}\mathbf{w}$.

Now $a_{11}\mathbf{v}$ gives the rate at which $\mathbf{U}$ turns towards $\mathbf{v}$ as (x, y) moves away from (x_0, y_0) along the normal curve in the direction $\mathbf{v}$. Since $\mathbf{v}$ and $\mathbf{U}$ are perpendicular and in the same plane, this is also the rate at which $\mathbf{v}(x, y)$ turns as we move in the direction $\mathbf{v}(x_0, y_0)$ away from (x_0, y_0). Since $\mathbf{v}$ is the unit tangent vector to the normal curve, this equals the curvature of the normal curve at (x_0, y_0), except that it is negative if U is turning *away from* $\mathbf{v}$ (towards $-\mathbf{v}$), i. e. if the normal curve is curving towards $\mathbf{U}$.

On the other hand, a_{12} is the rate at which $\mathbf{U}$ turns towards $\mathbf{w}$ as (x, y) moves along the normal curve in the $\mathbf{v}$ direction. Since $\mathbf{U}$ is perpendicular to the level surface, it is more or less self-evident that this represents the rate at which the surface is tilting.

With respect to the orthonormal vector basis $\mathbf{v}$, $\mathbf{w}$, $\mathbf{U}$, the Jacobean matrix for the derivative of $\mathbf{U}$ is $\begin{bmatrix} a_{11} & a_{12} & * \\ a_{21} & a_{22} & * \\ 0 & 0 & 0 \end{bmatrix}$. The divergence of $\mathbf{U}$ is the trace of this matrix, i. e. $a_{11} + a_{22}$.

Since $\mathbf{U}$ is a gradient field, it is known that this matrix is symmetric. Thus $a_{12} = a_{21}$. (Furthermore, this shows that the unspecified entries in the matrix indicated by asterisks must both be 0.)

Restating this more verbally, we get:

Theorem. If $\mathbf{U}$ is the gradient field of a function $g(x,y,z)$ and $\mathbf{U}$ has constant magnitude 1, then at any given point (x_0,y_0,z_0), the divergence of $\mathbf{U}$ is the sum of the curvatures of any two orthogonal normal curves on the level surface $g(x,y,z)=\text{const}$ through the point (x_0,y_0,z_0), provided that one treats the curvature as negative in case the normal curve curves towards (rather than away from) $\mathbf{U}$.

As noted in a Proposition above, the Jacobean matrix for the derivative of a vector field $\mathbf{U}=\nabla g$ with constant magnitude 1 has at any given point $\mathbf{U}$ itself as an eigenvector, corresponding to the eigenvalue 0, along with two other eigenvectors which are orthogonal to each other and lie in the tangent plane to the level surface $g(x,y,z)=\text{const}$. Let us choose $\mathbf{v}$ and $\mathbf{w}$ as these two eigenvectors. The discussion above shows that to say that $\mathbf{v}$ is an eigenvector for $\mathcal{J}$ is thus to say that when (x,y,z) moves away from the point (x_0,y_0,z_0) in the $\mathbf{v}$ direction, then the normal vector $\mathbf{U}(x,y,z)$ to the surface turns in the direction $\mathbf{v}$. The eigenvalue corresponding to $\mathbf{v}$ will be negative in the case that the normal curve through (x_0,y_0,z_0) in the direction $\mathbf{v}$ curves toward $\mathbf{U}$ and will be positive if the normal curve curves away from the direction of $\mathbf{U}$.

In terms of my previous analogy, the eigenvectors $\mathbf{v}$ and $\mathbf{w}$ correspond to the only two directions in which one can drive on the mountain and have the car tip only up or down in the direction one is driving without also tilting sideways.

Finally, there is a fact from differential geometry which I will not prove, namely that at any given point, the normal curves in the directions of the two eigenvectors of the Jacobean matrix will have the maximum and minimum normal curvatures of all normal curves through that point.

To summarize, we get the following theorem:

Theorem. If a three-dimensional vector field $\mathbf{F}$ is the gradient field of a function $g(x,y,z)$, and if $\mathbf{F}$ has constant magnitude, then the divergence of $\mathbf{F}$ at a point (x_0,y_0,z_0) is in absolute value of the product of $||\mathbf{F}||$ and the sum (or possibly difference) of the maximal and minimal normal curvatures of the level surface $g(x,y,z)=\text{const}$ at that point.

(One needs to take the difference rather than the sum of the two curvatures in case one normal curve curves towards $\mathbf{F}$ and the other curves away from it, in which case the surface will look vaguely like a saddle point at that point. The difficulty arises because I have attempted to state the theorem in terms that respect the standard convention that curvature is always positive. If one adopts the convention that I have used in most of the discussion, where curvature is a signed number, then the statement of the theorem is much more straightforward.)

The Bottom Line on Divergence.

From what we have seen for the special case where $\mathbf{F}$ is a gradient field, one can see the significance in general of $\nabla \bullet \mathbf{F}$ for any differentiable vector field $\mathbf{F}$. What we see is that the divergence of a vector field shows the rate at which that vector field is, um, diverging. And this happens in two ways: positive divergence will occur when $\mathbf{F}(x,y,z)$ increases in magnitude when (x,y,z) moves in the direction of the field, and also when the vector field "fans out" as it were, that is to say when $\mathbf{F}(x,y,z)$ turns towards the direction of the movement of (x,y,z) as (x,y,z) moves in a direction perpendicular to the direction of $\mathbf{F}(x_0,y_0,z_0)$. If one thinks of the field as giving the velocity of a fluid, this would mean that as one moves away from a given point P_0 in a direction perpendicular to the velocity of the fluid at that point, the nearby streams splay out from the direction of the fluid through P_0, somewhat the way water sprays out from a nozzle that is set on "spray."

If the verbal expression of this seems unclear, the technical discussion below should clarify it. We simply duplicate the same reasoning used above in the speciaL case where $\mathbf{F}$ was a gradient. As in that special case, the crucial fact is that as shown in an article on my web site) the directional derivative of a vector field $\mathbf{F}$ in any given direction can be decomposed into two orthogonal components, one of which is parallel to the vector field $\mathbf{F}$ and shows the rate at which the magnitude of $\mathbf{F}$ is increasing as one moves in the given direction, and the other is orthogonal to $\mathbf{F}$ and whose magnitude is the product of the magnitude of $\mathbf{F}$ and the rate at which $\mathbf{F}$ is turning as one moves in the given direction.

Consider a vector field $\mathbf{F}(x, y, z)$. Change coordinates so that $\mathbf{i}$ is in the direction $\mathbf{F}(x_0, y_0, z_0)$. Suppose that the Jacobean matrix for $\mathbf{F}$ at the point (x_0, y_0, z_0) is

$$\mathcal{J} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Then

$$\nabla \cdot \mathbf{F} = a_{11} + a_{22} + a_{33} = \mathbf{i} \cdot D_{\mathbf{i}}(\mathbf{F}) + \mathbf{j} \cdot D_{\mathbf{j}}(\mathbf{F}) + \mathbf{k} \cdot D_{\mathbf{k}}(\mathbf{F}).$$

The directional derivative of $\mathbf{F}$ in the direction $\mathbf{i}$ is the vector

$$D_{\mathbf{i}}(\mathbf{F}) = a_{11}\,\mathbf{i} + a_{21}\,\mathbf{j} + a_{31}\,\mathbf{k}.$$

Since the direction of $\mathbf{i}$ is the same as the direction of $\mathbf{F}$, a_{11} is the rate at which $||\mathbf{F}||$ increases (or decreases, if a_{11} is negative) as (x, y, z) moves away from (x_0, y_0, z_0) in the direction of $\mathbf{i}$ (i. e. the direction of $\mathbf{F}$). On the other hand, $a_{21}\,\mathbf{j} + a_{31}\,\mathbf{k}$, which is the component of $D_{\mathbf{i}}(\mathbf{F})$ perpendicular to the direction of $\mathbf{F}$, points in the direction in which $\mathbf{F}$ turns as (x, y, z) moves in the $\mathbf{i}$ direction and its magnitude, $|a_{21} + a_{31}|$, is the product of $||\mathbf{F}||$ and the rate at which $\mathbf{F}$ turns. This second component of $D_{\mathbf{i}}(\mathbf{F})$ is not relevant to the divergence.

Now the directional derivative of $\mathbf{F}$ in the $\mathbf{j}$-direction is $D_{\mathbf{j}}(\mathbf{F}) = a_{12}\,\mathbf{i} + a_{22}\,\mathbf{j} + a_{32}\,\mathbf{k}$. Since $\mathbf{i}$ is in the direction of $\mathbf{F}$, the component $a_{12}\,\mathbf{i}$, which does not contribute to the divergence, is in the direction of $\mathbf{F}$ and indicates the rate at which $||\mathbf{F}||$ changes as (x, y, z) moves in the $\mathbf{j}$-direction. And $a_{22}\,\mathbf{j} + a_{32}\,\mathbf{k}$ points in the direction in which $\mathbf{F}$ is turning as (x, y, z) moves away from (x_0, y_0, z_0) in the $\mathbf{j}$-direction, and its magnitude, i. e. $|a_{22}| + |a_{32}|$ (here we use the fact that $\mathbf{j}$ and $\mathbf{k}$ are orthogonal), gives the speed of this turning multiplied by $||\mathbf{F}||$.

Now we see that this vector corresponding to the turning is the sum of two components: $a_{22}\,\mathbf{j}$ and $a_{32}\,\mathbf{k}$. The first of these is determined by the turning of $\mathbf{F}$ towards the $\mathbf{j}$ direction, which is the direction the point (x, y, z) is moving in. This is the component which is significant for the divergence. On the other hand, the component $a_{32}\,\mathbf{k}$, which does not contribute to the divergence, is determined by the rate at which $\mathbf{F}$ turns toward the $\mathbf{k}$ direction as (x, y) moves in the $\mathbf{j}$ direction. In the language used previously, a_{22} is the product of $\mathbf{F}$ and the rate at which $\mathbf{F}$ tips forwards or backwards as (x, y, z) moves in the $\mathbf{j}$-direction, and a_{32} is the product of $\mathbf{F}$ and the rate at which $\mathbf{F}$ tilts sideways at (x, y, z) moves in that direction.

Likewise, a_{33} is determined by the rate at which $\mathbf{F}$ tips forwards or backwards and as (x, y, z) moves in the $\mathbf{k}$-direction, and a_{23} shows the rate at which $\mathbf{F}$ tilts sideways. We see that the tipping of $\mathbf{F}$ forwards or backwards as (x, y, z) moves in the $\mathbf{j}$ and $\mathbf{k}$ directions contributes to the divergence, but the sidewise tilting of $\mathbf{F}$ does not.

We can sum this all up as a theorem.

Theorem. Consider a vector field $\mathbf{F}(x, y, z)$ at a point (x_0, y_0, z_0). Let $\mathbf{u}$ be a unit vector in the direction of $\mathbf{F}(x_0, y_0, z_0)$ (i. e. $\mathbf{u} = \mathbf{F}/||\mathbf{F}||$) and let $\mathbf{v}$ and $\mathbf{w}$ be unit vectors orthogonal to $\mathbf{F}(x_0, y_0, z_0)$ and to each other. Then the divergence of $\mathbf{F}$ at the point (x_0, y_0, z_0) is the sum of the rate at which $||\mathbf{F}||$ is increasing as (x, y, z) moves away from (x_0, y_0, z_0) in the $\mathbf{u}$ direction (taken as negative if $||\mathbf{F}||$ decreases rather than increases) and the product of $||\mathbf{F}||$ with the sum of the rates at which $\mathbf{F}$ tips toward the direction of the movement of (x, y, z) as (x, y, z) moves away from (x_0, y_0, z_0) in the $\mathbf{v}$ and $\mathbf{w}$-directions (with the convention at these rates should be taken as negative if $\mathbf{F}$ tips away from the direction of movement of (x, y, z) rather than towards it).

The Bottom Line for the Curl of a Vector Field with Constant Magnitude. Now look at curl in terms of its basic definition. Consider a vector field $\mathbf{U}$ in three-space with constant magnitude 1 and fix a point $P_0 = (x_0, y_0, z_0)$. Change to a new coordinate system in which $\mathbf{i}$ is parallel to $\mathbf{U}(x_0, y_0, z_0)$ and $\mathbf{j}$ points in the direction in which $\mathbf{U}$ is turning when (x, y, z) moves away from P_0 in the direction $\mathbf{U}$. The direction of $\mathbf{k}$ will then be determined by the right-hand rule. Becuase $\mathbf{U}$ has constant magnitude, the directional derivatives of $\mathbf{U}$ at P_0 are orthogonal to $\mathbf{U}(x_0, y_0, z_0)$, and thus orthogonal to $\mathbf{i}$, which means that the Jacobean matrix $\mathcal{J}$ for $\mathbf{U}$ at P_0 will have zeros in the first row. (It then follows that $\mathcal{J}$ has an eigenvector corresponding to the eigenvalue 0.) Furthermore, the directional derivative of $\mathbf{U}$ in the $\mathbf{i}$ direction will point to the direction that $\mathbf{U}$ turns in when (x, y, z) moves in the $\mathbf{i}$ direction, i. e. the first column will be a multiple of $\mathbf{j}$, i. e. is zero except in the second entry.

Thus the Jacobean matrix for $\mathbf{U}$ at the point P_0 will look like

$$\mathcal{J} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{bmatrix}$$

where a_{21} is the speed at which $\mathbf{U}$ turns as (x, y, z) moves away from P_0 in the $\mathbf{i}$ direction, i. e. the direction of $\mathbf{U}$. Then

$$\nabla \times \mathbf{U} = (a_{32} - a_{23})\,\mathbf{i} + a_{21}\,\mathbf{k}\,.$$

Thus the curl for a vector field $\mathbf{U}$ which has constant magnitude 1 is the sum of two vector components. The first component is in the direction of the axis around which $\mathbf{U}(x, y, z)$ turns as (x, y, z) moves away from (x_0, y_0, z_0) the the direction of $\mathbf{U}$ (i. e. the direction $\mathbf{i}$) at unit speed, and its magnitude is the speed of this turning. This is what we found in the two-dimensional case. But for the three dimensional case, there is a second component to the curl in the direction of $\mathbf{U}$ itself. This second component, $(a_{32} - a_{23})\,\mathbf{i}$, is the curl of the two-dimensional field which is the projection of $\mathbf{U}$ onto the plane perpendicular to $\mathbf{U}(x_0, y_0, z_0)$. It vaguely (very vaguely) indicates the extent to which the tip of $\mathbf{U}$ swirls ("circulates" would be the technically accurate word) as (x, y, z) moves around (x_0, y_0, z_0) in the plane perpendicular to $\mathbf{U}(x_0, y_0, z_0)$.

The intuitive significance of this third component is not as nice as one might hope for, but it seems that it's the best one can get.