Formal definition of divergence in two dimensions

Google Classroom

Learn how line integrals are used to formalize the idea of divergence.

Background

- Divergence
- Flux in two dimensions

If you haven't already, you may also want to read "Why care about the formal definitions of divergence and curl" for motivation.

What we're building to

- In two dimensions, divergence is formally defined as follows:

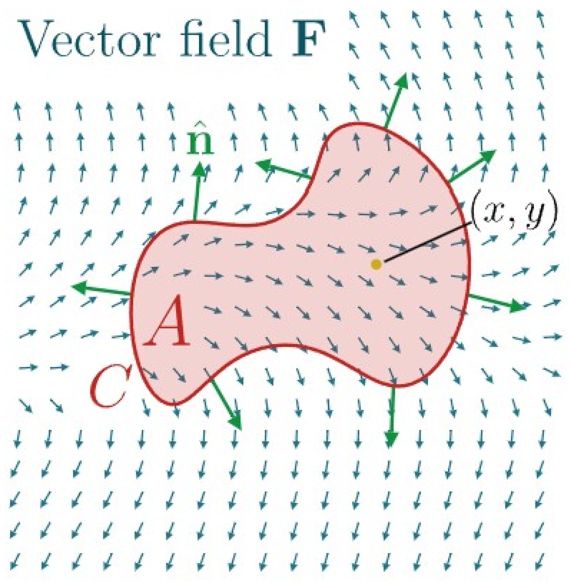
$$\text{div}\,\mathbf{F}(x,y) = \lim_{|A_{(x,y)}| \to 0} \underbrace{\frac{1}{|A_{(x,y)}|} \overbrace{\oint_C \mathbf{F} \cdot \hat{\mathbf{n}}\, ds}^{\text{2d-flux through } C}}_{\text{Flux per unit area}}$$

[Hide explanation]

- $F(x,y)$ is a two-dimensional vector field.
- (x,y) is a specific point on the xy-plane.
- $A_{(x,y)}$ is some region of the xy-plane which includes the point (x,y). For example, you can think of this as being a circle centered at (x,y).
- $|A_{(x,y)}|$ is the area of $A_{(x,y)}$.
- $\lim_{|A_{(x,y)}| \to 0}$ indicates that we are considering the limit as the area $|A_{(x,y)}|$ goes to zero. In other words, as $A_{(x,y)}$ shrinks around the point (x,y).
- C is the boundary of $A_{(x,y)}$.
- $\hat{\mathbf{n}}(x,y)$ is a function which gives an outward-facing unit normal vector at each point on C.

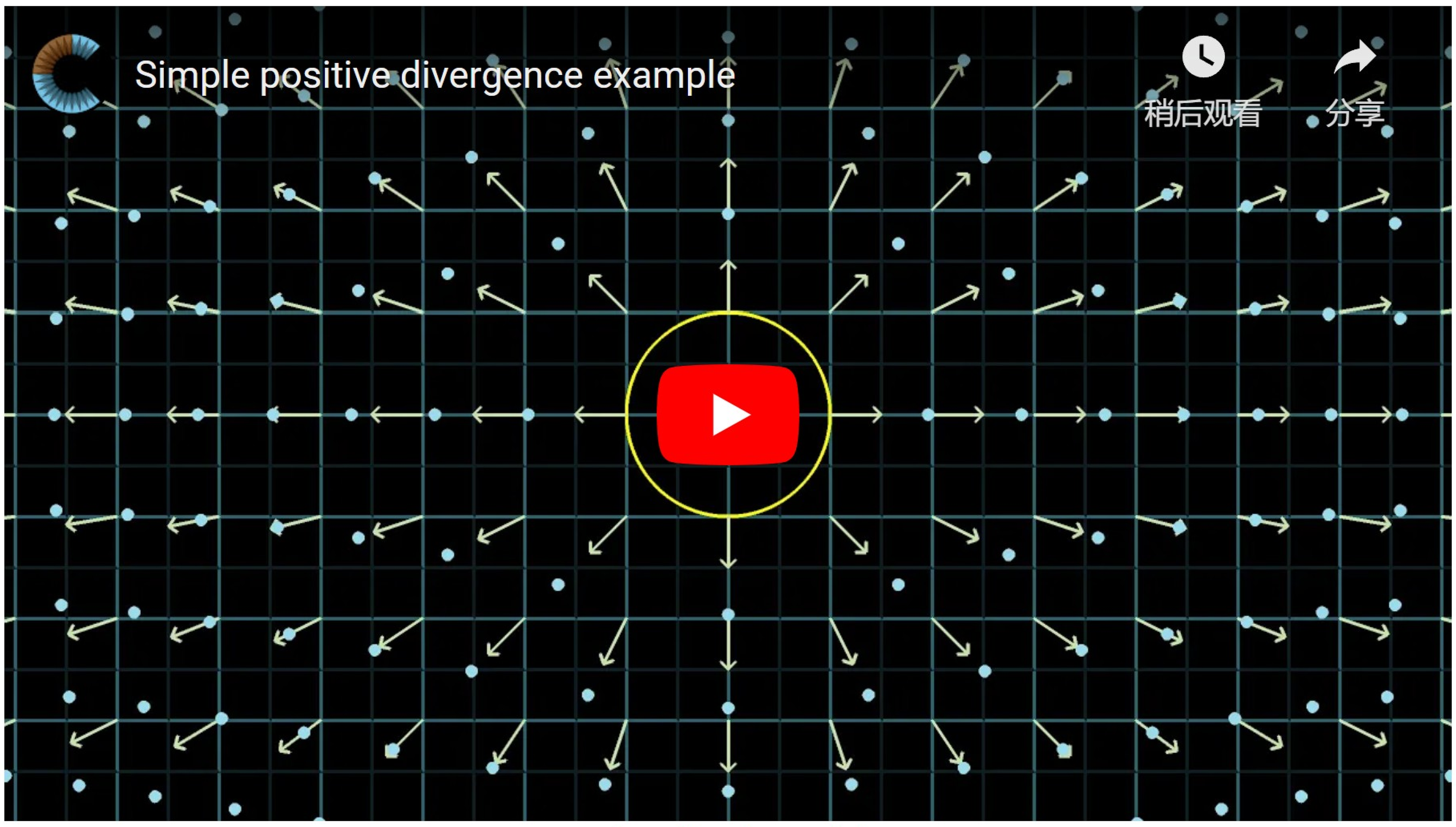
> - The line integral $\oint_C \mathbf{F} \cdot \hat{\mathbf{n}} \, ds$ represents the two-dimensional flux of F through C.

There is a lot going on in this definition, but we will build up to it one piece at a time. The bulk of the intuition comes from the background understanding of flux.



"Outward flow" at a point doesn't really make sense

By this point you should have some idea of what divergence is trying to measure. When a vector field $\mathbf{F}(x, y)$ represents a fluid flow, divergence measures the tendency for the fluid to flow away from each point.



See video transcript

However, **there's a disconnect between the idea of "outward flow" and divergence itself**:

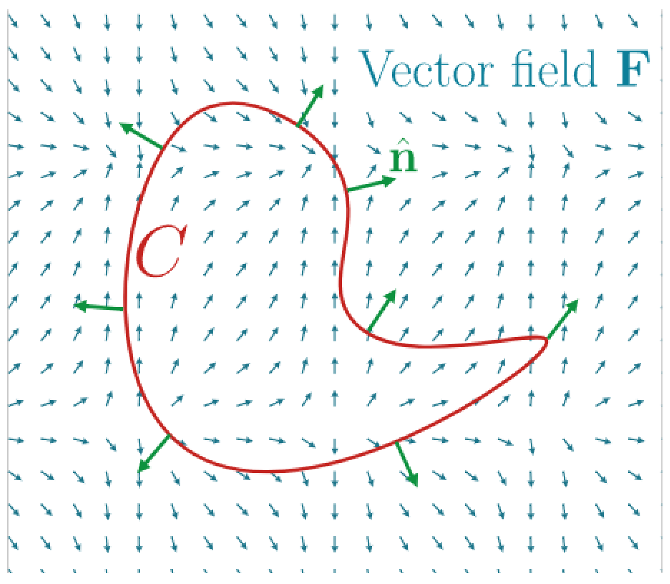
- Divergence is a function which takes in *individual points* in space.
- The idea of outward flow only makes sense with respect to a *region* in

space. You can ask if a fluid flows out of a given region or into it, but it doesn't make sense to talk about fluid flowing out of a single point.

Formally defining divergence will involve using a flux integral, which measures the outward flow in a region, then taking the appropriate limit as this region shrinks around a specific point.

From a region to a point

In the article on two-dimensional flux, we had the following setup:



- $\mathbf{F}(x, y)$ is a vector-valued function representing the velocity vector field of some fluid.

- C is a closed loop in the xy-plane.

- $\hat{\mathbf{n}}(x, y)$ is a function that gives the outward unit normal vector at all points on the curve C.

I talked about how if you were tracking the mass of fluid in the region enclosed by the curve C, you could compute the rate at which mass is leaving the region using the following line integral:

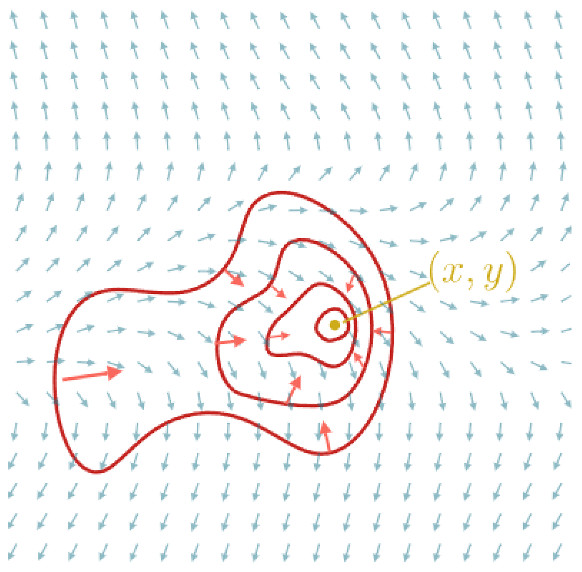
$$\underbrace{-\frac{d(\text{fluid mass in region})}{dt}}_{\text{Rate at which mass leaves region}} = \underbrace{\oint_C \mathbf{F} \cdot \hat{\mathbf{n}}\, ds}_{\text{Flux integral}}$$

This is called a "flux integral". If it is positive, fluid tends to be exiting the region, otherwise it tends to be entering the region. You can interpret this integral by imagining walking along the boundary C and measuring how much fluid tends to be exiting/entering the region at each point.

What if instead of measuring the change of mass, you wanted to know the change in *density*? Well, just divide this integral by the area of the region in question. Let's go ahead and give that region a name, A, and say that $|A|$ is the area of the region.

$$\underbrace{-\frac{d(\text{fluid density in region})}{dt}}_{\text{Change in mass per unit area in } A} = \frac{1}{|A|} \oint_C \mathbf{F} \cdot \hat{\mathbf{n}}\, ds$$

To formally define divergence of $\mathbf{F}$ at a point (x, y), we consider the limit of this change in density as the region shrinks around the point (x, y).



There is no set-in-stone notation for this, but here's what I'll go with:

- Rather than just writing A, write $A_{(x,y)}$ to emphasize that this region contains a specific point (x, y).

This is important because as we start letting the region shrink, we don't want it to wander away from the point.

The expression "$|A_{(x,y)}| \to 0$" will indicate that we are considering the limit as the area of $A_{(x,y)}$ goes to 0, meaning $A_{(x,y)}$ is shrinking around the point (x, y).

With all this, here's how we write the formal definition of divergence:

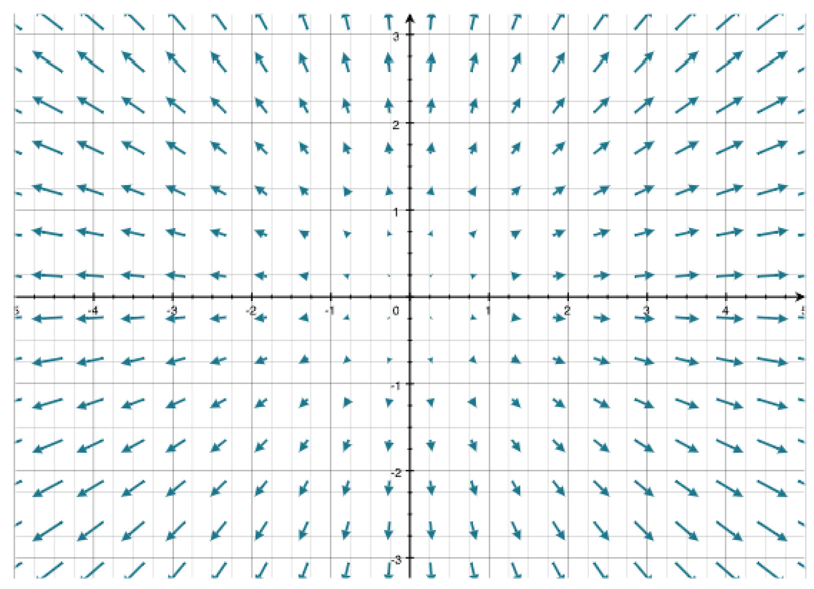
$$\text{div}\, \mathbf{F}(x, y) = \underbrace{\lim_{|A_{(x,y)}| \to 0}}_{\substack{\text{Region is shrinking} \\ \text{around } (x,y)}} \overbrace{\frac{1}{|A_{(x,y)}|} \underbrace{\oint_C \mathbf{F} \cdot \hat{\mathbf{n}}\, ds}_{\text{Flux through } C}}^{\substack{\text{Negative change of} \\ \text{fluid density in } A_{(x,y)}}}$$

"Simple" example: Constant divergence

Unlike other topics, the purpose of an example here is not to practice a skill that you will need. It is just to get a feel for what this relatively abstract definition actually looks like with a concrete function.

Let's use the quintessential "outward flowing" vector field in two dimensions:

$$\mathbf{F}(x, y) = \begin{bmatrix} x \\ y \end{bmatrix}$$



Concept check: Using the usual divergence formula, the one which arises from the notation $\nabla \cdot \mathbf{F}$, what is the divergence of $\mathbf{F}(x, y) = \begin{bmatrix} x \\ y \end{bmatrix}$?

$\nabla \cdot \mathbf{F}(x, y) =$ ____

Check

[Hide explanation]

Using the divergence formula, we can see that the divergence happens to be a constant 2 everywhere in the plane:

$$\nabla \cdot \mathbf{F}(x, y) = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

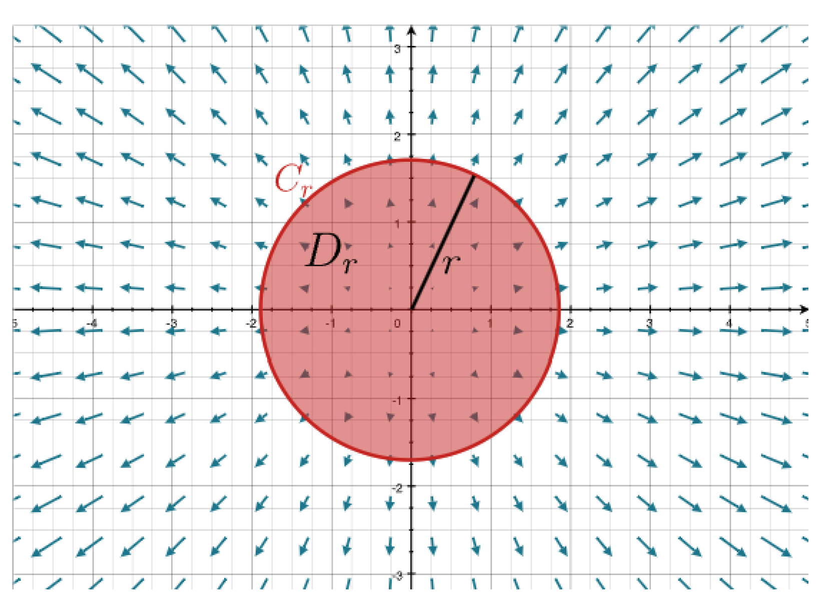
$$= \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y)$$

$$= 1 + 1$$
$$= 2$$

Now let's see how the formal definition of divergence works in this case. Let's focus on the origin.

$$(x, y) = (0, 0)$$

And for our shrinking regions around this point, consider circles. Let C_r denote a circle of radius r centered at the origin, and D_r represent the region enclosed by that circle, where D stands for "Disk".



Notice, for all values of r, the disk D_r will contain the point $(0, 0)$, so this is indeed a good family of regions to use.

The formal definition of divergence at $(0, 0)$ would then be written as follows:

$$\operatorname{div} \mathbf{F}(0,0) = \lim_{|D_r| \to 0} \frac{1}{|D_r|} \oint_{C_r} \mathbf{F} \cdot \hat{\mathbf{n}}\, ds$$

This is rather abstract, so let's start filling in the details of this integral.

Concept check: What is $|D_r|$?

$|D_r| =$ [ ]

Check

[Hide explanation]

> The area of a disk with radius r is
>
> $$|D_r| = \pi r^2$$

Concept check: Which of the following parameterizes C_r?

Choose 1 answer:

(A) $\begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$ as t ranges from 0 to 2π.

(B) $\begin{bmatrix} r\cos(t) \\ r\sin(t) \end{bmatrix}$ as t ranges from 0 to 2π.

Check

[Hide explanation]

The second answer choice is correct:

$$\begin{bmatrix} r\cos(t) \\ r\sin(t) \end{bmatrix} \text{ as } t \text{ ranges from } 0 \text{ to } 2\pi.$$

We start with the usual parameterization of a unit circle, using the expression:

$$\begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \text{ as } t \text{ ranges from } 0 \text{ to } 2\pi.$$

Then we scale it by the radius r.

Concept check: Using this parameterization, what should we replace ds with in the integral $\int_{C_r} \dots ds$?

$ds =$ [ ] dt

Check

[Hide explanation]

If you are unfamiliar with this step, consider reviewing the articles on line integrals or arc length.

The value ds represents a tiny bit of arc length on the curve C_r. Start by taking the derivative of the function parameterizing C_r:

$$\frac{d}{dt}\begin{bmatrix} r\cos(t) \\ r\sin(t) \end{bmatrix} = \begin{bmatrix} -r\sin(t) \\ r\cos(t) \end{bmatrix}$$

Then take its magnitude:

$$\begin{aligned} &\sqrt{(-r\sin(t))^2 + (r\cos(t))^2} \\ &= \sqrt{r^2(\sin^2(t) + \cos^2(t))} \\ &= r \end{aligned}$$

When you multiply this by dt, it gives you a little snippet of arc length:

$$ds = r\,dt$$

Concept check: Which of the following gives an outward facing unit normal vector $\hat{\mathbf{n}}$ to C_r?

Choose 1 answer:

(A) $\hat{\mathbf{n}}(x, y) = \begin{bmatrix} x/r \\ y/r \end{bmatrix}$

(B) $\hat{\mathbf{n}}(x, y) = \begin{bmatrix} x \\ y \end{bmatrix}$

(C) $\hat{\mathbf{n}}(x, y) = \begin{bmatrix} rx \\ ry \end{bmatrix}$

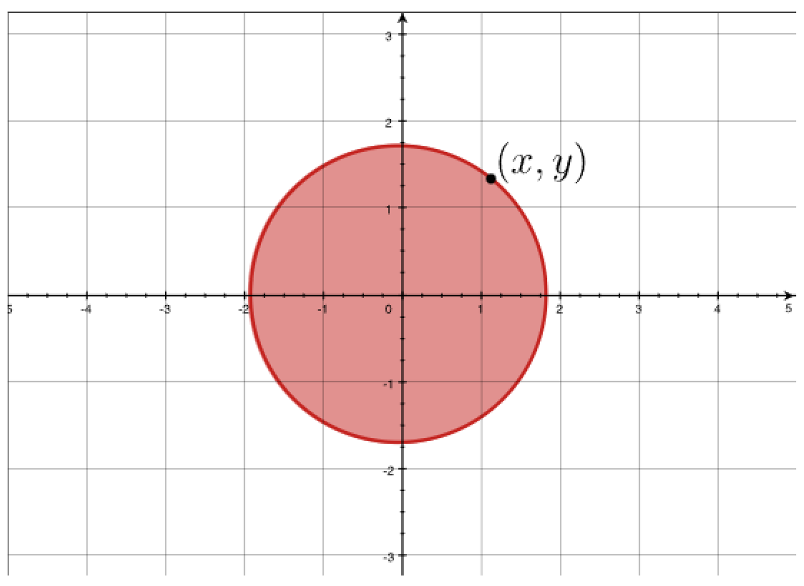
Check

[Hide explanation]

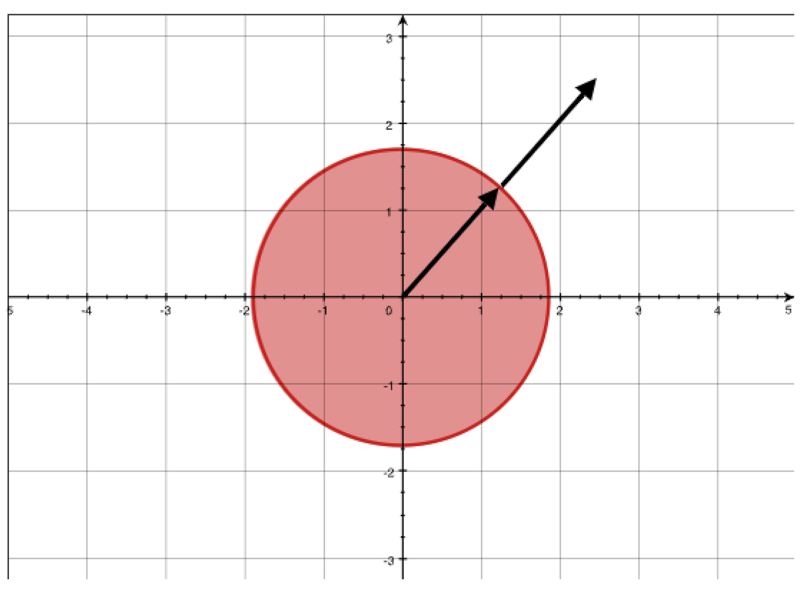
The first answer choice is correct:

$$\hat{\mathbf{n}}(x, y) = \begin{bmatrix} x/r \\ y/r \end{bmatrix}$$

Consider a point (x, y) on the circle C_r.



The vector $\begin{bmatrix} x \\ y \end{bmatrix}$ not only points from the origin to (x, y), it also gives a normal vector at that point.



However, since (x, y) is on the circle with radius r, the magnitude of this vector is r. To get a **unit** normal vector, we must divide each component by r:

$$\hat{\mathbf{n}}(x, y) = \begin{bmatrix} x/r \\ y/r \end{bmatrix}$$

Applying all these answers to the expression we had before, here's what we get:

$$
\begin{aligned}
&\operatorname{div} \mathbf{F}(0,0) \\
&\quad\Downarrow \\
&= \lim_{|D_r| \to 0} \frac{1}{|D_r|} \oint_{C_r} \mathbf{F} \cdot \hat{\mathbf{n}}\, ds \\
&= \lim_{r \to 0} \frac{1}{\pi r^2} \int_0^{2\pi} \mathbf{F}(r\cos(t), r\sin(t)) \cdot \hat{\mathbf{n}}(r\cos(t), r\sin(t))\, r\, dt \\
&= \lim_{r \to 0} \frac{1}{\pi r^2} \int_0^{2\pi} \begin{bmatrix} r\cos(t) \\ r\sin(t) \end{bmatrix} \cdot \begin{bmatrix} (r\cos(t))/r \\ (r\sin(t))/r \end{bmatrix} r\, dt \\
&= \lim_{r \to 0} \frac{1}{\pi r^2} \int_0^{2\pi} \underbrace{(r\cos^2(t) + r\sin^2(t))}_{r} r\, dt \\
&= \lim_{r \to 0} \frac{1}{\pi r^2} \int_0^{2\pi} r^2 dt \\
&= \lim_{r \to 0} \frac{1}{\pi r^2} 2\pi r^2 \\
&= 2
\end{aligned}
$$

So the formal definition *does* actually match the formula $\nabla \cdot \mathbf{F}$ that we know and love. Well, at least for this specific example anyway.

I think you'll agree, though, that this is much more labor-intensive to compute. But the point of this formal definition is *not* to use it for actual computations. The point is that it does a much better job reflecting the idea of "outward fluid flow" in a mathematical formula. Having such a solid grasp of that idea will be helpful when you learn about Green's divergence theorem.

What about higher dimensions?

In the next article, I'll show how you can do essentially the same thing to define three-dimensional divergence using three-dimensional flux, which involves a surface integral.

Summary

- Given a fluid flow, divergence tries to capture the idea of "outward flow" at a point. But this doesn't quite make sense, because you can only measure the change in fluid density of a region.

- When talking about a region, the idea of "outward flow" is the same as flux through that region's boundary.

- To adapt the idea of "outward flow in a region" to the idea of "outward flow at a point", start by considering the average outward flow per unit area in a region. This just means dividing the flux integral by the area of the region.

- Next, consider the limit of this outward flow per unit area as the region shrinks around a specific point.

- Putting this all into symbols, we get the following definition of divergence:

$$\operatorname{div} \mathbf{F}(x, y) = \lim_{\left|A_{(x,y)}\right| \to 0} \underbrace{\frac{1}{\left|A_{(x,y)}\right|} \overbrace{\oint_C \mathbf{F} \cdot \hat{\mathbf{n}}\, ds}^{\text{2d-flux}}}_{\text{Flux per unit area}}$$