Week 2 Worksheet Solutions
Math Review

Jacob Erlikhman

Exercise 1. a) What does the gradient tell you about a function? Why?
Hints: If $\nabla f(\mathbf{x}) = \mathbf{w}$, argue or show that

$$D_{\hat{v}} f(\mathbf{x}) = \hat{v} \cdot \mathbf{w},$$

where $D_{\hat{v}} f(\mathbf{x})$ is the directional derivative of f at $\mathbf{x}$ in the direction $\hat{v}$. It may help to consider the special cases $\hat{v} = (0, 1)$ and $\hat{v} = (1, 0)$ in the case that $f(x, y)$ is a function on the plane. In this scenario, what do you get if $\hat{v} = (a, b)$ (with $a^2 + b^2 = 1$)?

Remark. Notice that this result holds in *any dimension* $n \in \mathbb{N}$.

b) What does the curl tell you about a vector field? Why?
Hint: Draw and calculate the curls of some example vector fields, like $-y\hat{x} + x\hat{y}$ or $x\hat{y}$. Now, try the vector fields $x\hat{x} + y\hat{y} + z\hat{z}$, $\hat{z}$, and $z\hat{z}$.

c) What does the divergence tell you about a vector field? Why?
Hint: Using the same vector fields from the hint from (b), calculate the divergence for each of them.

d) Use (a) and (b) to give an intuitive explanation of why the curl of a gradient is always 0.

e) Use (b) and (c) to give an intuitive explanation of why the divergence of a curl is always 0.

f) Show that $\nabla \times \nabla f = 0$ directly.

g) Show that $\nabla \cdot \nabla \times \mathbf{v} = 0$ directly.

a) The gradient tells us the direction of greatest change in the function. Notice that this will follow from the formula in the hint. Indeed, if the gradient is given by the fixed vector $\mathbf{w}$ and the directional derivative in the direction given by $\hat{v}$ is $\hat{v} \cdot \mathbf{w}$, then this will be greatest when $\hat{v} = \hat{w}$. It remains to prove the formula. By definition,

$$D_{\mathbf{v}} f(\mathbf{x}) = \sum v_i D_i f(\mathbf{x}),$$

where $D_i f(\mathbf{x})$ is the i^{th} directional derivative of f at $\mathbf{x}$, i.e. it is

$$D_i f(\mathbf{x}) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(\mathbf{x})}{h}.$$

(I've written it in n-dimensional space, but the specification to 3-dimensions should be clear). But notice that these are just

$$D_i f(\mathbf{x}) = \frac{\partial f}{\partial x_i}(\mathbf{x}),$$

so that

$$D_{\mathbf{v}} f(\mathbf{x}) = \mathbf{v} \cdot \mathbf{w}.$$

b) The curl tells you how much a vector field "curls." The idea is that if you think of the vector field as defining a water current and place a floating bob with needles sticking out of it (parallel to the plane of the water), the curl tells you about which axis that bob would float, how fast it would do this (this is the magnitude of the curl), and which way it would float (this is the sign of the curl). This is given by the examples on pages 17-19 of Griffiths 4th edition. These calculations show what's going on, and they're done in Examples 1.4 and 1.5 on those pages.

c) Similarly, the divergence tells you how much a vector field "diverges." Its sign tells you whether a bob gets pulled in (negative) or pushed out (positive), and its magnitude tells you "how fast" that happens.

d) Suppose we had a gradient which went around in a circle. Now, imagine going along this circle and coming back. You'd find that the function should be larger after you return than when you started, which is impossible! Alternatively, suppose you had a function which defined a surface; further, suppose that this surface had a dip or a valley somehwere. Then the gradient would point in the direction that a ball would fall if placed on the surface (the ball would of course fall into the dip or valley). But if the gradient had a curl, what would happen to the ball? It would spin in circles rather than falling down! That doesn't make sense, so the gradient can't have a curl.

e) I initially thought that one could intuit this by using a vector field entirely contained in a plane, but I don't think this works, since, although the curl would be everywhere perpendicular to the plane, the divergence of such a vector field is not necessarily 0 (e.g. $\boldsymbol{\nabla} \cdot \mathbf{z} \neq 0$). So I'm not sure of an explanation for this! I would be grateful if someone could provide one. Sorry for the red herring!

f) This part and part (g) were on the homework, so you'll get solutions when those come out.

Exercise 2. ***Griffiths 1.13.*** Let $\mathbf{d}$ be the separation vector from a fixed point (x', y', z') to the point (x, y, z), and let d be its length. Show that

a) $\boldsymbol{\nabla}(d^2) = 2\mathbf{d}$,

b) $\boldsymbol{\nabla}(1/d) = -\hat{d}/d^2$.

c) What is the general formula for $\boldsymbol{\nabla}(d^n)$?

d) You computed these formulas in cartesian coordinates. Do they hold in other coordinate systems? Why or why not?

Remark. To prove this would require quite a bit of work or more tools than we have at our disposal. However, you should be able to come up with an intuitive argument.

a) Note that $d = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$, while $\mathbf{d} = \mathbf{x} - \mathbf{x}'$. Thus,

$$\begin{aligned}\nabla(d^2) =& 2(x-x')\hat{x} + 2(y-y')\hat{y} + 2(z-z')\hat{z} \\ =& 2\mathbf{d}.\end{aligned}$$

b) We have

$$\begin{aligned}\nabla\left(\frac{1}{d}\right) =& -\frac{1}{2}\cdot\frac{2\mathbf{d}}{d^3} \\ =& -\frac{\hat{d}}{d^2}\end{aligned}$$

c) Likewise,

$$\nabla(d^n) = \frac{n}{2}\cdot 2\mathbf{d}\cdot d^{n-2}.$$

d) The formulas we've computed are coordinate-free. To see this, note that d is a distance; hence, it cannot depend on our choice of coordinates. Similarly, our work from Exercise 1 tells us that the gradient is the direction of greatest change of a function. Its magnitude can't depend on our choice of coordinates, since it tells us a rate of change in a direction that also can't depend on our choice of coordinates. This is because the direction of greatest change is a coordinate-independent notion. Thus, we have that all the pieces of our calculation can't depend on coordinate choices, so our calculation is coordinate-independent.

Exercise 3. Challenge/Extra Problem. The Stokes' and divergence theorems have generalizations to any dimension $n \in \mathbb{N}$. In this problem, you'll get an idea of what those are.

a) The divergence theorem directly generalizes to any dimension. Write down the generalization.

b) Focus for the divergence theorem on the 1- and 2-dimensional cases (i.e. the "volume" we're integrating over is 1- or 2-dimensional). Do either of these look familiar?

c) To generalize Stokes' theorem requires more tools than we have at our disposal at the moment. However, we can readily "specialize" to dimension 1. Write down what Stokes' theorem should say in dimension 1.
Hint: Use the usual Stokes' theorem. Consider an integration region of the form $(-\varepsilon, \varepsilon) \times [a, b]$, and take $\varepsilon \to 0$.

d) Now, specialize to dimension 2. What do you obtain? It should look similar to your 2-dimensional result from (b).

e) Explain why in dimensions 1 and 2 the divergence and Stokes' theorems give the same results.

Remark. Although it isn't clear from this exercise, it is actually the generalized Stokes' theorem that is (more of) a generalization of the fundamental theorem of calculus, not the divergence theorem. One comment is that the generalized Stokes' theorem is used as such, while the divergence theorem is not.

a) Let

$$\nabla \cdot \mathbf{v} = \sum_{i=1}^{n} \frac{\partial v_i}{\partial x_i}$$

be the generalization of the divergence to n-vectors. The divergence theorem is then

$$\int_M \nabla \cdot \mathbf{v}\, \mathrm{d}V_n = \int_{\partial M} \mathbf{v} \cdot \hat{n}\, \mathrm{d}V_{n-1}.$$

Here, M is an n-dimensional "volume" that we're integrating over and $\mathrm{d}V_n$ is an infinitesimal piece of that volume. Similarly, ∂M is its boundary, which is $(n-1)$-dimensional, and $\mathrm{d}V_{n-1}$ is its infinitesimal piece of volume. Lastly, $\hat{n}$ is the normal vector of ∂M.

b) In 1 dimension, the divergence theorem says

$$\int_{[a,b]} f'(x)\, \mathrm{d}x = f(b) - f(a).$$

Indeed, the boundary of the interval is its endpoints, and taking the Riemann sum over a finite set is just taking the ordinary sum. Why do we get the minus sign? This is because the normal to the interval will point in opposite directions at the end points (this follows from the right hand rule). Thus, in dimension 1 we get the fundamental theorem of calculus.

In 2 dimensions, we have to be careful. The definition of the normal vector is not entirely obvious. If we have a 2-dimensional region which is bounded by a 1-dimensional line, i.e. a closed curve, what is the normal to this curve? Well, if the line element $\mathrm{d}\boldsymbol{\ell} = (\mathrm{d}x, \mathrm{d}y)$ is tangent to the curve, then something orthogonal to it will be $\hat{n} = (\mathrm{d}y, -\mathrm{d}x)$, since $\hat{n} \cdot \mathrm{d}\boldsymbol{\ell} = 0$. Thus, we have

$$\int_C \mathbf{v} \cdot \hat{n}\, \mathrm{d}\ell = \int_C (v_1\, \mathrm{d}y - v_2\, \mathrm{d}x),$$

where C is the bounding curve of a region D. This is the right hand side of the divergence theorem. For the other side, we get

$$\int_D \left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} \right) \mathrm{d}a.$$

This looks a bit funky, but, if we change $\mathbf{v} = (v_1, v_2) \mapsto (v_2, -v_1)$, then we get

$$\int_D \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) = \int_C (v_1\, \mathrm{d}x + v_2\, \mathrm{d}y),$$

which is more reminiscent of Stokes' theorem.[1] This theorem is known as Green's theorem.

c) Following the hint, let's consider a vector field $\mathbf{v} = (f, 0, 0)$ and integrate it over the suggested region. Since our end-result should be one-dimensional, we should set f to be independent of two of

[1]Note that we can do this without loss of generality. Indeed, both v_1 and v_2 are functions of *both* x and y.

the coordinates, but which ones? Before figuring this out, let's compute what the curl of $\mathbf{v}$ is. We get Thus, we have

$$\nabla \times \mathbf{v} = \left(0, -\frac{\partial f}{\partial z}, \frac{\partial f}{\partial y}\right).$$

Now, it's up to us if we want to set f to depend on y or z. Note that if we choose f to depend on x, Stokes' theorem will just say $0 = 0$. For simplicity, suppose f depends only on y, so that

$$\nabla \times \mathbf{v} = \left(0, 0, \frac{\partial f}{\partial y}\right).$$

Now,

$$\lim_{\varepsilon \to 0} \int_{(-\varepsilon,\varepsilon)\times[a,b]} \nabla \times \mathbf{v} \cdot \mathrm{d}\mathbf{a} = f(b) - f(a).$$

The right hand side of Stokes' theorem gives the line integral over the boundary of a rectangle which has an ever-shrinking width. Thus, its boundary is just the boundary of the interval, and the line integral over this is just evaluation at the boundary points. What do we get on the left hand side? Well,

$$\lim_{\varepsilon \to 0} \nabla \times \mathbf{v} \cdot \mathrm{d}\mathbf{a} = \frac{\partial f}{\partial y}\,\mathrm{d}y$$

since for $[a, b]$ along the y-direction and $(-\varepsilon, \varepsilon)$ along the x-direction, the normal to this region will point in the z–direction. Thus, we have that Stokes' theorem says

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a);$$

we got the fundamental theorem of calculus again!

d) Let's use the same strategy from (c). We'll integrate $\mathbf{v} = (f, g, 0)$ over a two-dimensional region S. We then get

$$\int_S \nabla \times \mathbf{v} \cdot \mathrm{d}a = \int_{\partial S} \mathbf{v} \cdot \mathrm{d}\boldsymbol{\ell}.$$

Now, we want to specialize to two-dimensions, so that means we should assume that the region we're integrating over is entirely contained in the (x, y)-plane. Thus, its normal will point in the z-direction. Similarly, f and g should be independent of z. Using this, the left hand side gives

$$\int_S \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) \mathrm{d}a.$$

On the other hand, the right hand side of Stokes' theorem gives

$$\int_{\partial S} (f\,\mathrm{d}x + g\,\mathrm{d}y).$$

Putting these together, we have the result,

$$\int_S \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) \mathrm{d}a = \int_{\partial S} (f \; \mathrm{d}x + g \, \mathrm{d}y).$$

This is Green's theorem again! This is identical to our result from (b)!

e) So what's going on? It appears that Stokes' theorem and the divergence theorem degenerate to the same theorem in dimensions 1 and 2. A possible explanation in two dimensions is that Stokes' theorem measures how much water is "trapped" inside a two-dimensional region, while the divergence theorem measures how much water "leaks out." It follows that in dimension 2 these two things are measuring two sides of the same coin, so there's no way we could obtain extra information from using one over the other. Of course, in 1 dimension there is no difference between curling and divergence, since there's only one derivative to take anyway.