MAT 218 FALL 2008
FEEDBACK ON PROBLEM SET 2

1. Some common mistakes.

Folland 2.2.2: Some of you wrote

$$df = (\partial_x f, \partial_y f, \partial_z f)$$

but this is not correct; the right hand side of the above expression is a tangent vector (more commonly denoted by ∇f) while the left hand side, df, should be a differential (or 1-form) instead. Indeed, by definition, df should be expressed in terms of dx, dy and dz as follows:

$$df = \partial_x f dx + \partial_y f dy + \partial_z f dz.$$

It looks as if we are just expressing the same thing in two different ways, but this is coincidental: as you will see when we progress through this course, a 1-form is a *linear functional* on the space of vectors, and shouldn't be confused with vectors themselves, especially if we will be dealing with curved spaces or surfaces. Put another way, ∇f is a tangent vector to $\mathbb{R}^3$, while df is an element in the dual space to the tangent space of $\mathbb{R}^3$. As you'll see, it is the latter that we shall be able to integrate over a curve, not the former. So let's keep this distinction in mind now and write, for instance,

$$df = (2x + x^2)e^{x-y+3z}dx - x^2e^{x-y+3z}dy + 3x^2e^{x-y+3z}dz$$

in part (a) of the question.

Folland 2.3.1: In say part (b) of the question, it was given that

$$w = f(x, u, v), \quad u = g(x, y), \quad \text{and} \quad v = h(x, z).$$

A number of you wrote something like

$$\partial_y w = \frac{\partial f}{\partial g}\frac{\partial g}{\partial y},$$

but the expression $\frac{\partial f}{\partial g}$ is not particularly well-defined: for instance, if now I change the question a little bit, and suppose you are given instead

$$w = f(x, u, v), \quad u = g(x, y) \quad \text{and} \quad v = g(x, z).$$

What should $\partial_y w$ be then? If one still writes $\frac{\partial f}{\partial g}\frac{\partial g}{\partial y}$, this is then ambiguous and confusing, because it will not be clear from the answer whether you are differentiating f with respect to the second or the third variable, now that g 'occurs twice' in the explicit expression of f. Therefore in place of such an ambiguous notation, one either writes $\partial_2 f$ or $\frac{\partial f}{\partial u}$ to indicate the derivative of f with respect to the second variable (evaluated at a suitable point). (Do you see why $\frac{\partial f}{\partial g}$ is ambiguous while $\frac{\partial f}{\partial u}$ is not?) Back to the original question, the answer should be either

$$\partial_y w = \partial_2 f \partial_2 g$$

or

$$\partial_y w = \frac{\partial f}{\partial u}\frac{\partial g}{\partial y};$$

more precisely, one could also write

$$\frac{\partial}{\partial y}\left(f(x,g(x,y),h(x,z))\right)\bigg|_{(x,y,z)} = \frac{\partial f}{\partial u}\bigg|_{(x,g(x,y),h(x,z))}\frac{\partial g}{\partial y}\bigg|_{(x,y)}.$$

Folland 2.2.3: One should distinguish the differential at a general point from that at a specific point. For instance, it is not correct to write

$$df = \frac{2xy^{\frac{3}{2}}z}{z+1}dx + \frac{3x^2y^{\frac{1}{2}}z}{2(z+1)}dy + \frac{x^2y^{\frac{3}{2}}}{(z+1)^2}dz = 40dx + 37.5dy + 50dz$$

by suddenly implicitly evaluating the differential at the point $(5,4,1)$; the last equality is simply false.

Folland 2.2.7(c): Many of you argued that since

$$\nabla f = \left(\frac{2xy(x^2+y^2)-2x(x^2y)}{(x^2+y^2)^2}, \frac{x^2(x^2+y^2)-2y(x^2y)}{(x^2+y^2)^2}\right)$$

and since the right hand side is not defined at $(0,0)$, f is not differentiable at $(0,0)$. But this argument is flawed. The quotient rule only asserts the following: if you differentiate a quotient of differentiable functions, whose denominator is non-zero at a certain point, then the quotient is actually differentiable, and you can apply the quotient rule formula to compute that derivative. It doesn't assert that if the quotient rule formula doesn't make sense at a certain point, then the function is not differentiable at that point. (Think about the example

$$f(x) = \frac{\sin x}{x}$$

on $\mathbb{R}$.) Also, if f were differentiable at a certain point, then we already know (by Theorem 2.17 of Folland or Theorem 2.7 of Spivak) that

$$\nabla f = (\partial_x f, \partial_y f)$$

at that point; in other words, the components of the gradient of f are just the *directional derivatives* of the function in the coordinate directions in this situation. So if we have already assumed by contradiction here that f were differentiable at $(0,0)$, we would already have known that

$$\begin{aligned}\nabla f(0,0) &= (\partial_{(1,0)}f(0,0), \partial_{(0,1)}f(0,0))\\ &= (\cos^2 0\sin 0, \cos^2\frac{\pi}{2}\sin\frac{\pi}{2})\\ &= (0,0)\end{aligned}$$

from part (b); there is no need to use quotient rule at all in the first place. This leads to the desired contradiction: because then by Theorem 2.23 of Folland, all directional derivatives will have to vanish, which is not the case as we have shown in part (b).

Some of you even confused the direction $(\cos\theta,\sin\theta)$ at the tangent space of $(0,0)$ with the actual point $(\cos\theta,\sin\theta)$ on the plane; the former is a direction that should be thought of as 'attached' to the origin (and that just indicates a direction), while the latter is a point on the plane that lies away from the origin (it actually is at a distance 1 from the origin). Again, this will be an important difference when we begin to work with curved space or surfaces.

Spivak 2-6: Some of you argued that $f(x,y) = \sqrt{|xy|}$ is not differentiable at 0 since it is the composition of the absolute value with the square root, both of which are not differentiable at 0. But this is not true; can you come up with a simple example in which compositions of non-differentiable functions give differentiable functions?

By the way, both Folland 2.2.7(c) and this question requires one to show the non-differentiability of a function, but this one is easier; indeed in this question, some partial derivatives of the function at the desired point doesn't even exist. See solution below, and also refer to the comments above on Folland 2.2.7(c) for some common mistakes.

Spivak 2-7: To show that f is differentiable at the origin, one has to choose first a linear function λ and then argue that

$$\lim_{h\to 0}\frac{|f(h)-f(0)-\lambda(h)|}{\|h\|}=0.$$

(Note here that h is a vector and one can never divide by a vector; so in the denominator it must be $\|h\|$, and not just h. Take care to distinguish between vectors and scalars, as this also helps you avoid many common mistakes.) Many of you just began by writing down the expression

$$\frac{f(h)-f(0)-\lambda(h)}{\|h\|},$$

work on it for a while, and say (or even conclude, which is logically incorrect) suddenly what λ is. This is not advisable, for such presentations are usually confusing and problematic. See solution below.

2. Solution to selected exercises.

Folland 2.2.7(c): If f were differentiable at $(0,0)$, then its gradient ∇f at $(0,0)$ must be given by its directional derivatives along the coordinate axes. Hence by (b), $\nabla f(0,0)=(0,0)$. But then other directional derivatives of f, being the dot product of $\nabla f(0,0)$ with unit vectors, must also be 0 at the origin. This contradicts what we have found in (b), so f could not be differentiable at $(0,0)$.

Spivak 2-6: To prove that f is not differentiable at the origin, it suffices to show that the directional derivative of f along the direction $(1,1)$ does not exist at $(0,0)$. Now for all real t,

$$\frac{f(t,t)-f(0,0)}{t}=\frac{\sqrt{|t^2|}}{t}=\frac{|t|}{t},$$

and this does not have a limit as $t\to 0$. Hence the directional derivative of f along the direction $(1,1)$ does not exist at the origin, and f couldn't be differentiable there.

Spivak 2-7: Define a linear map λ by $\lambda(h_1,\dots,h_n)=0$ for all $(h_1,\dots,h_n)\in\mathbb{R}^n$. Then we claim that

$$\lim_{(h_1,\dots,h_n)\to 0}\frac{|f(h_1,\dots,h_n)-f(0,\dots,0)-\lambda(h_1,\dots,h_n)|}{\|(h_1,\dots,h_n)\|}=0.$$

To see this, first observe that

$$|f(0,\dots,0)|\le\|(0,\dots,0)\|^2=0$$

so $f(0,\dots,0)=0$. Next,

$$\begin{aligned}\frac{|f(h_1,\dots,h_n)-f(0,\dots,0)-\lambda(h_1,\dots,h_n)|}{\|(h_1,\dots,h_n)\|}&=\frac{|f(h_1,\dots,h_n)|}{\|(h_1,\dots,h_n)\|}\\&\le\frac{\|(h_1,\dots,h_n)\|^2}{\|(h_1,\dots,h_n)\|}\\&=\|(h_1,\dots,h_n)\|.\end{aligned}$$

Hence as $(h_1, \ldots, h_n) \to 0$,

$$\frac{|f(h_1, \ldots, h_n) - f(0, \ldots, 0) - \lambda(h_1, \ldots, h_n)|}{\|(h_1, \ldots, h_n)\|} \to 0,$$

as desired. It follows that f is differentiable at $(0, \ldots, 0)$, and its derivative is $\nabla f(0, \ldots, 0) = (0, \ldots, 0)$.

Remark: You could have written the above proof using vector notations like h for $(h_1, \ldots, h_n)$ (and are encouraged to do so), but you should be able to switch between that and the explicit formalism that we have adopted above. This helps you keep in mind which is a vector and which is a scalar, and helps you check that all the expressions you write actually makes sense.