4.3 partial differentiation in $\mathbb{R}^3$ and $\mathbb{R}^n$

Definition 4.3.1.

Let $f : dom(f) \subseteq \mathbb{R}^3 \to \mathbb{R}$ be a function with $(x_o, y_o, z_o) \in dom(f)$. If the directional derivative below exists, then we define the **partial derivative** of f at $\vec{p}_o = (x_o, y_o, z_o)$ with respect to x, y, z by

$$\frac{\partial f}{\partial x}(\vec{p}_o) = (D_{\widehat{x}} f)(\vec{p}_o), \qquad \frac{\partial f}{\partial y}(\vec{p}_o) = (D_{\widehat{y}} f)(\vec{p}_o), \qquad \frac{\partial f}{\partial y}(\vec{p}_o) = (D_{\widehat{z}} f)(\vec{p}_o)$$

respective. We also use the notations $\frac{\partial f}{\partial x} = \partial_x f = f_x$, $\frac{\partial f}{\partial y} = \partial_y f = f_y$ and $\frac{\partial f}{\partial z} = \partial_z f = f_z$. Generally, if $f : dom(f) \subseteq \mathbb{R}^n \to \mathbb{R}$ is a function with $\vec{p}_o \in dom(f)$ and the limit below exists, then we define the **partial derivative** of f at $\vec{p}_o$ with respect to x_j by

$$\frac{\partial f}{\partial x_j}(\vec{p}_o) = (D_{\widehat{x}_j} f)(\vec{p}_o).$$

The notation $\frac{\partial f}{\partial x_j} = \partial_j f$ is at times useful.

Once more we have natural interpretations for these partial derivatives:

f_x **gives the rate of change in** f **in the** x**-direction.**
f_y **gives the rate of change in** f **in the** y**-direction.**
f_z **gives the rate of change in** f **in the** z**-direction.**

It is useful to rewrite the definition of the partial derivatives explicitly in terms of derivatives.

$$\frac{\partial f}{\partial x}(x_o, y_o, z_o) = \frac{d}{dt}\left[\ f(x_o + t, y_o, z_o)\ \right]\Bigg|_{t=0}$$

$$\frac{\partial f}{\partial y}(x_o, y_o, z_o) = \frac{d}{dt}\left[\ f(x_o, y_o + t, z_o)\ \right]\Bigg|_{t=0}$$

$$\frac{\partial f}{\partial z}(x_o, y_o, z_o) = \frac{d}{dt}\left[\ f(x_o, y_o, z_o + t)\ \right]\Bigg|_{t=0}.$$

Partial differentiation is just differentiation where we hold all but one of the **independent variables** constant. Notice that z in the context above is an independent variable. In constrast, when we studied $z = f(x, y)$ the variable z was a **dependent variable**. The symbols x, y, z are not reserved. They have multiple meanings in multiple contexts and you must have the correct conceptual framework if you are to make the correct computations. When z, x are independent we have $\frac{\partial z}{\partial x} = 0$. If z, x are dependent then it is generally some function. In any event, the following proposition should be entirely unsurprising at this point:

Proposition 4.3.2.

Assume f, g are functions from $\mathbb{R}^3$ to $\mathbb{R}$ whose partial derivatives exist. Then for $c \in \mathbb{R}$,

1. $(f+g)_x = f_x + g_x$ and $(f+g)_y = f_y + g_y$ and $(f+g)_z = f_z + g_z$.
2. $(cf)_x = cf_x$ and $(cf)_y = cf_y$ and $(cf)_z = cf_z$.
3. $(fg)_x = f_x g + f g_x$ and $(fg)_y = f_y g + f g_y$ and $(fg)_z = f_z g + f g_z$.

Moreover, if $h : dom(h) \subseteq \mathbb{R} \to \mathbb{R}$ is a differentiable and $x_1 = x,\ x_2 = y,\ x_3 = z$,

4. $$\frac{\partial}{\partial x_j}\big[h(f(x_1,x_2,x_3))\big] = \frac{dh}{dt}\Big|_{f(x_1,x_2,x_3)}\frac{\partial f}{\partial x_j} = \frac{dh}{df}\frac{\partial f}{\partial x_j}$$
5. $$\frac{\partial x_i}{\partial x_j} = \delta_{ij} \qquad \text{where } x_1 = x, x_2 = y, x_3 = z.$$

Proof: The proofs are nearly identical to those given in the $n = 2$ case. However, I will offer a proof of (5.) for arbitrary n. Suppose $f(x_1, x_2, \dots, x_n) = x_i = \vec{x} \bullet \widehat{x}_i$ and calculate

$$\frac{\partial f}{\partial x_j} = \lim_{t\to 0}\left[\frac{f(\vec{x}) - f(\vec{x} + t\,\widehat{x}_j)}{t}\right] = \lim_{t\to 0}\left[\frac{x_i - (\vec{x} + t\,\widehat{x}_j)\bullet\widehat{x}_i}{t}\right] = \lim_{t\to 0}\left[\frac{x_i - x_i + t\delta_{ij}}{t}\right] = \delta_{ij}.$$

Therefore, $\partial_j x_i = \delta_{ij}$ for all $i, j \in \mathbb{N}_n$. In particular, this result applies to the case $n = 3$ hence the proof of (5.) is complete. Naturally this proposition equally well applies to $f : \mathbb{R}^n \to \mathbb{R}$. The proofs are nearly identical to the $n = 2$ case, we just have a few sums to sort through. I leave those to the reader. $\square$

Example 4.3.3. *. 1*

$\boxed{\text{E52}}$ Let $g(x,y,z) = xy^2z^3 + \sin(xyz)$ then

$g_x = y^2z^3 + yz\cos(xyz)$: y, z treated as constants.

$g_y = 2xyz^3 + xz\cos(xyz)$: x, z treated as constants.

$g_z = 3xy^2z^2 + xy\cos(xyz)$: x, y treated as constants.

Example 4.3.4. . *2*

Let $f(x,y,z) = x/(y+z)$ find $f_z(3,2,1) = \frac{\partial f}{\partial z}\Big|_{(3,2,1)}$

$$\frac{\partial f}{\partial z}\Big|_{(3,2,1)} = \frac{-x}{(y+z)^2}\Big|_{(3,2,1)} = \frac{-3}{(2+1)^2} = \frac{-3}{9} = \boxed{\frac{-1}{3}}$$

Example 4.3.5. . *3*

$\boxed{E53}$ Suppose $r = \sqrt{x^2+y^2+z^2}$.

$$\frac{\partial r}{\partial x} = \frac{1}{2\sqrt{x^2+y^2+z^2}}\frac{\partial}{\partial x}\left[x^2+y^2+z^2\right] = \frac{x}{\sqrt{x^2+y^2+z^2}} = \frac{x}{r}$$

Likewise $\partial r/\partial y = y/r$ and $\partial r/\partial z = z/r$.

Example 4.3.6. . *4*

$u = \sqrt{x_1^2+x_2^2+\cdots+x_n^2}$. Let $1 \leq k \leq n$ then,

$$\frac{\partial u}{\partial x_k} = \frac{\partial}{\partial x_k}\left[\sqrt{x_1^2+x_2^2+\cdots+x_n^2}\right]$$

$$= \frac{1}{2\sqrt{x_1^2+x_2^2+\cdots+x_n^2}}\frac{\partial}{\partial x_k}\left[\cancel{x_1^2}^{\,0}+\cancel{x_2^2}^{\,0}+\cdots+\cancel{x_k^2}^{\,2x_k}+\cdots+\cancel{x_n^2}^{\,0}\right]$$

$$= \boxed{\frac{x_k}{\sqrt{x_1^2+\cdots+x_n^2}}}$$

Example 4.3.7. . 5

Verify $u = \frac{1}{\sqrt{x^2+y^2+z^2}}$ solves $u_{xx} + u_{yy} + u_{zz} = 0$.

if $W \equiv \sqrt{x^2+y^2+z^2}$ then $W_x = \frac{x}{W}$

$$\frac{\partial}{\partial x}(u) = \frac{\partial}{\partial x}\left[\frac{1}{W}\right] = \frac{-1}{W^2}\frac{\partial W}{\partial x} = \frac{-1}{W^2}\frac{x}{W} = \frac{-x}{W^3}$$

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left[\frac{-x}{W^3}\right] = \frac{-1W^3 + x \cdot 3W^2 W_x}{W^6} = \frac{-W^3 + 3x^2 W}{W^6} = u_{xx} = \frac{-W^2 + 3x^2}{W^5}$$

Likewise, u_{yy} and u_{zz} have same form with $x \to y$ or z,

$$u_{xx} + u_{yy} + u_{zz} = \frac{-W^2+3x^2}{W^5} + \frac{-W^2+3y^2}{W^5} + \frac{-W^2+3z^2}{W^5}$$

$$= \frac{-3W^2 + 3(x^2+y^2+z^2)}{W^5}$$

$$= \frac{\cancel{-3(x^2+y^2+z^2)} + \cancel{3(x^2+y^2+z^2)}}{W^5}$$

we'll explain later.

$$= 0. \quad \therefore \boxed{u \text{ solves } u_{xx} + u_{yy} + u_{zz} = \nabla^2 u = 0.}$$

Remark: I'm pretty-sure that introducing W makes life easier here.

4.3.1 directional derivatives and the gradient in $\mathbb{R}^3$ and $\mathbb{R}^n$

The idea of Example 4.2.21 equally well transfer to functions of three or more variables. We usually require the functions we analyze to be continuously differentiable since that avoids certain pathological examples:

Definition 4.3.8.

A function $f : dom(f) \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be **continuously differentiable** at $\vec{p}_o \in dom(f)$ iff the partial derivative functions $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots \frac{\partial f}{\partial x_n}$ are continuous at $\vec{p}_o$. We say $f \in C^1(\vec{p}_o)$. If all the second-order partial derivatives of f are continuous at $\vec{p}_o$ then we say $f \in C^2(\vec{p}_o)$. If continuous partial derivatives of arbitrary order exist at $\vec{p}_o$ then we say f is **smooth** and write $f \in C^\infty \vec{p}_o$.

We'll see an example in the next section where the formula below holds for a multivariate functions which is not even continuously differentiable, however the geometric analysis which flows from this formula is most meaningful for continuously differentiable functions.

Proposition 4.3.9.

Suppose $f : dom(f) \subseteq \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable at $\vec{p}_o \in \mathbb{R}^n$ then the directional derivative at $\vec{p}_o$ in the direction of the unit vector $\widehat{u}$ is given by:

$$D_{\widehat{u}} f(\vec{p}_o) = \Big\langle\ \partial_1 f(\vec{p}_o),\ \partial_2 f(\vec{p}_o), \dots, \partial_n f(\vec{p}_o)\ \Big\rangle \bullet \widehat{u}.$$

Proof: delayed until the next section. $\square$

At this point it is useful to introduce a convenient notation which groups all the partial derivatives together in a particular vector of functions. Notice that the length of the gradient vector depends on the context in which it is used.

Definition 4.3.10.

If the partial derivatives of f exist then we define

$$\nabla f = \langle \partial_1 f,\ \partial_1 f,\ \dots,\ \partial_n f \rangle = \widehat{x}_1 \frac{\partial f}{\partial x_1} + \widehat{x}_2 \frac{\partial f}{\partial x_2} + \cdots + \widehat{x}_n \frac{\partial f}{\partial x_n}.$$

we also use the notation $grad(f)$ and call this the **gradient** of f.

The upside-down triangle ∇ is also known as *nabla*. Identify that for $\mathbb{R}^3$ $\nabla = \widehat{x}\partial_x + \widehat{y}\partial_y + \widehat{z}\partial_z$. The operator ∇ takes a function f and produces a vector field ∇f. This is called the **gradient vector field** of f. For a continuously differentiable function we have the following beautiful formula for the directional derivative:

$$\boxed{D_{\widehat{u}} f(\vec{p}_o) = (\nabla f)(\vec{p}_o) \bullet \widehat{u}.}$$

Technically this isn't the definition, but pragmatically this is almost always what we use to work out problems. We can also write the dot-product in terms of lengths and the angle between the gradient vector $(\nabla f)(\vec{p}_o)$ and the unit-direction vector $\widehat{u}$:

$$\boxed{D_{\widehat{u}} f(\vec{p}_o) = ||(\nabla f)(\vec{p}_o)|| \cos\theta.}$$

Just like the $n = 2$ case we can use the gradient vector field to point us in the directions in which f either increases, decreases or simply stays constant.

Example 4.3.11. Problem: *Suppose $f(x, y, z) = x^2 + y^2 + z^2$. Does f increase at a rate of 10 in any direction at the point $(1, 2, 3)$?*

Solution: *Note $\nabla f(x, y, z) = \langle 2x, 2y, 2z\rangle$ thus $\nabla f(1, 2, 3) = \langle 2, 4, 6\rangle$. The magnitude of $\nabla f(1, 2, 3)$ is $||\nabla f(1, 2, 3)|| = \sqrt{4 + 16 + 36} = \sqrt{56}$ and that is the maximum rate possible. Therefore, the answer is no. This function increases at a rate of $\sqrt{56}$ in the direction $\frac{1}{\sqrt{14}}\langle 1, 2, 3\rangle$.*

Example 4.3.12. Problem: *Suppose $f(x, y, z) = 2x + y + 2z$. Does f increase at a rate of 2 in any direction at the point $(1, 1, 1)$?*

Solution: *Note $\nabla f(x, y, z) = \langle 2, 1, 2\rangle$ thus $\nabla f(1, 1, 1) = \langle 2, 1, 2\rangle$. The magnitude $||\nabla f(1, 1, 1)|| = \sqrt{9} = 3$ and that is the maximum rate possible. Therefore, the answer is yes. Now let's find the direction(s) in which this occurs. Solve:*

$$D_{\langle a,b,c\rangle} f(1, 1, 1) = \langle 2, 1, 2\rangle \bullet \langle a, b, c\rangle = 2a + b + 2c = 2$$

subject the unit-vector condition $a^2 + b^2 + c^2 = 1$. I'll eliminate c by solving the linear equation for $c = \frac{1}{2}(2 - 2a - b)$ and substituting:

$$a^2 + b^2 + \frac{1}{4}(2 - 2a - b)^2 = 1.$$

This give an ellipse in a, b-space. Apparently there is not just one direction where f increases at a rate of 2. There are infinitely many. For example, we can easily solve the ellipse equation for its b-intercepts by putting $a = 0$,

$$b^2 + \frac{1}{4}(2 - b)^2 = 1 \quad\Rightarrow\quad 4b^2 + 4 - 4b + b^2 = 4 \quad\Rightarrow\quad 5b^2 - 4b = 0 \quad\Rightarrow\quad b(5b - 4) = 0.$$

We find the points $(0, 0)$ and $(0, 4/5)$ are on the ellipse. Returning to the plane equation we find the c-value for these points by substituting them into the equation $c = \frac{1}{2}(2 - 2a - b)$:

$$(0, 0): \quad c = \frac{1}{2}(2 - 2a - b) = 1 \qquad \& \qquad (0, 4/5): \quad c = \frac{1}{2}(2 - 4/5) = \frac{1}{2}\cdot\frac{6}{5} = \frac{3}{5}.$$

Thus, we find the direction vectors $\langle 0, 0, 1\rangle$ and $\langle 0, \frac{4}{5}, \frac{3}{5}\rangle$ point where f increases at a rate of 2. You can probably see a few more possibilities by just thinking about $(\nabla f)(1, 1, 1) = \langle 2, 1, 2\rangle$ directly. For example, I see $\langle 1, 0, 0\rangle$ also works.

The two-dimensional analogue of this problem is much easier since we have to solve the intersection of a line and the unit-circle. In that case there are either $0, 1$ or 2 solutions. The three dimensional case is much more interesting. If f models the temperature at the point (x, y, z) then this calculation shows there are many directions in which the temperature increases at a rate of 2.

Example 4.3.13. .

find $(D_{\hat{u}} f)(P)$ at $P = (0,0,0)$ in $\vec{V} = \langle 5, 1, -2 \rangle$ direction.

$$f(x,y,z) = xe^y + ye^z + ze^x$$

$$\nabla f = \left\langle \frac{\partial}{\partial x}(xe^y + ye^z + ze^x), \frac{\partial}{\partial y}(xe^y + ye^z + ze^x), \frac{\partial}{\partial z}(xe^y + ye^z + ze^x) \right\rangle$$

$$= \langle e^y + ze^x, xe^y + e^z, ye^z + e^x \rangle$$

Thus we can calculate,

$$(\nabla f)(0,0,0) = \langle 1 + 0, 0+1, 0+1 \rangle = \langle 1, 1, 1 \rangle.$$

Finally, we need to normalize $\vec{V}$,

$$\hat{V} = \frac{1}{|\vec{V}|}\vec{V} = \frac{1}{\sqrt{25+1+4}}\langle 5, 1, -2 \rangle = \frac{1}{\sqrt{30}}\langle 5, 1, -2 \rangle.$$

Thus,

$$(D_{\hat{V}} f)(0,0,0) = (\nabla f)(0,0,0) \cdot \left[\frac{1}{\sqrt{30}}\langle 5, 1, -2 \rangle\right]$$

$$= \frac{1}{\sqrt{30}}\langle 1,1,1 \rangle \cdot \langle 5, 1, -2 \rangle = \boxed{\frac{4}{\sqrt{30}}}$$

Example 4.3.14. .

Let $\vec{V} = \langle 1,2,3 \rangle$ then $\hat{V} = \frac{1}{\sqrt{14}}\langle 1,2,3 \rangle$ has $|\hat{V}| = 1$.

Suppose $f(x,y,z) = x/(y+z)$. Find $(D_{\hat{V}} f)(4,1,1)$,

$$\nabla f = \left\langle \frac{1}{y+z}, \frac{-x}{(y+z)^2}, \frac{-x}{(y+z)^2} \right\rangle$$

$$(\nabla f)(4,1,1) = \langle 1/2, -4/4, -4/4 \rangle = 1/2\langle 1, -2, -2 \rangle$$

$$(\nabla f)(4,1,1) \cdot \hat{V} = \frac{1}{2}\frac{1}{\sqrt{14}}\langle 1,-2,-2 \rangle \cdot \langle 1, 2, 3 \rangle = \frac{1}{2\sqrt{14}}(1-4-6) = \boxed{\frac{-9}{2\sqrt{14}}}$$

Therefore we find $\boxed{(D_{\hat{V}} f)(4,1,1) = -9/2\sqrt{14}}$

Example 4.3.15. .

Let $f(x,y,z) = \sqrt{x^2+y^2+z^2}$ find the max. rate of change at $(3,6,-2)$ and the <u>direction</u> in which it occurs.

$$\nabla f = \langle f_x, f_y, f_z \rangle$$

$$= \left\langle \frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{z}{\sqrt{x^2+y^2+z^2}} \right\rangle$$

$$= \frac{1}{\sqrt{x^2+y^2+z^2}} \langle x, y, z \rangle.$$

Notice for $(3,6,-2)$ we have $\sqrt{x^2+y^2+z^2} = \sqrt{9+36+4} = 7$.

$$\nabla f(3,6,-2) = \underline{\frac{1}{7}\langle 3, 6, -2 \rangle}.$$

$$|\nabla f(3,6,-2)| = \frac{1}{7}\sqrt{3^2+6^2+2^2} = \boxed{1 = |\nabla f(3,6,-2)|}$$

Max Rate of Change

This occurs in the $\nabla f(3,6,-2)$ direction which is the $\boxed{\frac{1}{7}\langle 3, 6, -2 \rangle - \text{direction}}$

Example 4.3.16. .

Consider $f(x,y,z) = x^2+y^2+z^2$. Find the directional derivative of f at $(2,1,3)$ in the direction of the origin. That is the $(-2,-1,-3)$ direction, we need a unit vector so ÷ by length $\sqrt{4+1+9}$ to construct,

$$\hat{u} = \frac{1}{\sqrt{14}}(-2,-1,-3)$$

We find the gradient of f,

$$\nabla f = \langle 2x, 2y, 2z \rangle \Rightarrow (\nabla f)(2,1,3) = \langle 4, 2, 6 \rangle.$$

Hence,

$$(D_{\hat{u}} f)(2,1,3) = \frac{-1}{\sqrt{14}}\langle 4,2,6\rangle \cdot \langle 2,1,3 \rangle = \frac{-1}{\sqrt{14}}(8+2+18) = \frac{-28}{\sqrt{14}}$$

Since $\frac{28}{\sqrt{14}} = \frac{2(14)\sqrt{14}}{\sqrt{14}\sqrt{14}} = 2\sqrt{14}$ we find $\boxed{(D_{\hat{u}} f)(2,1,3) = -2\sqrt{14}}$

Moreover, we extend the definition of critical point to the general case in the obvious way:

Definition 4.3.17.

We say $\vec{p}_o$ is a **critical point** of f if $(\nabla f)(\vec{p}_o)$ does not exist or $(\nabla f)(\vec{p}_o) = \vec{0}$.

The function in Example 4.3.11 the origin $(0,0,0)$ is the only critical point. On the other hand, the function in Example 4.3.12 has no critical point.

4.3.2 gradient vector fields in $\mathbb{R}^3$ and $\mathbb{R}^n$

We can calculate the gradient vector field for functions on $\mathbb{R}^n$ with $n \geq 1$ but, visualization is beyond most of us if $n > 3$. I mainly focus on the $n = 3$ case here and we see how the gradient aids our understanding of non-cartesian coordinate systems. Then we examine how the gradient vector field naturally provides a normal vector field to a level surface.

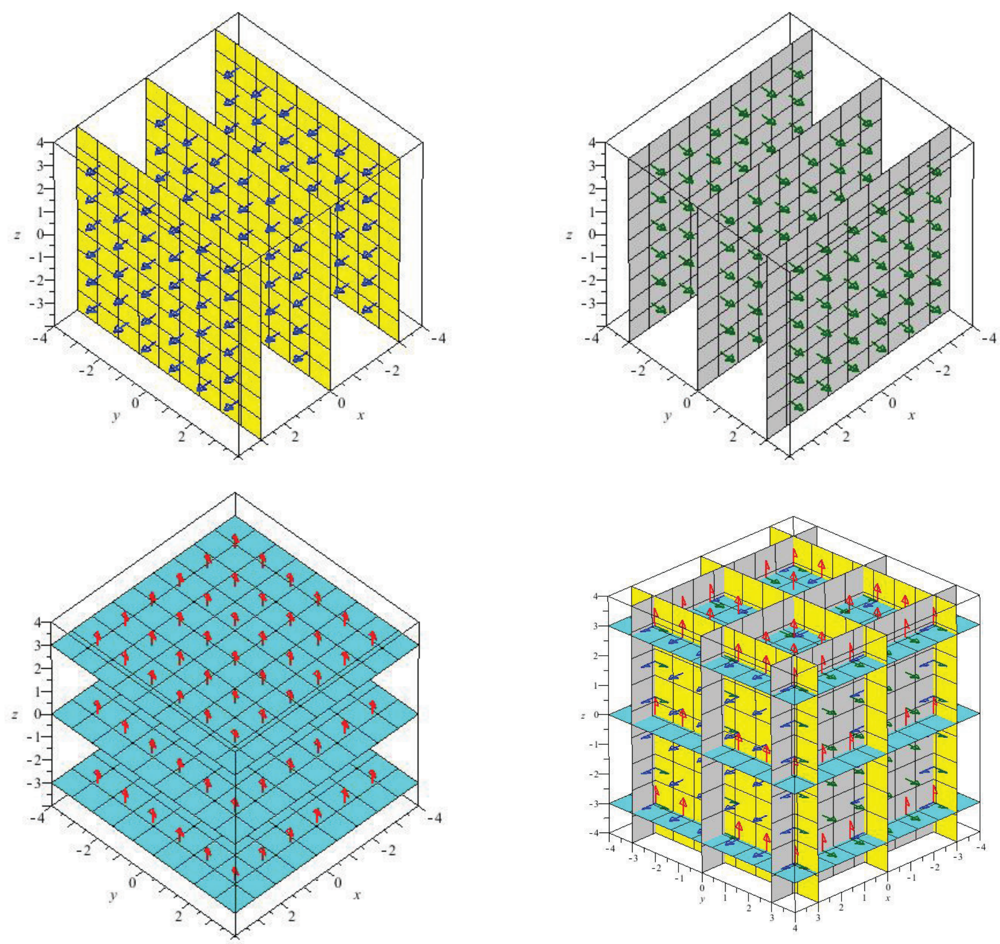
Definition 4.3.18.

If f is differentiable on $U \subseteq \mathbb{R}^n$ then ∇f defines the gradient vector field on U. We assign to each point $\vec{p} \in U$ the vector $\nabla f(\vec{p})$.

Example 4.3.19. *If x, y, z denote the coordinate functions on $\mathbb{R}^3$ then we find*

$$\nabla x = \langle 1, 0, 0, \rangle = \widehat{x},$$
$$\nabla y = \langle 0, 1, 0, \rangle = \widehat{y},$$
$$\nabla z = \langle 0, 0, 1, \rangle = \widehat{z}.$$

These define constant vector fields on $\mathbb{R}^3$.



Generally, the gradient vector fields of the coordinate functions of a non-cartesian coordinate system provide a vector fields which point in the direction of increasing coordinates. To obtain unit-vectors we simply normalize the gradient vector fields. In particular, if $y_1, y_2, \dots y_n$ are coordinates on $\mathbb{R}^n$ then there exist differentiable functions $f_1, f_2, \dots, f_n$ such that $y_j = f_j(x_1, x_2, \dots, x_n)$ for $j = 1, 2, \dots, n$. We can define:

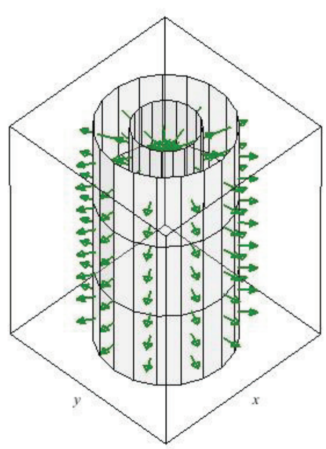
$$\widehat{y}_1 = \frac{\nabla f_1}{||\nabla f_1||} \quad \text{and} \quad \widehat{y}_2 = \frac{\nabla f_2}{||\nabla f_2||}, \ \dots, \quad \widehat{y}_n = \frac{\nabla f_n}{||\nabla f_n||}.$$

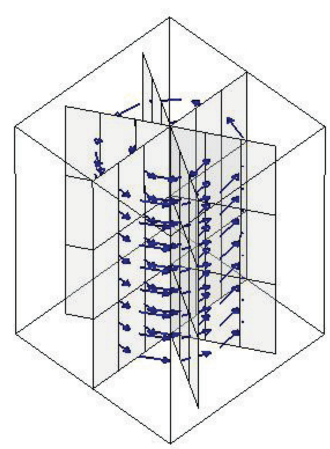
I mention this general idea for the interested reader. We are primarily interested in the cylindrical and spherical three dimensional coordinate systems. That's just a custom, we could easily extend these techniques to orthonormal coordinates based on ellipses or hyperbolas. If we are willing to give up on nice distance formulas we could even use coordinates based on tilted lines which meet at angles other than 90 degrees.

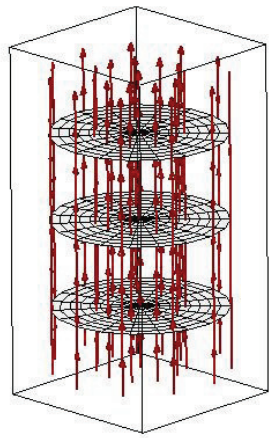
Example 4.3.20. *For cylindrical coordinates r, θ, z we can easily derive (following the same calculational steps as the polar two-dimensional case)*

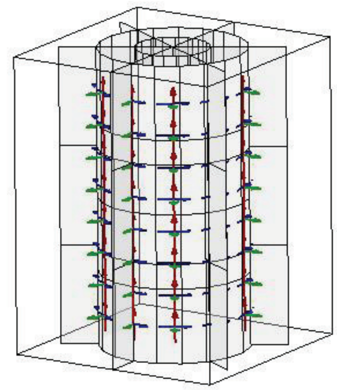
$$\begin{aligned} \widehat{r} &= \tfrac{1}{||\nabla r||}\nabla r = \widehat{r} = \langle \cos(\theta), \sin(\theta), 0\rangle \\ \widehat{\theta} &= \tfrac{1}{||\nabla \theta||}\nabla \theta = \langle -\sin(\theta), \cos(\theta), 0\rangle \\ \widehat{z} &= \tfrac{1}{||\nabla z||}\nabla z = \langle 0, 0, 1\rangle \end{aligned}$$

The difference between the calculations above and the polar coordinate case is that cylindrical coordinates are three dimensional and that means the gradient vector fields of the coordinate functions are three dimensional vector fields. I advocated a geometric derivation of these cylindrical unit vectors earlier in this course, but we now have computational method which requires almost no geometric intuition.









Example 4.3.21. *Suppose ρ, ϕ, θ denote spherical coordinates. Recall*[7]

$$\rho = \sqrt{x^2+y^2+z^2}, \quad \phi = \cos^{-1}\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right), \quad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$

You can calculate that

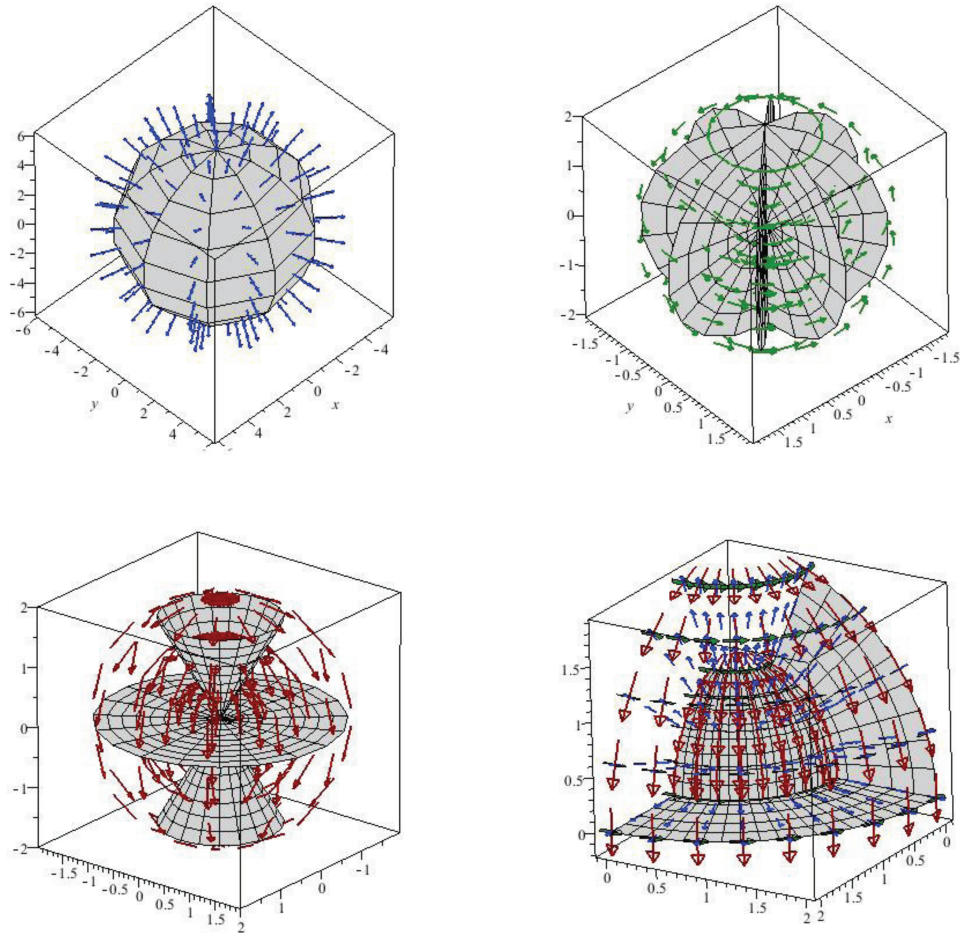
$$\begin{aligned} \widehat{\rho} &= \tfrac{1}{||\nabla\rho||}\nabla\rho = \sin(\phi)\cos(\theta)\,\widehat{x} + \sin(\phi)\sin(\theta)\,\widehat{y} + \cos(\phi)\,\widehat{z} \\ \widehat{\phi} &= \tfrac{1}{||\nabla\phi||}\nabla\phi = -\cos(\phi)\cos(\theta)\,\widehat{x} - \cos(\phi)\sin(\theta)\,\widehat{y} + \sin(\phi)\,\widehat{z} \\ \widehat{\theta} &= \tfrac{1}{||\nabla\theta||}\nabla\theta = -\sin(\theta)\,\widehat{x} + \cos(\theta)\,\widehat{y}. \end{aligned}$$

I'll walk you through the ρ calculation. To begin you can show that $\nabla\rho = \langle x/\rho,\ y/\rho,\ z/\rho\rangle$. But, we also know $x = \rho\cos\theta\sin\phi$, $y = \rho\sin\theta\sin\phi$ and $z = \rho\cos\phi$. Therefore,

$$\nabla\rho = \langle \cos\theta\sin\phi,\ \sin\theta\sin\phi,\ \cos\phi\rangle.$$

But, $||\nabla\rho|| = 1$. We derive that $\widehat{\rho} = \langle \cos\theta\sin\phi,\ \sin\theta\sin\phi,\ \cos\phi\rangle$. Perhaps I asked you to verify the formulas for $\widehat{\phi}, \widehat{\theta}$ in your homework. Making nice pictures of the spherical frame is an art I have yet to master... here's my best for now:

[7] these formulas only apply for certain octants, however, the ambiguity for the remaining octants only involves shifting the angular formulas by a constant. As you continue to read you'll notice that differentiation ultimately will kill any such constant so these formulas suffice.



Another nice application of the gradient involves level surfaces. Consider this: a level surface is the set of points which solves $f(x, y, z) = k$ for some value k. If we consider a point (x_o, y_o, z_o) on the level surface $f(x, y, z) = k$ then the gradient vector $(\nabla f)(x_o, y_o, z_o)$ will be perpendicular to the tangent plane of the level surface. Remember that when $\theta = \pi/2$ we find a direction in which $f(x, y, z)$ stays constant near (x_o, y_o, z_o). What does this mean? Let's summarize it:

The gradient vector field ∇f is normal to the level surface $f(x, y, z) = k$.

I use geometric intuition to make this claim here. We will offer a better proof later in this chapter. For now, let's try to appreciate the geometry.

Example 4.3.22. *Suppose $V(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$ represents the voltage due to a point-charge at the origin. Electrostatics states that the electric field $\vec{E} = -\nabla V$. Geometrically this has a simple meaning; the electric field points along the normal direction to the level-surfaces of the voltage function[8]. In other words, the electric field vectors are normal to the equipotential surfaces where they are attached. What is an "equipotential surface", it's a surface on which the voltage assumes a constant value. This is nothing more than a level-surface of the voltage function. For the given potential function, using $\rho = \sqrt{x^2+y^2+z^2}$,*

$$\begin{aligned}
\nabla V &= \langle \partial_x(1/\rho), \partial_y(1/\rho), \partial_z(1/\rho)\rangle \\
&= \langle (-1/\rho^2)\partial_x \rho,\ (-1/\rho^2)\partial_y \rho,\ (-1/\rho^2)\partial_z \rho\rangle \\
&= \frac{-1}{\rho^2}\langle \partial_x \rho,\ \partial_y \rho,\ \partial_z \rho\rangle \\
&= -\frac{1}{\rho^2}\widehat{\rho}.
\end{aligned}$$

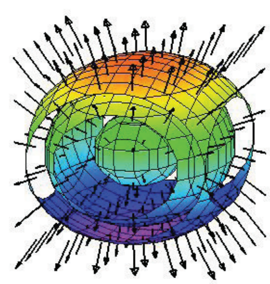
Equipotentials $V = V_o = 1/\rho$ are simply spheres $\rho = 1/V_o$ and the electric field is a purely radial field $\vec{E} = \frac{1}{r^2}\widehat{\rho}$.

Example 4.3.23. *Consider the ellipsoid $f(x,y,z) = x^2/a^2 + y^2/b^2 + z^2/c^2 = k$. At any point on the ellipse the vector field*

$$\nabla f = \frac{2x}{a^2}\,\widehat{x} + \frac{2y}{b^2}\,\widehat{y} + \frac{2z}{c^2}\,\widehat{z}$$

points in the normal direction to the ellipsoid.

It amazes me how easy it is to find a formula to assign a normal-vector to an arbitrary point on an ellipse. Imagine solving that problem without calculus.



[8]The voltage function is the electric potential or simply the potential function in this context

4.4 the general derivative

Thus far we have primarily discussed partial derivatives in their connection to the rate of change of a given function in a particular direction. However, we would like to characterize the change in the function as a whole. Moreover, even in the one-dimensional case the derivative was closely tied to the best linear approximation to the function. In the single variable case it is as simple as this: the best linear approximation to a differentiable function at a point is the linearization of the function at that point whose graph is the tangent line. The slope of the tangent line is the value of the derivative function at the point. How do these ideas generalize? I take an n-dimensional approach in the beginning of this section because little is gained by talking in lower dimensions for the basic definitions.

Definition 4.4.1.

Suppose that U is open and $\vec{F} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a mapping the we say that $\vec{F}$ is **differentiable** at $\vec{a} \in U$ iff there exists a linear mapping $\vec{L} : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\vec{h}\to 0} \frac{\vec{F}(\vec{a}+\vec{h}) - \vec{F}(\vec{a}) - \vec{L}(\vec{h})}{||\vec{h}||} = 0.$$

In such a case we call the linear mapping $\vec{L}$ the **differential at** $\vec{a}$ and we denote $\vec{L} = d\vec{F}_{\vec{a}}$. The matrix of the differential is called the **derivative of** $\vec{F}$ **at** $\vec{a}$ and we denote $[d\vec{F}_{\vec{a}}] = \vec{F}'(\vec{a}) \in \mathbb{R}^{\ m\times n}$ which means that $d\vec{F}_{\vec{a}}(\vec{v}) = \vec{F}'(\vec{a})\vec{v}$ for all $\vec{v} \in \mathbb{R}^n$.

4.4.1 matrix of the derivative

If we know a function is differentiable at a point then we can calculate the formula for $\vec{L}$ in terms of partial derivatives. In particular, if $\vec{F} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{a} \in U$ then the differential $d\vec{F}_{\vec{a}}$ has the derivative matrix $\vec{F}'(\vec{a})$ which has components expressed in terms of partial derivatives of the component functions:

$$\left[\ d\vec{F}_{\vec{a}}\ \right]_{ij} = \partial_j F_i = \frac{\partial F_i}{\partial x_j}(\,\vec{a}\,)$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$. This result is proved in advanced calculus. Let me expand this claim in detail for a few common cases: in each case we note $\vec{L}(\vec{a}+\vec{h}) \approx \vec{F}(a) + \vec{F}'(\vec{a})\vec{h}$

1. **function on** $\mathbb{R}$, $f : \mathbb{R} \to \mathbb{R}$, $\boxed{L(a+h) \approx f(a) + f'(a)h}$ the derivative matrix is just the derivative $f'(a)$ at the point.

2. **path into** $\mathbb{R}^n$, $\vec{r} : \mathbb{R} \to \mathbb{R}^n$, $\boxed{\vec{r}(a+h) \approx \vec{r}(a) + \vec{r}\,'(a)h}$. The derivative matrix is just the velocity vector $\vec{r}\,'(a)$ viewed as an $n \times 1$ matrix (it's a column vector).

3. **multivariate real-valued function**, $f : \mathbb{R}^n \to \mathbb{R}$, $\boxed{f(\vec{a}+\vec{h}) \approx f(\vec{a}) + (\nabla f)(\vec{a})\vec{h}}$. The derivative matrix is just the gradient vector $(\nabla f)(\vec{a})$ viewed as an $1 \times n$ matrix (it's a row vector).

4. **coordinate change mapping**, $\vec{T} : \mathbb{R}^3 \to \mathbb{R}^3$, $\boxed{\vec{T}(\vec{a}+\vec{h}) \approx \vec{T}(\vec{a}) + \vec{T}'(\vec{a})\vec{h}}$. The derivative matrix is a 3×3 matrix. In particular, if we denote $\vec{T} = \langle x, y, z\rangle$ and use u, v, w for cartesian coordinates in the domain of $\vec{T}$

$$\vec{T}'(\vec{a}) = [\partial_u \vec{T} | \partial_v \vec{T} | \partial_w \vec{T}] = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix}$$

For two-dimensional coordinate change, $\vec{T} : \mathbb{R}^2 \to \mathbb{R}^2$ we again write $\boxed{\vec{T}(\vec{a}+\vec{h}) \approx \vec{T}(\vec{a}) + \vec{T}'(\vec{a})\vec{h}}$ but the matrix $\vec{T}'(\vec{a})$ is just a 2×2 matrix

$$\vec{T}'(\vec{a}) = [\partial_u \vec{T} | \partial_v \vec{T}] = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$

Example 4.4.2. *Let $f(x) = \sqrt{x}$. The linearization at $x = 4$ is given by $L(x) = 2 + \frac{1}{4}(x-4)$ since $f'(4) = \frac{1}{2\sqrt{4}} = \frac{1}{4}$. We could also express L by $L(4+h) = 2 + h/4$. As an application, note the approximation $\sqrt{5} \approx 2 + 1/4 = 2.25$.*

Example 4.4.3. *Let $\vec{r}(t) = \langle t, t^2, sin(10t)\rangle$ for $t \in [0, 2]$. The linearization of $\vec{r}$ at $t = 1$ is given by $\vec{L}(1+h) = \vec{r}(1) + h\vec{r}\,'(1)$. In particular,*

$$\vec{L}(1+h) = \langle 1+h, 1+2h, sin(10) + 10hcos(10)\rangle.$$

Example 4.4.4. .

Find linearization of $f(x,y) = x/y$ at $(6,3)$.

Notice $\frac{\partial f}{\partial x} = 1/y$ and $\frac{\partial f}{\partial y} = -x/y^2$. These are continuous at $(6,3)$ so $f(x,y)$ is differentiable at $(6,3)$.

$$\begin{aligned} L(x,y) &= f(6,3) + \frac{\partial f}{\partial x}\Big|_{(6,3)}(x-6) + \frac{\partial f}{\partial y}\Big|_{(6,3)}(y-3) \\ &= 6/3 + \frac{1}{3}(x-6) - \frac{6}{9}(y-3) \\ &= \boxed{\frac{1}{3}x - \frac{2}{3}y + 2 = L(x,y)} \end{aligned}$$

Example 4.4.5. .

$\boxed{E72}$ Find the linearization of $f(x,y) = x^2+y^2$ at $(1,2)$. Then approximate $f(2,2)$ and compare to the real-value. We found the tangent plane's eq$^{\underline{n}}$ in $\boxed{E71}$ so we already know

$$\boxed{L(x,y) = 5 + 2(x-1) + 4(y-2)}$$

We approximate f via L,

$$f(2,2) \cong L(2,2) = 5 + 2(2-1) + 4(0) = 7$$

Of course we can just evaluate $f(2,2) = 2^2+2^2 = 8$ to see we have an absolute error of $8-7 = 1$.

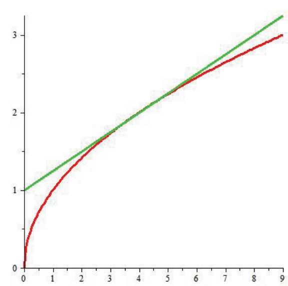
4.4.2 tangent space as graph of linearization

In the section after this I wrestle with why these are good definitions. For now I'll state them without justification.

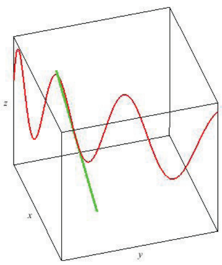
1. $f : \mathbb{R} \to \mathbb{R}$ has tangent line at $(a, f(a))$ with equation $y = f(a) + f'(a)(x-a)$.
2. $\vec{r} : \mathbb{R} \to \mathbb{R}^n$ has tangent line at $\vec{r}(a)$ with natural parametrization $\vec{l}(h) = \vec{r}(a) + \vec{r}\,'(a)h$.
3. $f : \mathbb{R}^2 \to \mathbb{R}$ has tangent plane at $(a, b, f(a,b)$ with equation $z = f(a,b) + (\nabla f)(a,b) \bullet \langle x - a, y - b\rangle$.

These are the cases of interest, in case 2 we usually deal with $n = 2$ or $n = 3$ in this course. The following triple of examples mirror those given in the last section. The overall theme is simple: the tangent space to a graph of a function is the graph of the linearization of that function. There are several other viewpoints on the tangent space of a surface and we devote an entire section to that a little later in this chapter. Here I just want you to get what we mean when we say a derivative *gives the best linear approximation to a function.*

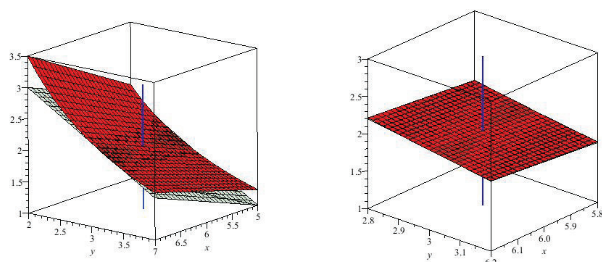
Example 4.4.6. *We continue Example 4.4.2,* $f(x) = \sqrt{x}$ *and the linearization at* $x = 4$ *is given by* $L(x) = 2 + \frac{1}{4}(x-4)$. *The tangent line is the graph* $y = L(x)$ *which is in green, whereas the* $y = f(x)$ *is in red.*



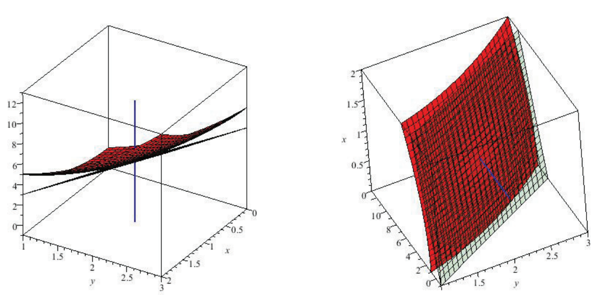
Example 4.4.7. *We continue Example 4.4.3, $\vec{r}(t) = \langle t, t^2, sin(10t)\rangle$ for $t \in [0,2]$ and the linearization of $\vec{r}$ at $t = 1$ is given by $\vec{L}(1+h) = \langle 1+h, 1+2h, sin(10) + 10hcos(10)\rangle$. Once more we plot the curve in red and the tangent line parametrized by $\vec{L}$ in green:*



Example 4.4.8. *Continue Example 4.4.4, $f(x,y) = x/y$ and the tangent plane to $z = x/y$ at $(6,3)$ is the solution set of $z = x/3 - 2y/3 + 2$. Below I illustrate the tangent plane, the blue line goes through the point of tangency. See how the surface is locally flat, note the right picture is zoomed further in towards the point of tangency.*



Example 4.4.9. *Continue Example 4.4.5, $f(x,y) = x^2 + y^2$ and the tangent plane to $z = x^2 + y^2$ at $(1,2)$ is the solution set of $z = 5 + (x-1) + 4(y-2)$. Below I illustrate the tangent plane, the blue line goes through the point of tangency. See how the surface is locally flat, these are just two views of the same scale, I put a rotating animation of this on the webpage, take a look.*



4.4.3 existence and connections to directional differentiation

Existence is usually more troublesome than calculation. But, that is no reason to ignore it. In this subsection I attempt to give you a better sense of what it means for a function to be differentiable at a point. Geometrically we eventually come to the simple realization that a function is differentiable iff it is well-approximated by its linearization. This in turn is tied to the proper definition of the tangent plane. We already gave formulas for important cases in the last subsection, my goal here is to explain why we use those definitions and not something else. Before we get to those more subtle topics, I begin by demonstrating the general derivative recovers single-variable differentiation:

Example 4.4.10. *Suppose $f : dom(f) \subseteq \mathbb{R} \to \mathbb{R}$ is differentiable at x. It follows that there exists a linear function $df_x : \mathbb{R} \to \mathbb{R}$ such that*[9]

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

Since $df_x : \mathbb{R} \to \mathbb{R}$ is linear there exists a constant matrix m such that $df_x(h) = mh$. In this silly case the matrix m is a 1×1 matrix which otherwise known as a real number. Note that

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - df_x(h)}{|h|} = 0 \quad \Leftrightarrow \quad \lim_{h \to 0^{\pm}} \frac{f(x+h) - f(x) - df_x(h)}{|h|} = 0.$$

In the left limit $h \to 0^-$ we have $h < 0$ hence $|h| = -h$. On the other hand, in the right limit $h \to 0^+$ we have $h > 0$ hence $|h| = h$. Thus, differentiability suggests that $\lim_{h \to 0^{\pm}} \frac{f(x+h)-f(x)-df_x(h)}{\pm h} = 0$. But we can pull the minus out of the left limit to obtain $\lim_{h \to 0^-} \frac{f(x+h)-f(x)-df_x(h)}{h} = 0$. Therefore,

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - df_x(h)}{h} = 0.$$

We seek to show that $\lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = m$.

$$m = \lim_{h \to 0} \frac{mh}{h} = \lim_{h \to 0} \frac{df_x(h)}{h}$$

A theorem from calculus I states that if $\lim(f - g) = 0$ and $\lim(g)$ exists then so must $\lim(f)$ and $\lim(f) = \lim(g)$. Apply that theorem to the fact we know $\lim_{h \to 0} \frac{df_x(h)}{h}$ exists and

$$\lim_{h \to 0} \left[\frac{f(x+h) - f(x)}{h} - \frac{df_x(h)}{h} \right] = 0.$$

It follows that

$$\lim_{h \to 0} \frac{df_x(h)}{h} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Consequently,

$$df_x(h) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \quad \textit{defined } f'(x) \textit{ in calc. I.}$$

[9] unless we state otherwise, $\mathbb{R}^n$ is assumed to have the euclidean norm, in this case $||x||_{\mathbb{R}} = \sqrt{x^2} = |x|$.

Therefore, $\boxed{df_x(h) = f'(x)h}$. *In other words, if a function is differentiable in the sense we defined at the beginning of this section then it is differentiable in the terminology we used in calculus I. Moreover, the derivative at x is precisely the matrix of the differential. If we use the notation $y = f(x)$ and $h = dx$ then we recover formula for the differential often taught in first semester calculus:*

$$dy_x(dx) = \frac{dy}{dx}(x)dx$$

Or, more compactly, $dy = \frac{dy}{dx}dx$ where dy is the change in y corresponding to the change dx in x. These seemingly heuristic statements take a rigorous meaning in the boxed equation above.

Of course, what really makes the general derivative interesting is its ability to tackle problems such as given below:

Example 4.4.11. *Suppose $F : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by $F(x, y) = (xy,\ x^2,\ x+3y)$ for all $(x, y) \in \mathbb{R}^2$. Consider the difference function $\triangle F$ at (x, y):*

$$\triangle F = F((x, y) + (h, k)) - F(x, y) = F(x + h, y + k) - F(x, y)$$

Calculate,

$$\triangle F = \big((x + h)(y + k),\ (x + h)^2,\ x + h + 3(y + k)\big) - \big(xy,\ x^2,\ x + 3y\big)$$

Simplify by cancelling terms which cancel with $F(x, y)$:

$$\triangle F = \big(xk + hy,\ 2xh + h^2,\ h + 3k)\big)$$

Identify the linear part of $\triangle F$ as a good candidate for the differential. I claim that:

$$L(h, k) = \big(xk + hy,\ 2xh,\ h + 3k\big).$$

is the differential for f at (x,y). Observe first that we can write

$$L(h, k) = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

therefore $L : \mathbb{R}^2 \to \mathbb{R}^3$ is manifestly linear. Use the algebra above to simplify the difference quotient below:

$$\lim_{(h,k)\to(0,0)} \left[\frac{\triangle F - L(h, k)}{||(h, k)||} \right] = \lim_{(h,k)\to(0,0)} \left[\frac{(0, h^2, 0)}{||(h, k)||} \right]$$

Note $||(h, k)|| = \sqrt{h^2 + k^2}$ therefore we fact the task of showing that $(0,\ h^2/\sqrt{h^2 + k^2},\ 0) \to (0, 0, 0)$ as $(h, k) \to (0, 0)$. Recall from our study of limits that we can prove the vector tends to $(0, 0, 0)$ by showing the each component tends to zero. The first and third components are obviously zero however the second component requires study. Observe that

$$0 \le \frac{h^2}{\sqrt{h^2 + k^2}} \le \frac{h^2}{\sqrt{h^2}} = |h|$$

Clearly $\lim_{(h,k)\to(0,0)}(0) = 0$ *and* $\lim_{(h,k)\to(0,0)} |h| = 0$ *hence the squeeze theorem for multivariate limits shows that* $\lim_{(h,k)\to(0,0)} \frac{h^2}{\sqrt{h^2+k^2}} = 0$. *Therefore,*

$$dF_{(x,y)}(h,k) = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

Fortunately we can usually avoid explicit limit calculations due to the nice proposition below.

Example 4.4.12. *Again consider* $F(x,y) = (xy,\ x^2,\ x+3y)$. *Identify* $F_1(x,y) = xy$, $F_2(x,y) = x^2$ *and* $F_3(x,y) = x + 3y$. *Calculate,*

$$[F'(x,y)] = \begin{bmatrix} \partial_x F_1 & \partial_y F_1 \\ \partial_x F_2 & \partial_y F_2 \\ \partial_x F_3 & \partial_y F_3 \end{bmatrix} = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix}$$

In single-variable calculus we learn that differentiability implies continuity. However, continuity does not imply differentiability at a given point. The same is true for multivariate functions.

Proposition 4.4.13.

If $\vec{F} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{a} \in U$ then $\vec{F}$ is continuous at $\vec{a}$.

The proof is given in advanced calculus. It's not too difficult. $\square$

The general derivative also reproduces all the directional derivatives we previously discussed.

Proposition 4.4.14.

If $\vec{F} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{a} \in U$ then the directional derivative $D_{\vec{v}}\vec{F}(\vec{a})$ exists for each $\vec{v} \in \mathbb{R}^n$ and $D_{\vec{v}}\vec{F}(\vec{a}) = d\vec{F}_{\vec{a}}(\vec{v})$.

The proof is given in advanced calculus. It's not terribly difficult. $\square$

We should consider the example below. It may challenge some of your misconceptions. It shows that directional differentiation at a point does not give us enough to build the derivative. In fact, the example below has **all** directional derivatives and yet the function is not even continuous.

Example 4.4.15. *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be defined by* $f(x,y) = \frac{x^2 y}{x^4+y^2}$ *for* $(x,y) \neq (0,0)$ *and* $f(0,0) = 0$. *We proved in Example 3.3.4 that this function is not continuous at* $(0,0)$. *Given the proposition above we also may infer the function is not differentiable at* $(0,0)$. *You might expect this indicates at least some directional derivative fails to exist. Let's investigate. We turn to the problem of* *calculating the directional derivative of this function in the unit-vector $\langle a, b\rangle$ direction, suppose $b \neq 0$ to begin,*

$$D_{\langle a,b\rangle} f(0,0) = \frac{d}{dt}\Big[f(at,bt)\Big]\Big|_{t=0} = \frac{d}{dt}\left[\frac{a^2bt^3}{a^4t^4+b^2t^2}\right]\Big|_{t=0} = \left[\frac{a^2b(a^4t^2+b^2) - a^2bt(2ta^4)}{(a^4t^2+b^2)^2}\right]\Big|_{t=0} = \frac{a^2}{b}.$$

On the other hand, if $b = 0$ then we know $a \neq 0$ since $\langle a, b\rangle$ is a unit-vector[10] *hence $f(at, bt) = \frac{a^2bt^3}{a^4t^4+b^2t^2} = 0$ and it follows $D_{\langle a,0\rangle} f(0,0) = 0$. We find the directional derivatives of f exist in all directions.*

Notice that the directional derivatives do jump from one value to another as we travel around the unit-circle. In particular, as we traverse the arc of the circle through the point $\langle 1, 0\rangle$ we have $\langle a, b\rangle$ go from vectors with $b > 0$ which have $\frac{a^2}{b} \to \infty$ to vectors with $b < 0$ which have $\frac{a^2}{b} \to -\infty$. In the middle, we hit $\langle 1, 0\rangle$ where $D_{\langle a,0\rangle} f(0,0) = 0$. These directional derivatives may exist but they certainly do not continuously paste together. It turns out that continuity of the directional derivatives in the coordinate directions is a sufficient condition to eliminate the trouble of the previous example.

Definition 4.4.16.

A mapping $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is **continuously differentiable** at $a \in U$ iff all the partial derivative mappings $\partial F_i/\partial x_j$ exist on an open set containing a and are continuous at a.

Continuous differentiability is typically easier than differentiability to check. The reason is that partial derivatives are straightforward to calculate. On the other hand, it is sometimes challenging to find the linearization and actually check the appropriate limit vanishes. It follows that the proposition below is welcome news:

Proposition 4.4.17.

If F is continuously differentiable at a then F is differentiable at a

The proof is somewhat involved. The main construction involves breaking a vector into a sum of vector components. Then continuity of the partial derivatives paired with a mean value theorem argument goes to prove the differentiability of the mapping. Again, details are given in my advanced calculus notes (or any good text on the subject). $\square$

There do exist functions which are differentiable at a point and yet fail to be continuously differentiable at that point. In single variable calculus I usually present the example: Let $f(0) = 0$ and

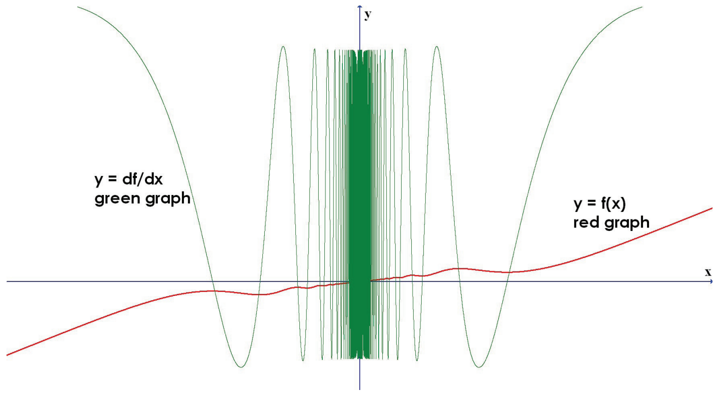
$$f(x) = \frac{x}{2} + x^2 \sin\frac{1}{x}$$

for all $x \neq 0$. I can be shown that the derivative $f'(0) = 1/2$. Moreover, we can show that $f'(x)$ exists for all $x \neq 0$, we can calculate:

$$f'(x) = \frac{1}{2} + 2x\sin\frac{1}{x} - \cos\frac{1}{x}$$

[10] if $a = 0$ and $b = 0$ then $||\langle a, b\rangle|| = 0 \neq 1$

Notice that $dom(f') = \mathbb{R}$. Note then that the tangent line at $(0,0)$ is $y = x/2$.



The lack of continuity for the derivative means that the tangent line at the origin does not well-approximate the graph near the point of tangency. In other words, the linearization is not a good approximation near the point of tangency. This is not just a single-variable phenonmenon. Pathological multivariate examples exist. For example,

Example 4.4.18. *Let $f(0, y) = 0$ and*

$$f(x, y) = x^2 \sin \frac{1}{x}$$

for all $(x, y) \in \mathbb{R}^2$ such that $x \neq 0$. You can show that $D_{\widehat{u}}f(0,0) = 0$ for all unit vectors u. This means that the tangent vectors to any path $t \to (at, bt, f(at, bt))$ reside in the xy-plane. It appears the set of all tangent vectors fill out the xy-plane. However, I'm not sure what happens with non-linear paths in the domain. I suspect the curves on the graph $z = f(x, y)$ built from composing a smooth, but non-linear, path $\gamma : \mathbb{R} \to \mathbb{R}^2$ with f might result in a path $f \circ \gamma$ which is not even differentiable at the origin.

Let's investigate the differentiability of f at $(0,0)$. Given the triviality of all the directional derivatives we suspect $L(h, k) = 0$. Consider,

$$\frac{|f(h,k) - f(0,0) - L(h,k)|}{||(h,k)||} = \frac{|h^2 \sin(1/h)|}{\sqrt{h^2 + k^2}} = \frac{|h \sin(1/h)|}{\sqrt{1 + k^2/h^2}} \leq |h \sin(1/h)| \leq |h|.$$

It follows that f is differentiable at $(0,0)$ since we have $|h| \to 0$ as $(h,k) \to (0,0)$ along any path. Therefore, my suspicion was incorrect. Even nonlinear paths composed with f yield a differentiable path. However, this does give us another example of a function which is differentiable at $(0,0)$ but is not continuously differentiable. If you're wondering it is clear that f_x is not continuous along the entire y-axis. Given our experience in the single variable case we suspect the linearization does not approximate the function in a natural way as we leave the point of tangency. We need the continuity of the partial derivatives to insure the function does not wildly misbehave in the locality of the tangent point.

I haven't proved it yet but I suspect the function below is not differentiable. It gives and example of a function which is continuous but is not differentiable at zero. However, both partial derivatives exist at $(0,0)$, they're just not continuous.

Example 4.4.19. *Let us define $f(0,0) = 0$ and*

$$f(x,y) = \frac{x^2 y}{x^2 + y^2}$$

for all $(x,y) \neq (0,0)$ in $\mathbb{R}^2$. It can be shown[11] *that f is continuous at $(0,0)$. Moreover, since $f(x,0) = f(0,y) = 0$ for all x and all y it follows that f vanishes identically along the coordinate axis. Thus the rate of change in the $\widehat{x}$ or $\widehat{y}$ directions is zero. We can calculate that*

$$\frac{\partial f}{\partial x} = \frac{2xy^3}{(x^2+y^2)^2} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = \frac{x^4 - x^2 y^2}{(x^2+y^2)^2}$$

Consider the path to the origin $t \mapsto (t,t)$ gives $f_x(t,t) = 2t^4/(t^2+t^2)^2 = 1/2$ hence $f_x(x,y) \to 1/2$ along the path $t \mapsto (t,t)$, but $f_x(0,0) = 0$ hence the partial derivative f_x is not continuous at $(0,0)$. Therefore, this function has discontinuous partial derivatives. It is not continuously differentiable.

Let's return to the question of directional derivatives and differentiability. It is tempting to think that the reason the function in Example 4.4.15 failed to be differentiable is that the tangent vectors to the curves $t \mapsto (at, bt, f(at, bt))$ failed to fill out a plane. This suspicion is further encouraged by Example 4.4.18 where we see the function is differentiable and the tangent vectors to the curves $t \mapsto (at, bt, f(at, bt))$ do fill out the xy-plane. However, this suspicion is false. Think back to our experience with multivariate limits in Example 3.3.2. Differentiability also concerns a multivariate limit so intuitively we may expect something could be hidden if we only think about straight-line approaches to the limit point. I suspect that if we had that the tangents to $t \mapsto (\vec{r}(t), f(\vec{r}(t)))$ fill out a plane for all differentiable paths $\vec{r}$ with $\vec{r}(0) = \langle 0, 0 \rangle$ then it would follow f is differentiable. I don't have a proof of this claim in the notes at the present time.

Why all this fuss? Let me try to clarify the confusion which pushed me to this discussion:

1. some authors define the tangent plane to be the union of all tangent vectors at a point.

2. other authors say the tangent plane is a plane which well-approximates the graph of the function near the point of tangency.

Item (2.) begs some questions, what exactly do we mean by "well-approximates". Is the nearness to the graph the concept captured by mere differentiability or is it the stronger version captured by continuous differentiability? Item (1.) is dangerous since it would *seem* that looking at all possible directional derivatives should give a complete picture of the tangent vectors at a point. We just

[11] you did this one in homework... or at least you were supposed to...

argued this is not the case[12]. It is possible for all tangents to curves built from linear paths to exist whereas the tangent vectors to a path built from a nonlinear path may not even exist. If we are to use item (1.) as a definition we must clarify it a bit:

> **The tangent plane to the graph $z = f(x, y)$ is formed by the union of all possible tangent vectors of curves $f \circ \vec{\gamma}$ where $\vec{\gamma}$ is a smooth curve in $dom(f)$ which pass at $t = 0$ through the xy-coordinates of the point of tangency. If there exists a smooth curve $\vec{\gamma}$ such that $f \circ \vec{\gamma}$ is not differentiable at $t = 0$ then the tangent plane fails to exist.**

This is just my comment here, I haven't seen this elsewhere. Most authors don't bother with these details or deliberations. In fact, many authors assume continuous differentiability in their definitions. In any event, it seems clear to me that we should prefer a slightly more careful version of (2.) since it has far less technical trouble. With all of this in mind we define (I expand on the most important case to this course after this general definition),

Definition 4.4.20. *general tangent space to a graph.*

Suppose that U is open and $\vec{F} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a mapping which is differentiable at $\vec{a} \in U$ then the linear mapping $\vec{L} : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\vec{h} \to 0} \frac{\vec{F}(\vec{a} + \vec{h}) - \vec{F}(\vec{a}) - \vec{L}(\vec{h})}{||\vec{h}||} = 0.$$

defines the **tangent space** at $(\vec{a}, \vec{F}(\vec{a}))$ to $graph(\vec{F}) = \{(\vec{x}, \vec{F}(\vec{x})) \mid dom(\vec{F})\}$ with equations $\vec{z} = \vec{F}(\vec{a}) + \vec{L}(\vec{x} - \vec{a})$ in $\mathbb{R}^n \times \mathbb{R}^m$. We use the notation $\vec{z} \in \mathbb{R}^m$ whereas $\vec{x}, \vec{a} \in \mathbb{R}^n$ in the equation above.

In particular, for $f : \mathbb{R}^2 \to \mathbb{R}$ we have $L(x - x_o, y - y_o) = (\nabla f)(x_o, y_o) \bullet \langle x - x_o, y - y_o \rangle$ and the tangent plane has equation:

$$\boxed{z = f(x_o, y_o) + (x - x_o) f_x(x_o, y_o) + (y - y_o) f_y(x_o, y_o).}$$

The assumption of differentiability of f at (x_o, y_o) insures that the tangent plane $z = f(x_o, y_o) + L(x, y) \approx f(x, y)$ for points near (x_o, y_o). In other words, the graph $z = f(x, y)$ looks like a plane if we zoom in close to the point $(x_o, y_o, f(x_o, y_o))$. In fact, many authors simply define differentiability in view of this concept:

A function is differentiable at $\vec{p}$ iff it has a tangent plane at $\vec{p}$.

This is less than satisfactory if the text you're reading nowhere defines the tangent plane. I won't name names. The boxed statement is true, but it is not a definition. Not here at least.

[12] I have an example if you ask

In the case a function is differentiable but not continuously differentiable we have the situation that there is a tangent plane, but it fails to well-approximate the graph near the point of tangency.

Continuous differentiability is needed for many of the calculations we perform in the remainder of this course. I conclude this section with an example of how it may happen that $f_{xy} \neq f_{yx}$ at a point which is merely differentiable. On the other hand, Clairaut's Theorem states that $f_{xy} = f_{yx}$ for continuously differentiable functions.

Example 4.4.21. .

A CURIOUS EXAMPLE: WHY $f_{xy} \neq f_{yx}$ ALWAYS.

$$f(x,y) = \begin{cases} (x^3y - xy^3)/(x^2+y^2) & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

When $(x,y) \neq (0,0)$ its a simple matter to differentiate,

$$f_x = \frac{(3x^2y - y^3)(x^2+y^2) - 2x(x^3y - xy^3)}{(x^2+y^2)^2} = \frac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}$$

$$f_y = \frac{x^5 - 4y^2x^3 - y^4x}{(x^2+y^2)^2}$$

$$f_{xy} = \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2+y^2)^3} = f_{yx}(x,y) \quad \text{for } (x,y) \neq 0.$$

At the origin we need to use the def^{n} of partial differentiation,

$$f_x(0,0) = \lim_{h\to 0}\left[\frac{f(h,0) - f(0,0)}{h}\right] = \lim_{h\to 0}\left[\frac{0-0}{h}\right] = 0.$$

$$f_y(0,0) = \lim_{h\to 0}\left[\frac{f(0,h) - f(0,0)}{h}\right] = \lim_{h\to 0}\left[\frac{0-0}{h}\right] = 0.$$

$$f_{xy}(0,0) \equiv \frac{\partial f_x}{\partial y}(0,0) = \lim_{h\to 0}\left[\frac{f_x(0,h) - f_x(0,0)}{h}\right] = \lim_{h\to 0}\left[\frac{-h^5/(h^2)^2 - 0}{h}\right] = -1.$$

$$f_{yx}(0,0) \equiv \frac{\partial f_y}{\partial x}(0,0) = \lim_{h\to 0}\left[\frac{f_y(h,0) - f_y(0,0)}{h}\right] = \lim_{h\to 0}\left[\frac{h^5/h^4 - 0}{h}\right] = 1.$$

Therefore $f_{xy} \neq f_{yx}$ since at $(0,0)$ they disagree. You might object that this to picky on our part, well sorry its math. The trouble here is that f_{xy} is not continuous at $(0,0)$, everywhere else it is and in all those places $f_{xy}(x,y) = f_{yx}(x,y)$ $\forall (x,y) \neq (0,0)$.

Theorem 4.4.22. *Clairaut's Theorem:*

If $f : dom(f) \subseteq \mathbb{R}^2 \to \mathbb{R}$ is a function where $dom(f)$ contains an open disk D centered at (a,b) and the function f_{xy} and f_{yx} are both continuous on D then

$$\boxed{f_{xy}(a,b) = f_{yx}(a,b).}$$

The proof is found in most advanced calculus texts. Finally, I should mention that the concerns and examples of this section readily generalize to functions from $\mathbb{R}^m$ to $\mathbb{R}^n$.

4.4.4 properties of the derivative

Suppose $\vec{F}_1 : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $\vec{F}_2 : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ are differentiable at $\vec{a} \in U$ then $\vec{F}_1 + \vec{F}_2$ is differentiable at $\vec{a}$ and $d(\vec{F}_1 + \vec{F}_2)_a = (d\vec{F}_1)_a + (d\vec{F}_2)_a$ which means for the Jacobian matrices we also have $(\vec{F}_1 + \vec{F}_2)'(\vec{a}) = \vec{F}_1'(\vec{a}) + \vec{F}_2'(\vec{a})$. Likewise, if $c \in \mathbb{R}$ then $d(c\vec{F}_1)_a = c(d\vec{F}_1)_a$ hence for the Jacobian matrices we have $(c\vec{F}_1)'(\vec{a}) = c(\vec{F}_1'(\vec{a})$. Nothing terribly surprising here. What is much more fascinating is the following general version of the chain rule:

Proposition 4.4.23.

If $\vec{F} : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at $\vec{a}$ and $\vec{G} : V \subseteq \mathbb{R}^p \to \mathbb{R}^m$ is differentiable at $\vec{F}(\vec{a}) \in V$ then $\vec{G} \circ \vec{F}$ is differentiable at $\vec{a}$ and $d(\vec{G} \circ \vec{F})_{\vec{a}} = (d\vec{G})_{\vec{F}(\vec{a})} \circ d\vec{F}_{\vec{a}}$. Moreover, in Jacobian matrix notation,

$$(\vec{G} \circ \vec{F})'(\vec{a}) = \vec{G}'(\vec{F}(\vec{a}))\vec{F}'(\vec{a}).$$

In words, the Jacobian matrix of the composite of $\vec{G}$ with $\vec{F}$ is simply the matrix product of the Jacobian matrices of $\vec{G}$ with the Jacobian matrix of $\vec{F}$. Unfortunately, not all students really learned matrix algebra in highschool so this statement lacks the power it should have in your mind. This proposition builds the foundation for the multivariate version of u-substution. All the chain rules in the next section are derivable from this general proposition. For this reason I offer no proofs in the next section. The calculations in the next section all follow from the calculation below[13]:

Proof: $\approx$ Suppose $\vec{F} : dom(\vec{F}) \subseteq \mathbb{R}^n \to \mathbb{R}^p$ and $\vec{G} : dom(\vec{G}) \subseteq \mathbb{R}^p \to \mathbb{R}^m$. Let $\vec{x}_o \in \mathbb{R}^n$ for which $\vec{F}(\vec{x}_o) = \vec{y}_o \in dom(\vec{G})$ and suppose that $\vec{F}$ is differentiable at $\vec{x}_o$ and $\vec{G}$ is differentiable at $\vec{y}_o$. We seek to show that $\vec{G} \circ \vec{F}$ is differentiable at $\vec{x}_o$ with Jacobian matrix $\vec{G}'(\vec{y}_o)\vec{F}'(\vec{x}_o)$. Observe that the existence of $\vec{G}'(\vec{y}_o) \in \mathbb{R}^{\,m\times p}$ and $\vec{F}'(\vec{x}_o) \in \mathbb{R}^{\,p\times n}$ follow from the differentiability of $\vec{G}$ at $\vec{y}_o$ and $\vec{F}$ at $\vec{x}_o$. In particular, if $||\vec{k}|| \approx 0$ then

$$\vec{G}(\vec{y}_o + \vec{k}) \approx \vec{G}(\vec{y}_o) + \vec{G}'(\vec{y}_o)\vec{k}.$$

Likewise, if $||\vec{h}|| \approx 0$ then

$$\vec{F}(\vec{x}_o + \vec{h}) \approx \vec{F}(\vec{x}_o) + \vec{F}'(\vec{x}_o)\vec{h}.$$

Suppose $\vec{h}$ is given such that $||\vec{h}|| \approx 0$. It follows that $\vec{F}'(\vec{x}_o)\vec{h} \approx 0$. Let $\vec{k} = \vec{F}'(\vec{x}_o)\vec{h}$ and note that

$$\underbrace{\vec{G}(\vec{F}(\vec{x}_o + \vec{h}) \approx \vec{G}(\vec{F}(\vec{x}_o) + \vec{F}'(\vec{x}_o)\vec{h})}_{continuity\ of\ G\ at\ y_o} = \vec{G}(\vec{y}_o + \vec{k}) \approx \vec{G}(\vec{y}_o) + \vec{G}'(\vec{y}_o)\vec{k}$$

[13] this is a plausibility argument, not a formal proof, all the $\approx$ symbols are shorthands for a more detailed estimation which is not given in these notes, however, you guessed it, can be found in a good advanced calculus text.

Therefore, for $||\vec{h}|| \approx 0$,

$$\vec{G}(\vec{F}(\vec{x}_o + \vec{h})) \approx \vec{G}(\vec{F}(\vec{x}_o)) + \vec{G}\,'(\vec{F}(\vec{x}_o))\vec{F}\,'(\vec{x}_o)\vec{h}.$$

Thus $\vec{G}(\vec{F}(\vec{x}_o + \vec{h})) - \vec{G}(\vec{F}(\vec{x}_o)) - \vec{G}\,'(\vec{F}(\vec{x}_o))\vec{F}\,'(\vec{x}_o)\vec{h} \approx 0$. In fact, if we worked out the careful details we could show that

$$\lim_{\vec{h} \to 0} \frac{\vec{G}(\vec{F}(\vec{x}_o + \vec{h})) - \vec{G}(\vec{F}(\vec{x}_o)) - \vec{G}\,'(\vec{F}(\vec{x}_o))\vec{F}\,'(\vec{x}_o)\vec{h}}{||\vec{h}||} = 0$$

and it follows that $(\vec{G} \circ \vec{F})'(\vec{x}_o) = \vec{G}\,'(\vec{F}(\vec{x}_o))\vec{F}\,'(\vec{x}_o)$. Technically, this is not a proof, but perhaps it makes the rule a bit more plausible. The chain rule is primarily a consequence of matrix multiplication when we look at it the right way. $\Box$.

Example 4.4.24. . *14*

$\boxed{E69}$ Suppose $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is differentiable. Define the polar coordinate change map; $\mathbb{X}(r,\theta) \equiv (r\cos\theta, r\sin\theta)$ this means $\mathbb{X} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ it takes $(r,\theta) \longmapsto (x(r,\theta), y(r,\theta))$. Consider $g = f \circ \mathbb{X}$. Then $Dg = Df \circ D\mathbb{X}$ where $\mathbb{X} = (x,y)$

$$Df = \left[\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right] \qquad D\mathbb{X} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

Hence if $w = f \circ \mathbb{X} = g$

$$Dg(r,\theta) = \left[\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right] \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

$$||$$

$$\left[\frac{\partial w}{\partial r}, \frac{\partial w}{\partial \theta} \right] = \left[\cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}, -r\sin\theta \frac{\partial f}{\partial x} + r\cos\theta \frac{\partial f}{\partial y} \right]$$

Thus,

$$\frac{\partial w}{\partial r} = \cos\theta \frac{\partial w}{\partial x} + \sin\theta \frac{\partial w}{\partial y} \longrightarrow \boxed{\frac{\partial}{\partial r} = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}}$$

$$\frac{\partial w}{\partial \theta} = -r\sin\theta \frac{\partial w}{\partial x} + r\cos\theta \frac{\partial w}{\partial y} \longrightarrow \boxed{\frac{\partial}{\partial \theta} = -r\sin\theta \frac{\partial}{\partial x} + r\cos\theta \frac{\partial}{\partial y}}$$

Example 4.4.25. . *13*

$\boxed{E70}$ Let $z = f(x,y) = x^2 - 3y^2$ and let $x = uv$ & $y = u + v^2$
Calculate $\partial z/\partial u$ and $\partial z/\partial v$.

$$\frac{\partial z}{\partial u} = \frac{\partial}{\partial u}\Big[f(x(u,v), y(u,v))\Big] = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} = (2x)(v) - 6y(1).$$

$$\frac{\partial z}{\partial v} = \frac{\partial}{\partial v}\Big[f(x(u,v), y(u,v))\Big] = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} = (2x)(u) - 6y(2v).$$

this is simple enough, you can use a tree-diagram if you like, but I've never needed them, you just identify the intermediate variables and sort-of "conserve partials". Lets see how this is done in the matrix / Jacobian formalism. We define

$$X(u,v) \equiv (x(u,v), y(u,v)) = (uv, u+v^2).$$

Thus, notice $x_1 = u$ and $x_2 = v$ while $x = f_1$, $y = f_2$ and $X = f$

$$DX = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \quad \text{while} \quad Df = \left[\frac{\partial f}{\partial x} \;\; \frac{\partial f}{\partial y}\right]$$

Then $z = f \circ X$ so $z = z(u,v)$

$$Dz = \left[\frac{\partial z}{\partial u}, \frac{\partial z}{\partial v}\right] = (Df)(DX) = \left[\frac{\partial f}{\partial x} \;\; \frac{\partial f}{\partial y}\right]$$

$$= \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$

$$= \Big[\underbrace{\frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}}_{\frac{\partial z}{\partial u}}, \underbrace{\frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}}_{\frac{\partial z}{\partial v}}\Big]$$