Using our expertise with functions of one variable, we have defined partial derivatives; rate of change in the two directions: horizontal and vertical or equivalently in the x-direction and y-direction respectively.

There are several peculiarities that take place. We shall just try to understand some so that they serve as warning. However, the main focus of our course is to develop smooth calculus and not really to spend time on pathologies that can occur. They are also interesting, but not part of our course.

Conditions on f_1, f_2 for continuity of f:

Let f be a function defined on an open set $\Omega \subset R^2$ and $f : \Omega \to R$ and $a = (a_1, a_2) \in \Omega$. Then

$$D_x f(a) = \lim_{h \to 0} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}$$

and

$$D_y f(a) = \lim_{k \to 0} \frac{f(a_1, a_2 + k) - f(a_1, a_2)}{k}$$

whenever these limits exist. However existence of these derivatives does not even imply that the function is continuous at the point a. For example the function

$$f(x, y) = \frac{xy}{x^2 + y^2}, \quad (x, y) \neq (0, 0); \quad f(0, 0) = 0.$$

has partial derivatives at the point $(0, 0)$ but is not continuous at that point.

Here is one condition when the existence of these partial derivatives implies continuity. Suppose that the partial derivatives are bounded in Ω. Then the functioc is continuous in Ω.

Theorem: if the partial derivatives are bounded in Ω then f is continuous in Ω.

Obviously, you see that, as in the one dimensional case, continuity is a local property. Recall, though unnecessary, this means the following: if f

and g are defined in Ω and both agree in $B(a,r)$ for some $r > 0$, then continuity of f at a is equivalent to continuity of g at a. In particular, if $a \in \Omega$ and if there is a number $r > 0$ such that f is continuous in $B(a,r)$ then f is continuous at the point a. Thus if the partial derivatives are bounded in $B(a,r)$ for some $r > 0$, then f is continuous at a.

The proof of the theorem is simple. Take $a \in R$ and $\epsilon > 0$. First fix $M > 0$ a bound for $|f_x|$ and $|f_y|$. Fix a small rectangle or square, say, $Q = (a_1 - \delta, a_1 + \delta) \times (a_2 - \delta, a_2 + \delta) \subset \Omega$ with $\delta < \epsilon/2M$. We show that

$$x \in Q \Rightarrow |f(x) - f(a)| < \epsilon.$$

we can write $x = (x_1, x_2)$ where $x_1 = a_1 + h_1$ and $x_2 = a_2 + h_2$ with $|h_1| < \delta$ and $|h_2| < \delta$

$$|f(x)-f(a)| \leq |f(a_1+h_1, a_2+h_2)-f(a_1+h_1, a_2)|+|f(a_1+h_1, a_2)-f(a_1, a_2)|.$$

Note that all the points that appear on the right side are in the rectangle Q and hence so are the lines joining the first two points (vertical line) and the last two points (horizontal line). The mean value theorem that we learnt for functions of one variable tells us that the first term on right side equals $|h_2||f_y(P_1)|$ for some point P_1 on the vertical line. But this is smaller than $(\epsilon/2M)M = \epsilon/2$. similarly the second term is also smaller than $\epsilon/2$. This completes the proof of continuity.

equality of mixed derivatives:

It is also possible that f has partial derivatives everywhere giving us the functions f_x and f_y everywhere. It is also quite possible that they again have partial derivatives

$$f_{xx} = (f_x)_x, \quad f_{xy} = (f_x)_y, \quad f_{yx} = (f_y)_x, \quad f_{yy} = (f_y)_y.$$

One would like to have $f_{xy} = f_{yx}$, that is the order in which we differentiate should not matter. But it is a peculiarity that this may not happen. For example, if you take the function

$$f(x,y) = \frac{x^3y - y^3x}{x^2 + y^2}, \quad (x,y) \neq (0,0); \quad f(0,0) = 0.$$

Then

$$f_1(x,y) = \frac{3x^2y - y^3}{x^2 + y^2} - \frac{(x^3y - y^3x)2x}{(x^2 + y^2)^2}, \quad (x,y) \neq (0,0);$$

and

$$f_1(0,0) = 0.$$

$$f_2(x,y) = \frac{x^3 - 3y^2x}{x^2+y^2} - \frac{(x^3y - y^3x)2y}{(x^2+y^2)^2}, \quad (x,y) \neq (0,0);$$

and

$$f_2(0,0) = 0.$$

$$f_{12}(0,0) = \lim_{k \to 0} \frac{f_1(0,k) - f_1(0,0)}{k} = -1.$$

while

$$f_{21}(0,0) = \lim_{h \to 0} \frac{f_2(h,0) - f_2(0,0)}{h} = +1.$$

This is rather unpleasant.

However if f_{12} and f_{21} are continuous, then they are equal.

Before you loose track where we are heading, let me repeat the following. the concept of derivative is rather delicate, if you are not careful. Many unpleasant things happen. We are seeing those and along the way try to remedy the situation; if you assume suitable hypothesis then it does not happen. For every ill, if you use separate hypothesis to rectify it then life becomes difficult. We soon get out of all these warning signals and assume *one* nice hypothesis that allows good things to happen.

Returning to the problem we are discussing, we shall now show that if f_{12} and f_{21} are continuous then they are equal. So let us take a point $a \in \Omega$. Need to show $f_{12}(a) = f_{21}(a)$. You fix any $\epsilon > 0$. We show that

$$|f_{12}(a) - f_{21}(a)| < \epsilon \qquad (\diamondsuit).$$

You agree that this would suffice.

Using continuity of the 'mixed' derivatives, fix a rectangle, or square, $Q = (a-\delta, a+\delta) \subset \Omega$ such that

$$x \in Q \Rightarrow |f_{12}(a) - f_{12}(x)| < \epsilon/2; \quad |f_{21}(a) - f_{21}(x)| < \epsilon/2. \quad (\heartsuit)$$

Here we have used an abbreviation:

$$(a-\delta, a+\delta) = (a_1 - \delta, a_1 + \delta) \times (a_2 - \delta, a_2 + \delta).$$

Let us fix $0 < h < \delta$ and consider the quantity

$$\Delta = f(a_1 + h, a_2 + h) - f(a_1 + h, a_2) - f(a_1, a_2 + h) + f(a_1, a_2).$$

The plan is to show

$$(\exists\ P_1, P_2 \in Q) \quad f_{12}(P_1) = \Delta = f_{21}(P_2). \qquad (\dagger)$$

If this is done then ($\heartsuit$) with $x = P_1$ in the first inequality and $x = P_2$ in the second inequality will give us ($\diamondsuit$) as required.

Towards proof of ($\dagger$), consider the functions of one variable,

$$\varphi(x) = f(x, a_2 + h) - f(x, a_2); \qquad \Psi(y) = f(a_1 + h, y) - f(a_1, y).$$

Apply mean value theorem to get $0 < \theta_1 < 1$ and $0 < \eta_2 < 1$ such that

$$\varphi(a_1 + h) - \varphi(a_1) = h\varphi'(a_1 + \theta_1 h); \quad \Psi(a_2 + h) - \Psi(a_2) = h\Psi'(a_2 + \eta_2 h).$$

Observe

$$\varphi(a_1 + h) - \varphi(a_1) \ = \ \Delta \ = \ \Psi(a_2 + h) - \Psi(a_2).$$

and also that

$$\varphi'(x) = f_1(x, a_2 + h) - f_1(x, a_2); \qquad \Psi'(y) = f_2(a_1 + h, y) - f_2(a_1, y).$$

Thus

$$h[f_1(a_1 + \theta_1 h, a_2 + h) - f_1(a_1 + \theta_1 h, a_2)] = \Delta,$$

and

$$h[f_2(a_1 + h, a_2 + \eta_2 h) - f_2(a_1, a_2 + \eta_2 h)] = \Delta.$$

Now applying mean value theorem to the left sides of the two equations above we get $0 < \theta_2 < 1$ and $0 < \eta_1 < 1$ such that

$$h^2 f_{12}(a_1 + \theta_1 h, a_2 + \theta_2 h) = \Delta = h^2 f_{21}(a_1 + \eta_1 h, a_2 + \eta_2 h).$$

Take

$$P_1 = (a_1 + \theta_1 h, a_2 + \theta_2 h), \quad P_2 = (a_1 + \eta_1 h, a_2 + \eta_2 h),$$

to complete the proof of ($\dagger$).

differentiability:

As discussed last time, let us turn to the problem of understanding a function near a point — either in the sense of geometry like drawing a tangent plane at that point or in the sense of approximating by simpler functions.

Let $f : \Omega \to R$ where $\Omega \subset R^2$ is an open set and $a \in \Omega$. What is the simplest function that approximates f near a. First we need to make it precise. we are looking for a function $\varphi(x)$ on Ω so that $f(x) - \varphi(x) \to 0$ as x approaches a, that is as $||x - a|| \to 0$. The answer is simple, take the function whose value at every point is the number $f(a)$. Then this is the simplest function we can think of and it satisfies the requirement.

Next, as in the case of R, let us demand that $\varphi(a) = f(a)$ and $f(x) - \varphi(x) \to 0$ faster than $||x - a||$. So what is meant by faster. well even the ratio $|f(x) - \varphi(x)|/||x - a|| \to 0$ as $||x - a|| \to 0$. But we are now allowed linear functions. What are linear functions? Just like in the case of R, they are of the form $\varphi(x) = a_1 x_1 + a_2 x_2 + b$ for some numbers a_1, a_2, b. This can be succinctly expressed as $\varphi(x) = \alpha \cdot x + \beta$ where $\alpha \in R^2$ and dot is the inner product.

If that happens, then

$$f(a) = \varphi(a) = \alpha \cdot a + \beta; \qquad \text{that is} \qquad \beta = f(a) - \alpha \cdot a.$$

Thus the function is $\varphi(x) = \alpha \cdot x + f(a) - \alpha \cdot a$. Of course, we still do not low what is the vector α. It should satisfy, $|f(x) - \varphi(x)|/||x - a|| \to 0$ as $||x - a|| \to 0$.

Let us agree to say that the function f is differentiable at the point a in case there is a vector $\alpha \in R^2$ such that

$$\lim_{||x-a|| \to 0} \frac{f(x) - f(a) - \alpha \cdot (x - a)}{||x - a||} \to 0. \qquad (\bullet)$$

In such a case, we can refer to the vector α as the derivative of f at a.

Let us see what could be the vector α. Taking the sequence $x_n = (a_1 + \frac{1}{n}, a_2)$ we see that $\alpha_1 = f_1(a)$. Similarly, taking $x_n = (a_1, a_2 + \frac{1}{n})$ we see $\alpha_2 = f_2(a)$.

Thus if the function is differentiable then

$$\alpha = (f_1(a), f_2(a)).$$

But is this enough, that is, if the partial derivatives exist then will ($\bullet$) hold? Not always.

In fact, existence of partial derivatives need not even imply that the function is continuous. We shall return to this in a minute, but let us also see the geometric picture of the derivative.

Just as the graph $\{(x, y) : y = mx+c\}$ of a linear function $\varphi(x) = mx+c$ is a straight line (never mind, we have missed y-axis by loosely representing line as above), the graph $\{(x, y, z) : z = ax + by + c\}$ of linear function $\varphi(x, y) = ax + by + c$ is a plane. Just as graphs of functions on R to R are called curves in the plane R^2, graphs of functions from R^2 to R are called surfaces in R^3.

How do you imagine surfaces. Imagine the ground to be the plane and think of a tent that has height $f(x, y)$ at the point (x, y) on the ground. You can imagine it as a tent or bowl or inverted bowl etc, whatever you are comfortable with. of course specific functions have specific shapes.

A tangent plane to the surface $z = f(x)$ at the point $a = (a_1, a_2) \in \Omega$ is the graph of a map $\varphi(x) = \alpha \cdot x + \beta$ that passes through the point $(a, f(a)) \in R^3$ on the surface but makes 'stronger' contact with the surface than the constant function $\Psi(x) \equiv f(a)$. What does this mean? We mean, the ratio $[f(x) - \varphi(x)]/||x - a|| \to 0$ as $x \to a$; not simply $f(x) - \varphi(x) \to 0$.

This concept of tangent plane again leads to the same conclusion as earlier, namely solving ($\bullet$) for vector α and you end up with the same answer as above.

continuity of f_1, f_2 implies differentiability:

We saw that existence of partial derivatives does not imply ($\bullet$) holds. However if the partial derivatives are continuous then the function is differentiable and at any point a, the vector $(f_1(a), f_2(a))$ is indeed the derivative.

Theorem: Let $\Omega \subset R^2$ be open set and $f : \Omega \to R$ be such that the partial derivatives f_1 and f_2 are continuous. Then f is differentiable and derivative of f at a point $a \in \Omega$ is the vector $(f_1(a), f_2(a))$.

Fix $a \in \Omega$ and denote $\alpha = (f_1(a), f_2(a))$. Need to show

$$\lim_{||x-a|| \to 0} \frac{f(x) - f(a) - \alpha \cdot (x-a)}{||x-a||} \to 0.$$

Fix $\epsilon > 0$. Using continuity of f_1 and f_2 choose $\delta > 0$ so that $B(a, \delta) \subset \Omega$ and

$$||x - a|| < \delta \Rightarrow |f_1(x) - f_1(a)| < \epsilon/2; \quad |f_2(x) - f_2(a)| < \epsilon/2.$$

Let us take any point $x = (x_1, x_2) \in B(a, \delta)$. Then using the mean value theorem (for functions of one variable) get points P_1 and P_2 on the appropriate horizontal and vertical line segments such that

$$f(x) - f(a) = [f(x_1, x_2) - f(a_1, x_2)] + [f(a_1, x_2) - f(a_1, a_2)]$$

$$= (x_1 - a_1) f_1(P_1) + (x_2, a_2) f_2(P_2) = (x - a) \cdot v$$

where v is the vector $v = (f_1(P_1), f_2(P_2)$. Thus

$$|f(x) - f(a) - \alpha \cdot (x - a)| = |(x - a) \cdot (v - \alpha)|$$

$$\leq ||x - a|| \, ||v - \alpha|| \leq ||x - a|| \sqrt{2\epsilon^2/4} \leq \epsilon ||x - a||.$$

Thus for $x \in B(a, \delta)$ we have

$$\left| \frac{f(x) - f(a) - \alpha \cdot (x - a)}{||x - a||} \right| < \epsilon.$$

proving the result.

differentiability $\not\Rightarrow$ f_1, f_2 conitnuous:

We have shown that if the partial derivatives f_1, f_2 are continuous then f is differentiable. However differentiability is not equivalent to the continuity of partial derivatives.

Let us consider the function,

$$f(x, y) = (x^2 + y^2) \sin \frac{1}{x^2 + y^2}, \quad (x, y) \neq (0, 0); \qquad f(0, 0) = 0.$$

Then

$$f_1(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - 0}{h} = \lim_{h \to 0} h \sin \frac{1}{h^2} = 0.$$

When $(x, y) \neq (0, 0)$

$$f_1(x, y) = 2x \sin \frac{1}{x^2 + y^2} - \frac{2x}{x^2 + y^2} \cos \frac{1}{x^2 + y^2}.$$

You can see that f_1 is not continuous at $(0, 0)$. Indeed the first term in f_1 converges to zero. Argue that the second term is unbounded. However, f is differentiable at $(0, 0)$ simply because

$$\frac{f(x, y)}{||(x, y)||} \to 0.$$

showing that f is differentiable and the derivative is the vector $(0, 0)$. you can show that f_2 is also not continuous at $(0, 0)$.

bringing some order:

I hope you do understand how confusing is the situation. so to avoid all bad things, we are going to assume that the functions we deal with are all C^1 functions. A function is C^1 if its first partial derivatives are continuous functions. Recall, then the function is differentiable and derivative at point a is the vector $(f_1(a), f_2(a))$. Derivative is denoted by $\nabla f(a)$., read it as gradient or grad or nabla.

Let f be a C^1 function on $\Omega \subset R^2$ and $a \in \Omega$. Denoting points in R^2 by $x = (x_1, x_2)$; the hyperplane $x_3 = f(a) + \nabla f(a) \cdot (x - a)$ is the tangent plane to the surface $x_3 = f(x_1, x_2)$ at the point a.

We shall study properties of the derivatives. But before we do this, let us get back to a statement regarding partial derivatives. We did recognise the possibility of rate of change, not only in the x and y directions, but also in any direction at the point. Let us take-up this idea now.

What is meant by direction? it is simply a unit vector u. It is the direction pointed by that vector. More precisely, join origin to u and consider the full line — extended both ways, but remember the direction is from origin to u. In common parlance, it is customary to say the line extended in both directions. we would not use it, because it would be confusing. There are no two directions, there is only one direction, namely *that pointed by* u. When one uses the phrase, the line extended in ‘both directions’, one is only referring to the act of drawing the line to ‘both sides’ — join origin to u and extend beyond u and also beyond zero.

Now, let f be a C^1 function on $\Omega \subset R^2$ and $a \in \Omega$. If we take the unit vector $e_1 = (1, 0)$ then the quantity

$$\lim_{t \to 0} \frac{f(a + te_1) - f(a)}{t}$$

gives precisely the partial derivative $f_1(a)$. similarly, if we take the unit vector $e_2 = (0, 1)$, then

$$\lim_{t \to 0} \frac{f(a + te_2) - f(a)}{t}$$

gives precisely $f_2(a)$. We use the same idea to define the directional derivative in the direction of u;

$$D_u f(a) = \lim_{t \to 0} \frac{f(a + tu) - f(a)}{t}.$$

You agree that this does indeed give rate of change of the function in the direction of u. You already know that

$$f_1(a) = \nabla f(a) \cdot e_1; \qquad f_2(a) = \nabla f(a) \cdot e_2.$$

It is natural to expect that $D_u f(a) = \nabla f(a) \cdot u$. This is indeed so. We see this soon. But note that in the new notation, f_1 and f_2 become $D_{e_1} f$ and $D_{e_2} f$ respectively.

It is also natural to think of direction as an angle θ, $0 \leq \theta \leq 2\pi$. of course angles zero and 2π correspond to the same direction. Thus x direction corresponds to $\theta = 0$, y direction corresponds to $\theta = \pi/2$. The negative x direction corresponds to $\theta = \pi$ and the negative y direction corresponds to $\theta = 3\pi/2$.

The reason we did not use angles is because, in higher dimensions you need several angles. Of course unit vector still specifies a direction. all the concepts and results that we discussed have analogues in R^d as well.