Derivatives–Part One

Reminder: In the context of the x,y plane, "vertical" means perpendicular to the x axis, or parallel to the y axis, and "horizontal" means parallel to the x axis, or perpendicular to the y axis. In the context of x,y,z space, "vertical" means perpendicular to the x,y plane, or parallel to the z axis, and "horizontal" means parallel to the x,y plane, or perpendicular to the z axis.

Suppose we have a function $z = f(x,y)$. Its graph is a surface, S, which passes the Vertical Line Test (i.e., any vertical line intersects the surface at no more than one point). Let $P = (x_0,y_0)$ be any point in the domain of f, and let $z_0 = f(x_0,y_0)$. Thus, (x_0,y_0,z_0) is a point on S. If we view the x,y plane as a subset of x,y,z space (i.e., as the plane $z = 0$), then P would be the point $(x_0,y_0,0)$, which would align vertically with the point (x_0,y_0,z_0).

Let L be any line in the x,y plane passing through the point (x_0,y_0). Let $< a,b >$ be a direction vector for this line and also a unit vector. The parametric equations of the line are $x = x_0 + at,\ y = y_0 + bt$. Alternatively, we can view this line in the context of x,y,z space, in which case it lies in the plane $z = 0$, passes through the point $(x_0,y_0,0)$, has direction vector $< a,b,0 >$ (which is still a unit vector), and has parametric equations $x = x_0 + at,\ y = y_0 + bt,$ $z = 0 + 0t$ (or simply $z = 0$).

Let V be the vertical line passing through the point $(x_0,y_0,0)$. The standard basis vector $\mathbf{k} =$ $< 0,0,1 >$ can serve as the direction vector for this line. If we parameterize V using this point and this direction vector, and using the variable w as our parameter, then its parametric equations will be $x = x_0 + 0w,\ y = y_0 + 0w,\ z = 0 + 1w,$ or simply $x = x_0,\ y = y_0,$ $z = w$.

Let $\wp$ be the orthogonal projection of line L into x,y,z space. $\wp$ is thus a vertical plane containing the lines L and V, as well as the points $(x_0,y_0,0)$ and (x_0,y_0,z_0).

We shall impose a two-dimensional coordinate system on plane $\wp$, using line L as the horizontal axis and line V as the vertical axis. We shall refer to the former line as the t axis and to the latter as the w axis. The point where these lines cross, $(x_0,y_0,0)$, is the point generated when $t = 0$ and when $w = 0$, so it serves as the origin, $(0,0)$, in the t,w coordinate system. Since the lines L and V were parameterized using direction vectors that were *unit* vectors, t and w qualify as arclength parameters–i.e., any value of t generates a point on line L whose distance from $(x_0,y_0,0)$ is $|t|$, and any value of w generates a point on V whose distance from $(x_0,y_0,0)$ is $|w|$. Hence, any ordered pair (t_1,w_1) corresponds to a point in the t,w plane whose directed distance from the t axis is t_1 and whose directed distance from the w axis is w_1.

The intersection of surface S and plane $\wp$ is a curve, C, which passes the Vertical Line Test (i.e., any vertical line intersects the curve at no more than one point). Hence, this curve is the graph of a function in the t,w plane. Let us call this function g. This function can be represented algebraically by an equation expressing w in terms of t, $w = g(t)$. We can obtain this equation from the equation $z = f(x,y)$, if we substitute w in place of z, $x_0 + at$ in place of x, and $y_0 + bt$ in place of y, giving us $w = f(x_0 + at, y_0 + bt)$. Notice that $g(0) = f(x_0, y_0) = z_0$. Hence, the point (x_0, y_0, z_0) on surface S is the point $(0, z_0)$ on the graph of the function $w = g(t)$, i.e., it is the w intercept of the function.

If the graph of $w = g(t)$ has a nonvertical tangent line at the point $(0, z_0)$, then the slope of that tangent line is $g'(0)$. Let us refer to this tangent line as T. In the t, w coordinate system, we may write the equation of line T in slope-intercept form: $w = g'(0)t + z_0$.

Line T can be interpreted as a tangent line in two ways. On the one hand, as a line in plane $\wp$, it is tangential to the curve C at the point $(0, z_0)$. On the other hand, as a line in x,y,z space, it is tangential to the surface S at the point (x_0, y_0, z_0).

Can we write a vector equation for line T in x,y,z space? To accomplish this, we will need a direction vector for T, and to find such a vector, we will need a second point on T. We can easily find such a point in the t, w coordinate system. Using the equation $w = g'(0)t + z_0$, we may substitute $t = 1$, giving us $w = g'(0) + z_0$. So $(1, g'(0) + z_0)$ is a point on T. Now we convert these t, w coordinates into x,y,z coordinates, using the equations $x = x_0 + at$, $y = y_0 + bt$, $z = w$. The result is $(x_0 + a, y_0 + b, g'(0) + z_0)$. So we may use the vector $< x_0 + a - x_0, y_0 + b - y_0, g'(0) + z_0 - z_0 > = < a, b, g'(0) >$ as a direction vector for line T. Hence, the vector equation of line T is $\mathbf{r}(t) = < x_0, y_0, z_0 > + t < a, b, g'(0) >$.

T may be referred to as **the tangent line at** (x_0, y_0) **in the direction of** $< a, b >$. This refers to a point and a unit direction vector *in the domain* of the function $z = f(x,y)$. This is simply a compact way of saying that when the line through the point (x_0, y_0) with unit direction vector $< a, b >$ is orthogonally projected into x,y,z space, the resulting vertical plane intersects the surface $z = f(x,y)$ to give us a curve whose tangent line at the point (x_0, y_0, z_0) is T.

To illustrate all these concepts, suppose $z = f(x,y) = x^2 + y^2$, whose graph is a circular paraboloid. Consider the point $(2,3)$ in the x,y plane. $f(2,3) = 13$, so the point $(2,3,13)$ lies on the graph of f. It lies directly above the point $(2,3,0)$ in the plane $z = 0$.

If a line L in the x,y plane passes through the point $(2,3)$ and has unit direction vector $< a, b >$, then its parametric equations are $x = 2 + at$, $y = 3 + bt$. In the context of x,y,z space, line L lies in the plane $z = 0$, passes through the point $(2,3,0)$, has unit direction vector $< a, b, 0 >$, and has parametric equations $x = 2 + at$, $y = 3 + bt$, $z = 0$.

Let V be the vertical line passing through the point $(2,3,0)$. Its parametric equations are $x = 2$, $y = 3$, $z = w$.

Let us analyze three different choices for line L.

Case One: In the x,y plane, let L_1 be the horizontal line (i.e., the line parallel to the x axis) passing through the point $(2,3)$. The two-dimensional equation of this line is $y = 3$. This is also the three-dimensional equation of the line's orthogonal projection into x,y,z space, which is a vertical plane. Notice that the plane $y = 3$ is parallel to the x,z plane, i.e., the plane $y = 0$. The standard basis vector $\mathbf{i} = < 1, 0 >$ can serve as the direction vector for L_1.

The line's parametric equations are $x = 2 + 1t,\ y = 3 + 0t$, or simply $x = 2 + t,\ y = 3$. (We add the third equation $z = 0$ if we view line L_1 in the context of x,y,z space.) The intersection of the surface $z = x^2 + y^2$ with the plane $y = 3$ is a curve C_1, which is the graph of the function $w = g_1(t) = f(2 + t, 3) = (2 + t)^2 + 3^2 = 4 + 4t + t^2 + 9 = t^2 + 4t + 13$. Curve C_1 is an upward-opening parabola with w intercept $(0, 13)$ and with vertex $(-2, 9)$. Let T_1 be the tangent line at the point $(0, 13)$. $g_1{}'(t) = 2t + 4$ and $g_1{}'(0) = 4$. In the t, w coordinate system, T_1 has slope 4, and its slope-intercept equation is $w = 4t + 13$. In x,y,z space, T_1 has vector equation $\mathbf{r}_1(t) = < 2, 3, 13 > + t < 1, 0, 4 >$. T_1 may be referred to as the tangent line at $(2, 3)$ in the direction of $\mathbf{i}$.

Case Two: In the x,y plane, let L_2 be the vertical line (i.e., the line parallel to the y axis) passing through the point $(2, 3)$. The two-dimensional equation of this line is $x = 2$. This is also the three-dimensional equation of the line's orthogonal projection into x,y,z space, which is a vertical plane. Notice that the plane $x = 2$ is parallel to the y,z plane, i.e., the plane $x = 0$. The standard basis vector $\mathbf{j} = < 0, 1 >$ can serve as the direction vector for L_2. The line's parametric equations are $x = 2 + 0t,\ y = 3 + 1t$, or simply $x = 2,\ y = 3 + t$. (We add the third equation $z = 0$ if we view line L_2 in the context of x,y,z space.) The intersection of the surface $z = x^2 + y^2$ with the plane $x = 2$ is a curve C_2, which is the graph of the function $w = g_2(t) = f(2, 3 + t) = 2^2 + (3 + t)^2 = 4 + 9 + 6t + t^2 = t^2 + 6t + 13$. Curve C_2 is an upward-opening parabola with w intercept $(0, 13)$ and with vertex $(-3, 4)$. Let T_2 be the tangent line at the point $(0, 13)$. $g_2{}'(t) = 2t + 6$ and $g_2{}'(0) = 6$. In the t, w coordinate system, T_2 has slope 6, and its slope-intercept equation is $w = 6t + 13$. In x,y,z space, T_2 has vector equation $\mathbf{r}_2(t) = < 2, 3, 13 > + t < 0, 1, 6 >$. T_2 may be referred to as the tangent line at $(2, 3)$ in the direction of $\mathbf{j}$.

Case Three: In the x,y plane, let L_3 be the oblique line passing through the point $(2, 3)$ with unit direction vector $\mathbf{u} = < 0.6, 0.8 >$. The two-dimensional equation of this line is $4x - 3y = -1$. This is also the three-dimensional equation of the line's orthogonal projection into x,y,z space, which is a vertical plane. The line's parametric equations are $x = 2 + 0.6t$, $y = 3 + 0.8t$. (We add the third equation $z = 0$ if we view line L_3 in the context of x,y,z space.) The intersection of the surface $z = x^2 + y^2$ with the plane $4x - 3y = -1$ is a curve C_3, which is the graph of the function $w = g_3(t) = f(2 + 0.6t, 3 + 0.8t) = (2 + 0.6t)^2 + (3 + 0.8t)^2 = 4 + 2.4t + 0.36t^2 + 9 + 4.8t + 0.64t^2 = t^2 + 7.2t + 13$. Curve C_3 is an upward-opening parabola with w intercept $(0, 13)$ and with vertex $(-3.6, 0.04)$. Let T_3 be the tangent line at the point $(0, 13)$. $g_3{}'(t) = 2t + 7.2$ and $g_3{}'(0) = 7.2$. In the t, w coordinate system, T_3 has slope 7.2, and its slope-intercept equation is $w = 7.2t + 13$. In x,y,z space, T_3 has vector equation $\mathbf{r}_3(t) = < 2, 3, 13 > + t < 0.6, 0.8, 7.2 >$. T_3 may be referred to as the tangent line at $(2, 3)$ in the direction of $\mathbf{u} = < 0.6, 0.8 >$.

The concept of slope applies only to lines in a two-dimensional coordinate system; it is not applicable to lines in three-dimensional space. Thus, when we are dealing with a tangent line to a surface in x,y,z space, if we refer to the slope of the tangent line, it is always understood in the context of the *vertical plane* containing that tangent line. In the above examples, the lines T_1, T_2, and T_3 are all tangential to the surface $z = x^2 + y^2$ at the point $(2, 3, 13)$. When we specify that their slopes are 4, 6, and 7.2, respectively, it is understood that the first slope is in the context of the vertical plane $y = 3$, which contains T_1, the second slope is in the context of the vertical plane $x = 2$, which contains T_2, and the third slope is in the context of the vertical plane $4x - 3y = -1$, which contains T_3.

Above, we have found three different tangent lines for the surface $z = x^2 + y^2$ at the point $(2,3,13)$, corresponding to three different unit direction vectors that may be placed at the point $(2,3)$ in the x,y plane–namely, $\mathbf{i}$, $\mathbf{j}$, and $< 0.6, 0.8 >$. But since there are *infinitely many* different unit direction vectors that may be placed at the point $(2,3)$, there are *infinitely many* different tangent lines for the surface $z = x^2 + y^2$ at the point $(2,3,13)$.

In the above example, the lines T_1, T_2, and T_3 intersect at the point $(2,3,13)$. Furthermore, the lines are *coplanar*–i.e., they all lie in one plane. How do we know this? Three lines intersecting at one point are coplanar if and only if their three *direction vectors* are coplanar, and the direction vectors are coplanar if and only if their box product is 0. The direction vectors of T_1, T_2, and T_3 are $< 1,0,4 >$, $< 0,1,6 >$, and $< 0.6, 0.8, 7.2 >$. You may confirm for yourself that their box product is 0.

In fact, for the function $f(x,y) = x^2 + y^2$, *all* tangent lines at the point $(2,3,13)$ are coplanar. In other words, there exists a unique plane, which we will call $\Im$, such that *every* tangent line at the point $(2,3,13)$ lies in this plane. We refer to this plane as the **tangent plane** for the surface $z = x^2 + y^2$ at the point $(2,3,13)$

To write an equation for the plane $\Im$, we need a point in the plane and a normal vector. We already have the point, namely, $(2,3,13)$. To find a normal vector, we can compute the cross product of the direction vectors for T_1 and T_2. $< 0,1,6 > \times < 1,0,4 > = < 4,6,-1 >$. Hence, $\Im$ has equation $4(x-2) + 6(y-3) - 1(z-13) = 0$, or $4x + 6y - z = 13$. Notice that in this last equation, the coefficient of x is the slope of T_1, the coefficient of y is the slope of T_2, and the right side of the equation is the value of z_0. (As we shall see later on, the right side of the equation of the tangent plane does not *always* turn out to be z_0. We're getting such a nice result here because $f(x,y) = x^2 + y^2$ is such a nice, symmetrical function.)

Although in this example, all tangent lines at the point $(2,3,13)$ are coplanar, this is not necessarily true in all situations. It is possible that a surface $z = f(x,y)$ could have a point (x_0,y_0,z_0) for which its infinite collection of tangent lines may *not* be coplanar. In other words, there might not be one plane containing all the tangent lines. In this case, we would say the surface *has no tangent plane* at the point (x_0,y_0,z_0), or we can say that the tangent plane *does not exist* (or is *undefined*) at this point. If all tangent lines at (x_0,y_0,z_0) *are* coplanar, i.e., if the tangent plane *does* exist at (x_0,y_0,z_0), then we say the function f is **differentiable** at this point. Alternatively, we can say f is differentiable at (x_0,y_0), referring to the point in the *domain* of f that gives rise to the point (x_0,y_0,z_0) on the *graph* of f.

Thus, the function $f(x,y) = x^2 + y^2$ is differentiable at the point $(2,3,13)$ on its graph, or at the point $(2,3)$ in its domain.

In Calculus I, we learned the following principle: Say we have a function $y = f(x)$. Let x_0 be a point in its domain. Suppose the function has a nonvertical tangent line at x_0. Then the slope of this tangent line is known as the *derivative* of f at x_0, and is denoted $f'(x_0)$. It represents the *instantaneous rate of change* of the function at x_0.

This idea can be adapted to Calculus III. Say we have a function $z = f(x,y)$. Let (x_0,y_0) be a point in its domain. Suppose the function has a nonvertical tangent line at (x_0,y_0) in the direction of a unit vector $\mathbf{u} = < a,b >$. Then the slope of this tangent line is known as the **derivative** of f at (x_0,y_0) **in the direction of** $\mathbf{u}$, and is denoted $D_{\mathbf{u}}f(x_0,y_0)$. It represents the **instantaneous rate of change** of the function at (x_0,y_0) **in the direction of** $\mathbf{u}$. This kind of derivative–i.e., a derivative in a specified direction–is known as a **directional derivative.**

Returning to our original example, for the function $f(x,y) = x^2 + y^2$,

- Since T_1 was the tangent line at $(2,3)$ in the direction of $\mathbf{i}$, and since its slope was 4, we can say that the derivative of f at $(2,3)$ in the direction of $\mathbf{i}$ is 4, i.e., $D_{\mathbf{i}}f(2,3) = 4$.
- Since T_2 was the tangent line at $(2,3)$ in the direction of $\mathbf{j}$, and since its slope was 6, we can say that the derivative of f at $(2,3)$ in the direction of $\mathbf{j}$ is 6, i.e., $D_{\mathbf{j}}f(2,3) = 6$.
- Since T_3 was the tangent line at $(2,3)$ in the direction of $\mathbf{u} = < 0.6, 0.8 >$, and since its slope was 7.2, we can say that the derivative of f at $(2,3)$ in the direction of $\mathbf{u} = < 0.6, 0.8 >$ is 7.2, i.e., $D_{\mathbf{u}}f(2,3) = 7.2$.

In our discussion so far, we have found all our directional derivatives by means of algebraic substitution involving the variables t and w. This is not the most efficient method for finding directional derivatives. Our next step is to study efficient techniques.

xxxxxxxxxxxxxxxxxxxxx
In Case One, the function $w = g_1(t)$ was obtained by passing a line through the point $(2,3)$ with unit direction vector $\mathbf{i} = < 1,0 >$. In Case Two, the function $w = g_2(t)$ was obtained by passing a line through the point $(2,3)$ with unit direction vector $\mathbf{j} = < 0,1 >$. In Case Three, the function $w = g_3(t)$ was obtained by passing a line through the point $(2,3)$ with unit direction vector $\mathbf{u} = < 0.6, 0.8 >$. In each case, we obtained a quadratic function with w intercept $(0,13)$. Each of these three quadratic functions has a tangent line at this point.

$g'(0)$ is the slope of the tangent line to the curve C at its z intercept in the t,z plane. But this curve was obtained by intersecting the graph of the function $f(x,y)$ with a vertical plane determined by a point (x_0,y_0) and a unit vector $\mathbf{u} = < a,b >$. Consequently, we may refer to $g'(0)$ as the **directional derivative of** $f(x,y)$ **at the point** (x_0,y_0) **in the direction of the unit vector** $\mathbf{u}$. This is denoted $D_{\mathbf{u}} f(x_0,y_0)$.

To illustrate, consider the function $z = f(x,y) = x^2 + y^2$, whose graph is a circular paraboloid and whose domain is the entire x,y plane. Consider the point $(1,8)$ in the x,y plane. $f(1,8) = 65$, so the point $(1,8,65)$ lies on the graph of f.

We shall consider three different directional derivatives for this function at the point $(1,8)$, based on three different unit vectors.

First, consider the unit vector $\mathbf{i} = < 1,0 >$. The line through $(1,8)$ with direction vector $\mathbf{i}$ has the vector equation $\mathbf{r}(t) = < 1+t, 8 >$. For points on this line, $x = 1+t$ and $y = 8$. By substitution, $f(x,y) = (1+t)^2 + (8)^2 = t^2 + 2t + 65 = g(t)$. This is an upward-opening parabola in the t,z plane, whose z intercept is $(0,65)$ and whose vertex is the second-quadrant point $(-1,64)$. $g'(t) = 2t+2$, so $g'(0) = 2$. Thus, $D_{\mathbf{i}} f(1,8) = 2$.

Second, consider the unit vector $\mathbf{j} = < 0,1 >$. The line through $(1,8)$ with direction vector $\mathbf{j}$ has the vector equation $\mathbf{r}(t) = < 1, 8+t >$. For points on this line, $x = 1$ and $y = 8+t$. By substitution, $f(x,y) = (1)^2 + (8+t)^2 = t^2 + 16t + 65 = g(t)$. This is an upward-opening parabola in the t,z plane, whose z intercept is $(0,65)$ and whose vertex is the second-quadrant point $(-8,1)$. $g'(t) = 2t+16$, so $g'(0) = 16$. Thus, $D_{\mathbf{j}} f(1,8) = 16$.

Third, consider the unit vector $\mathbf{u} = < \frac{3}{5}, \frac{4}{5} >$. The line through $(1,8)$ with direction vector $\mathbf{u}$ has the vector equation $\mathbf{r}(t) = < 1 + \frac{3}{5}t, 8 + \frac{4}{5}t >$. For points on this line, $x = 1 + \frac{3}{5}t$, and $y = 8 + \frac{4}{5}t$. By substitution, $f(x,y) = (1 + \frac{3}{5}t)^2 + (8 + \frac{4}{5}t)^2 = t^2 + 14t + 65 = g(t)$. This is an upward-opening parabola in the t,z plane, whose z intercept is $(0,65)$ and whose vertex is the second-quadrant point $(-7,16)$. $g'(t) = 2t+14$, so $g'(0) = 14$. Thus, $D_{\mathbf{u}} f(1,8) = 14$.

The above analysis illustrates the basic concept, but there is another way to obtain the same results, which is generally much easier. Instead of setting up a new coordinate system and then using substitution, we can instead use a process known as **partial differentiation**.

The directional derivative of $f(x,y)$ at a point (x_0,y_0) in the direction of $\mathbf{i}$ (or, in other words, parallel to the x axis) is known as the **partial derivative of** $f(x,y)$ **with respect to** x **at** (x_0,y_0). It is denoted $f_x(x_0,y_0)$ or $\frac{\partial f}{\partial x}\big|_{(x_0,y_0)}$. Here is how we find it:

1. Start with the formula expressing $f(x,y)$ in terms of the two independent variables x and y.
2. Treat y as if it were a constant, and differentiate the formula with respect to x.
3. Evaluate the resulting formula at the point (x_0,y_0), i.e., substitute x_0 in place of x and y_0 in place of y, then do the math.

The directional derivative of $f(x,y)$ at a point (x_0,y_0) in the direction of $\mathbf{j}$ (or, in other words, parallel to the y axis) is known as the **partial derivative of** $f(x,y)$ **with respect to** y **at** (x_0,y_0). It is denoted $f_y(x_0,y_0)$ or $\frac{\partial f}{\partial y}\big|_{(x_0,y_0)}$. Here is how we find it:

1. Start with the formula expressing $f(x,y)$ in terms of the two independent variables x and y.
2. Treat x as if it were a constant, and differentiate the formula with respect to y.
3. Evaluate the resulting formula at the point (x_0,y_0), i.e., substitute x_0 in place of x and y_0 in place of y, then do the math.

Note that in the case of both partial derivatives, we must *differentiate before we evaluate*. This is exactly the same principle we learned in Calculus I. For instance, say you want to find the slope of the tangent line to the curve $y = x^3$ at the point $(4,64)$. You first differentiate y with respect to x, giving you $y' = 3x^2$. Then you substitute 4 for x, giving you $y' = 3(4)^2 = 48$. You cannot substitute 4 in place of x until *after* you have differentiated!

When finding either partial derivative, the result of Step 2 is a formula which, in general, will involve both x and y. In any particular case, it is possible that either variable could drop out, leaving a formula involving only one variable. It is even possible that both variables will drop out, leaving a constant. However, the general case is a formula involving both x and y.

- When partially differentiating with respect to x, the result of Step 2 is denoted $f_x(x,y)$ or $\frac{\partial f}{\partial x}$.
- When partially differentiating with respect to y, the result of Step 2 is denoted $f_y(x,y)$ or $\frac{\partial f}{\partial y}$.

For $f(x,y) = x^2 + y^2$, $f_x(x,y) = \frac{\partial f}{\partial x} = 2x$ and $f_y(x,y) = \frac{\partial f}{\partial y} = 2y$. At the point $(1,8)$, we get $f_x(1,8) = \frac{\partial f}{\partial x}\big|_{(1,8)} = 2(1) = 2$ and $f_y(1,8) = \frac{\partial f}{\partial y}\big|_{(1,8)} = 2(8) = 16$.

Next, we must learn how do we find a directional derivative for any unit vector other than $\mathbf{i}$ or $\mathbf{j}$. But before we do so, we must introduce a key concept, known as the *gradient vector*.

For any function $f(x,y)$, its **gradient vector** is denoted ∇f, and is defined as $\nabla f = < f_x(x,y),\, f_y(x,y) > = < \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} >$. If this vector is evaluated at a point (x_0,y_0), we obtain $\nabla f(x_0,y_0) = < f_x(x_0,y_0),\, f_y(x_0,y_0) >$.

For $f(x,y) = x^2 + y^2$, $\nabla f = < 2x, 2y >$, and $\nabla f(1,8) = < 2, 16 >$.

For any unit vector $\mathbf{u}$, if we wish to find $D_{\mathbf{u}} f(x_0,y_0)$, simply compute the dot product of $\mathbf{u}$ and $\nabla f(x_0,y_0)$. In other words, $D_{\mathbf{u}} f(x_0,y_0) = \mathbf{u} \cdot \nabla f(x_0,y_0)$.

For instance, if $f(x,y) = x^2 + y^2$ and $\mathbf{u} = < \frac{3}{5}, \frac{4}{5} >$, then $D_{\mathbf{u}} f(1,8) = < \frac{3}{5}, \frac{4}{5} > \cdot < 2, 16 > = \frac{6}{5} + \frac{64}{5} = \frac{70}{5} = 14$.

We already had this result. Let's find a directional derivative where we don't already know the answer. If $f(x,y) = x^2 + y^2$ and $\mathbf{u} = < \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} >$, then $D_{\mathbf{u}} f(1,8) = < \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} > \cdot < 2, 16 > = \frac{2}{\sqrt{2}} + \frac{-16}{\sqrt{2}} = \frac{-14}{\sqrt{2}}$, or $-7\sqrt{2}$.

In all the examples considered so far, we have focused on the point $(1,8)$. There is nothing special about this point. Our formulas apply equally well at any other point. For instance, let us continue to address the function $f(x,y) = x^2 + y^2$, but shift our attention to the point $(-7,13)$. Then:

- $f_x(-7,13) = \frac{\partial f}{\partial x} \,|_{(-7,13)} = -14$
- $f_y(-7,13) = \frac{\partial f}{\partial y} \,|_{(-7,13)} = 26$
- $\nabla f(-7,13) = < -14, 26 >$
- For $\mathbf{u} = < \frac{3}{5}, \frac{4}{5} >$, $D_{\mathbf{u}} f(-7,13) = < \frac{3}{5}, \frac{4}{5} > \cdot < -14, 26 > = \frac{-42}{5} + \frac{104}{5} = \frac{62}{5}$
- For $\mathbf{u} = < \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} >$, $D_{\mathbf{u}} f(-7,13) = < \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} > \cdot < -14, 26 > = \frac{-14}{\sqrt{2}} + \frac{-26}{\sqrt{2}} = \frac{-40}{\sqrt{2}}$, or $-20\sqrt{2}$.

Let us now consider a completely fresh example:
$z = f(x,y) = 3x^5 - 7x^2y^4 + 9y^2 + 4x - 6y + 12$.

- $f_x(x,y) = \frac{\partial f}{\partial x} = 15x^4 - 14xy^4 + 4$
- $f_y(x,y) = \frac{\partial f}{\partial y} = -28x^2y^3 + 18y - 6$
- $f_x(6,-2) = \frac{\partial f}{\partial x} \,|_{(6,-2)} = 18,100$
- $f_y(6,-2) = \frac{\partial f}{\partial y} \,|_{(6,-2)} = 8,022$
- $\nabla f = < 15x^4 - 14xy^4 + 4,\; -28x^2y^3 + 18y - 6 >$
- $\nabla f(6,-2) = < 18,100,\;\; 8,022 >$
- For $\mathbf{u} = < \frac{1}{2}, \frac{\sqrt{3}}{2} >$, $D_{\mathbf{u}} f(6,-2) = < \frac{1}{2}, \frac{\sqrt{3}}{2} > \cdot < 18,100,\;\; 8,022 > = 9,050 + 4,011\sqrt{3}$

In general, since there are infinitely many different unit vectors, a function $z = f(x,y)$ has infinitely many different directional derivatives at a given point (x_0,y_0). Each of these directional derivatives is the slope of a tangent line to the graph of the function f. In other words, the graph of the function is a surface, which we may name S. If $z_0 = f(x_0,y_0)$, then (x_0,y_0,z_0) is a point on the surface S. At this point, there are infinitely many tangent lines–i.e., there are infinitely many lines tangential to the surface (since we can approach (x_0,y_0) along infinitely many different linear paths).

Under a special condition known as **differentiability** (to be discussed shortly), all of these tangent lines lie in a unique plane, which is known as the **tangent plane** to the surface at the point (x_0,y_0,z_0). The equation of the tangent plane is
$z = \nabla f(x_0,y_0) \cdot < x - x_0, y - y_0 > + z_0$, or
$z = f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) + z_0$.

The tangent plane may be thought of as the graph of a linear function, $L(x,y) = f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) + z_0$. This is known as the **linearization** of $f(x,y)$ at the point (x_0,y_0). You can also call it the **linear approximation** of the function at (x_0,y_0).

Notice that $L(x_0,y_0) = 0 + 0 + z_0 = z_0$. Thus, $L(x_0,y_0) = f(x_0,y_0)$. When $(x,y) \neq (x_0,y_0)$, $L(x,y)$ serves as an *approximation* to $f(x,y)$. The approximation is generally good when (x,y) is close to (x_0,y_0), and is generally poor when (x,y) is far away from (x_0,y_0).

Let dx be the deviation of x from x_0, and let dy be the deviation of y from y_0. In other words, $dx = x - x_0$ and $dy = y - y_0$. It follows that $x = x_0 + dx$ and $y = y_0 + dy$, and so $(x,y) = (x_0 + dx,\ y_0 + dy)$.

When (x,y) changes from (x_0,y_0) to $(x_0 + dx,\ y_0 + dy)$, $f(x,y)$ changes from $f(x_0,y_0) = z_0$ to $f(x_0 + dx,\ y_0 + dy)$. We denote this change as Δf.
$\Delta f = f(x_0 + dx,\ y_0 + dy) - f(x_0,y_0) = f(x_0 + dx,\ y_0 + dy) - z_0$.

When (x,y) changes from (x_0,y_0) to $(x_0 + dx,\ y_0 + dy)$, $L(x,y)$ changes from $L(x_0,y_0) = z_0$ to $L(x_0 + dx,\ y_0 + dy)$. We denote this change as ΔL.
$\Delta L = L(x_0 + dx,\ y_0 + dy) - L(x_0,y_0) = L(x_0 + dx,\ y_0 + dy) - z_0$.

Just as $L(x,y) \approx f(x,y)$, likewise $\Delta L \approx \Delta f$.

$L(x_0 + dx,\ y_0 + dy) = f_x(x_0,y_0)(x_0 + dx - x_0) + f_y(x_0,y_0)(y_0 + dy - y_0) + z_0 =$
$f_x(x_0,y_0)dx + f_y(x_0,y_0)dy + z_0$.

So $\Delta L = f_x(x_0,y_0)dx + f_y(x_0,y_0)dy + z_0 - z_0 = f_x(x_0,y_0)dx + f_y(x_0,y_0)dy$.

We define this quantity to be the **differential** of the function f, denoted df. By definition, $df = \Delta L$. Hence $df \approx \Delta f$.

Since we have $z = f(x,y)$, we may write dz in place of df.

All of this is analogous to what we do in Calculus I...

Say we have a function, $y = f(x)$. At x_0, the slope of the tangent line is $f'(x_0)$. If $y_0 = f(x_0)$, then the tangent line has the equation $y - y_0 = f'(x_0)(x - x_0)$, or $y = f'(x_0)(x - x_0) + y_0$. We may think of this as a linear function, $L(x) = f'(x_0)(x - x_0) + y_0$, known as the linearization of $f(x)$ at the point x_0.

Let dx be the deviation of x from x_0. $dx = x - x_0$, so $x = x_0 + dx$.

When x changes from x_0 to $x_0 + dx$, $f(x)$ changes from $f(x_0) = y_0$ to $f(x_0 + dx)$. We denote this change as Δf. $\Delta f == f(x_0 + dx) - f(x_0) = f(x_0 + dx) - y_0$.

When x changes from x_0 to $x_0 + dx$, $L(x)$ changes from $L(x_0) = y_0$ to $L(x_0 + dx)$. We denote this change as ΔL. $\Delta L = L(x_0 + dx) - L(x_0) = L(x_0 + dx) - y_0$. But
$L(x_0 + dx) = f'(x_0)(x_0 + dx - x_0) + y_0 = f'(x_0)dx + y_0$, so
$\Delta L = f'(x_0)dx + y_0 - y_0 = f'(x_0)dx$.

We define this quantity to be the differential of the function f, denoted df, i.e., $df = f'(x_0)dx$. By definition, $df = \Delta L$. Hence $df \approx \Delta f$.

Since we have $y = f(x)$, we may write dy in place of df.