The Tangent Plane, Linearization, And Differentials

The Tangent Plane:

Say we have a function $z = f(x,y)$, whose graph is a surface S, and point (x_0,y_0) in the domain of f. If $z_0 = f(x_0,y_0)$, then (x_0,y_0,z_0) is a point on the surface S. Let $\mathbf{u} = < a,b >$ be a unit vector. Let T be the tangent line at (x_0,y_0) in the direction of $\mathbf{u}$. In other words, T is a line passing through the point (x_0,y_0,z_0), tangential to surface S. The slope of line T is the derivative of f at (x_0,y_0) in the direction of $\mathbf{u}$, $D_{\mathbf{u}} f(x_0,y_0)$. Let us refer to this slope as m. Then the vector equation of T is $\mathbf{r}(t) = < x_0,y_0,z_0 > + t < a,b,m >$.

Let us focus on two particular tangent lines at (x_0,y_0), one in the direction of $\mathbf{i} = < 1,0 >$ and the other in the direction of $\mathbf{j} = < 0,1 >$. We shall refer to these tangent lines as $T_{\mathbf{i}}$ and $T_{\mathbf{j}}$, respectively.

- The slope of $T_{\mathbf{i}}$ is $f_x(x_0,y_0)$, so the vector equation of $T_{\mathbf{i}}$ is $\mathbf{r}(t) = < x_0,y_0,z_0 > + t < 1,0,f_x(x_0,y_0) >$.
- The slope of $T_{\mathbf{j}}$ is $f_y(x_0,y_0)$, so the vector equation of $T_{\mathbf{j}}$ is $\mathbf{r}(t) = < x_0,y_0,z_0 > + t < 0,1,f_y(x_0,y_0) >$.

The direction vectors of $T_{\mathbf{i}}$ and $T_{\mathbf{j}}$ are $< 1,0,f_x(x_0,y_0) >$ and $< 0,1,f_y(x_0,y_0) >$. If these vectors are positioned at the common tail (x_0,y_0,z_0), they determine a unique plane, which is the tangent plane, $\Im$, assuming f is differentiable at (x_0,y_0). To find a normal vector for this plane, we compute the cross product of the direction vectors of $T_{\mathbf{i}}$ and $T_{\mathbf{j}}$.

$$\det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_x(x_0,y_0) \\ 0 & 1 & f_y(x_0,y_0) \end{bmatrix} = \begin{vmatrix} 0 & f_x(x_0,y_0) \\ 1 & f_y(x_0,y_0) \end{vmatrix}\mathbf{i} - \begin{vmatrix} 1 & f_x(x_0,y_0) \\ 0 & f_y(x_0,y_0) \end{vmatrix}\mathbf{j} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\mathbf{k} =$$
$$-f_x(x_0,y_0)\mathbf{i} - f_y(x_0,y_0)\mathbf{j} + 1\mathbf{k} = < -f_x(x_0,y_0), -f_y(x_0,y_0), 1 >.$$

Although this vector could serve as the normal vector for plane $\Im$, it'll be simpler if we use the opposite vector, which is $< f_x(x_0,y_0), f_y(x_0,y_0), -1 >$ (which we would have obtained if we had computed the cross product in the reverse order).

Now we can write the equation of plane $\Im$. Since it contains the point (x_0,y_0,z_0) and has normal vector $< f_x(x_0,y_0), f_y(x_0,y_0), -1 >$, its equation must be $f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) + (-1)(z-z_0) = 0$. We rewrite this as follows:

$f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) - z + z_0 = 0$

$z = f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) + z_0$ Call this Equation #1

$z = f_x(x_0,y_0)x - f_x(x_0,y_0)x_0 + f_y(x_0,y_0)y - f_y(x_0,y_0)y_0 + z_0$

$z = f_x(x_0,y_0)x + f_y(x_0,y_0)y + z_0 - f_x(x_0,y_0)x_0 - f_y(x_0,y_0)y_0$ Call this Equation #2

As previously discussed, the standard form for the equation of a plane is $Ax + By + Cz = D$. For a nonvertical plane (where $C \neq 0$), we can solve for z in terms of x and y, giving us

$z = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$. Equation #2 is in this form.

The standard form equation for $\Im$ is $f_x(x_0,y_0)x + f_y(x_0,y_0)y - z = f_x(x_0,y_0)x_0 + f_y(x_0,y_0)y_0 - z_0$. Call this Equation #3. Here we have:

- $A = f_x(x_0,y_0)$
- $B = f_y(x_0,y_0)$
- $C = -1$
- $D = f_x(x_0,y_0)x_0 + f_y(x_0,y_0)y_0 - z_0$

Let's return our attention to Equation #1. The form of this equation has a special significance that you might not realize. To see the significance, let's go back for a moment to basic algebra. Recall that in the x,y plane, a line with slope m and passing through the point (x_0,y_0) has the equation $y - y_0 = m(x - x_0)$. This equation is said to be in *point, slope form*. It could be rewritten into the form $y = mx + b$, which is *slope, y intercept form*. However, there are times when it's preferable to keep the equation in point, slope form, but to modify that form as $y = m(x - x_0) + y_0$. This was seen in Calculus I. Given a function $f(x)$, its tangent line at the point (x_0,y_0) has slope $f'(x_0)$, so the equation of the tangent line is $y = f'(x_0)(x - x_0) + y_0$. (This concept was generalized in Calculus II when we studied Taylor polynomials. For instance, at the point (x_0,y_0), the function has a tangent *parabola* whose equation is $y = f''(x_0)(x - x_0)^2 + f'(x_0)(x - x_0) + y_0$. We could go on to formulate tangent cubics, tangent quartics, and so on.)

Anyway, if we take the equation $y = m(x - x_0) + y_0$ and "crank it up" an extra dimension, we get $z = m_1(x - x_0) + m_2(y - y_0) + z_0$. This new equation represents a plane rather than a line. Call this plane $\wp$. Just as (x_0,y_0) was a point on the line $y = m(x - x_0) + y_0$, (x_0,y_0,z_0) is a point on plane $\wp$. What is the significance of the coefficients m_1 and m_2, if any? The concept of slope is not directly applicable to a plane, but it is *indirectly* applicable. If we intersect $\wp$ with the vertical plane $y = y_0$ (which is parallel to the x,z plane), we obtain a line whose equation is $z = m_1(x - x_0) + z_0$, and m_1 is the slope of this line. On the other hand, if we intersect $\wp$ with the vertical plane $x = x_0$ (which is parallel to the y,z plane), we obtain a line whose equation is $z = m_2(y - y_0) + z_0$, and m_2 is the slope of this line. Thus, m_1 and m_2 are the slopes of two traces (or cross sections) of the plane $\wp$. On the basis of this insight, it makes sense for us to refer to the equation $z = m_1(x - x_0) + m_2(y - y_0) + z_0$ as **point**, **slope form** for the equation of the plane.

Now we see that Equation #1 is the equation of the tangent plane in point, slope form.

While we're at it, let's take a further look at the equation of a nonvertical plane in the form $z = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$. This is the three-dimensional version of the two-dimensional equation $y = mx + b$, which is the equation of a nonvertical line in the x,y plane. Technically, the y intercept of this line is the point $(0,b)$, but, speaking casually, we can say the y intercept is b. That's why $y = mx + b$ is referred to as the slope, y intercept form of the equation. Analogously, the z intercept of the plane $z = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$ is the point $(0,0,\frac{D}{C})$, but, speaking casually, we can say the z intercept is $\frac{D}{C}$. What is the significance of the coefficients $-\frac{A}{C}$ and $-\frac{B}{C}$, if any? Let us refer to the plane $z = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$ as $\wp$. If we intersect $\wp$ with the vertical plane $y = 0$ (which is the x,z plane), we obtain a line whose equation is $z = (-\frac{A}{C})x + \frac{D}{C}$, and $-\frac{A}{C}$ is the slope of this line. On the other hand, if we intersect $\wp$ with the vertical plane $x = 0$ (which is the y,z plane), we obtain a line whose equation is $z = (-\frac{B}{C})y + \frac{D}{C}$, and $-\frac{B}{C}$ is the slope of this line. Thus, $-\frac{A}{C}$ and $-\frac{B}{C}$ are the slopes of two traces (or cross sections) of the plane $\wp$. On the basis of this insight, it makes sense for us to refer to the equation $z = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$ as **slope**, z **intercept form** for the equation of the plane.

In summary, the equation of the tangent plane can be written in three major forms:

- $z = f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) + z_0$ is **point**, **slope form**.
- $z = f_x(x_0,y_0)x + f_y(x_0,y_0)y + z_0 - f_x(x_0,y_0)x_0 - f_y(x_0,y_0)y_0$ is **slope**, z **intercept form**.
- $f_x(x_0,y_0)x + f_y(x_0,y_0)y - z = f_x(x_0,y_0)x_0 + f_y(x_0,y_0)y_0 - z_0$ is **standard form**.

All three are worthwhile, but point, slope form is the preferred form.

As previously discussed, the function $f(x,y) = x^2 + y^2$ has a tangent plane at $(2,3)$ and its equation is $4x + 6y - z = 13$ in standard form. We have noted that the left side of the equation is $f_x(2,3)x + f_y(2,3)y - z$, which is consistent with our general formula, where the left side is $f_x(x_0,y_0)x + f_y(x_0,y_0)y - z$. The general formula says the right side of the equation should be $f_x(x_0,y_0)x_0 + f_y(x_0,y_0)y_0 - z_0$, i.e., $f_x(2,3)2 + f_y(2,3)3 - f(2,3)$, which is $(4)2 + (6)3 - 13$, which does work out to be 13.

For the function $f(x,y) = x^2 + y^2$, the tangent plane at $(2,3)$ has point, slope equation $z = 4(x-2) + 6(y-3) + 13$, and it has slope, z intercept equation $z = 4x + 6y - 13$.

The right side of the tangent plane's equation in standard form can be expressed as $\nabla f(x_0,y_0) \cdot < x_0,y_0 > -z_0$. The left side can be expressed as $\nabla f(x_0,y_0) \cdot < x,y > -z$. Hence, the standard form equation can be written as $\nabla f(x_0,y_0) \cdot < x,y > -z = \nabla f(x_0,y_0) \cdot < x_0,y_0 > -z_0$. In fact, we could rewrite this as follows:

$\nabla f(x_0,y_0) \cdot < x,y > - \nabla f(x_0,y_0) \cdot < x_0,y_0 > = z - z_0$

$\nabla f(x_0,y_0) \cdot (< x,y > - < x_0,y_0 >) = z - z_0$

$\nabla f(x_0,y_0) \cdot < x-x_0, y-y_0 > = z - z_0$ Call this the **gradient vector form**.

In the case of the function $f(x,y) = x^2 + y^2$, the gradient vector form for the equation of the tangent plane at $(2,3)$ is $< 4,6 > \cdot < x-2, y-3 > = z - 13$.

As previously discussed, the function $f(x,y) = x^2 + y^2$ has gradient vector $< -14,26 >$ at the point $(-7,13)$. Since $z_0 = f(-7,13) = 218$, the tangent plane at $(-7,13)$ has the following equations:

- $< -14,26 > \cdot < x+7, y-13 > = z - 218$ in gradient vector form.
- $z = -14(x+7) + 26(y-13) + 218$ in point, slope form.
- $z = -14x + 26y - 218$ in slope, z intercept form.
- $-14x + 26y - z = 218$ in standard form.

Linearization:

Recall that in two dimensions, a linear function is a function whose graph is a nonvertical line with slope m, and whose equation, in slope, y intercept form, is $y = mx + b$. In three dimensions, a linear function is a function whose graph is a nonvertical plane, and whose equation is $z = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$ in slope, z intercept form, or $z = m_1(x - x_0) + m_2(y - y_0) + z_0$ in point, slope form. A linear function in three dimensions is commonly named $L(x,y)$. Thus, we may write $L(x,y) = (-\frac{A}{C})x + (-\frac{B}{C})y + \frac{D}{C}$, or $L(x,y) = m_1(x - x_0) + m_2(y - y_0) + z_0$.

If the function $z = f(x,y)$ is differentiable at (x_0,y_0), then it has a tangent plane at this point, $\Im$, which is the graph of a linear function, $L(x,y)$. We call this function the **linearization** of f at (x_0,y_0). We have:

- $L(x,y) = f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) + z_0$
- $L(x,y) = f_x(x_0,y_0)x + f_y(x_0,y_0)y + z_0 - f_x(x_0,y_0)x_0 - f_y(x_0,y_0)y_0$

The linearization of f at (x_0,y_0) is also called the **linear approximation** of the function at (x_0,y_0).

The linearization of $f(x,y) = x^2 + y^2$ at $(2,3)$ is $L(x,y) = 4(x - 2) + 6(y - 3) + 13$, or $L(x,y) = z = 4x + 6y - 13$.

For the function $f(x,y) = 7x^2 - 5xy + 2y^3$, $f_x(x,y) = 14x - 5y$ and $f_y(x,y) = -5x + 6y^2$. At the point $(2,1)$, we have $z_0 = f(2,1) = 20$ and $\nabla f(2,1) = < 23,-4 >$, so the tangent plane's equation is $< 23,-4 > \cdot < x - 2, y - 1 > = z - 20$, or $z = 23(x - 2) - 4(y - 1) + 20$, or $z = 23x - 4y - 22$. Hence, the function's linearization at $(2,1)$ is $L(x,y) = 23(x - 2) - 4(y - 1) + 20$, or $L(x,y) = 23x - 4y - 22$.

Differentials:

Say we have a function $z = f(x,y)$, whose linearization at (x_0,y_0) is $L(x,y) = f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) + z_0$. By definition, $z_0 = f(x_0,y_0)$. Notice that $L(x_0,y_0) = 0 + 0 + z_0 = z_0$. Thus, $L(x_0,y_0) = f(x_0,y_0)$. If we refer to the graph of f as surface S, and to the graph of L as plane $\Im$, then the equation $L(x_0,y_0) = f(x_0,y_0)$ means that S and $\Im$ coincide at the point (x_0,y_0,z_0). This is actually quite trivial. All we are saying is that the graph of the function and its tangent plane coincide at the point of tangency.

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).

For any point (x,y) different 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 $\Delta L = f_x(x_0,y_0)dx + f_y(x_0,y_0)dy + z_0 - z_0 = f_x(x_0,y_0)dx + f_y(x_0,y_0)dy$. Note that this could also be expressed as $\nabla f(x_0,y_0) \cdot < dx, dy >$.

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.

To illustrate, consider $f(x,y) = x^2 + y^2$, whose linearization at $(2,3)$ is $L(x,y) = 4(x-2) + 6(y-3) + 13$. Both functions have a value of 13 at $(2,3)$. At $(1,5)$, the values of f and L will differ. $f(1,5) = 26$, whereas $L(1,5) = 21$. 21 is a poor approximation to 26, but that is because $(1,5)$ is relatively far away from $(2,3)$–the distance is $\sqrt{5} \approx 2.24$. Anyway, when (x,y) varies from $(2,3)$ to $(1,5)$, we have $\Delta f = 26 - 13 = 13$ and

$\Delta L = 21 - 13 = 8$. Again, 8 is a poor approximation to 13, but this is because of the relatively large distance between $(2,3)$ and $(1,5)$. Here we have $dx = -1$ and $dy = 2$. Note that $df = \Delta L = \nabla f(2,3) \cdot < -1,2 > = < 4,6 > \cdot < -1,2 > = -4 + 12 = 8$.

Now suppose we have a smaller deviation from $(2,3)$, let's say to the point $(1.8, 3.4)$. $f(1.8, 3.4) = 14.8$, whereas $L(1.8, 3.4) = 14.6$. 14.6 is a good approximation to 14.8. When (x,y) varies from $(2,3)$ to $(1.8, 3.4)$, we have $\Delta f = 14.8 - 13 = 1.8$ and $\Delta L = 14.6 - 13 = 1.6$. 1.6 is a good approximation to 1.8. Here we have $dx = -0.2$ and $dy = 0.4$. Note that $df = \Delta L = \nabla f(2,3) \cdot < -0.2, 0.4 > = < 4,6 > \cdot < -0.2, 0.4 > = -0.8 + 2.4 = 1.6$.