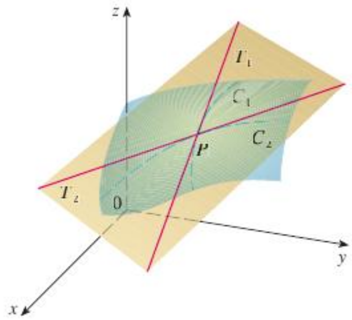


Let T_1 and T_2 be tangent lines to the curves C_1 and C_2 at the point $P(a, b, f(a, b))$. The **tangent plane** to the surface S at P is the plane containing T_1 and T_2.

The tangent plane to S at P consists of all possible tangent lines at P to the curves C that lie on S and pass through P. This plane approximates S at P most closely.

Write the z-coordinate of P as c. Then $P = (a, b, c)$. Equation of any plane passing through P is $z - c = A(x - a) + B(y - b)$. When $y = b$, the tangent plane represents the tangent to the intersected curve at P. Thus, $A = f_x(a, b)$, the slope of the tangent line. Similarly, $B = f_y(a, b)$. Hence equation of the tangent plane to the surface $z = f(x, y)$ at the point $P(a, b, c)$ on S is

$$z - c = f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

provided that f_x, f_y are continuous at (a, b).

Example 1.18. Find the equation of the tangent plane to the elliptic paraboloid $z = 2x^2 + y^2$ at $(1, 1, 3)$.

Here, $z_x = 4x$, $z_y = 2y$. So, $z_x(1, 1) = 4$, $z_y(1, 1) = 2$. Then the equation of the tangent plane is $z - 3 = 4(x - 1) + 2(y - 1)$. It simplifies to $z = 4x + 2y - 3$.

The tangent plane gives a linear approximation to the surface at that point. Why?
Write the equation as $f(x, y) - f(a, b) = f_x(a, b)(x - a) + f_y(a, b)(y - b)$. Then

$$f(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

This formula holds true for all points $(x, y, f(x, y))$ on the tangent plane at $(a, b, f(a, b))$. For approximating $f(x, y)$ for (x, y) close to (a, b), we may take

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

The RHS is called the *standard linear approximation* of $f(x, y, z)$.
Writing in the increment form,

$$f(a + h, b + k) \approx f(a, b) + f_x(a, b)h + f_y(a, b)k.$$

This gives rise to the **total increment** $f(a + h, b + k) - f(a, b)$.

The total increment can be written in a more suggestive form. Towards this, write

$$\Delta f := f(a + h, b + k) - f(a + h, b) + f(a + h, b) - f(a, b).$$

By MVT, there exist $c \in [a, a+h]$ and $d \in [b, b+k]$ such that

$$\begin{aligned} f(a+h,b) - f(a,b) &= h[f_x(c,b) - f_x(a,b)] + hf_x(a,b) \\ f(a+h,b+k) - f(a+h,b) &= k[f_y(a+h,d) - f_y(a,b)] + kf_y(a,b) \end{aligned}$$

Write $\epsilon_1 = f_x(d,b) - f_x(a,b)$ and $\epsilon_2 = f_y(a+h,c) - f_y(a,b)$. When both $h \to 0, k \to 0$, we see that $c \to a$ and $d \to b$. Since f_x and f_y are assumed to be continuous, we have $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$. Then the total increment can be written as

$$\Delta f = f(a+h,b+k) - f(a,b) = hf_x(a,b) + kf_y(a,b) + \epsilon_1 h + \epsilon_2 k,$$

where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as both $h \to 0, k \to 0$.

We also write the increments h, k in x, y as Δx, Δy respectively.

From the above rewriting of Δf it is also clear that $f(x,y)$ is a continuous function. Let us note down what we have proved.

Theorem 1.5. **(Increment Theorem)** *Let D be a region in $\mathbb{R}^2$. Let the function $f : D \to \mathbb{R}$ have continuous first order partial derivatives on D. Then $f(x,y)$ is continuous on D and the total increment $\Delta f = f(a+\Delta x, b+\Delta y) - f(a,b)$ at $(a,b) \in D$ can be written as*

$$\Delta f = f_x(a,b)\Delta x + f_y \Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y,$$

where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as both $\Delta x \to 0$ and $\Delta y \to 0$.

Recall that for a function g of one variable, its differential is defined as $dg = g'(t)dt$.

Let $f(x,y)$ be a given function. The **differential** of f, also called the **total differential**, is

$$df = f_x(x,y)dx + f_y(x,y)dy.$$

Here, $dx = \Delta x$ and $dy = \Delta y$ are the increments in x and y, respectively. Then df is a linear approximation to the *total increment* Δf.

Example 1.19. The dimensions of a rectangular box are measured to be 75cm, 60cm, and 40 cm, and each measurement is correct to within 0.2cm. Use differentials to estimate the largest possible error when the volume of the box is calculated from these measurements.

The volume of the box is $V = xyz$. So,

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz.$$

Given that $|\Delta x|, |\Delta y|, |\Delta z| \leq 0.2$cm, the largest error in cubic cm is

$$|\Delta V| \approx |dV| = 60 \times 40 \times 0.2 + 40 \times 75 \times 0.2 + 75 \times 60 \times 0.2 = 1980.$$

Notice that the relative error is $1980/(75 \times 60 \times 40)$ which is about 1%.

Remark: Let D be a region in $\mathbb{R}^2$. A function $f : D \to \mathbb{R}$ is called *differentiable* at a point $(a,b) \in D$ if the total increment $\Delta z = f(a+\Delta x, b+\Delta y) - f(a,b)$ in f with respect to increments $\Delta x, \Delta y$ in x, y, can be written as

$$\Delta z = f_x(a,b)\Delta x + f_y(a,b)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as both $\Delta x \to 0$ and $\Delta y \to 0$.

The following statements give some connections between differentiability, continuity and the partial derivatives.

- Let D be a region in $\mathbb{R}^2$. Let $f : D \to R$ be such that both f_x and f_y exist on D and at least one of them is continuous at $(a,b) \in D$. Then f is differentiable at (a,b).
- Let D be a region in $\mathbb{R}^2$. Let $f : D \to R$ be differentiable at $(a,b) \in D$. Then f is continuous at (a,b).

Notice that the first statement strengthens the increment theorem. Instead of increasing the load on terminology, we will continue with the increment theorem. Note that whenever we assume that f_x and f_y are continuous, you may replace this with the weaker assumption: " $f(x,y)$ is differentiable".

Remember that we formulate and discuss our results for a function $f(x,y)$ of two variables. Analogously, all the notions and the results can be formulated for a function $f(x_1,\ldots,x_n)$ of n variables for $n \geq 2$.

1.6 Chain Rules

We apply the increment theorem to partially differentiate composite functions.

Theorem 1.6. (Chain Rule 1) *Let $x(t)$ and $y(t)$ be differentiable functions. Let $f(x,y)$ have continuous first order partial derivatives. Then*

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Proof: Using the increment theorem (Theorem 1.5) at a point P we obtain

$$\frac{\Delta f}{\Delta t} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t} + \epsilon_1 \frac{\Delta x}{\Delta t} + \epsilon_2 \frac{\Delta y}{\Delta t}.$$

As $\Delta t \to 0$, we have $\Delta x \to 0, \Delta y \to 0, \epsilon_1 \to 0, \epsilon_2 \to 0$. Then the result follows. □

For example, if $z = xy$ and $x = \sin t$, $y = \cos t$, then

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}x'(t) + \frac{\partial z}{\partial y}y'(t) = \cos^2 t - \sin^2 t.$$

Check: $z(t) = \sin t \cos t = \frac{1}{2}\sin 2t$. So, $z'(t) = \cos 2t = \cos^2 t - \sin^2 t$.

Theorem 1.7. **(Chain Rule 2)** *Let $f(x, y)$ have continuous first order partial derivatives. Suppose $x = x(s, t)$ and $y = y(s, t)$ are functions such that x_s, x_t, y_s and y_t are also continuous. Then*

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}, \quad \frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}.$$

Proof of this follows a similar line to that of Chain Rule - 1. The pattern is clearer if you use the subscript notation:

$$f_s = f_x x_s + f_y y_s, \quad f_t = f_x x_t + f_y y_t.$$

Example 1.20. Let $z = e^x \sin y,\ x = st^2,\ y = s^2 t$. Then

$$\frac{\partial z}{\partial s} = (e^x \sin y)t^2 + (e^x \cos y)2st = te^{st^2}(t\sin(s^2t) + 2s\cos(s^2t)).$$

$$\frac{\partial z}{\partial t} = (e^x \sin y)2st + (e^x \cos y)s^2 = se^{st^2}(2t\sin(s^2t) + s\cos(s^2t)).$$

Substitute expressions for x and y to get $z = z(s, t)$ and then check that the results are correct.

Example 1.21. Given that $z = f(x, y)$ has continuous second order partial derivatives and that $x = r^2 + s^2,\ y = 2rs$, find z_{rr}.

We have $x_r = 2r,\ y_r = 2s$. Then

$$\begin{aligned}
z_r &= 2rz_x + 2sz_y. \\
z_{xr} &= z_{xx}x_r + z_{xy}y_r = 2rz_{xx} + 2sz_{xy}. \\
z_{yr} &= z_{yx}x_r + z_{yy}y_r = 2rz_{yx} + 2sz_{yy}. \\
z_{rr} &= \frac{\partial z_r}{\partial r} = \frac{\partial}{\partial r}(2rz_x + 2sz_y) = 2z_x + 2rz_{xr} + 2sz_{yr} \\
&= 2z_x + 2r(2rz_{xx} + 2sz_{xy}) + 2s(2rz_{yx} + 2sz_{yy}) \\
&= 2z_x + 4r^2z_{xx} + 8rsz_{xy} + 4s^2z_{yy}.
\end{aligned}$$

Functions can be differentiated implicitly. If F is defined within a sphere S containing a point (a, b, c), where $F(a, b, c) = 0,\ F_z(a, b, c) \neq 0$, and F_x, F_y, F_z are continuous inside the sphere, then the equation $F(x, y, z) = 0$ defines a function $z = f(x, y)$ in a sphere containing (a, b, c) and contained in the sphere S. Moreover, the function $z = f(x, y)$ can now be differentiated partially with $z_x = -F_x/F_z,\ z_y = -F_y/F_z$.

It is easier to differentiate implicitly than remembering the formula.

Example 1.22. Find z_x and z_y if $x^3 + y^3 + z^3 + 6xyz = 1$.

We differentiate 'the equation' with respect to x and y as follows:

$$3x^2 + 3z^2z_x + 6y(z + xz_x) = 0 \Rightarrow z_x = -\frac{(x^2 + 2yz)}{z^2 + 2xy}.$$

$$3y^2 + 3z^2z_y + 6x(z + xz_y) = 0 \Rightarrow z_y = -\frac{(y^2 + 2xz)}{z^2 + 2xy}.$$

Example 1.23. Find $\dfrac{dy}{dx}$ if $y = y(x)$ is given by $y^2 = x^2 + \sin(xy)$.

$$2y\frac{dy}{dx} - 2x - \cos(xy)(y + x\frac{dy}{dx}) = 0 \Rightarrow \frac{dy}{dx} = \frac{2x + y\cos(xy)}{2y - x\cos(xy)}.$$

Example 1.24. Find w_x if $w = x^2 + y^2 + z^2$ and $z = x^2 + y^2$.

As it looks,

$$\frac{\partial w}{\partial x} = 2x.$$

However, since $z = x^2 + y^2$, we have $w = x^2 + y^2 + (x^2 + y^2)^2$. Then

$$\frac{\partial w}{\partial x} = 2x + 4x^3 + 4xy^2.$$

Notice that, here we take z as the dependent variable and x, y as independent variables. But suppose we know that x and z are the independent variables and y is the dependent variable. Then the second equation says that $y^2 = z - x^2$. Then $w = x^2 + (z - x^2) + z^2 = z + z^2$. Thus

$$\frac{\partial w}{\partial x} = 0.$$

The correct procedure to get $\partial w/\partial x$ is :

1. w must be dependent variable and x must be independent variable.
2. Decide which of the other variables are dependent or independent.
3. Eliminate the dependent variables from w using the constraints.
4. Then take the partial derivative $\partial w/\partial x$.

Example 1.25. Given that $w = x^2 + y^2 + z^2$ and $z(x, y)$ satisfies $z^3 - xy + yz + y^3 = 1$, evaluate $\partial w/\partial x$ at $(2, -1, 1)$.

It is now clear that z, w are dependent variables and x, y are independent variables. So,

$$\frac{\partial w}{\partial x} = 2x + 2z\frac{\partial z}{\partial x}, \quad 3z^2\frac{\partial z}{\partial x} - y + y\frac{\partial z}{\partial x} = 0.$$

These two together give $\frac{\partial w}{\partial x} = 2x + \frac{2yz}{y+3z^2}$. Evaluating it at $(2, -1, 1)$ gives $\frac{\partial w}{\partial x}(2, -1, 1) = 3$.

A function $f(x, y)$ is called **homogeneous** of degree n in a region $D \subseteq \mathbb{R}^2$ if for all $(x, y) \in D$, and for each positive λ, $f(\lambda x, \lambda y) = \lambda^n f(x, y)$.

For example, $f(x, y) = x^{1/3}y^{-4/3}\tan^{-1}(y/x)$ is homogeneous of degree -1 in the region D, which is any quadrant without the axes.

$f(x, y) = (\sqrt{x^2 + y^2})^3$ is homogeneous of degree 3 in the whole plane.

Theorem 1.8. (Euler) *Let D be a region in $\mathbb{R}^2$. Let $f : D \to \mathbb{R}$ have continuous first order partial derivatives. Then f is a homogeneous function of degree n iff $xf_x + yf_y = nf$.*

Proof: Differentiate $f(\lambda x, \lambda y) - \lambda^n f(x, y) = 0$ partially with respect to λ to obtain:

$$x f_x(\lambda x, \lambda y) + y f_y(\lambda x, \lambda y) = n\lambda^{n-1} f(x, y).$$

Then set $\lambda = 1$ to get $x f_x(x, y) + y f_y(x, y) = n f(x, y)$.
Conversely, let $(a, b) \in D$. Define $\phi(\lambda) = f(\lambda a, \lambda b)$. Differentiate with respect to λ to get

$$\lambda\phi'(\lambda) = a f_x(\lambda a, \lambda b) + b f_y(\lambda a, \lambda b).$$

However,

$$n f(\lambda a, \lambda b) = a f_x(\lambda a, \lambda b) + b f_y(\lambda a, \lambda b) = \lambda a f_x(\lambda a, \lambda b) + \lambda b f_y(\lambda a, \lambda b).$$

That is,

$$\lambda\phi'(\lambda) = n\phi(\lambda).$$

Now, differentiate $\lambda^{-n}\phi(\lambda)$ with respect to λ to obtain

$$[\phi(\lambda)\lambda^{-n}]' = \phi'(\lambda)\lambda^{-n} - n\phi(\lambda)\lambda^{-n-1} = 0.$$

Therefore, $\phi(\lambda)\lambda^{-n} = c$ for some constant c. Set $\lambda = 1$ to get $c = f(a, b)$. Then

$$f(\lambda a, \lambda b) = \lambda^n f(a, b).$$

Since (a, b) is any arbitrary point in D, we have $f(\lambda x, \lambda y) = \lambda^n f(x, y)$. □

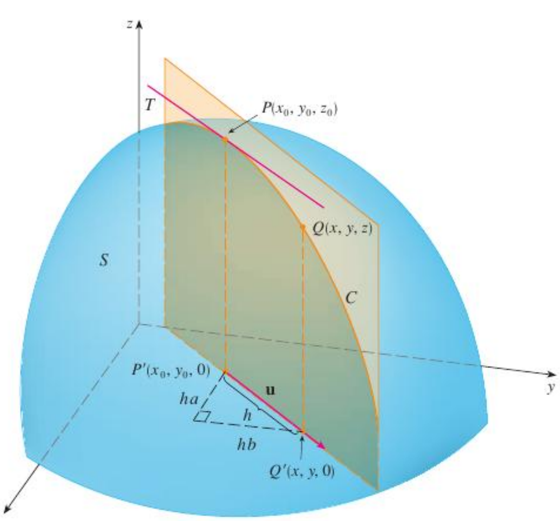
For our earlier examples, you can check that

$$x\frac{\partial[x^{1/3}y^{-4/3}\tan^{-1}(y/x)]}{\partial x} + y\frac{\partial[x^{1/3}y^{-4/3}\tan^{-1}(y/x)]}{\partial y} + x^{1/3}y^{-4/3}\tan^{-1}(y/x) = 0.$$

$$x\frac{\partial[(\sqrt{x^2+y^2})^3]}{\partial x} + y\frac{\partial[(\sqrt{x^2+y^2})^3]}{\partial y} = 3[(\sqrt{x^2+y^2})^3].$$

1.7 Directional Derivative

Recall that if $f(x, y)$ is a function, then $f_x(x_0, y_0)$ is the rate of change in f with respect to change in x, at (x_0, y_0), that is, in the direction $\hat{\imath}$. Similarly, $f_y(x_0, y_0)$ is the rate of change at (x_0, y_0) in the direction $\hat{\jmath}$. How do we find the rate of change of $f(x, y)$ at (x_0, y_0) in the direction of any unit vector $\hat{u}$?



Consider the surface S with the equation $z = f(x, y)$. Let $z_0 = f(x_0, y_0)$. The point $P(x_0, y_0, z_0)$ lies on S. The vertical plane that passes through P in the direction of $\hat{u}$ (containing $\hat{u}$) intersects S in a curve C. The slope of the tangent line T to the curve C at the point P is the rate of change of z in the direction of $\hat{u}$.

Let $f(x, y)$ be a function defined in a region D. Let $(x_0, y_0) \in D$. The **directional derivative** of $f(x, y)$ in the direction of a unit vector $\hat{u} = a\hat{\imath} + b\hat{\jmath}$ at (x_0, y_0) is given by

$$(D_u f)(x_0, y_0) = \left(\frac{df}{ds}\right)_u \Big|_{(x_0,y_0)} = \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}.$$

Example 1.26. Find the derivative of $z = x^2 + y^2$ at $(1, 2)$ in the direction $\hat{u} = (1/\sqrt{2})\hat{\imath} + (1/\sqrt{2})\hat{\jmath}$.

$$D_u z(1, 2) = \lim_{h\to 0} \frac{f(1 + h/\sqrt{2}, 2 + h/\sqrt{2}) - f(1, 2)}{h} = \lim_{h\to 0} \frac{2h/\sqrt{2} + 2 \cdot 2h/\sqrt{2}}{h} = \frac{6}{\sqrt{2}}.$$

Notice that $f_x(1, 2)(1/\sqrt{2}) + f_y(1, 2)(1/\sqrt{2}) = (2 + 2(2)) \cdot (1/\sqrt{2}) = 6/\sqrt{2}$.

Theorem 1.9. *Let $f(x, y)$ have continuous first order partial derivatives. Then $f(x, y)$ has a directional derivative at (x, y) in any direction $\hat{u} = a\hat{\imath} + b\hat{\jmath}$; and it is given by*

$$D_u f(x, y) = f_x(x, y)a + f_y(x, y)b.$$

Proof: Let (x_0, y_0) be a point in the domain of definition of $f(x, y)$. Define the function $g(\cdot)$ by $g(h) = f(x_0 + ah, y_0 + bh)$. Then $g(h)$ is a continuously differentiable function of h. Now,

$$g'(h) = f_x \frac{dx}{dh} + f_y \frac{dy}{dh} = f_x\, a + f_y\, b.$$

Then $g'(0) = f_x(x_0, y_0)\, a + f_y(x_0, y_0)\, b$. Also,

$$g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h} = D_u f(x_0, y_0).$$

Hence $D_u f(x_0, y_0) = g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$. □

Example 1.27. Find the directional derivative of $f(x, y) = x^3 - 3xy + 4y^2$ in the direction of the line that makes an angle of $\pi/6$ with the x-axis.

Here, the direction is given by the unit vector $\hat{u} = \cos(\pi/6)\hat{\imath} + \sin(\pi/6)\hat{\jmath} = \frac{\sqrt{3}}{2}\hat{\imath} + \frac{1}{2}\hat{\jmath}$. Thus

$$D_u f(x, y) = \frac{\sqrt{3}}{2} f_x + \frac{1}{2} f_y = \frac{\sqrt{3}}{2}(3x^2 - 3y) + \frac{1}{2}(-3x + 8y) = \frac{1}{2}\Big[3\sqrt{3}x^2 - 3x + (8 - 3\sqrt{3})y\Big].$$

The formula for the directional derivative in the direction of the unit vector $\hat{u} = a\hat{\imath} + b\hat{\jmath}$ can be written as

$$D_u f = f_x a + f_y b = (f_x\hat{\imath} + f_y\hat{\jmath}) \cdot (a\hat{\imath} + b\hat{\jmath}).$$

The vector operator $\nabla := \frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath}$ is called the **gradient** and the **gradient of** $f(x, y)$ is

$$\nabla f := \text{grad } f := \frac{\partial f}{\partial x}\hat{\imath} + \frac{\partial f}{\partial y}\hat{\jmath}.$$

Therefore, $D_u f = \text{grad } f \cdot \hat{u}$. That is, at (x_0, y_0), the directional derivative is given by

$$D_u f|_{(x_0,y_0)} = \text{grad } f|_{(x_0,y_0)} \cdot \hat{u}.$$

For example, for the function $f(x, y) = xe^y + \cos(xy)$, grad $f|_{(2,0)} = \hat{\imath} + 2\hat{\jmath}$. Thus, the directional derivative of f in the direction of $3\hat{\imath} - 4\hat{\jmath}$ is grad $f|_{(1,2)} \cdot ((3/5)\hat{\imath} - (4/5)\hat{\jmath}) = -1$.

Caution: To apply this formula, we have assumed that f_x, f_y are continuous at (x_0, y_0), and $\hat{u}$ is a unit vector.

Example 1.28. How much the value of $y \sin x + 2yz$ change if the point (x, y, z) moves 0.1 units from $(0, 1, 0)$ toward $(2, 2, -2)$?

Let $f(x, y, z) = y \sin x + 2yz$. $P(0, 1, 0)$, $Q(2, 2, -2)$. $\vec{v} = \overrightarrow{PQ} = 2\hat{\imath} + \hat{\jmath} - 2\hat{k}$. The unit vector in the direction of $\vec{v}$ is $\hat{u} = \frac{1}{3}\vec{v}$. We find D_u at P which requires grad f.

$$\text{grad } f = (y \cos x)\hat{\imath} + (\sin x + 2z)\hat{\jmath} + 2y\,\hat{k}.$$

Then

$$D_u(P) = \text{grad } f|_{(0,1,0)} \cdot \vec{u} = (\hat{\imath} + 2\hat{k}) \cdot \hat{u} = -\frac{2}{3}.$$

The change df in the direction of $\vec{u}$ in moving $ds = 0.1$ units is approximately

$$df \approx D_u(P)\, ds = -\frac{2}{3}\,(0.1) = -0.067 \text{ units}.$$

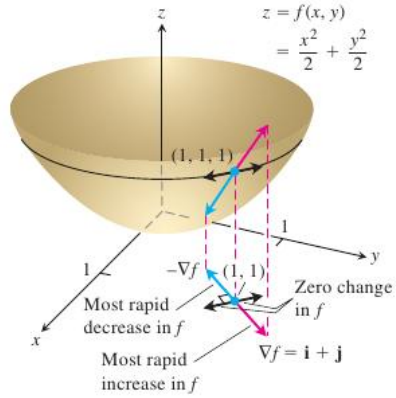
Theorem 1.10. *Let $f(x, y)$ have continuous first order partial derivatives. The maximum value of the directional derivative $D_u f(x, y)$ is* $|\text{grad } f|$ *and it is achieved when the unit vector $\hat{u}$ has the same direction as that of* grad f.

This is obvious since $D_u f = \text{grad } f \cdot \hat{u}$ says that the directional derivative is the scalar projection of the gradient in the direction of $\hat{u}$.

Proof: $D_u f = \text{grad } f \cdot \hat{u} = |\text{grad } f|\,|\hat{u}|\,\cos\theta = |\text{grad } f| \cos\theta$, where θ is the angle between grad f and $\hat{u}$. Since maximum of $\cos\theta$ is 1, maximum of $D_u f$ is $|\text{grad } f|$. The maximum is achieved when $\theta = 0$, that is, when the directions of grad f and $\hat{u}$ coincide. □

This also says the following:

$f(x, y)$ increases most rapidly in the direction of its gradient.
$f(x, y)$ decreases most rapidly in the opposite direction of its gradient.
$f(x, y)$ remains constant in any direction orthogonal to its gradient.



Example 1.29. Find the directions in which the function $f(x, y) = (x^2 + y^2)/2$ changes most, least, and not at all, at the point $(1, 1)$.

Note: When we ask for a direction, we mean a unit vector.

$$\text{grad } f = f_x\hat{\imath} + f_y\hat{\jmath} = x\hat{\imath} + y\hat{\jmath}. \quad (\text{grad } f)(1, 1) = \hat{\imath} + \hat{\jmath}.$$

Thus the function $f(x, y)$ increases most at $(1, 1)$ in the direction $(\hat{\imath} + \hat{\jmath})/\sqrt{2}$. It decreases most at $(1, 1)$ in the direction $-(\hat{\imath} + \hat{\jmath})/\sqrt{2}$. And it does not change at $(1, 1)$ in the directions $\pm(\hat{\imath} - \hat{\jmath})/\sqrt{2}$.

1.8 Normal to Level Curve and Tangent Planes

Let $z = f(x, y)$ be a given surface. Assume that f_x and f_y are continuous. Recall that a level curve to this surface is a curve in the plane where $f(x, y)$ is a constant. Fix some constant c in the range of f. On the corresponding level curve, $f(x, y)$ takes the constant value c. Suppose $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath}$ is a parametrization of this level curve.
Differentiating, we have $\frac{d}{dt}f(x(t), y(t)) = 0$. Or,

$$f_x\frac{dx}{dt} + f_y\frac{dy}{dt} = \text{grad } f \cdot \frac{d\vec{r}(t)}{dt} = 0.$$

Since $d\vec{r}/dt$ is the tangent to the curve, grad f is the normal to the level curve. That is,

Let $f(x, y)$ have continuous first order partial derivatives. At any point (x_0, y_0) in the domain of $f(x, y)$, its gradient grad f is the normal to the level curve that passes through (x_0, y_0), provided grad f is nonzero at (x_0, y_0).

In higher dimensions, if $f(x_1, \ldots, x_n)$ is a function of n independent variables defined on $D \subseteq \mathbb{R}^n$, then its gradient at any point is

$$\text{grad } f = \Big(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\Big).$$

The directional derivative at any point $\vec{x}$ in the direction of a unit vector $\hat{u} = (u_1, \ldots, u_n)$ is

$$D_u f = \lim_{h\to 0}\frac{f(\vec{x} + h\hat{u}) - f(\vec{x})}{h} = \text{grad } f \cdot \hat{u} = f_{x_1}u_1 + \cdots + f_{x_n}u_n.$$

The algebraic rules for the gradient are as follows:

1. Constant multiple: grad $(kf) = k(\text{grad } f)$ for $k \in \mathbb{R}$.
2. Sum: grad $(f + g) = \text{grad } f + \text{grad } g$.
3. Difference: grad $(f - g) = \text{grad } f - \text{grad } g$.
4. Product: grad $(fg) = f(\text{grad } g) + g(\text{grad } f)$.
5. Quotient: grad $(f/g) = \dfrac{g(\text{grad } f) - f(\text{grad } g)}{g^2}$.

In $\mathbb{R}^3$, let $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$ be a smooth curve on the level surface $f(x, y, z) = c$. Then $f(x(t), y(t), z(t)) = c$ for all t. Differentiating this we get

$$\text{grad } f \cdot \vec{r}\,'(t) = 0.$$

Look at all such smooth curves that pass through a point P on the level surface. The velocity vectors $\vec{r}\,'(t)$ to all these smooth curves are orthogonal to the gradient at P.

Let $f(x, y, z)$ have continuous partial derivatives f_x, f_y, and f_z. The **tangent plane** at $P(x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$ is the plane through P which is orthogonal to grad f at P. Its equation is

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

The **normal line** to the level surface $f(x, y, z) = c$ at $P(x_0, y_0, z_0)$ is the line through P parallel to grad f. Its parametric equation is

$$x = x_0 + f_x(x_0, y_0, z_0)\,t,\;\; y = y_0 + f_y(x_0, y_0, z_0)\,t,\;\; z = z_0 + f_z(x_0, y_0, z_0)\,t.$$

The equation of the tangent plane to the surface $z = f(x, y)$ at (a, b) can be obtained as follows: Write the surface as $F(x, y, z) = 0$, where $F(x, y, z) = f(x, y) - z$. Then $F_x = f_x$, $F_y = f_y$, $F_z = -1$. Then the equation of the tangent plane is

$$f_x(a, b)(x - a) + f_y(a, b)(y - b) - (z - f(a, b)) = 0.$$

Example 1.30. Find the tangent plane and the normal line of the surface $x^2 + y^2 + z - 9 = 0$ at the point $(1, 2, 4)$.

First, check that the point $(1, 2, 4)$ lies on the surface. Next, $f_x(1, 2, 4) = 2$, $f_y(1, 2, 4) = 4$ and $f_z(1, 2, 4) = 1$. The tangent plane is given by

$$2(x - 1) + 4(y - 2) + (z - 4) = 0.$$

The normal line at $(1, 2, 4)$ is given by

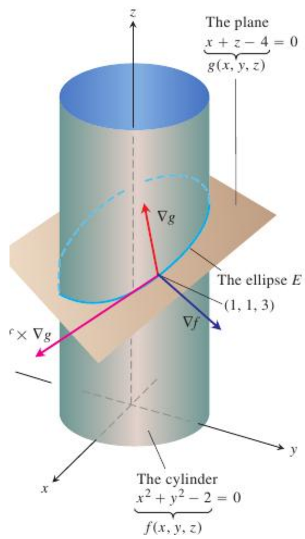
$$x = 1 + 2t,\;\; y = 2 + 4t,\;\; z = 4 + t.$$

Example 1.31. Find the tangent plane to the surface $z = x\cos y - ye^x$ at the origin.

$f_x(0, 0) = 1$, $f_y(0, 0) = -1$. The tangent plane is

$$x - y - z = 0.$$

Example 1.32. Find the tangent line to the curve of intersection of the surfaces $f(x, y, z) := x^2 + y^2 - 2 = 0$ and $g(x, y, z) := x + z - 4 = 0$ at the point $(1, 1, 3)$.



The tangent line is orthogonal to both grad f and grad g at $(1, 1, 3)$. So, it is parallel to

$$\text{grad } f \times \text{grad } g = (2\hat{\imath} + 2\hat{\jmath}) \times (\hat{\imath} + \hat{k}) = 2\hat{\imath} - 2\hat{\jmath} - 2\hat{k}.$$

Thus the tangent line is $x = 1 + 2t,\ y = 1 - 2t,\ z = 3 - 2t$.

1.9 Taylor's Theorem

For a function of one variable, a polynomial approximation is given by the Taylor's formula. Observe that it is a generalization of the Mean value theorem.

Theorem 1.11. (Taylor's Formula for one variable) *Let $n \in \mathbb{N}$. Suppose that $f^{(n)}(x)$ is continuous on $[a, b]$ and is differentiable on (a, b). Then there exists a point $c \in (a, b)$ such that*

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + \frac{f^{(n+1)}(c)}{(n + 1)!}(x - a)^{n+1}.$$

Proof: For $x = a$, the formula holds. So, let $x \in (a, b]$. For any $t \in [a, x]$, let

$$p(t) = f(a) + f'(a)(t - a) + \frac{f''(a)}{2!}(t - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(t - a)^n.$$

Here, we treat x as a certain point, not a variable; and t as a variable. Write

$$g(t) = f(t) - p(t) - \frac{f(x) - p(x)}{(x - a)^{n+1}}\,(t - a)^{n+1}.$$

We see that $g(a) = 0,\ g'(a) = 0,\ g''(a) = 0,\ \ldots, g^{(n)}(a) = 0,$ and $g(x) = 0$.

By Rolle's theorem, there exists $c_1 \in (a, x)$ such that $g'(c_1) = 0$. Since $g(a) = 0$, apply Rolle's theorem once more to get a $c_2 \in (a, c_1)$ such that $g''(c_2) = 0$.

Continuing this way, we get a $c_{n+1} \in (a, c_n)$ such that $g^{(n+1)}(c_{n+1}) = 0$.

Since $p(t)$ is a polynomial of degree at most n, $p^{(n+1)}(t) = 0$. Then

$$g^{(n+1)}(t) = f^{(n+1)}(t) - \frac{f(x) - p(x)}{(x-a)^{n+1}}\,(n+1)!.$$

Evaluating at $t = c_{n+1}$ we have $f^{(n+1)}(c_{n+1}) - \dfrac{f(x) - p(x)}{(x-a)^{n+1}}\,(n+1)! = 0$. That is,

$$\frac{f(x) - p(x)}{(x-a)^{n+1}} = \frac{f^{(n+1)}(c_{n+1})}{(n+1)!}.$$

Consequently, $g(t) = f(t) - p(t) - \dfrac{f^{(n+1)}(c_{n+1})}{(n+1)!}\,(t-a)^{n+1}$.
Evaluating it at $t = x$ and using the fact that $g(x) = 0$, we get

$$f(x) = p(x) + \frac{f^{(n+1)}(c_{n+1})}{(n+1)!}\,(x-a)^{n+1}.$$

Since x is an arbitrary point in $(a, b]$, this completes the proof. □

We have a similar result for functions of several variables.

Theorem 1.12. (Taylor) *Let $D \subseteq \mathbb{R}^2$ be a region. Let (a, b) be an interior point of D. Let $f : D \to \mathbb{R}$ have continuous partial derivatives of order up to $n+1$ in some open disk D_0 centered at (a, b) and contained in D. Then for any $(a+h, b+k) \in D_0$, there exists $\theta \in [0, 1]$ such that*

$$\begin{aligned} f(a+h, b+k) &= f(a,b) + \sum_{m=1}^{n} \frac{1}{m!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^m f(a,b) \\ &+ \frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n+1} f(a+\theta h, b+\theta k). \end{aligned}$$

For example, $m = 2$ on the right gives $\frac{1}{2!}(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy})$.
Proof: Let $\phi(t) = f(a+th, b+tk)$. For any $t \in [0, 1]$,

$\phi'(t) = f_x(a+th, b+tk)h + f_y(a+th, b+tk)k = (h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})f(a+th, b+tk)$.

$\phi^{(2)}(t) = (f_{xx}h + f_{xy}k)h + (f_{yx}h + f_{yy}k)k = (h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^2 f(a+th, b+tk)$.

By induction, it follows that

$$\phi^{(m)}(t) = (h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^m f(a+th, b+tk).$$

Using Taylor's formula for the single variable function $\phi(t)$, we have

$$\phi(1) = \phi(0) + \sum_{m=1}^{n} \frac{\phi^{(m)}(0)}{m!} + \frac{\phi^{(n+1)}(\theta)}{(n+1)!} \quad \text{for some } \theta \in [0, 1].$$

Substituting the expressions for $\phi(1)$, $\phi(0)$, $\phi^{(m)}(0)$ and $\phi^{(n+1)}(\theta)$, we get the required result. □

Example 1.33. Let $f(x,y) = x^2 + xy - y^2,\ a = 1,\ b = -2$.
Here, $f(1,-2) = -5,\ f_x(1,-2) = 0,\ f_y(1,-2) = 5,\ f_{xx} = 2,\ f_{xy} = 1,\ f_{yy} = -2$. Then

$$f(x,y) = -5 + 5(y+2) + \frac{1}{2}\left[2(x-1)^2 + 2(x-1)(y+2) - 2(y+2)^2\right].$$

This becomes exact, since third (and more) order derivatives are 0.

Recall that the standard linearization (linear approximation) of $f(x,y)$ at (a,b) is

$$L(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

The **error** in the standard linearization at (a,b) can now be written as

$$E(x,y) = f(x,y) - L(x,y) = \frac{1}{2!}((x-a)^2 f_{xx} + 2(x-a)(y-b)f_{xy} + (y-b)^2 f_{yy})\Big|_{(c,d)}$$

where $c = a + \theta(x-a),\ d = b + \theta(y-b)$ for some $\theta \in [0,1]$.

Theorem 1.13. *Let $D \subseteq \mathbb{R}^2$ be a region. Let $f : D \to \mathbb{R}$ have continuous first and second order partial derivatives. Let R be a rectangle centered at (a,b) and contained in D. Suppose there exists an $M \in \mathbb{R}$ such that $|f_{xx}|, |f_{xy}|, |f_{yy}| \le M$ for all points in R. Then*

$$|E(x,y)| \le \frac{1}{2}M(|x-a| + |y-b|)^2.$$

Proof: Taylor's formula says that $f(x,y) = L(x,y) + E(x,y)$, where

$$E(x,y) = \frac{1}{2}\Big[(x-a)^2 f_{xx}(c,d) + 2(x-a)(y-b)f_{xy}(c,d) + (y-b)^2 f_{yy}(c,d)\Big].$$

for some c in between x and a, and some d in between y and b. Since $|f_{xx}| \le M$, $|f_{xy}| \le M$, and $|f_{yy}| \le M$ for all points in R,

$$|E(x,y)| \le \frac{M}{2}|(x-a)^2 + 2(x-a)(y-b) + (y-b)^2| \le \frac{M}{2}(|(x-a| + |y-b|)^2. \qquad \square$$

Example 1.34. Find the standard linearization of $f(x,y) = x^2 - xy + y^2/2 + 3$ at $(3,2)$. Also find an upper bound of the error in the linearization in the rectangle $|x-3| \le 0.1,\ |y-2| \le 0.1$.

The standard linearization (linear approximation) of $f(x,y)$ at (a,b) is

$$L(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Now, $f(3,2) = 8,\ f_x(3,2) = (2x-y)|_{(3,2)} = 4$ and $f_y(3,2) = (-x+y)|_{(3,2)} = -1$. Thus

$$L(x,y) = 8 + 4(x-3) - (y-2) = 4x - y - 2.$$

The error in this linearization is

$$E(x,y) = f(x,y) - L(x,y) = x^2 - xy + y^2/2 + 3 - 4x + y + 2.$$

The rectangle is $R:\ |x-3| \le 0.1, |y-2| \le 0.1$. Here, $f_{xx} = 2,\ f_{xy} = -1,\ f_{yy} = 1$.
So, we take $M = 2$ as an upper bound for their absolute values. Then

$$|E(x,y)| \le |x-3|^2 + |y-2|^2 \le (0.1 + 0.1)^2 = 0.04.$$