Chain Rule and Total Differentials

1. Find the total differential of $w = x^3yz + xy + z + 3$ at $(1, 2, 3)$.

<u>Answer:</u> The total differential at the point (x_0, y_0, z_0) is

$$dw = w_x(x_0, y_0, z_0)\, dx + w_y(x_0, y_0, z_0)\, dy + w_z(x_0, y_0, z_0)\, dz.$$

In our case,

$$w_x = 3x^2yz + y, \quad w_y = x^3z + x, \quad w_z = x^3y + 1.$$

Substituting in the point $(1, 2, 3)$ we get: $w_x(1, 2, 3) = 20,\ w_y(1, 2, 3) = 4,\ w_z(1, 2, 3) = 3$.
Thus,

$$dw = 20\, dx + 4\, dy + 3\, dz.$$

2. Suppose $w = x^3yz + xy + z + 3$ and

$$x = 3\cos t, \quad y = 3\sin t, \quad z = 2t.$$

Compute $\dfrac{dw}{dt}$ and evaluate it at $t = \pi/2$.

<u>Answer:</u> We do not substitute for x, y, z before differentiating, so we can practice the chain rule.

$$\begin{aligned} \frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\ &= (3x^2yz + y)(-3\sin t) + (x^3z + x)(3\cos t) + (x^3y + 1)(2). \end{aligned}$$

At $t = \pi/2$ we have $x = 0,\ y = 3,\ z = \pi, \qquad \sin\pi/2 = 1, \cos\pi/2 = 0.$
Thus,

$$\left.\frac{dw}{dt}\right|_{\pi/2} = 3(-3) + 3(0) + (1)2 = -7.$$

3. Show how the tangent approximation formula leads to the chain rule that was used in the previous problem.

<u>Answer:</u> The approximation formula is

$$\Delta w \approx \left.\frac{\partial f}{\partial x}\right|_o \Delta x + \left.\frac{\partial f}{\partial y}\right|_o \Delta y + \left.\frac{\partial f}{\partial z}\right|_o \Delta z.$$

If x, y, z are functions of time then dividing the approximation formula by Δt gives

$$\frac{\Delta w}{\Delta t} \approx \left.\frac{\partial f}{\partial x}\right|_o \frac{\Delta x}{\Delta t} + \left.\frac{\partial f}{\partial y}\right|_o \frac{\Delta y}{\Delta t} + \left.\frac{\partial f}{\partial z}\right|_o \frac{\Delta z}{\Delta t}.$$

In the limit as $\Delta t \to 0$ we get the chain rule.

Note: we use the regular 'd' for the derivative $\frac{dw}{dt}$ because in the chain of computations

$$t \to x, y, z \to w$$

the dependent variable w is ultimately a function of exactly one independent variable t. Thus, the derivative with respect to t is not a partial derivative.

Chain rule

Now we will formulate the chain rule when there is more than one independent variable.

We suppose w is a function of x, y and that x, y are functions of u, v. That is,

$$w = f(x, y) \text{ and } x = x(u, v),\ y = y(u, v).$$

The use of the term chain comes because to compute w we need to do a chain of computations

$$(u, v) \to (x, y) \to w.$$

We will say w is a *dependent* variable, u and v are *independent* variables and x and y are *intermediate* variables.

Since w is a function of x and y it has partial derivatives $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial w}{\partial y}$.

Since, ultimately, w is a function of u and v we can also compute the partial derivatives $\dfrac{\partial w}{\partial u}$ and $\dfrac{\partial w}{\partial v}$. The chain rule relates these derivatives by the following formulas.

$$\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} \\
\frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v}.
\end{aligned}$$

Example: Given $w = x^2y + y^2 + x$, $x = u^2v$, $y = uv^2$ find $\dfrac{\partial w}{\partial u}$.

<u>**Answer:**</u> First we compute

$$\frac{\partial w}{\partial x} = 2xy + 1,\ \ \frac{\partial w}{\partial y} = x^2 + 2y,\ \ \frac{\partial x}{\partial u} = 2uv,\ \ \frac{\partial y}{\partial u} = v^2,\ \ \frac{\partial x}{\partial v} = u^2,\ \ \frac{\partial y}{\partial v} = 2uv.$$

The chain rule then implies

$$\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} \\
&= (2xy + 1)2uv + (x^2 + 2y)v^2 \\
\frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} \\
&= (2xy + 1)u^2 + (x^2 + 2y)2uv.
\end{aligned}$$

Often, it is okay to leave the variables mixed together. If, for example, you wanted to compute $\frac{\partial w}{\partial u}$ when $(u, v) = (1, 2)$ all you have to do is compute x and y and use these values, along with u, v, in the formula for $\dfrac{\partial w}{\partial u}$.

$$x = 2,\ y = 4 \ \Rightarrow\ \frac{\partial w}{\partial u} = (5)(4) + (12)(4) = 68.$$

If you actually need the derivatives expressed in just the variables u and v then you would have to substitute for x, y and z.

Proof of the chain rule:
Just as before our argument starts with the tangent approximation at the point (x_0, y_0).

$$\Delta w \approx \left.\frac{\partial w}{\partial x}\right|_o \Delta x + \left.\frac{\partial w}{\partial y}\right|_o \Delta y.$$

Now hold v constant and divide by Δu to get

$$\frac{\Delta w}{\Delta u} \approx \left.\frac{\partial w}{\partial x}\right|_o \frac{\Delta x}{\Delta u} + \left.\frac{\partial w}{\partial y}\right|_o \frac{\Delta y}{\Delta u}.$$

Finally, letting $\Delta u \to 0$ gives the chain rule for $\dfrac{\partial w}{\partial u}$.

Ambiguous notation

Often you have to figure out the dependent and independent variables from context.

Thermodynamics is a big player here. It has, for example, the variables P, T, V, U, S. and *any* two can be taken to be independent and the others are functions of those two.

We will do more with this topic in the future.

Chain rule with more variables

1. Let $w = xyz$, $x = u^2v$, $y = uv^2$, $z = u^2 + v^2$.

a) Use the chain rule to find $\dfrac{\partial w}{\partial u}$.

b) Find the total differential dw in terms of du and dv.

c) Find $\dfrac{\partial w}{\partial u}$ at the point $(u, v) = (1, 2)$.

Answer: a) The chain rule says

$$\begin{aligned} \frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u} \\ &= (yz)(2uv) + (xz)(v^2) + (xy)(2u). \end{aligned}$$

b) Using the formulas given we get

$$dw = yz\,dx + xz\,dy + xy\,dz$$

and

$$dx = 2uv\,du + u^2\,dv, \quad dy = v^2\,du + 2uv\,dv, \quad dz = 2u\,du + 2v\,dv.$$

Substituting for dx, dy, dz in the equation for dw gives

$$\begin{aligned} dw &= (yz)(2uv\,du + u^2\,dv) + (xz)(v^2\,du + 2uv\,dv) + (xy)(2u\,du + 2v\,dv). \\ &= (2yzuv + xzv^2 + 2xyu)\,du + (yzu^2 + 2xzuv + 2xyv)\,dv. \end{aligned}$$

Therefore

$$\frac{\partial w}{\partial u} = 2yzuv + xzv^2 + 2xyu \ \text{ and } \ \frac{\partial w}{\partial v} = yzu^2 + 2xzuv + 2xyv.$$

c) We do the chain of computations to compute the partial.

$$(u, v) = (1, 2) \ \Rightarrow \ (x, y, z) = (2, 4, 5) \ \Rightarrow \ \frac{\partial w}{\partial u} = (20)(4) + (10)(4) + (8)(2) = 136.$$

Gradient: definition and properties

Definition of the gradient

If $w = f(x, y)$, then $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial w}{\partial y}$ are the rates of change of w in the $\mathbf{i}$ and $\mathbf{j}$ directions.

It will be quite useful to put these two derivatives together in a vector called the *gradient* of w.

$$\operatorname{grad} w = \left\langle \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \right\rangle.$$

We will also use the symbol $\boldsymbol{\nabla} w$ to denote the gradient. (You read this as 'gradient of w' or 'grad w'.)

Of course, if we specify a point $P_0 = (x_0, y_0)$, we can evaluate the gradient at that point. We will use several notations for this

$$\operatorname{grad} w(x_0, y_0) = \boldsymbol{\nabla}\, w|_{P_0} = \boldsymbol{\nabla}\, w|_o = \left\langle \left.\frac{\partial w}{\partial x}\right|_o, \left.\frac{\partial w}{\partial y}\right|_o \right\rangle.$$

Note well the following: (as we look more deeply into properties of the gradient these can be points of confusion).

1. The gradient takes a scalar function $f(x, y)$ and produces a vector $\boldsymbol{\nabla} f$.
2. The vector $\boldsymbol{\nabla} f(x, y)$ lies in the plane.

For functions $w = f(x, y, z)$ we have the gradient

$$\operatorname{grad} w = \boldsymbol{\nabla} w = \left\langle \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z} \right\rangle.$$

That is, the gradient takes a scalar function of three variables and produces a three dimensional vector.

The gradient has many geometric properties. In the next session we will prove that for $w = f(x, y)$ the gradient is perpendicular to the level curves $f(x, y) = c$. We can show this by direct computation in the following example.

Example 1: Compute the gradient of $w = (x^2 + y^2)/3$ and show that the gradient at $(x_0, y_0) = (1, 2)$ is perpendicular to the level curve through that point.

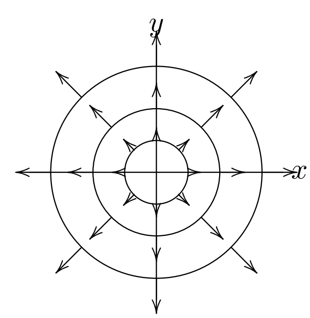
Answer: The gradient is easily computed

$$\boldsymbol{\nabla} w = \langle 2x/3, 2y/3 \rangle = \frac{2}{3}\langle x, y \rangle.$$

At $(1, 2)$ we get $\boldsymbol{\nabla} w(1, 2) = \frac{2}{3}\langle 1, 2 \rangle$. The level curve through $(1, 2)$ is

$$(x^2 + y^2)/3 = 5/3,$$

which is identical to $x^2 + y^2 = 5$. That is, it is a circle of radius $\sqrt{5}$ centered at the origin. Since the gradient at (1,2) is a multiple of $\langle 1, 2 \rangle$, it points radially outward and hence is perpendicular to the circle. Below is a figure showing the gradient field and the level curves.



$z = (x^2 + y^2)/3$

Example 2: Consider the graph of $y = \mathrm{e}^x$. Find a vector perpendicular to the tangent to $y = \mathrm{e}^x$ at the point $(1, \mathrm{e})$.

Old method: Find the slope take the negative reciprocal and make the vector.

New method: This graph is the level curve of $w = y - \mathrm{e}^x = 0$.

$\boldsymbol{\nabla} w = \langle -\mathrm{e}^x, 1\rangle \;\Rightarrow\;$ (at $x = 1$) $\boldsymbol{\nabla} w(1, e) = \langle -\mathrm{e}, 1\rangle$ is perpendicular to the tangent vector to the graph, $\mathbf{v} = \langle 1, \mathrm{e}\rangle$.

Higher dimensions

Similarly, for $w = f(x, y, z)$ we get level surfaces $f(x, y, z) = c$. The gradient is perpendicular to the level surfaces.

Example 3: Find the tangent plane to the surface $x^2 + 2y^2 + 3z^2 = 6$ at the point $P = (1, 1, 1)$.

Answer: Introduce a new variable

$$w = x^2 + 2y^2 + 3z^2.$$

Our surface is the level surface $w = 6$. Saying the gradient is perpendicular to the surface means exactly the same thing as saying it is normal to the tangent plane. Computing

$$\boldsymbol{\nabla} w = \langle 2x, 4y, 6z\rangle \;\Rightarrow\; \boldsymbol{\nabla} w|_P = \langle 2, 4, 6\rangle.$$

Using point normal form we get the equation of the tangent plane is

$$2(x-1) + 4(y-1) + 6(z-1) = 0, \qquad \text{or} \qquad 2x + 4y + 6z = 12.$$

Gradient: proof that it is perpendicular to level curves and surfaces

Let $w = f(x, y, z)$ be a function of 3 variables. We will show that at any point $P = (x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$ (so $f(x_0, y_0, z_0) = c$) the gradient $\nabla f|_P$ is perpendicular to the surface.

By this we mean it is perpendicular to the tangent to any curve that lies on the surface and goes through P. (See figure.)

This follows easily from the chain rule: Let

$$\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$$

be a curve on the level surface with $\mathbf{r}(t_0) = \langle x_0, y_0, z_0 \rangle$. We let $g(t) = f(x(t), y(t), z(t))$. Since the curve is on the level surface we have $g(t) = f(x(t), y(t), z(t)) = c$. Differentiating this equation with respect to t gives

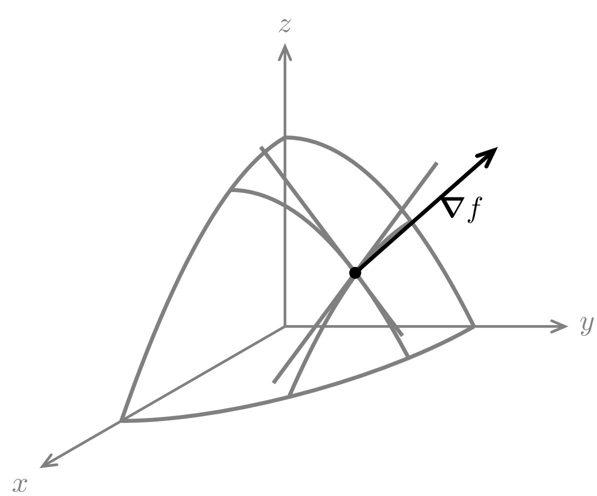
$$\frac{dg}{dt} = \left.\frac{\partial f}{\partial x}\right|_P \left.\frac{dx}{dt}\right|_{t_0} + \left.\frac{\partial f}{\partial y}\right|_P \left.\frac{dy}{dt}\right|_{t_0} + \left.\frac{\partial f}{\partial z}\right|_P \left.\frac{dz}{dt}\right|_{t_0} = 0.$$

In vector form this is

$$\left\langle \left.\frac{\partial f}{\partial x}\right|_P, \left.\frac{\partial f}{\partial y}\right|_P, \left.\frac{\partial f}{\partial z}\right|_P \right\rangle \cdot \left\langle \left.\frac{dx}{dt}\right|_{t_0}, \left.\frac{dy}{dt}\right|_{t_0}, \left.\frac{dz}{dt}\right|_{t_0} \right\rangle = 0$$

$$\Leftrightarrow \quad \nabla f|_P \cdot \mathbf{r}'(t_0) = 0.$$

Since the dot product is 0, we have shown that the gradient is perpendicular to the tangent to any curve that lies on the level surface, which is exactly what we needed to show.



Tangent Plane to a Level Surface

1. Find the tangent plane to the surface $x^2 + 2y^2 + 3z^2 = 36$ at the point $P = (1, 2, 3)$.

<u>Answer:</u> In order to use gradients we introduce a new variable

$$w = x^2 + 2y^2 + 3z^2.$$

Our surface is then the the level surface $w = 36$. Therefore the normal to surface is

$$\boldsymbol{\nabla} w = \langle 2x, 4y, 6z \rangle.$$

At the point P we have $\boldsymbol{\nabla} w|_P = \langle 2, 8, 18 \rangle$. Using point normal form, the equation of the tangent plane is

$$2(x-1) + 8(y-2) + 18(z-3) = 0, \text{ or equivalently } 2x + 8y + 18z = 72.$$

2. Use gradients and level surfaces to find the normal to the tangent plane of the graph of $z = f(x, y)$ at $P = (x_0, y_0, z_0)$.

<u>Answer:</u> Introduce the new variable

$$w = f(x, y) - z.$$

The graph of $z = f(x, y)$ is just the level surface $w = 0$. We compute the normal to the surface to be

$$\boldsymbol{\nabla} w = \langle f_x, f_y, -1 \rangle.$$

At the the point P the normal is $\langle f_x(x_0, y_0), f_y(x_0, y_0), -1 \rangle$, so the equation of the tangent plane is

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

We can write this in a more compact form as

$$(z - z_0) = \left.\frac{\partial f}{\partial x}\right|_0 (x - x_0) + \left.\frac{\partial f}{\partial y}\right|_0 (y - y_0),$$

which is exactly the formula we saw earlier for the tangent plane to a graph.

Directional Derivatives

Directional derivative

Like all derivatives the *directional derivative* can be thought of as a ratio. Fix a unit vector $\mathbf{u}$ and a point P_0 in the *plane*. The **directional derivative** of w at P_0 in the direction $\mathbf{u}$ is defined as

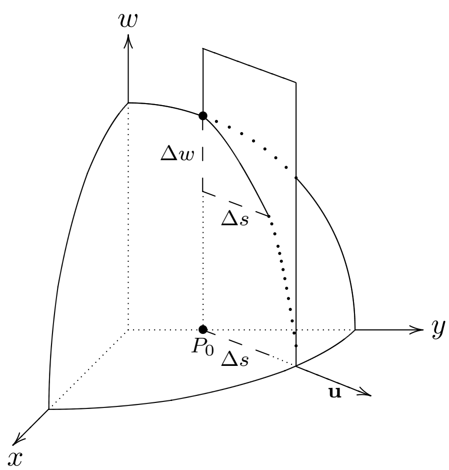
$$\left.\frac{dw}{ds}\right|_{P_0,\mathbf{u}} = \lim_{\Delta s \to 0} \frac{\Delta w}{\Delta s}.$$

Here Δw is the change in w caused by a step of length Δs in the direction of $\mathbf{u}$ (all in the xy-plane).

Below we will show that

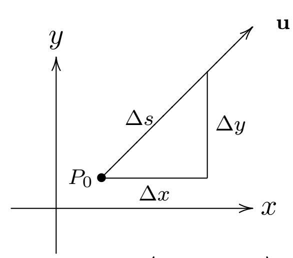
$$\left.\frac{dw}{ds}\right|_{P_0,\mathbf{u}} = \boldsymbol{\nabla} w(P_0) \cdot \mathbf{u}. \tag{1}$$

We illustrate this with a figure showing the graph of $w = f(x, y)$. Notice that Δs is measured in the plane and Δw is the change of w on the graph.



Proof of equation 1

The figure below represents the change in position from P_0 resulting from taking a step of size Δs in the $\mathbf{u}$ direction.



Since $(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2$ we have that $\left\langle \frac{\Delta x}{\Delta s}, \frac{\Delta y}{\Delta s} \right\rangle$. is a unit vector, so

$$\mathbf{u} = \left\langle \frac{\Delta x}{\Delta s}, \frac{\Delta y}{\Delta s} \right\rangle.$$

The tangent plane approximation at P_0 is

$$\Delta w \approx \left.\frac{\partial w}{\partial x}\right|_{P_0} \Delta x + \left.\frac{\partial w}{\partial y}\right|_{P_0} \Delta y$$

Dividing this approximation by Δs gives

$$\frac{\Delta w}{\Delta s} \approx \left.\frac{\partial w}{\partial x}\right|_{P_0} \frac{\Delta x}{\Delta s} + \left.\frac{\partial w}{\partial y}\right|_{P_0} \frac{\Delta y}{\Delta s}.$$

We can rewrite this as a dot product

$$\frac{\Delta w}{\Delta s} \approx \left\langle \left.\frac{\partial w}{\partial x}\right|_{P_0}, \left.\frac{\partial w}{\partial y}\right|_{P_0} \right\rangle \cdot \left\langle \frac{\Delta x}{\Delta s}, \frac{\Delta y}{\Delta s} \right\rangle.$$

In the dot product the first term is $\nabla w|_{P_0}$ and the second is just $\mathbf{u}$, so,

$$\frac{\Delta w}{\Delta s} \approx \nabla w|_{P_0} \cdot \mathbf{u}.$$

Now taking the limit we get equation (1).

Example: (Algebraic example) Let $w = x^3 + 3y^2$.

Compute $\frac{dw}{ds}$ at $P_0 = (1,2)$ in the direction of $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j}$.

<u>Answer:</u> We compute all the necessary pieces:

i) $\nabla w = \langle 3x^2, 6y\rangle \Rightarrow \nabla w|_{(1,2)} = \langle 3, 12\rangle$.

ii) $\mathbf{u}$ must be a unit vector, so $\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \langle \frac{3}{5}, \frac{4}{5}\rangle$.

iii) $\left.\frac{dw}{ds}\right|_{P_0,\mathbf{u}} = \nabla w|_{(1,2)} \cdot \mathbf{u} = \langle 3, 12\rangle \cdot \langle \frac{3}{5}, \frac{4}{5}\rangle = \boxed{\frac{57}{5}.}$

Example: (Geometric example) Let $\mathbf{u}$ be the direction of $\langle 1, -1\rangle$.

Using the picture at right estimate $\left.\frac{\partial w}{\partial x}\right|_P$, $\left.\frac{\partial w}{\partial y}\right|_p$, and $\left.\frac{dw}{ds}\right|_{P,\mathbf{u}}$.

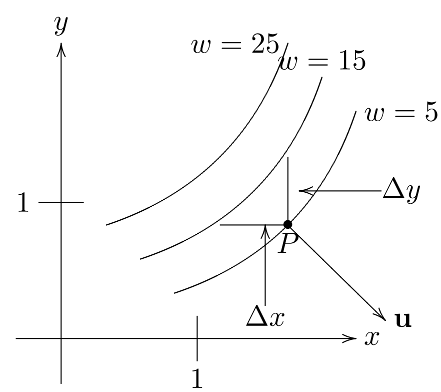
By measuring from P to the next in level curve in the x direction we see that $\Delta x \approx -.5$.

$\Rightarrow \boxed{\left.\frac{\partial w}{\partial x}\right|_P \approx \frac{\Delta w}{\Delta x} \approx \frac{10}{-.5} = -20.}$

Similarly, we get $\boxed{\left.\frac{\partial w}{\partial y}\right|_P \approx 20.}$

Measuring in the $\mathbf{u}$ direction we get $\Delta s \approx -.3$

$\Rightarrow \boxed{\left.\frac{dw}{ds}\right|_{P,\mathbf{u}} \approx \frac{\Delta w}{\Delta s} \approx \frac{10}{.3} = -33.3.}$



Direction of maximum change:

The direction that gives the maximum rate of change is in the same direction as ∇w. The proof of this uses equation (1). Let θ be the angle between ∇w and $\mathbf{u}$. Then the geometric form of the dot product says

$$\left.\frac{dw}{ds}\right|_{\mathbf{u}} = \nabla w \cdot \mathbf{u} = |\nabla w||\mathbf{u}|\cos\theta = |\nabla w|\cos\theta.$$

Chain rule and total differentials

1. Find the total differential of $w = z\mathrm{e}^{(x+y)}$ at $(0,0,1)$.

Answer: The total differential at the point (x_0, y_0, z_0) is

$$dw = w_x(x_0, y_0, z_0)dx + w_y(x_0, y_0, z_0)dy + w_z(x_0, y_0, z_0)dz.$$

In our case,

$$w_x = z\mathrm{e}^{(x+y)}, \quad w_y = z\mathrm{e}^{(x+y)}, \quad w_z = \mathrm{e}^{(x+y)}$$

Substituting in the point $(0,0,1)$ we get: $w_x(0,0,1) = 1, \quad w_y(0,0,1) = 1, \quad w_z(0,0,1) = 1.$ Thus,

$$dw = dx + dy + dz.$$

2. Suppose $w = z\mathrm{e}^{(x+y)}$ and $x = t, \quad y = t^2, \quad z = t^3$. Compute $\dfrac{dw}{dt}$ and evaluate it when $t = 2$.

Answer: We use the chain rule:

$$\begin{aligned} \frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\ &= (z\mathrm{e}^{(x+y)})(1) + (z\mathrm{e}^{(x+y)})(2t) + (\mathrm{e}^{(x+y)})(3t^2). \end{aligned}$$

At $t = 2$ we have $x = 2,\ y = 4,\ z = 8$. Thus,

$$\left.\frac{dw}{dt}\right|_2 = 8\mathrm{e}^6 + 8\mathrm{e}^6(4) + \mathrm{e}^6(12) = 52\mathrm{e}^6.$$

Chain Rule

1. The temperature on a hot surface is given by

$$T = 100\,\mathrm{e}^{-(x^2+y^2)}.$$

A bug follows the trajectory $\mathbf{r}(t) = \langle t\cos(2t), t\sin(2t)\rangle$.

a) What is the rate that temperature is changing as the bug moves?

b) Draw the level curves of T and sketch the bug's trajectory.

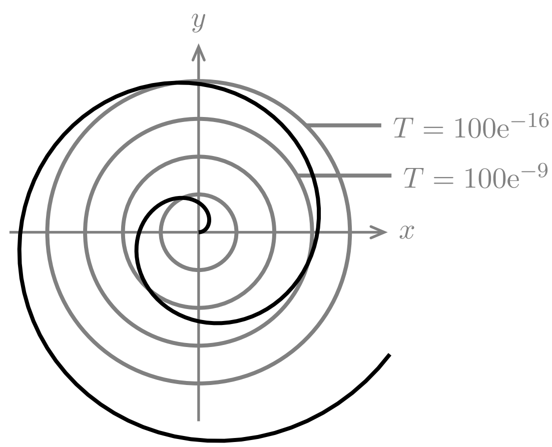
Answer: a) The chain rule says

$$\begin{aligned}\frac{dT}{dt} &= \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt} \\ &= -200x\mathrm{e}^{-(x^2+y^2)}(\cos(2t) - 2t\sin(2t)) - 200y\mathrm{e}^{-(x^2+y^2)}(\sin(2t) + 2t\cos(2t)).\end{aligned}$$

You could stop here, or substitute $x = t\cos(2t)$ and $y = t\sin(2t)$. After simplification you get

$$\frac{dT}{dt} = -200\,t\,\mathrm{e}^{-t^2}.$$

b) The level curves of T are the curves $x^2 + y^2 =$ constant, i.e., circles. The bug moves in a spiral.



2. Suppose $w = f(x, y)$ and $x = t^2$, $y = t^3$. Suppose also that at $(x, y) = (1, 1)$ we have $\dfrac{\partial w}{\partial x} = 3$ and $\dfrac{\partial w}{\partial y} = 1$. Compute $\dfrac{dw}{dt}$ at $t = 1$.

Answer: At $t = 1$ we have $(x, y) = (1, 1)$, $\left.\frac{dx}{dt}\right|_1 = 2$, $\left.\frac{dy}{dt}\right|_1 = 3$. Therefore the chain rule says

$$\left.\frac{dw}{dt}\right|_1 = \left.\frac{\partial f}{\partial x}\right|_{(1,1)} \left.\frac{dx}{dt}\right|_1 + \left.\frac{\partial f}{\partial y}\right|_{(1,1)} \left.\frac{dy}{dt}\right|_1 = 3(2) + 1(3) = 9.$$

Problems: Chain Rule Practice

One application of the chain rule is to problems in which you are given a function of x and y with inputs in polar coordinates. For example, let $w = (x^2 + y^2)xy$, $x = r\cos\theta$ and $y = r\sin\theta$.

1. Use the chain rule to find $\dfrac{\partial w}{\partial r}$.

<u>**Answer:**</u> We apply the chain rule.

$$\begin{aligned}\frac{\partial w}{\partial r} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} \\ &= (3x^2y + y^3)\cos\theta + (3xy^2 + x^3)\sin\theta.\end{aligned}$$

We could check our work by substituting to get $w = r^4\cos\theta\sin\theta$ and calculating $\frac{dw}{dr}$ directly. We practice using the chain rule because such substitutions are not always practical.

2. Find the total differential dw in terms of dr and $d\theta$.

<u>**Answer:**</u> We know $dw = w_x\,dx + w_y\,dy$. In terms of r and θ, $dx = x_r\,dr + x_\theta\,d\theta = \cos\theta\,dr - r\sin\theta\,d\theta$. Similarly, $dy = \sin\theta\,dr + r\cos\theta\,d\theta$. Thus,

$$\begin{aligned}dw &= w_x(\cos\theta\,dr - r\sin\theta\,d\theta) + w_y(\sin\theta\,dr + r\cos\theta\,d\theta) \\ &= (w_x\cos\theta + w_y\sin\theta)\,dr + (w_y r\cos\theta - w_x r\sin\theta)\,d\theta.\end{aligned}$$

We could stop here or go on to compute:

$$\begin{aligned}dw &= [(3x^2y + y^3)\cos\theta + (3xy^2 + x^3)\sin\theta]\,dr + [(3xy^2 + x^3)r\cos\theta - (3x^2y + y^3)r\sin\theta]\,d\theta \\ &= 4r^3\cos\theta\sin\theta\,dr + r^4(\cos^2\theta - \sin^2\theta)\,d\theta.\end{aligned}$$

Note that the answer to (1) appears in the dr component of dw. In practice, the best format for the answer is the one that is easiest to use.

3. Find $\dfrac{\partial w}{\partial r}$ at the point $(r, \theta) = (2, \pi/4)$.

<u>**Answer:**</u> Recall that $\frac{\partial w}{\partial r} = (3x^2y + y^3)\cos\theta + (3xy^2 + x^3)\sin\theta$. We need only compute $x = \sqrt{2}$ and $y = \sqrt{2}$ and plug in values.

$$\begin{aligned}(3x^2y + y^3)\cos\theta + (3xy^2 + x^3)\sin\theta &= 4(\sqrt{2})^3(\frac{\sqrt{2}}{2}) + 4(\sqrt{2})^3(\frac{\sqrt{2}}{2}) \\ &= 16.\end{aligned}$$

Problems: Elliptic Paraboloid

1. Compute the gradient of $w = x^2 + 5y^2$.

Answer: :

$$\nabla w = \left\langle \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \right\rangle = \langle 2x, 10y \rangle.$$

2. Show that ∇w is perpendicular to the level curves of w at the points $(x_0, 0)$.

Answer: At $(x_0, 0)$, $\nabla w = \langle 2x_0, 0 \rangle$.

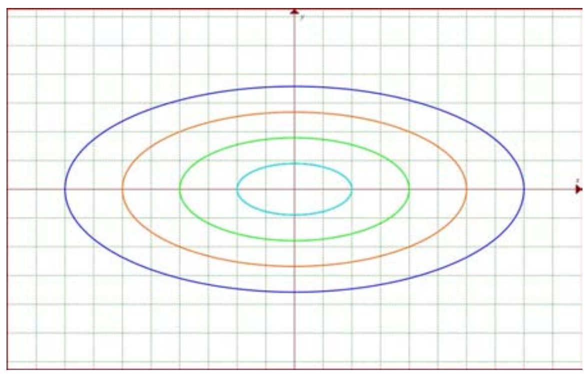


Figure 1: The level curves of $w = x^2 + 5y^2$.

In general, the level curves of w have equation $x^2 + 5y^2 = k$; each one is an ellipse whose major axis coincides with the x axis. Hence, the horizontal vector $\nabla w = \langle 2x_0, 0 \rangle$ will be normal to the level curve at the point $(x_0, 0)$.

The gradient is perpendicular to the level curves

1. Here is a challenging problem. Use the chain rule to show the slope of the gradient is the negative reciprocal of the slope of the level curves. (This is another way of saying the gradient is perpendicular to the level curves.)

Note, this problem is strictly about 2D functions $w = f(x, y)$ and their gradients and level curves. Also note, for a 2D vector $\langle a, b\rangle$ the slope is b/a.

Answer: Suppose $w = f(x, y)$ and we have a level curve $f(x, y) = c$. Implicitly this gives a relation between x and y, which means y can be thought of as a function of x, say $y = y(x)$. We then rewrite the equation of the level curve as

$$f(x, y(x)) = c.$$

The chain rule gives

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0 \ \Rightarrow\ \frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y}.$$

This last expression is the slope of the level curve.

Now, $\boldsymbol{\nabla} f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle$, which has slope

$$\frac{\partial f/\partial y}{\partial f/\partial x}.$$

This slope is clearly the negative reciprocal of the slope of the level curve computed above.

Problems: Equation of a Tangent Plane

Find the equation of the plane which is tangent to the surface $x^3 + y^3 + z^3 = 32$ at the point $(1, 2, 3)$.

Answer: : Let $w = x^3 + y^3 + z^3$. We're interested in the surface $w = 32$.
The vector $\boldsymbol{\nabla} w = \langle 3x^2, 3y^2, 3z^2 \rangle$ is normal to this surface, so the normal vector at $(1, 2, 3)$ is $\langle 3, 12, 27 \rangle$.

Applying point normal form for the equation of a plane tells us that:

$$3(x - 1) + 12(y - 2) + 27(z - 3) = 0 \text{ or } 3x + 12\text{y} + 27z = 108$$

is the equation of the tangent plane to $x^3 + y^3 + z^3 = 32$ at $(1, 2, 3)$.

Problems: Directional Derivatives

The function $T = x^2 + 2y^2 + 2z^2$ gives the temperature at each point in space.

1. At the point $P = (1, 1, 1)$, in which direction should you go to get the most rapid decrease in T? What is the directional derivative in this direction?

Answer: We know that the fastest *increase* is in the direction of $\boldsymbol{\nabla} T = \langle 2x, 4y, 4z\rangle$. At P, the fastest *decrease* is in the direction of $-\boldsymbol{\nabla} T|_{(1,1,1)} = -\langle 1, 2, 2\rangle$. The unit vector in this direction is $\widehat{\mathbf{u}} = -\langle 1/3, 2/3, 2/3\rangle$.

The rate of change in this direction is $-|\boldsymbol{\nabla} T| = -3$. Equivalently, you could compute:

$$\left.\frac{dT}{ds}\right|_{P,\widehat{\mathbf{u}}} = \boldsymbol{\nabla} T|_P \cdot \widehat{\mathbf{u}} = -3.$$

2. At P, about how far should you go in the direction found in part (1) to get a decrease of 0.3?

Answer: The directional derivative is a true derivative describing the limit of a ratio. In this case it equals $\lim\limits_{\Delta s \to 0} \dfrac{\Delta T}{\Delta s}$, where Δs is the distance moved in the $\widehat{\mathbf{u}}$ direction. Thus, we can write $\dfrac{\Delta T}{\Delta s} \approx \dfrac{dT}{ds}$.

In this problem we have $\dfrac{dT}{ds} = -3$ and $\Delta T = -0.3$.

$$\frac{-0.3}{\Delta s} \approx -3 \;\Rightarrow\; \Delta s \approx 0.1.$$

Partial Differential Equations

An important application of the higher partial derivatives is that they are used in partial differential equations to express some laws of physics which are basic to most science and engineering subjects. In this section, we will give examples of a few such equations. The reason is partly cultural, so you meet these equations early and learn to recognize them, and partly technical: to give you a little more practice with the chain rule and computing higher derivatives.

A **partial differential equation**, PDE for short, is an equation involving some unknown function of several variables and one or more of its partial derivatives. For example,

$$x\,\frac{\partial w}{\partial x} \;-\; y\,\frac{\partial w}{\partial y} \;=\; 0$$

is such an equation. Evidently here the unknown function is a function of two variables

$$w \;=\; f(x,y) \;;$$

we infer this from the equation, since only x and y occur in it as independent variables. In general a **solution** of a partial differential equation is a differentiable function that satisfies it. In the above example, the functions

$$w \;=\; x^n y^n \qquad \text{any } n$$

all are solutions to the equation. In general, PDE's have many solutions, far too many to find all of them. The problem is always to find the one solution satisfying some extra conditions, usually called either *boundary conditions* or *initial conditions* depending on their nature.

Our first important PDE is the **Laplace equation** in three dimensions:

$$\frac{\partial^2 w}{\partial x^2} \;+\; \frac{\partial^2 w}{\partial y^2} \;+\; \frac{\partial^2 w}{\partial z^2} \;=\; 0\;. \tag{1}$$

Any steady-state temperature distribution in three-space

$$w \;=\; T(x,y,z), \qquad\qquad T = \text{temperature at the point } \;(x,y,z) \tag{2}$$

satisfies Laplace's equation. (Here *steady-state* means that it is unchanging over time, here reflected in the fact that T is not a function of time. For example, imagine a solid object made of some uniform heat-conducting material (say a solid metal ball), and imagine a steady temperature distribution on its surface is maintained somehow (say with some arrangement of wires and thermostats). Then after a while the temperature at each point inside the ball will come to equilibrium — reach a steady state — and the resulting temperature function (2) inside the ball will then satisfy Laplace's equation.

As another example, the *gravitational potential*

$$w \;=\; \phi(x,y,z)$$

resulting from some arrangement of masses in space satisfies Laplace's equation in any region R of space not containing masses. The same is true of the *electrostatic potential* resulting from some collection of electric charges in space: (1) is satisfied in any region which is free of charge. This potential function measures the work done (against the field) carrying a unit test mass (or charge) from a fixed reference point to the point (x, y, z) in the gravitational (or electrostatic) field. Knowing ϕ, the field itself can be recovered as its negative gradient:

$$\mathbf{F} = -\nabla\phi .$$

All of this is just to stress the fundamental character of Laplace's equation — we live our lives surrounded by its solutions.

The *two-dimensional* Laplace equation is similar — you just drop the term involving z. The steady-state temperature distribution in a flat metal plate would satisfy the two-dimensional Laplace equation, if the faces of the plate were kept insulated and a steady-state temperature distribution maintained around the edges of the plate.

If in the temperature model we include also heat sources and sinks in the region, unchanging over time, the temperature function satisfies the closely related **Poisson equation**

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} = f(x, y, z), \tag{3}$$

where f is some given function related to the sources and sinks.

Another important PDE is the **wave equation**; given below are the one-dimensional and two-dimensional versions; the three dimensional version would add a similar term in z to the left:

$$\frac{\partial^2 w}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 w}{\partial t^2}; \qquad \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 w}{\partial t^2}. \tag{4}$$

Here $x, y, \ldots$ are the space variables, t is the time, and c is the velocity with which the wave travels — this depends on the medium and the type of wave (light, sound, etc.). A solution, respectively

$$w = w(x, t), \qquad w = w(x, y, t),$$

gives for each moment t_0 of time the shape $w(x, t_0)$, $w(x, y, t_0)$ of the wave.

The third PDE goes by two names, depending on the context: **heat equation** or **diffusion equation**. The one- and two-dimensional versions are respectively

$$\frac{\partial^2 w}{\partial x^2} = \frac{1}{a^2}\frac{\partial w}{\partial t}; \qquad \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \frac{1}{a^2}\frac{\partial w}{\partial t}. \tag{5}$$

It looks a lot like the wave equation (4), but the right-hand side this time involves only the first derivative, which gives it mathematically and physically an entirely different character.

When it is called the (one-dimensional) heat equation, a solution $w(x, t)$ represents a time-varying temperature distribution in say a uniform conducting metal rod, with insulated sides. In the same way, $w(x, y, t)$ would be the time-varying temperature distribution in a flat metal plate with insulated faces. For each moment t_0 in time, $w(x, y, t_0)$ gives the temperature distribution at that moment.

For example, if we assume the distribution is steady-state, i.e., not changing with time, then

$$\frac{\partial w}{\partial t} = 0 \qquad \text{(steady-state condition)}$$

and the two-dimensional heat equation would turn into the two-dimensional Laplace equation (1).

When (5) is referred to as the *diffusion equation*, say in one dimension, then $w(x, t)$ represents the concentration of a dissolved substance diffusing along a uniform tube filled with liquid, or of a gas diffusing down a uniform pipe.

Notice that all of these PDE's are second-order, that is, involve derivatives no higher than the second. There is an important fourth-order PDE in elasticity theory (the bilaplacian equation), but by and large the general rule seems to be either that Nature is content with laws that only require second partial derivatives, or that these are the only laws that humans are intelligent enough to formulate.

18.02SC Multivariable Calculus
Fall 2010