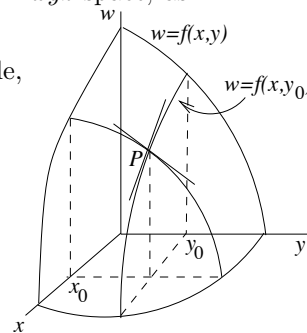# Partial derivatives

Partial derivatives

Let $w = f(x, y)$ be a function of two variables. Its graph is a surface in xyz-space, as pictured.

Fix a value $y = y_0$ and just let x vary. You get a function of *one* variable,

$$(1) \qquad w = f(x, y_0), \qquad \text{the \textbf{partial function} for } y = y_0.$$

Its graph is a curve in the vertical plane $y = y_0$, whose slope at the point P where $x = x_0$ is given by the derivative

$$(2) \qquad \left. \frac{d}{dx} f(x, y_0) \right|_{x_0}, \quad \text{or} \quad \left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)}.$$

We call (2) the **partial derivative** of f with respect to x at the point (x_0, y_0); the right side of (2) is the standard notation for it. The partial derivative is just the ordinary derivative of the partial function — it is calculated by holding one variable fixed and differentiating with respect to the other variable. Other notations for this partial derivative are

$$f_x(x_0, y_0), \qquad \left. \frac{\partial w}{\partial x} \right|_{(x_0, y_0)}, \qquad \left(\frac{\partial f}{\partial x} \right)_0, \qquad \left(\frac{\partial w}{\partial x} \right)_0;$$

the first is convenient for including the specific point; the second is common in science and engineering, where you are just dealing with relations between variables and don't mention the function explicitly; the third and fourth indicate the point by just using a single subscript.

Analogously, fixing $x = x_0$ and letting y vary, we get the partial function $w = f(x_0, y)$, whose graph lies in the vertical plane $x = x_0$, and whose slope at P is the *partial derivative of f with respect to y*; the notations are

$$\left. \frac{\partial f}{\partial y} \right|_{(x_0, y_0)}, \qquad f_y(x_0, y_0), \qquad \left. \frac{\partial w}{\partial y} \right|_{(x_0, y_0)}, \qquad \left(\frac{\partial f}{\partial y} \right)_0, \qquad \left(\frac{\partial w}{\partial y} \right)_0.$$

The partial derivatives $\partial f / \partial x$ and $\partial f / \partial y$ depend on (x_0, y_0) and are therefore functions of x and y.

Written as $\partial w / \partial x$, the partial derivative gives the rate of change of w with respect to x alone, at the point (x_0, y_0): it tells how fast w is increasing as x increases, when y is held constant.

For a function of three or more variables, $w = f(x, y, z, \dots)$, we cannot draw graphs any more, but the idea behind partial differentiation remains the same: to define the partial derivative with respect to x, for instance, hold all the other variables constant and take the ordinary derivative with respect to x; the notations are the same as above:

$$\frac{d}{dx} f(x, y_0, z_0, \dots) = f_x(x_0, y_0, z_0, \dots), \qquad \left(\frac{\partial f}{\partial x} \right)_0, \qquad \left(\frac{\partial w}{\partial x} \right)_0.$$

1

Partial derivatives

1. Let $f(x, y) = e^{(x^2+y^2)} + x^2 + y^2 + xy + 2y + 3$.

a) Compute $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$.

b) Show the second partials can be computed in any order. That is,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

c) Find $\dfrac{\partial f}{\partial x}(1, 3)$.

<u>Answer:</u> a) $\dfrac{\partial f}{\partial x} = 2xe^{(x^2+y^2)} + 2x + y$, $\qquad \dfrac{\partial f}{\partial y} = 2ye^{(x^2+y^2)} + 2y + x + 2$.

b) To compute $\dfrac{\partial^2 f}{\partial x \partial y}$ we compute the partial with respect to x of $\dfrac{\partial f}{\partial y}$.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(2ye^{(x^2+y^2)} + 2y + x + 2\right) = 4xye^{(x^2+y^2)} + 1.$$

Likewise

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}\left(2xe^{(x^2+y^2)} + 2x + y\right) = 4xye^{(x^2+y^2)} + 1.$$

We have shown the order of differentiation didn't matter.

c) Evaluating $\dfrac{\partial f}{\partial x}(1, 3) = 2e^{10} + 5$.

The Tangent Approximation

1. The tangent plane.

For a function of one variable, $w = f(x)$, the tangent line to its graph at a point (x_0, w_0) is the line passing through (x_0, w_0) and having slope $\left(\dfrac{dw}{dx}\right)_0$.

For a function of two variables, $w = f(x, y)$, the natural analogue is the *tangent plane* to the graph, at a point

(x_0, y_0, w_0). What's the equation of this tangent plane? Referring to the picture at right (this figure was also used when we introduced partial derivatives), we see that the tangent plane
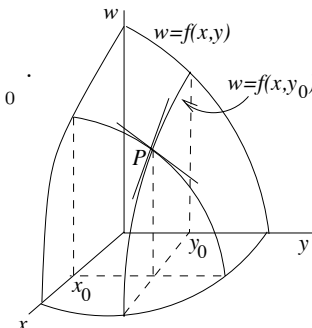
(i) must pass through (x_0, y_0, w_0), where $w_0 = f(x_0, y_0)$;

(ii) must contain the tangent lines to the graphs of the two partial functions — this will hold if the plane has the same slopes in the **i** and **j** directions as the surface does.

Using these two conditions, it is easy to find the equation of the tangent plane. The general equation of a plane through (x_0, y_0, w_0) is

$$A(x - x_0) + B(y - y_0) + C(w - w_0) \;=\; 0 \;.$$

Assume the plane is not vertical; then $C \neq 0$, so we can divide through by C and solve for $w - w_0$, getting

(3) $\qquad\qquad w - w_0 \;=\; a(x - x_0) + b(y - y_0), \qquad a = A/C, \; b = B/C.$

The plane passes through (x_0, y_0, w_0); what values of the coefficients a and b will make it also tangent to the graph there? We have

$$
\begin{aligned}
a \;&=\; \text{slope of plane (3) in the } \mathbf{i}\text{-direction} \qquad (\text{by putting } y = y_0 \text{ in (3)});\\
&=\; \text{slope of graph in the } \mathbf{i}\text{-direction}, \qquad\;\; (\text{by (ii) above})\\
&=\; \left(\frac{\partial w}{\partial x}\right)_0 ; \qquad\qquad (\text{by the definition of partial derivative}); \quad \text{similarly,}\\
b \;&=\; \left(\frac{\partial w}{\partial y}\right)_0 .
\end{aligned}
$$

Therefore the equation of the **tangent plane** to $w = f(x, y)$ at (x_0, y_0) is

(4) $\qquad\qquad w - w_0 \;=\; \left(\dfrac{\partial w}{\partial x}\right)_0 (x - x_0) + \left(\dfrac{\partial w}{\partial y}\right)_0 (y - y_0)$

2. The approximation formula.

The most important use for the tangent plane is to give an approximation that is the basic formula in the study of functions of several variables — almost everything follows in one way or another from it.

1

The intuitive idea is that if we stay near (x_0, y_0, w_0), the graph of the tangent plane (4) will be a good approximation to the graph of the function $w = f(x, y)$. Therefore if the point (x, y) is close to (x_0, y_0),

$$(5) \qquad f(x, y) \quad \approx \quad w_0 + \left(\frac{\partial w}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial w}{\partial y}\right)_0 (y - y_0)$$

$$\text{height of graph} \quad \approx \qquad \text{height of tangent plane}$$

The function on the right side of (5) whose graph is the tangent plane is often called the **linearization** of $f(x, y)$ at (x_0, y_0): it is the linear function which gives the best approximation to $f(x, y)$ for values of (x, y) close to (x_0, y_0).

An equivalent form of the approximation (5) is obtained by using Δ notation; if we put

$$\Delta x = x - x_0, \qquad \Delta y = y - y_0, \qquad \Delta w = w - w_0,$$

then (5) becomes

$$(6) \qquad \Delta w \approx \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y, \qquad \text{if} \quad \Delta x \approx 0, \ \Delta y \approx 0 .$$

This formula gives the approximate change in w when we make a small change in x and y. We will use it often.

The analogous approximation formula for a function $w = f(x, y, z)$ of three variables would be

$$(7) \qquad \Delta w \approx \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y + \left(\frac{\partial w}{\partial z}\right)_0 \Delta z, \qquad \text{if} \quad \Delta x, \Delta y \Delta z \approx 0 .$$

Unfortunately, for functions of three or more variables, we can't use a geometric argument for the approximation formula (7); for this reason, it's best to recast the argument for (6) in a form which doesn't use tangent planes and geometry, and therefore can be generalized to several variables. This is done at the end of this Chapter TA; for now let's just assume the truth of (7) and its higher-dimensional analogues.

Here are two typical examples of the use of the approximation formula. Other examples are in the Exercises. In the rest of your study of partial differentiation, you will see how the approximation formula is used to derive the important theorems and formulas.

Example 1. Give a reasonable square, centered at $(1, 1)$, over which the value of $w = x^3 y^4$ will not vary by more than $\pm .1$.

Solution. We use (6). We calculate for the two partial derivatives

$$w_x = 3x^2 y^4 \qquad\qquad w_y = 4x^3 y^3$$

and therefore, evaluating the partials at $(1, 1)$ and using (6), we get

$$\Delta w \approx 3\Delta x + 4\Delta y.$$

Thus if $|\Delta x| \le .01$ and $\Delta y| \le .01$, we should have
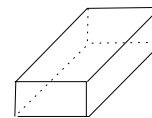
$$|\Delta w| \le 3|\Delta x| + 4|\Delta y| \le .07,$$

which is within the bounds. So the answer is the square with center at (1,1) given by

$$|x - 1| \le .01, \qquad |y - 1| \le .01 .$$

Example 2.   The sides a, b, c of a rectangular box have lengths measured to be respectively 1, 2, and 3. To which of these measurements is the volume V most sensitive?

Solution. $V = abc$, and therefore by the approximation formula (7),

$$\begin{aligned} \Delta V \;&\approx\; bc\,\Delta a + ac\,\Delta b + ab\,\Delta c \\ &\approx\; 6\,\Delta a + 3\,\Delta b + 2\,\Delta c, \qquad\qquad \text{at } (1, 2, 3); \end{aligned}$$

thus it is most sensitive to small changes in side a, since Δa occurs with the largest coefficient. (That is, if one at a time the measurement of each side were changed by say .01, it is the change in a which would produce the biggest change in V, namely .06 .)

> The result may seem paradoxical — the value of V is most sensitive to the length of the *shortest* side — but it's actually intuitive, as you can see by thinking about how the box looks.

Sensitivity Principle   *The numerical value of* $w = f(x, y, \dots)$, *calculated at some point* $(x_0, y_0, \dots)$, *will be most sensitive to small changes in that variable for which the corresponding partial derivative* $w_x, w_y, \dots$ *has the largest absolute value at the point.*

The Tangent approximation

4. Critique of the approximation formula.

First of all, the approximation formula for functions of two or three variables

(6) $$\Delta w \;\approx\; \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y, \qquad \text{if}\quad \Delta x \approx 0,\; \Delta y \approx 0 \;.$$

(7) $$\Delta w \;\approx\; \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y + \left(\frac{\partial w}{\partial z}\right)_0 \Delta z, \qquad \text{if}\quad \Delta x, \Delta y \Delta z \approx 0 \;.$$

is not a precise mathematical statement, since the symbol $\approx$ does not specify exactly how close the quantitites on either side of the formula are to each other. To fix this up, one would have to specify the error in the approximation. (This can be done, but it is not often used.)

A more fundamental objection is that our discussion of approximations was based on the assumption that the tangent plane is a good approximation to the surface at (x_0, y_0, w_0). Is this really so?

Look at it this way. The tangent plane was determined as the plane which has the same slope as the surface in the **i** and **j** directions. This means the approximation (6) will be good if you move away from (x_0, y_0) in the **i** direction (by taking $\Delta y = 0$), or in the **j** direction (putting $\Delta x = 0$). But does the tangent plane have the same slope as the surface in all the other directions as well?

Intuitively, we should expect that this will be so if the graph of $f(x, y)$ is a "smooth" surface at (x_0, y_0) — it doesn't have any sharp points, folds, or look peculiar. Here is the mathematical hypothesis which guarantees this.

Smoothness hypothesis. We say $f(x, y)$ is **smooth** at (x_0, y_0) if

(8) $\qquad f_x$ and f_y are continuous in some rectangle centered at (x_0, y_0).

If (8) holds, the approximation formula (6) will be valid.

Though pathological examples can be constructed, in general the normal way a function fails to be smooth (and in turn (6) fails to hold) is that one or both partial derivatives fail to exist at (x_0, y_0). This means of course that you won't even be able to write the formula (6), unless you're sleepy. Here is a simple example.

Example 3. Where is $w = \sqrt{x^2 + y^2}$ smooth? Discuss.

Solution. Calculating formally, we get

$$\frac{\partial w}{\partial x} \;=\; \frac{x}{\sqrt{x^2 + y^2}}\,, \qquad\qquad \frac{\partial w}{\partial y} \;=\; \frac{y}{\sqrt{x^2 + y^2}}.$$

These are continuous at all points except $(0,0)$, where they are undefined. So the function is smooth except at the origin; the approximation formula (6) should be valid everywhere except at the origin.

1
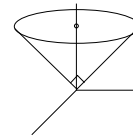
Indeed, investigating the graph of this function, since $w = \sqrt{x^2 + y^2}$ says that

$$\text{height of graph over } (x, y) \;=\; \text{distance of } (x, y) \text{ from } w\text{-axis,}$$

the graph is a right circular cone, with vertex at $(0, 0)$, axis along the w-axis, and vertex angle a right angle. Geometrically the graph has a sharp point at the origin, so there should be no tangent plane there, and no valid approximation formula (6) — there is no linear function which approximates a cone at its vertex.

A non-geometrical argument for the approximation formula

We promised earlier a non-geometrical approach to the approximation formula (6) that would generalize to higher-dimensions, in particular to the 3-variable formula (7). This approach will also show why the hypothesis (8) of *smoothness* is needed. The argument is still imprecise, since it uses the symbol $\approx$, but it can be refined to a proof (which you will find in your book, though it's not easy reading).
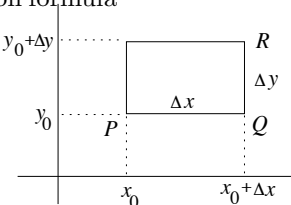
It uses the one-variable approximation formula for a differentiable function $w = f(u)$:

(9) $\Delta w \approx f'(u_0)\Delta u, \qquad \text{if } \Delta u \approx 0 .$

We wish to justify — without using reasoning based on 3-space — the approximation formula

(6) $\Delta w \;\approx\; \left(\dfrac{\partial w}{\partial x}\right)_0 \Delta x + \left(\dfrac{\partial w}{\partial y}\right)_0 \Delta y, \qquad \text{if} \quad \Delta x \approx 0,\ \Delta y \approx 0 .$

We are trying to calculate the change in w as we go from P to R in the picture, where $P = (x_0, y_0)$, $R = (x_0 + \Delta x, y_0 + \Delta y)$. This change can be thought of as taking place in two steps:

(10) $\Delta w \;=\; \Delta w_1 \;+\; \Delta w_2,$

the first being the change in w as you move from P to Q, the second the change as you move from Q to R. Using the one-variable approximation formula (9) :

(11) $\Delta w_1 \;\approx\; \left.\dfrac{d}{dx}\, f(x, y_0)\right|_{x_0} \cdot \Delta x \;=\; f_x(x_0, y_0)\,\Delta x;$

similarly,

$$\Delta w_2 \;\approx\; \left.\dfrac{d}{dy}\, f(x_0 + \Delta x, y)\right|_{y_0} \cdot \Delta y \;=\; f_y(x_0 + \Delta x, y_0)\,\Delta y$$

(12) $\approx f_y(x_0, y_0)\,\Delta y,$

if we assume that f_y is continuous (i.e., f is smooth), since the difference between the two terms on the right in the last two lines will then be like $\epsilon\,\Delta y$, which is negligible compared with either term itself. Substituting the two approximate values (11) and (12) into (10) gives us the approximation formula (6). $\qquad\square$

To make this a proof, the error terms in the approximations have to be analyzed, or more simply, one has to replace the $\approx$ symbol by equalities based on the Mean-Value Theorem of one-variable calculus.

This argument readily generalizes to the higher-dimensional approximation formulas, such as (7); again the essential hypothesis would be smoothness: the three partial derivatives w_x, w_y, w_z should be continuous in a neighborhood of the point (x_0, y_0, z_0).

Tangent approximation

1. a) Find the equation tangent plane to the graph of $z = x^2 + y^2$ at the point (2,1,5).

b) Give the tangent approximation for z near the point $(x_0, y_0) = (2, 1)$.

Answer: a) $\dfrac{\partial z}{\partial x} = 2x$ and $\dfrac{\partial z}{\partial y} = 2y \ \Rightarrow \ \dfrac{\partial z}{\partial x}(2, 1) = 4$ and $\dfrac{\partial z}{\partial y}(2, 1) = 2.$

The tangent plane at (2,1,5) is

$$(z - 5) = \left.\frac{\partial z}{\partial x}\right|_0 (x - 2) + \left.\frac{\partial z}{\partial y}\right|_0 (y - 1) = 4(x - 2) + 2(y - 1).$$

b) The tangent approximation is the same formula, with the interpretation that for a fixed (x_0, y_0) the value of z on the graph of the function is near that of z on the tangent plane. Thus, for $(x_0, y_0) \approx (2, 1)$ we have

$$\Delta z \approx 4\Delta x + 2\Delta y.$$

Tangent approximation

1. Find the equation of the tangent plane to the graph of $z = xy^2$ at the point (1,1,1).

Answer: $\dfrac{\partial z}{\partial x} = y^2$ and $\dfrac{\partial z}{\partial y} = 2xy \Rightarrow \dfrac{\partial z}{\partial x}(1,1) = 1$ and $\dfrac{\partial z}{\partial y}(1,1) = 1$.

The tangent plane at (1,1,1) is

$$(z - 1) = \frac{\partial z}{\partial x}\bigg|_0 (x - 1) + \frac{\partial z}{\partial y}\bigg|_0 (y - 1) = (x - 1) + 2(y - 1).$$

2. Give the linearization of $f(x, y) = e^x + x + y$ at (0,0).

Answer: The tangent approximation formula at the point (x_0, y_0, z_0) is

$$f(x, y) - f(x_0, y_0) \approx f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

(We usually abbreviate this as $\Delta z \approx f_x|_0 \Delta x + f_y|_0 \Delta y$.)

Linearization is just the following form of the tangent approximation formula

$$f(x, y) \approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

In our case,

$$f_x(x, y) = e^x + 1 \text{ and } f_y(x, y) = 1 \Rightarrow f(0,0) = 1, f_x(0,0) = 2, f_y(0,0) = 1$$

Thus, for $(x, y) \approx (0, 0)$ we have

$$f(x, y) \approx 1 + 2x + y.$$

# Critical Points

Critical points:
A standard question in calculus, with applications to many fields, is to find the points where a function reaches its relative maxima and minima.

Just as in single variable calculus we will look for maxima and minima (collectively called *extrema*) at points (x_0, y_0) where the first derivatives are 0. Accordingly we define a *critical point* as any point (x_0, y_0) where
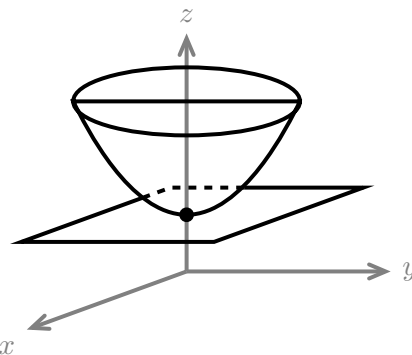
$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \text{ and } \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Often we will abbreviate this as $f_x = 0$ and $f_y = 0$.

Our first job is to verify that relative maxima and minima occur at critical points. The figures below illustrates that they occur at places where the tangent plane is horizontal.
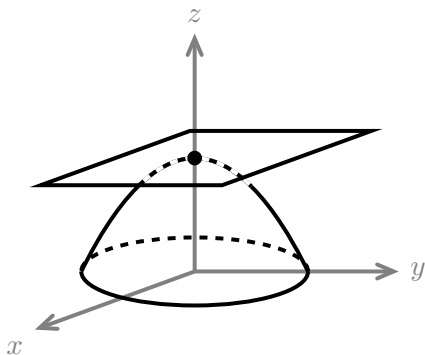
Max. with horizontal tang. plane   Min. with horizontal tang. plane

Since horizontal planes are of the form $z = $ constant. and the equation of the tangent plane at (x_0, y_0, z_0) is

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

we see it is horizontal when

$$f_x(x_0, y_0) = 0 \text{ and } f_y(x_0, y_0) = 0.$$

Thus, extrema occur at critical points. But, just as in single variable calculus, not all critical points are extrema.

**Example:** Find the critical points of $z = x^2 + y^2 + .5$.

Answer: $\dfrac{\partial z}{\partial x} = 2x$ and $\dfrac{\partial z}{\partial y} = 2y$. Clearly the only point where both derivatives are 0 is $(0, 0)$. Thus, there is a single critical point at $(0, 0)$. The figure shows it is clearly the point where z reaches a minimum value. (See the figure above on the right.)

Example: Find the critical points of $z = 1 - x^2 - y^2$.

Answer: $\dfrac{\partial z}{\partial x} = -2x$ and $\dfrac{\partial z}{\partial y} = -2y$. Clearly the only point where both derivatives are 0 is $(0, 0)$. Thus, there is a single critical point at $(0, 0)$. The figure shows it is clearly the point where z reaches a maximum value. (See the figure above on the left.)

Example: Find the critical points of $z = -x^2 + y^2$.

Answer: $\dfrac{\partial z}{\partial x} = -2x$ and $\dfrac{\partial z}{\partial y} = 2y$. Clearly the only point where both derivatives are 0 is $(0,0)$. Thus, there is a single critical point at $(0,0)$. The figure shows it is neither a minimum or a maximum.
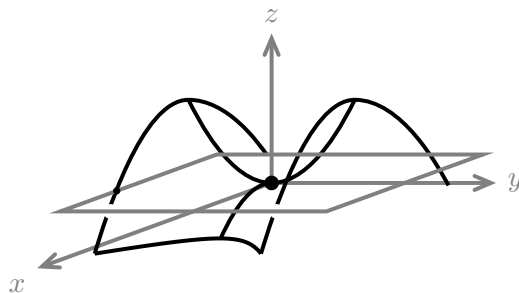
Saddle with horizontal tang. plane

Example: Making a box with minimum material.

A box is made of cardboard with double thick sides, a triple thick bottom, single thick front and back and no top. It's volume $= 3$.

What dimensions use the least amount of cardboard?

Answer: The box shown has dimensions x, y, and z.

The area of one side $= yz$. There are two double thick sides $\Rightarrow$ cardboard used $= 4yz$.

The area of the front (and back) $= xz$. It is single thick $\Rightarrow$ cardboard used $= 2xz$.

The area of the bottom $= xy$. It is triple thick $\Rightarrow$ cardboard used $= 3xy$.

Thus, the total cardboard used is

$$w = 4yz + 2xz + 3xy.$$

The volume $= 3 = xyz \Rightarrow z = \frac{3}{xy}$. Substituting this in the formula for w gives

$$w = \frac{12}{x} + \frac{6}{y} + 3xy.$$

We find the critical points of w.

$$w_x = -\frac{12}{x^2} + 3y = 0, \quad w_y = -\frac{6}{y^2} + 3x = 0.$$

The first equation implies $y = \dfrac{4}{x^2}$. Substituting this in the second equation gives $-\dfrac{6}{16}x^4 + 3x = 0$. Thus, $x = 0$ or 2. We reject 0 since then y is undefined. Using $x = 2$ we find $y = 1$. Thus, there there is one critical point at $(2,1)$. and at this point we have $z = 3/2$.

This point gives the box with minimum cardboard used because physically we know it must have a minimum somewhere. Later we will learn to check this with the second derivative test.
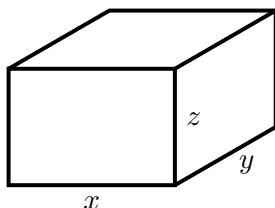
Critical points

1. Find all the critical points of

$$f(x,y) = x^6 + y^3 + 6x - 12y + 7.$$

Answer: Taking the first partials:

$$\frac{\partial z}{\partial x} = 6x^5 + 6 \text{ and } \frac{\partial z}{\partial y} = 3y^2 - 12.$$

Setting these equal to 0 gives

$$x^5 = -1 \implies x = -1 \text{ and } 3y^2 = 12 \implies y = \pm 2.$$

The critical points are $(-1, 2)$ and $(-1, -2)$.

18.02SC Multivariable Calculus
Fall 2010