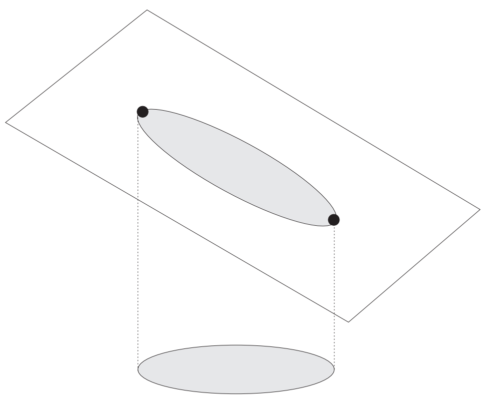
Lagrange multipliers.

The method of **Lagrange multipliers** deals with the problem of finding the maxima and minima of a function subject to a **side condition**, or **constraint**.

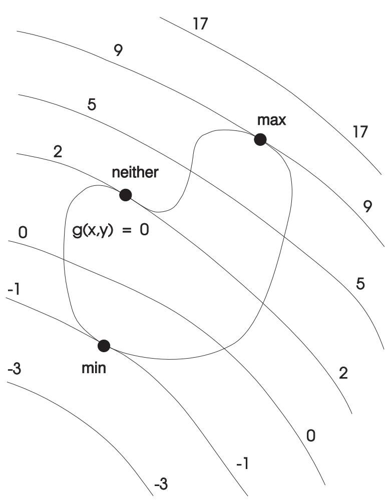
Example. Find the highest and lowest points on the plane $x + 2y + 3z = 6$ which lie above the circle $x^2 + y^2 = 1$.



The plane as a whole has no "highest point" and no "lowest point". If we confine ourselves to the part of the plane lying above the circle, there is a highest point and a lowest point. These **constrained extrema** are the highest and lowest points on the oval in the plane that lies above the circle. $\square$

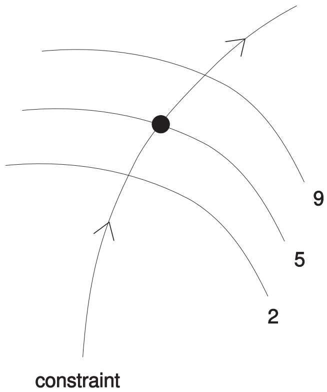
We want to know how to locate such **constrained extrema**. We will provide a heuristic argument based on level curves. (The same argument works in higher dimensions.)

Suppose we are trying to find the largest and smallest value of $z = f(x, y)$ subject to the constraint $g(x, y) = 0$. The equation $g(x, y) = 0$ defines a curve in the x-y-plane; we draw the constraint curve, together with the level curves for f.



The constraint curve is the dented oval; the numbers are the "altitude numbers" for the respective level curves. As you walk around the dented oval, the highest altitude you encounter is 9. It occurs at the point labelled "max". The lowest altitude you encounter is -1. It occurs at the point labelled "min".

Something interesting happens at the max and at the min. You can see from the picture that at those points, *the constraint curve is tangent to a level curve.* Why should this happen?



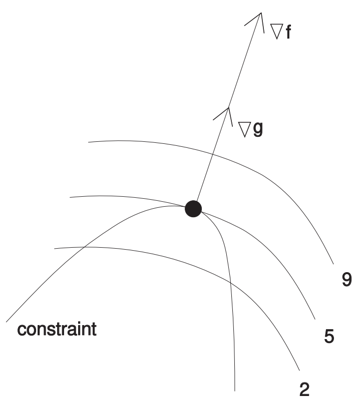
Imagine walking along the constraint curve. As you walk, you note the altitude numbers of the level curves that you pass through, hoping to find a constrained max. Suppose you're at a point where the constraint *passes through* a level curve (as opposed to being *tangent* to it). Then (as the picture shows) as you pass through the level curve, you will encounter level curves *with higher altitude numbers* "on the other side". Therefore, you haven't reached the max yet.

In the context of this picture, the only way you can reach a max is if the constraint curve "turns around" at some point. At the place where it "turns around", it will be tangent to a level curve.

This heuristic argument suggests that we should locate points where the constraint curve is tangent to a level curve. We need a way of locating such points of tangency algebraically.

> Note that some points of this type will be neither maxima nor minima. An example of such a point is labelled "neither" in the picture above.

What happens, then, at a point of tangency? Recall that the gradient vector at a point is perpendicular to the level curve passing through the point. Hence, ∇f is perpendicular to the level curve at the point. Moreover, the ∇g is perpendicular to $g(x,y) = 0$ at the point, because $g(x,y) = 0$ is a level curve of the function $z = g(x,y)$.



Since the curves are tangent at the point, and since the gradient vectors are perpendicular to the respective curves, the gradient vectors must lie along the same line. Vectors which lie along the same line are multiples. Therefore,

$$\nabla f = \lambda \nabla g$$

for some number λ. The number λ is the **Lagrange multiplier**; in some books, you will see L or l used instead.

In other words, to find the points where a constrained max or min could occur, we should locate all points which satisfy

$$\nabla f = \lambda \nabla g \quad \text{and} \quad g = 0.$$

The constraint equation is included, because any solution to our problem must satisfy the constraint.

When you're setting up the equations to be solved, here's the best way to do it. Suppose f and g are functions of x and y. Then the equations are:

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \quad g(x, y) = 0.$$

The first two equations represent, in component form, the multiplier equation $\nabla f = \lambda \nabla g$. The third equation is the constraint.

You can put λ on either the f side or the g side of the first two equations (i.e. on the f side in both or on the g side in both) — whichever is most convenient. For example, this is also okay:

$$\lambda \frac{\partial f}{\partial x} = \frac{\partial g}{\partial x}, \quad \lambda \frac{\partial f}{\partial y} = \frac{\partial g}{\partial y}, \quad g(x, y) = 0.$$

Note that for the purposes of computing the partial derivatives, your constraint must be written in the form $g = 0$. For instance, if you constraint is given as $x^2 = 4 - y^2$, then rewrite it as $x^2 + y^2 - 4 = 0$ and take $g(x, y) = x^2 + y^2 - 4$.

Example. Find the largest and smallest values of $z = x^2 y$ subject to the constraint $x^2 + 8y^2 = 24$.

Our constraint is $g(x, y) = x^2 + 8y^2 - 24 = 0$. We have

(1) $2xy = 2x\lambda$, $x(y - \lambda) = 0$, (2) $x^2 = 16y\lambda$, (3) $x^2 + 8y^2 = 24$

Notice that we were careful in the first equation *not to divide by x*!

> One of the most common mistakes in solving simultaneously is dividing by something with a variable in it. When you do this, you are liable to drop roots.

We will organize our work in **tree form**. It helps us avoid this problem. It will also help you avoid going in circles, and forgetting solutions by forgetting to take both positive and negative square roots.

The rules for setting up a solution tree are roughly as follows:

- Use one numbered equation at a time. Don't reintroduce a numbered equation into a given branch of the tree.

- A branch **forks** when you take cases. You take cases when you must take square roots, because you must take positive and negative square roots. You also take cases when you have something of the form

$$\text{blip} \cdot \text{blap} = 0.$$

  In this situation, the tree forks into branches with blip $= 0$ and with blap $= 0$.

- When you've solved for the variables, combine them to get solution points. However, you are only allowed to use information in the tree "below" you, i.e. along your branch toward the root. *You may not mix values from different branches, because different branches represent different cases.*

This is easier to show than to explain, so look at the work below.

$$
\begin{array}{cccccccccc}
 & & & (1)\quad x(y-\lambda)=0 & & & & & & \\
 & & \swarrow & & \searrow & & & & & \\
 & x=0 & & & & y=\lambda & & & & \\
(3) & x^2+8y^2=24 & & & (2) & x^2=16\lambda y & & & & \\
 & 8y^2=24 & & & & x^2=16y^2 & & & & \\
 & y^2=3 & & & (3) & x^2+8y^2=24 & & & & \\
 & \downarrow & \searrow & & & 24y^2=24 & & & & \\
 & y=\sqrt{3} & & y=-\sqrt{3} & & y^2=1 & & & & \\
 & (0,\sqrt{3}) & & (0,-\sqrt{3}) & & \downarrow & \searrow & & & \\
 & & & & & y=1 & & y=-1 & & \\
 & & & & & x^2=16 & & x^2=16 & & \\
 & & & & \swarrow & \downarrow & & \downarrow & \searrow & \\
 & & & x=4 & & x=-4 & & x=4 & & x=-4 \\
 & & & (4,1) & & (-4,1) & & (4,-1) & & (-4,-1)
\end{array}
$$

Once we've obtained our candidate points, we test them by plugging them back into the function we're trying to maximize/minimize. The largest values are the maxima; the smallest values are the minima.

(x,y)	$(0,\sqrt{3})$	$(0,-\sqrt{3})$	$(4,1)$	$(4,-1)$	$(-4,1)$	$(-4,-1)$
$f(x,y)$	0	0	16	-16	16	-16
			max	min	max	min

It is possible for a constrained optimization problem to have no solutions. We will normally avoid this difficulty by working with constraints which satisfy a technical condition called **compactness**. □

In the same way, if you want to optimize $w = f(x,y,z)$ subject to a constraint $g(x,y,z) = 0$, the multiplier equations are

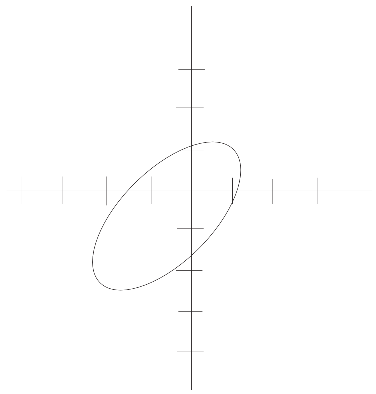
$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \quad \frac{\partial f}{\partial z} = \lambda \frac{\partial g}{\partial z}, \quad g(x,y,z) = 0.$$

You may put λ on either side of the first three equations (i.e. on the f side in all three or on the g side in all three). For example, we can use these equations instead of those above:

$$\lambda \frac{\partial f}{\partial x} = \frac{\partial g}{\partial x}, \quad \lambda \frac{\partial f}{\partial y} = \frac{\partial g}{\partial y}, \quad \lambda \frac{\partial f}{\partial z} = \frac{\partial g}{\partial z}, \quad g(x,y,z) = 0.$$

Problems on Lagrange multipliers.

1. The picture below shows a constraint curve $g(x, y) = 0$ (the oval). Approximate the largest and smallest values of $z = x + y$ subject to the constraint.

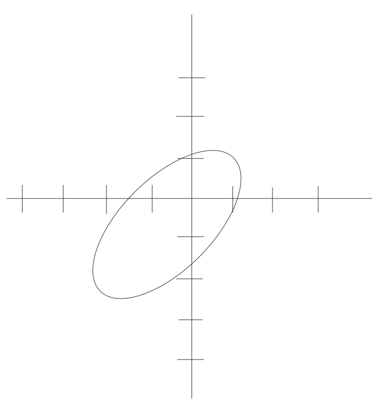


2. Find the highest and lowest points on the surface $f(x, y) = x^2 + 2xy + \frac{5}{2}y^2$ which lie above the circle $x^2 + y^2 = 20$.

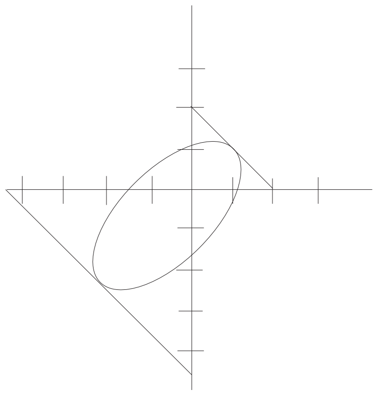
3. Find the largest and smallest values of $w = x + 2y^2 - z^2$ subject to the constraint $x^2 + 4y^2 + 2z^2 = 17$.

Solutions to the Problems on Lagrange multipliers.

1. The picture below shows a constraint curve $g(x, y) = 0$ (the oval). Approximate the largest and smallest values of $z = x + y$ subject to the constraint.



The level curves of $z = x + y$ are lines making an angle of 135° with the positive x-axis. Since the constrained extrema occur where the level curves are tangent to the oval, we see by sketching that the max is approximately $z = 2$ and the min is approximately $z = -4.5$. □



2. Find the highest and lowest points on the surface $f(x,y) = x^2 + 2xy + \frac{5}{2}y^2$ which lie above the circle $x^2 + y^2 = 20$.

We want to find the maxima and minima of f subject to the constraint $g(x,y) = x^2 + y^2 - 20 = 0$.

(1) $2x + 2y = 2x\lambda$, $x + y = \lambda x$, (2) $2x + 5y = 2y\lambda$, (3) $x^2 + y^2 = 20$

(1) $x + y = \lambda x$

$\swarrow$ $\searrow$

$x = 0$
$y = 0$
(3) $x^2 + y^2 = 20$
$0 = 20$
$\times$

$x \neq 0$
$\lambda = \dfrac{x+y}{x}$
(2) $2x + 5y = 2y\lambda$
$2x + 5y = \dfrac{2y(x+y)}{x}$
$2x^2 + 3xy - 2y^2 = 0$
$(2x - y)(x + 2y) = 0$

$\swarrow$ $\searrow$

$y = 2x$
(3) $x^2 + y^2 = 20$
$5x^2 = 20$
$x^2 = 4$

$\swarrow$ $\downarrow$

$x = 2$
$y = 4$
$(2, 4)$

$x = -2$
$y = -4$
$(-2, -4)$

$x = -2y$
(3) $x^2 + y^2 = 20$
$5y^2 = 20$
$y^2 = 4$

$\swarrow$ $\downarrow$

$y = 2$
$x = -4$
$(-4, 2)$

$y = -2$
$x = 4$
$(4, -2)$

(x,y)	$(2,4)$	$(-2,-4)$	$(-4,2)$	$(4,-2)$
$f(x,y)$	60	60	10	10
	max	max	min	min

3. Find the largest and smallest values of $w = x + 2y^2 - z^2$ subject to the constraint $x^2 + 4y^2 + 2z^2 = 17$.

The constraint is $g(x,y,z) = x^2 + 4y^2 + 2z^2 - 17 = 0$. The multiplier equations are

(1) $\lambda = 2x$, (2) $4\lambda y = 8y$, $y(\lambda - 2) = 0$, (3) $-2\lambda z = 4z$, $z(\lambda + 2) = 0$, (4) $x^2 + 4y^2 + 2z^2 = 17$

$$
\begin{array}{ll}
(1) & \lambda = 2x \\
(2) & y(\lambda - 2) = 0 \\
 & y(2x - 2) = 0
\end{array}
$$

$\swarrow$ $\qquad\qquad$ $\searrow$

$$
\begin{array}{ll}
 & y = 0 \\
(3) & z(\lambda + 2) = 0 \\
 & z(2x + 2) = 0
\end{array}
\qquad\qquad
\begin{array}{ll}
 & x = 1 \\
(3) & z(\lambda + 2) = 0 \\
 & z(2x + 2) = 0
\end{array}
$$

From $y = 0$:

$\downarrow$ $\qquad$ $\searrow$

$$
\begin{array}{ll}
 & z = 0 \\
(4) & x^2 + 4y^2 + 2z^2 = 17 \\
 & x^2 = 17
\end{array}
\qquad
\begin{array}{ll}
 & x = -1 \\
(4) & x^2 + 4y^2 + 2z^2 = 17 \\
 & z^2 = 8
\end{array}
$$

From $x^2 = 17$: $\swarrow$ $\downarrow$

$$
\begin{array}{c}
x = \sqrt{17} \\
(\sqrt{17}, 0, 0)
\end{array}
\qquad
\begin{array}{c}
x = -\sqrt{17} \\
(-\sqrt{17}, 0, 0)
\end{array}
$$

From $z^2 = 8$: $\downarrow$ $\searrow$

$$
\begin{array}{c}
z = \sqrt{8} \\
(-1, 0, \sqrt{8})
\end{array}
\qquad
\begin{array}{c}
z = -\sqrt{8} \\
(-1, 0, -\sqrt{8})
\end{array}
$$

From $x = 1$:

$\downarrow$ $\qquad$ $\searrow$

$$
\begin{array}{c}
x = -1 \\
\times
\end{array}
\qquad
\begin{array}{ll}
 & z = 0 \\
(4) & x^2 + 4y^2 + \\
 & 2z^2 = 17 \\
 & \downarrow \\
 & \downarrow \\
 & \downarrow \\
 & y^2 = 4
\end{array}
$$

From $y^2 = 4$: $\swarrow$ $\downarrow$

$$
\begin{array}{c}
y = 2 \\
(1, 2, 0)
\end{array}
\qquad
\begin{array}{c}
y = -2 \\
(1, -2, 0)
\end{array}
$$

(x, y, z)	$(\sqrt{17}, 0, 0)$	$(-\sqrt{17}, 0, 0)$	$(-1, 0, \sqrt{8})$	$(-1, 0, -\sqrt{8})$	$(1, 2, 0)$	$(1, -2, 0)$
$w(x, y, z)$	$\sqrt{17}$	$-\sqrt{17}$	-9	-9	9	9
			min	min	max	max

