March 2, 4

14.6 The Chain Rule

More precisely: the chain rules. For functions of several variables there are many versions of the chain rule depending on what is a function of what. To get the idea, let's have a few special cases.

Case 1: Let $z = f(x,y)$, $x = g(t)$, $y = h(t)$.

So z is a function of x, y, and each x and y are functions of some variable t. Then z is a function of t:

$$z = f(g(t), h(t)),$$

What is $\frac{dz}{dt}$? If f, g, and h are all differentiable we have:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}.$$

If we want $\left.\frac{dz}{dt}\right|_{t=p}$, denote $(a,b) = (x(p), y(p))$:

$$\left.\frac{dz}{dt}\right|_{t=p} = \left.\frac{\partial z}{\partial x}\right|_{(a,b)} \cdot \left.\frac{dx}{dt}\right|_{p} + \left.\frac{\partial z}{\partial y}\right|_{(a,b)} \cdot \left.\frac{dy}{dt}\right|_{p}$$

Why? Here is some justification. We want to convince ourselves that

$$\left.\frac{dz}{dt}\right|_{t=p} = f_x(a,b)\cdot g'(p) + f_y(a,b)\cdot h'(p) \qquad (*)$$

We have

$$\left.\frac{dz}{dt}\right|_{t=p} = \lim_{\Delta t\to 0}\frac{\Delta z}{\Delta t} \qquad \text{so} \qquad \left.\frac{dz}{dt}\right|_{t=p} \cong \frac{\Delta z}{\Delta t}$$

How can we estimate:

$$\frac{\Delta z}{\Delta t} = \frac{\Delta f}{\Delta t} \; ?$$

A change Δt produces changes Δx and Δy in x and y where:

$$\Delta x = g(p+\Delta t) - g(p) \cong g'(p)\cdot\Delta t, \quad \Delta y = h(p+\Delta t) - h(p) \cong h'(p)\cdot\Delta t$$

Changes Δx and Δy produce a change in f approximately equal:

$$\Delta f = f(a+\Delta x, b+\Delta y) - f(a,b) \cong f_x(a,b)\cdot\Delta x + f_y(a,b)\cdot\Delta y =$$

$$\cong f_x(a,b)\cdot g'(p)\cdot\Delta t + f_y(a,b)\,h'(p)\cdot\Delta t$$

So

$$\frac{\Delta f}{\Delta t} \cong f_x(a,b)\cdot g'(p) + f_y(a,b)\cdot h'(p)$$

This gives us some hint why $(*)$ is true.

Ex: Let $z = f(x,y) = x\sin(y)$, $x = t^2$, $y = 2t+1$.

Find

(a) $\frac{dz}{dt}$ (b) $\frac{dz}{dt}\Big|_{t=0}$.

We use the chain rule:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial z}{\partial y}\cdot\frac{dy}{dt} = \sin(y)\cdot 2t + x\cos(y)\cdot 2 =$$

$$= \sin(2t+1)\cdot 2t + t^2\cos(2t+1)\cdot 2$$

$$\frac{dz}{dt}\Big|_{t=0} = 0.$$

Or we can leave:

$$\frac{dz}{dt} = \sin(y)\cdot 2t + x\cos(y)\cdot 2$$

and reason as follows. When $t=0$, $x=0$, $y=1$. So

$$\frac{dz}{dt}\Big|_{t=0} = \sin(1)\cdot 0 + 0\cdot\cos(1)\cdot 2 = 0.$$

A more interesting problem:

Ex: Let $z = (x+y)e^{y}$. Suppose $x = g(t)$, $y = h(t)$. Given that $g(0) = 2$, $h(0) = 0$, $g'(0) = -0.3$, $h'(0) = 2.5$, find $\left.\frac{dz}{dt}\right|_{t=0}$

When $t = 0$, $x = 2$ and $y = 0$; So

$$\left.\frac{dz}{dt}\right|_{t=0} = \left.\frac{\partial z}{\partial x}\right|_{(2,0)} \cdot \left.\frac{dx}{dt}\right|_{t=0} + \left.\frac{\partial z}{\partial y}\right|_{(2,0)} \cdot \left.\frac{dy}{dt}\right|_{t=0} =$$

$$= \left.\frac{\partial z}{\partial x}\right|_{(2,0)} \cdot g'(0) + \left.\frac{\partial z}{\partial y}\right|_{(2,0)} \cdot h'(0)$$

$$\left.\frac{\partial z}{\partial x}\right|_{(2,0)} = \left.e^{y}\right|_{(2,0)} = 1, \quad \left.\frac{\partial z}{\partial y}\right|_{(2,0)} = \left.e^{y} + (x+y)e^{y}\right|_{(2,0)} = 3$$

Thus:

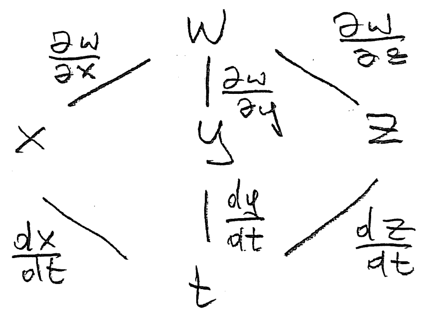
$$\left.\frac{dz}{dt}\right|_{t=0} = 1 \cdot (-0.3) + 3 \cdot 2.5 = 7.2$$

Case 1a: $w = f(x, y, z)$, $x = x(t)$, $y = y(t)$, $z = z(t)$.

Then:

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial w}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial w}{\partial z} \cdot \frac{dz}{dt}.$$

The book gives a mnemotechnical way of obtaining versions of the chain rule in all possible cases. Draw a diagram containing all the variables and edges which represent which variable depends on which. For example:



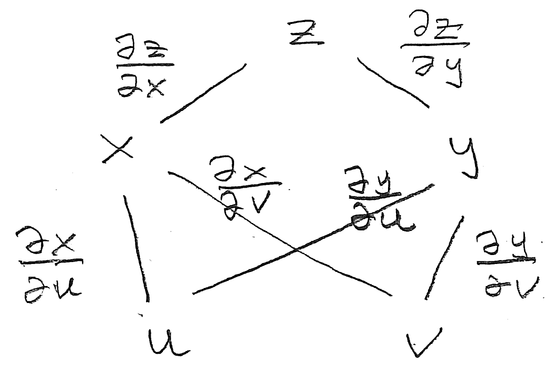
Label each edge with the corresponding derivative. To get $\frac{dw}{dt}$ look at all paths connecting w and t, multiply the derivatives along each path, and add contributions from each path.

<u>Case 2</u>: $z = f(x,y)$, $x = g(u,v)$, $y = h(u,v)$.

Then z depends on u, v:

$$z = f(x,y) = f(g(u,v), h(u,v)).$$

Find $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$.



$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial u}$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial v}$$

Ex: $z = x^2 e^y$, $x = 4u$, $y = 3u^2 - 2v$.

Find $\left.\frac{\partial z}{\partial u}\right|_{(u,v)=(1,2)}$.

$$\frac{\partial z}{\partial u} = (2xe^y) \cdot 4 + x^2 e^y \cdot 6u$$

$(u,v) = (1,2) \longrightarrow x = 4, \; y = 3 - 4 = -1$

$$\left.\frac{\partial z}{\partial u}\right|_{(1,2)} = 8e^{-1} \cdot 4 + 16 \cdot e^{-1} \cdot 6 = (32 + 96)e^{-1} = \frac{128}{e}$$

We will do an applied example later.

14.7 Second Partials

Given $z = f(x,y)$ each partial derivative

$$\frac{\partial z}{\partial x} = f_x(x,y) \quad , \quad \frac{\partial z}{\partial y} = f_y(x,y)$$

is again a function of (x,y). So we can take their partial derivatives. Partial derivatives of partial derivatives are called second partial derivatives. We denote them:

$$(f_x)_x = f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}[f_x(x,y)]$$

$$(f_x)_y = f_{\underset{\text{first}}{x}\underset{\text{second}}{y}} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\underset{\text{second}}{\partial y}\,\underset{\text{first}}{\partial x}} = \frac{\partial}{\partial y}[f_x(x,y)]$$

$$(f_y)_x = f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}[f_y(x,y)]$$

$$(f_y)_y = f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}[f_y(x,y)].$$

Four partials

$$f_{xy} \, , \, f_{yx}$$

are called mixed partials.

Ex: $z = f(x,y)$, $f(x,y) = x^3 + x^2y^2 - y^4$.

Find second partials.

$f_x = 3x^2 + 2xy^2$, $f_y = 2x^2y - 4y^3$

$f_{xx} = 6x + 2y^2$, $f_{xy} = 4xy$

$f_{yy} = 2x^2 - 12y^2$, $f_{yx} = 4xy$

$f_{xy} = f_{yx}$

Is it always so?

Th: If $f_{xy}(x,y)$ and $f_{yx}(x,y)$ are continuous,

then $f_{xy}(x,y) = f_{yx}(x,y)$.

Second partials important, among others, in the context of

- Taylor approximations of order 2
- Second Derivative Test for Local Extrema

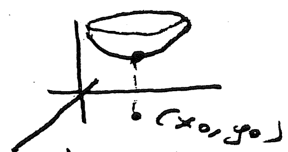
Remark: We are not going to talk much about conditions for differentiability of a function $z = f(x,y)$ but let us state at least the following:

Th: If $f_x(x,y)$, $f_y(x,y)$ exist and are continuous in a disk centered at (a,b), then $f(x,y)$ is differentiable at (a,b).

15.1 Local Extrema of a Function $z=f(x,y)$

<u>Def</u>: (a) $f(x,y)$ has a <u>local minimum</u> at (x_0,y_0) if for all (x,y) in some open disk around (x_0,y_0) we have:

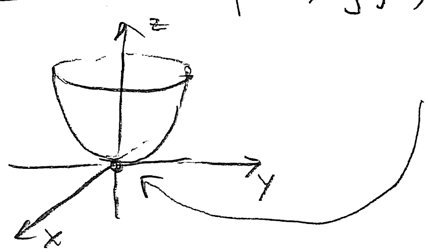
$$f(x,y) \geq f(x_0,y_0).$$



(b) $f(x,y)$ has a <u>local maximum</u> at (x_0,y_0) if for all (x,y) in some open disk around (x_0,y_0) we have

$$f(x,y) \leq f(x_0,y_0).$$

<u>Ex</u>: $z=f(x,y)$, $f(x,y)=x^2+y^2$.



$(0,0)$ a local minimum (and global minimum).

If (x_0,y_0) is a local extremum of $f(x,y)$, then both cross-sections $f(x,y_0)$ and $f(x_0,y)$ have a local extremum at x_0 and y_0, respectively. Thus, their derivatives are 0 at at x_0 and y_0, respectively or the derivatives do not exist. Thus:

<u>Th</u>: If $f(x,y)$ has a local extremum at (x_0,y_0), then

$f_x(x_0,y_0)=0$ or $f_x(x_0,y_0)$ does not exist and

$f_y(x_0,y_0)=0$ or $f_y(x_0,y_0)$ does not exist.

Note if the tangent plane at $(x_0, y_0, f(x_0,y_0))$ exists and (x_0, y_0) is a local extremum, the tangent plane must be horizontal $f_x(x_0,y_0)=f_y(x_0,y_0)=0$.

Def: (x_0, y_0) is a critical point of $f(x,y)$ if

$f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ or at least one of $f_x(x_0, y_0)$, $f_y(x_0, y_0)$ does not exist.

Th: If $f(x,y)$ has a local extremum at (x_0, y_0), then (x_0, y_0) is a critical point.

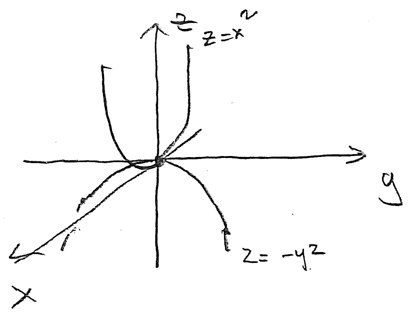
Note if (x_0, y_0) is a critical point and both partials exist, then they are both 0. Thus $\nabla f(x_0, y_0) = \vec{0}$. The gradient is the zero vector at a point of local extremum.

Question: Let $z = f(x, y)$, (x_0, y_0) be a critical point of $f(x,y)$. Does (x_0, y_0) have to be a local extremum?

No: Consider

Ex: $f(x,y) = x^2 - y^2$. $f_x(x,y) = 2x$, $f_y(x,y) = 2y$

At $(x_0, y_0) = (0, 0)$ (and only there) both partials are 0, so $(0, 0)$ is a critical point. Do we have a local minimum or maximum at $(0,0)$?



$f(x,y) = x^2 - y^2$

Cross-section with $x=0$:

$f(0,y) = -y^2$

Cross-section with $y=0$:

$f(x,y) = x^2$

So $f(0,0)=0$ is neither the smallest not the largest value no matter how small disk about $(0,0)$ we take.

The graph is saddle-shaped, the contour diagram shows the characteristic saddle behavior:

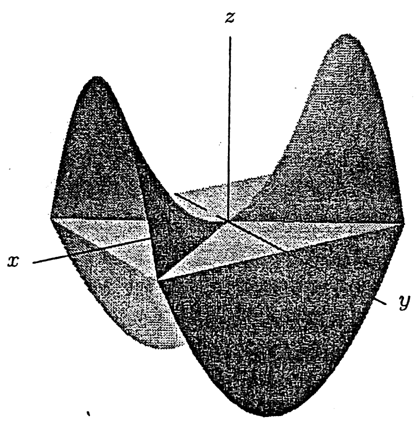


Figure 8.55: Graph of $f(x,y) = x^2 - y^2$ showing plane $z=0$

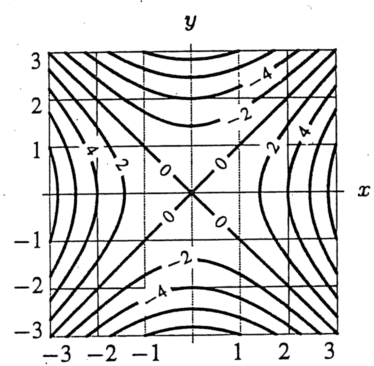


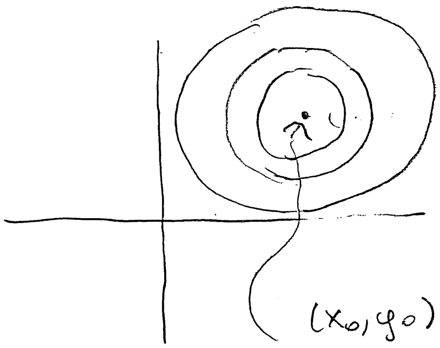
Figure 8.54: Contour map of $f(x,y) = x^2 - y^2$

Def: (x_0, y_0) is a saddle point of $f(x,y)$ if (x_0,y_0) is a critical point of $f(x,y)$ and in every open disk around (x_0,y_0) there exist points (x_1,y_1), (x_2,y_2) such that:

$$f(x_1,y_1) < f(x_0,y_0), \quad f(x_2,y_2) > f(x_0,y_0).$$

Remark: If $f(x,y)$ has a critical point at (x_0,y_0), then (x_0,y_0) is a local minimum, a local maximum or a saddle point.

We see above a typical appearance of a contour diagram around a saddle point. If (x_0, y_0) is a local minimum or maximum, the contour diagram looks something like:

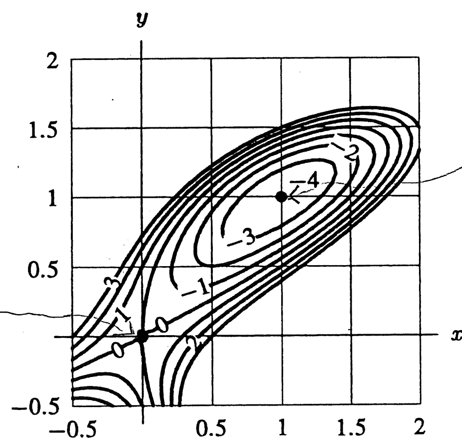


Values of contours <u>increase</u> when you move away from (x_0, y_0) if (x_0, y_0) is a <u>minimum</u>.

Values of contours <u>decrease</u> when we move away from (x_0, y_0) if (x_0, y_0) is a maximum.

<u>Ex</u>: If the contour map of $f(x,y)$ is:

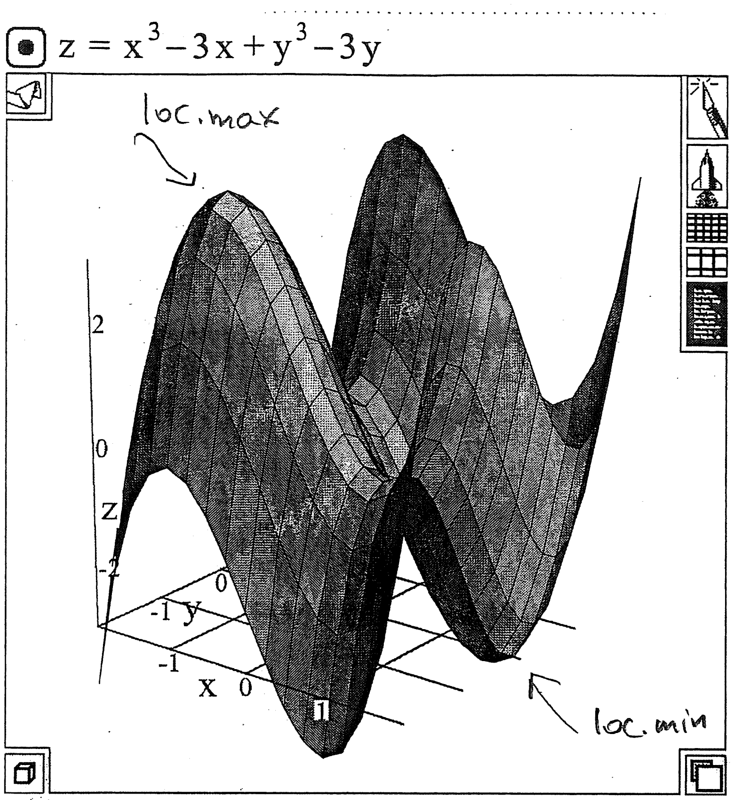
(0,0) saddle



(1,1) local minimum

Where do you see likely critical points? Are they minima, maxima or saddle points?

Ex :

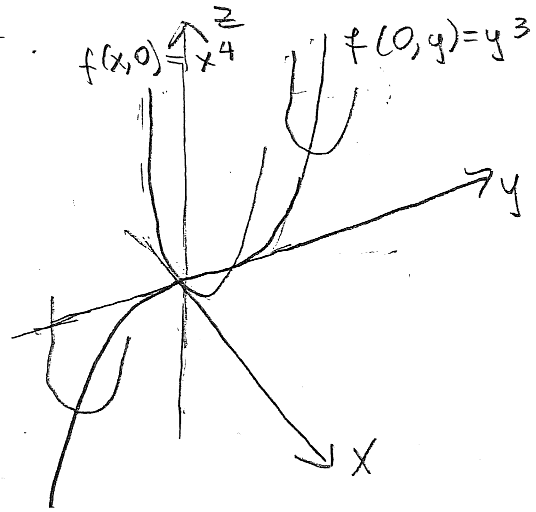


Two saddle points that we can see.

Note: The graph around a saddle point doesn't have to look like a "riding saddle".

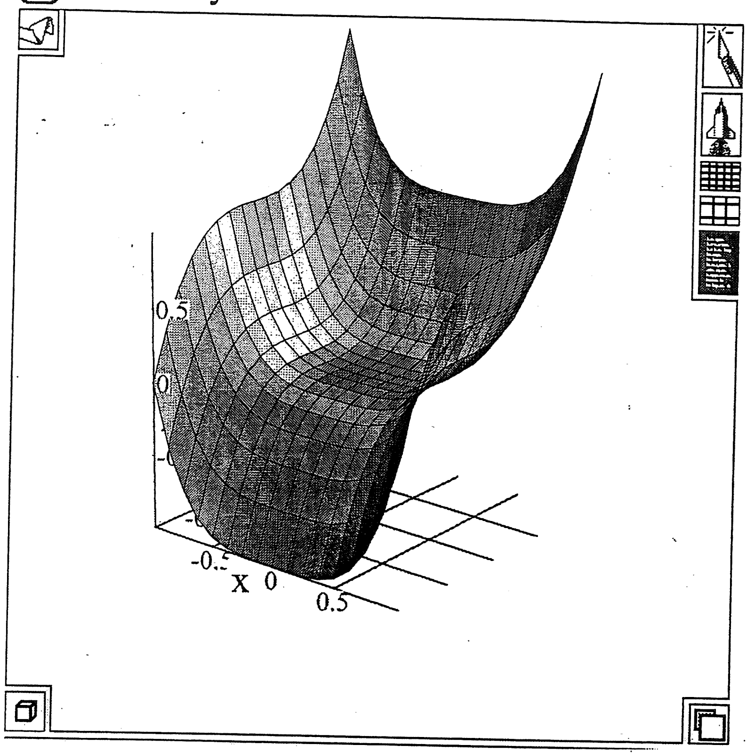
Ex : $f(x,y) = x^4 + y^3$

$f_x = 4x^3$, $f_y = 3y^2$, $(0,0)$ the only critical point.



A saddle point as $f(0,y) = y^3$ takes positive and negative values in any disk about $(0,0)$.

$z = x^4 + y^3$



A "chair".

Is a given critical point a minimum a maximum or a saddle point? The key is the discriminant and the Second Derivative Test.

Let $f(x,y)$ be given and have continuous second partials. We define the discriminant of $f(x,y)$ to be:

$$D(x,y) = f_{xx}(x,y) \cdot f_{yy}(x,y) - f_{xy}(x,y)^2$$

Second Derivative Test for Functions of Two Variables

Suppose (x_0, y_0) is a point where $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$. Let

$$D = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2.$$

- If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then f has a local minimum at (x_0, y_0).
- If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then f has a local maximum at (x_0, y_0).
- If $D < 0$, then f has a saddle point at (x_0, y_0).
- If $D = 0$, anything can happen at (x_0, y_0).

Ex: $f(x,y) = x^3 - 3x + y^3 - 3y$

$f_x = 3x^2 - 3$, $f_y = 3y^2 - 3$

Crt. pts:

$\begin{cases} 3x^2 - 3 = 0 \\ 3y^2 - 3 = 0 \end{cases}$ $x^2 = 1$ and $y^2 = 1$

Crt. pts: $(-1,-1)$, $(-1,1)$, $(1,-1)$, $(1,1)$.

$f_{xx} = 6x$, $f_{yy} = 6y$, $f_{xy} = 0$

$D(x,y) = 36xy$

$D(1,1) > 0$, $f_{xx}(1,1) > 0$ – loc. min

$D(1,-1) < 0$, $D(-1,1) < 0$ – saddle points

$D(-1,-1) > 0$, $f_{xx}(-1,-1) < 0$ – loc. max.

Look at the graph on page 13.