HW 3 due Mon 11:40 am

Quiz 3 on Tues 11:40 am

Today : Linear Approximation & optimization of multivariable functions.

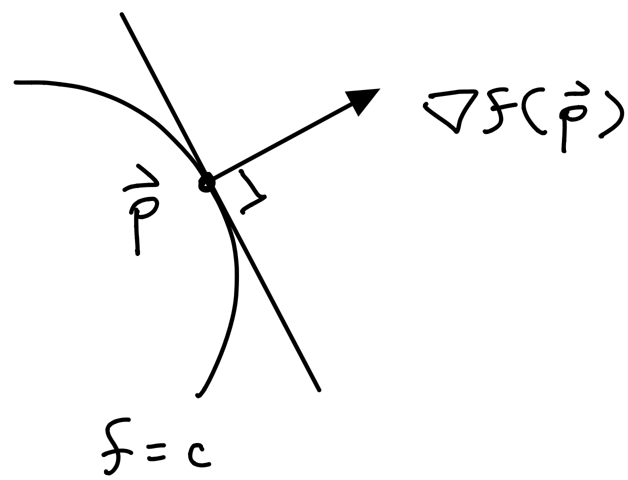
Chapter 4 : Gradients !

There are at least 3 points of view on the gradient. Given a scalar field $f : \mathbb{R}^n \to \mathbb{R}$ we get a vector field $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$.

1) Physics. If f is "height" then the vector field ∇f points "uphill" at each point.

2) Geometry. If $f(\vec{p}) = c$ for some point $\vec{p}$ then the vector

$\nabla f(\vec{p})$ is $\perp$ to the "level set" defined by $f = c =$ constant :



Proof : Suppose $\vec{r}(t)$ is a path that stays in the level set $f = c$, so that

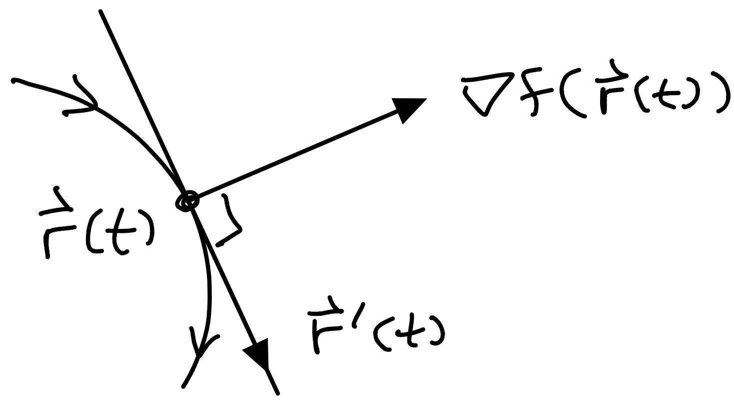
$$f(\vec{r}(t)) = c = \text{constant}$$

for all times t. Then the multivariable chain rule gives

$$[f(\vec{r}(t))]' = c'$$

$$\nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t) = 0,$$

so the velocity $\vec{r}\,'(t)$ is always $\perp$ to the gradient vector $\nabla f(\vec{p}(t))$:



3) Linear Approximation of multivariable functions. We will use the other form of the chain rule:

$$\left[f(\vec{r}(t)) \right]' = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$$

$$\frac{df}{dt} = \left\langle \frac{df}{dx_1}, \dots, \frac{df}{dx_n} \right\rangle \cdot \left\langle \frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right\rangle$$

$$\boxed{\frac{df}{dt} = \frac{df}{dx_1}\frac{dx_1}{dt} + \dots + \frac{df}{dx_n} \cdot \frac{dx_n}{dt}}$$

This tells us about linear approximation. Recall from Calc I & II:

$$f(x) \approx f(a) + f'(a)(x-a)$$

when

$$x \approx a$$

The two variable case says that

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

when $(x,y) \approx (a,b)$. This can be visualized in terms of the tangent plane to the surface $z = f(x,y)$ at the point (a,b,c) where $c = f(a,b)$. To compute the tangent plane, let $F(x,y,z) = f(x,y) - z$ so that

$$z = f(x,y) \iff F(x,y,z) = 0$$

The tangent plane to $F = 0$ at the point (a, b, c) is

$$\nabla F(a,b,c) \cdot \langle x-a, y-b, z-c \rangle = 0$$

Next we observe:

$$dF/dx = \frac{d}{dx}(f(x,y) - z) = df/dx$$

$$dF/dy = \frac{d}{dy}(f(x,y) - z) = df/dy$$

$$dF/dz = \frac{d}{dz}(f(x,y) - z) = -1$$

so that

$$\nabla F(a,b,c) = \langle f_x(a,b), f_y(a,b), -1 \rangle.$$

Hence the tangent plane is

$$f_x(a,b)(x-a) + f_y(a,b)(y-b) - 1(z-c) = 0$$

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

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

So the linear approximation stated

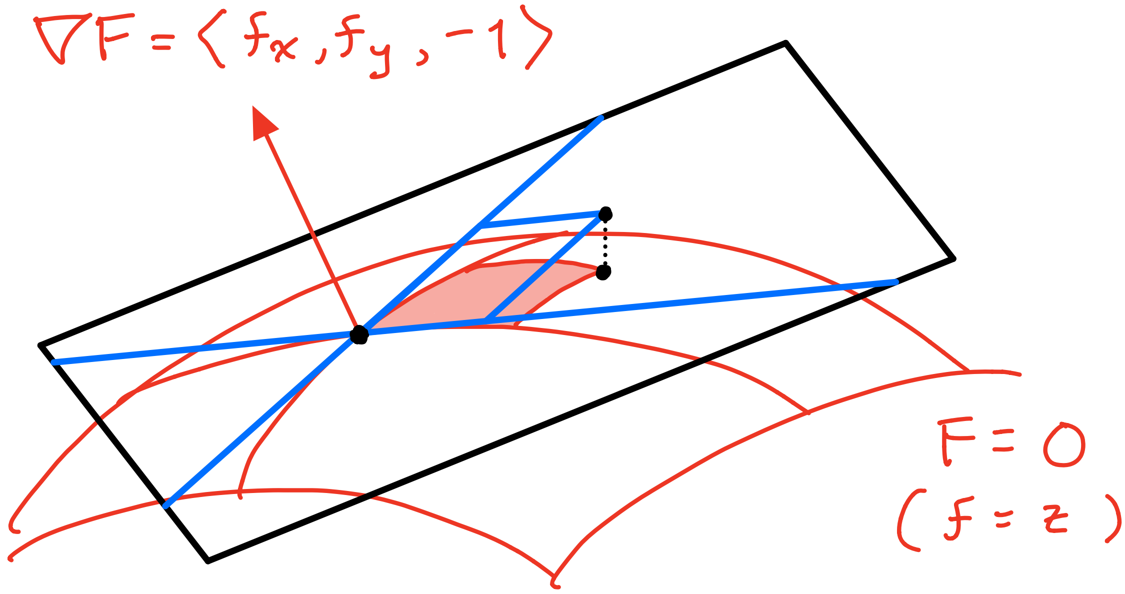
above says that

height of the surface $\approx$ height of the tangent plane at (a,b)

when

$$(x,y) \approx (a,b)$$

Picture :

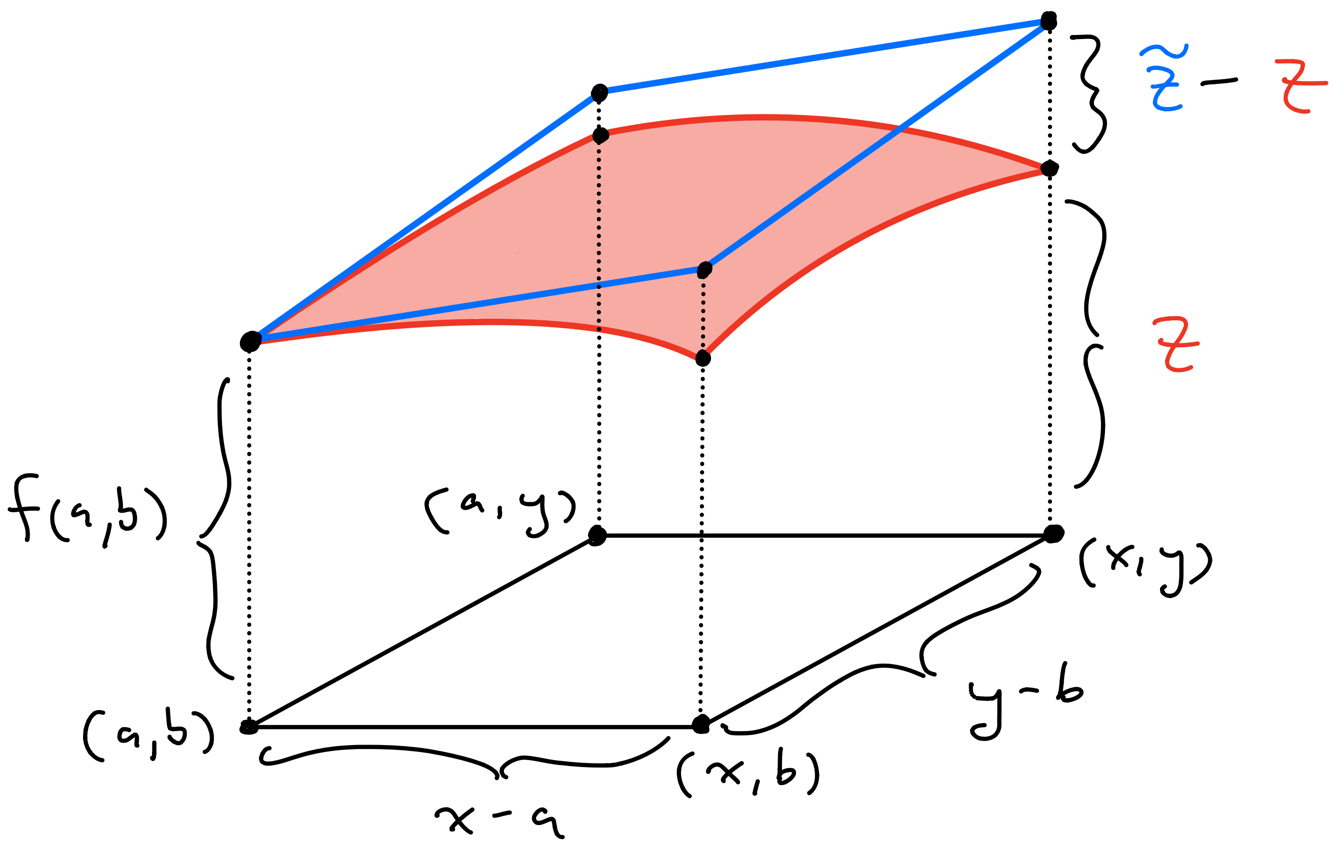


To apply more labels to this picture we need to zoom in. Let

$$\tilde{z} = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$

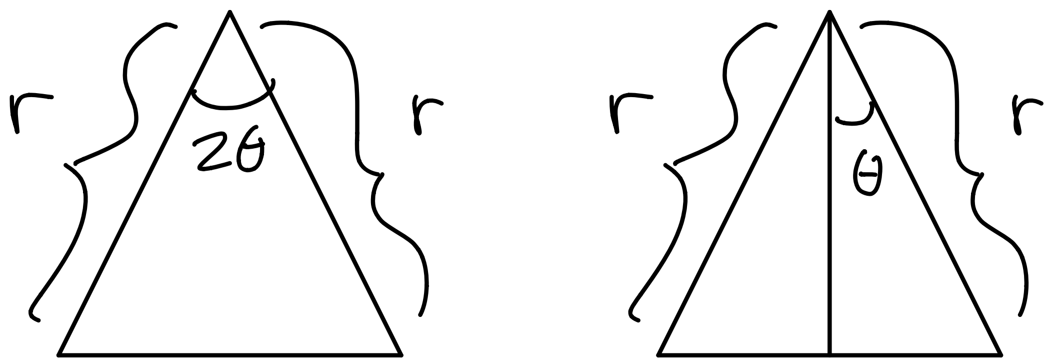
denote the height of the

tangent plane above a general point (x,y). We want to compare $\tilde{z}$ to the height of the surface, which is $z = f(x,y)$:



Idea : If (x,y) is close to (a,b) [so $x \approx a$ and $y \approx b$] then the difference $\tilde{z} - z$ in the picture is small !

Example: We want to estimate the area of the following isoceles triangle when $r \approx 1$ & $\theta \approx \pi/6$:



According to the diagram on the right, the height is $r\cos\theta$ and the base is $2r\sin\theta$, so the area of the triangle is

$$A = (r\cos\theta)(2r\sin\theta)/2$$

$$A = r^2\cos\theta\sin\theta$$

We will compute the linear approx. of $A(r,\theta)$ for $(r,\theta) \approx (1, \pi/6)$:

$$A(r,\theta) \approx A(1,\tfrac{\pi}{6}) + A_r(1,\tfrac{\pi}{6})(r-1) + A_\theta(1,\tfrac{\pi}{6})(\theta - \tfrac{\pi}{6})$$

Let's compute:

$$A(1,\pi/6) = 1^2 \cos(\tfrac{\pi}{6}) \sin(\tfrac{\pi}{6}) = \frac{\sqrt{3}}{2} \cdot \frac{1}{2} = \sqrt{3}/4$$

$$A_r(r,\theta) = 2r \cos\theta \sin\theta$$

$$A_r(1,\tfrac{\pi}{6}) = 2 \cos(\tfrac{\pi}{6}) \sin(\tfrac{\pi}{6}) = \sqrt{3}/2$$

$$A_\theta(r,\theta) = r^2 [\cos\theta \cos\theta - \sin\theta \sin\theta] = r^2 \cos(2\theta)$$

$$A_\theta(1,\tfrac{\pi}{6}) = \cos(\pi/3) = 1/2.$$

Conclusion: For $(r,\theta) \approx (1,\pi/6)$,

$$A(r,\theta) \approx \frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{2}(r-1) + \frac{1}{2}(\theta - \frac{\pi}{6})$$

This is only approximate, but it is easy to compute by hand. :)

Last topic for Chapter 4:

Optimization

Given a scalar field $f : \mathbb{R}^n \to \mathbb{R}$ we want to find the points $\vec{p}$ where $f(\vec{p})$ is a local maximum or minimum. [We want to "climb to the top of the hill".] To do this we look for so-called "critical points" $\vec{p}$ where the gradient vector is zero:

$$\nabla f(\vec{p}) = \langle 0, 0, \ldots, 0 \rangle$$

[Idea: If you are standing at a critical point then the gradient tells you not to move.]

Example: Find the maxima & minima of $f(x,y) = x^2 - x^4 - y^2 + 1$.

Compute the gradient:

$$\nabla f = \langle 2x - 4x^3, -2y \rangle$$

If $\nabla f = \langle 0, 0 \rangle$ then we must have

$$-2y = 0 \rightsquigarrow y = 0$$

$$2x - 4x^3 = 0 \longrightarrow 2x(1 - 2x^2) = 0$$

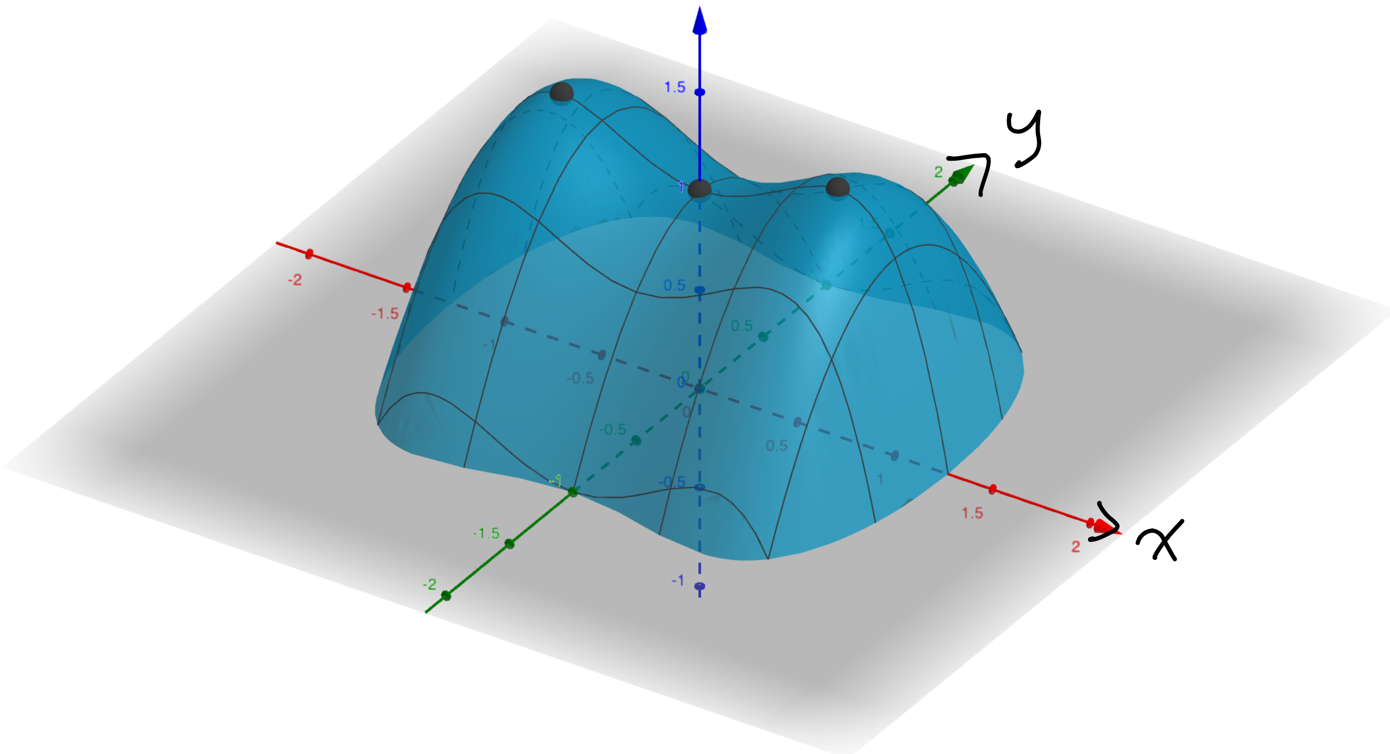
$$\longrightarrow x = 0 \text{ or } x = \pm 1/\sqrt{2}$$

So we obtain 3 critical points:

$$(x,y) = (0,0), \left(\tfrac{1}{\sqrt{2}}, 0\right), \left(-\tfrac{1}{\sqrt{2}}, 0\right)$$

Are these maxima, minima, neither?

If we think of $f(x,y)$ as "height above the point (x,y)" then we can draw a picture:



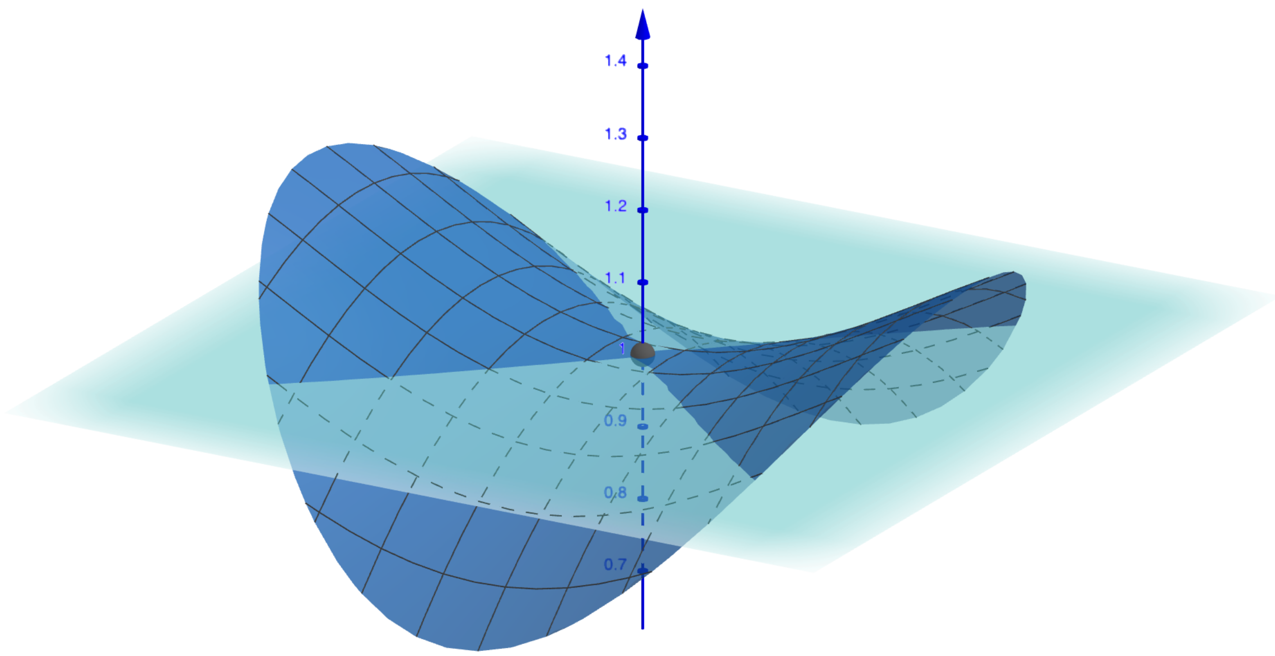
From the picture we can see that f attains a local maximum above

$$(x,y) = (\pm 1/\sqrt{2}, 0)$$

and this maximum height is

$$f\left(\pm \frac{1}{\sqrt{2}}, 0\right) = 5/4 = 1.25$$

But the critical point $\langle 0,0\rangle$ is not a local max or min. It is called a "saddle point". To get a better look we zoom in:



The tangent plane is indeed horizontal (as $\nabla f = \langle 0,0\rangle$ indicates) but there is a problem because the surface is "curving in two different directions". To see this without drawing a picture, we use the "2nd derivative test".

Two Variable 2nd Derivative Test:

Suppose that $\vec{p} = (a,b)$ is a critical point of $f(x,y)$, so that

$$\nabla f(\vec{p}) = \langle f_x(a,b), f_y(a,b) \rangle = \langle 0,0 \rangle$$

To get more information we compute the 2×2 matrix of 2nd derivatives:

$$Hf = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

where, for example, f_{xy} means

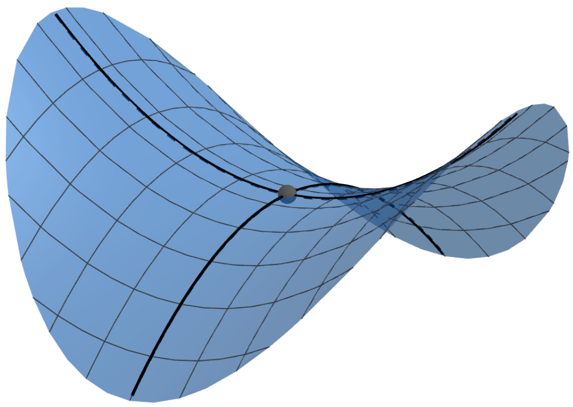
$$f_{xy} = \frac{d}{dy}(f_x) = \frac{d}{dy}\left(\frac{df}{dx}\right)$$

Then we compute the determinant

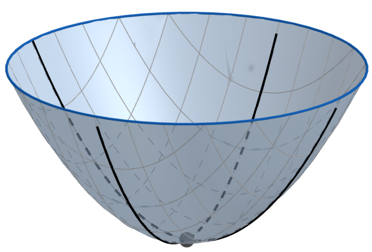
$$\det(Hf) = f_{xx} f_{yy} - f_{xy} f_{yx}$$

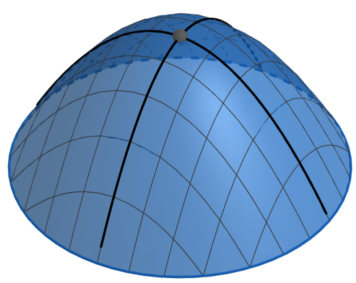
Now there are three basic cases:

- If $\det(Hf)(\vec{p}) < 1$ then the surface is curving in two different directions so that $\vec{p}$ is a saddle point:



- If $\det(Hf)(\vec{p}) > 0$ then any two slices are curving in the same direction, either up or down:





To tell which is which we can check if the "x-slice" is curving up or down:

- If $f_{xx}(\vec{p}) > 0$ then all slices are curving up so $\vec{p}$ is a local <u>min</u>

- If $f_{xx}(\vec{p}) < 0$ then all slices are curving down so $\vec{p}$ is a local <u>max</u>.

- If $\det(Hf)(\vec{p}) = 0$ then we don't have enough information. To get more information we need to compute triple derivatives, etc. [This case won't occur in our class.]

To complete the example

$$f(x,y) = x^2 - x^4 - y^2 + 1,$$

we compute the second derivatives:

$$f_x = 2x - 4x^3$$

$$f_{xx} = 2 - 12x^2$$

$$f_{xy} = 0$$

$$f_y = -2y$$

$$f_{yy} = -2$$

$$f_{yx} = 0.$$

Thus we have

$$Hf = \begin{pmatrix} 2-12x^2 & 0 \\ 0 & -2 \end{pmatrix}$$

$$\det(Hf) = -4 + 24x^2$$

The critical point $(0,0)$ is a saddle because

$$\det(Hf)(0,0) = -4 + 0 < 0.$$

The critical points $(\pm 1/\sqrt{2}, 0)$ are either maxima or minima because

$$\det(Hf)\left(\pm\frac{1}{\sqrt{2}}, 0\right) = -4 + 12 = 8 > 0.$$

To see if these are min or max, we observe that the "y-slice" is always curving down:

$$f_{yy}(x,y) = -2 < 0 \quad \text{for } \underline{\text{any}} \ (x,y).$$

So they are both local <u>maxima</u>.

[You will compute a very similar example on HW 3.3. But it is slightly harder because the "mixed partials" f_{xy}, f_{yx} are not zero.]