- In special relativity, the mass of an object actually depends on its speed:

$$m(v) = \frac{m_0}{\sqrt{1 - v^2/c^2}}$$

where m_0 is the mass at rest, and c is the speed of light

- The kinetic energy of an object in special relativity is

$$K(v) = m(v)\,c^2 - m_0 c^2,$$

whereas in classical physics:

$$K(v) = \frac{1}{2} m_0 v^2$$

- What gives? Lets find the 2nd order Taylor poly. of $m(v)$ at $v_0 = 0$: $\boxed{m(0) = m_0}$

$$m'(v) = \frac{m_0 v}{c^2}\left(1 - \frac{v^2}{c^2}\right)^{-3/2} \Rightarrow \boxed{m'(0) = 0},$$

$$m''(v) = \frac{m_0}{c^2}\left(1 - \frac{v^2}{c^2}\right)^{-3/2} + \frac{3 m_0 v^2}{c^4}\left(1 - \frac{v^2}{c^2}\right)^{-5/2}$$

$$\Rightarrow \boxed{m''(0) = \frac{m_0}{c^2}}$$

$$\Rightarrow T_2(v) = m_0 + 0\cdot(v-0) + \frac{m_0}{2c^2}(v-0)^2$$

$$= m_0 + \frac{m_0}{2c^2} v^2$$

- Substituting into $K(v)$:

$$K(v) \approx T_2(v)\,c^2 - m_0 c^2$$

$$= \left[m_0 + \frac{m_0}{2c^2} v^2\right] c^2 - m_0 c^2$$

$$= \frac{1}{2} m_0 v^2$$

- The classical physics version of kinetic energy is the 2nd-order Taylor polynomial of the special relativity version.

• The big question: how do we generalize to functions of more than one variable?

- The 1st-order Taylor polynomial is straightforward: If f is a scalar function of m variables, i.e. $f: \mathbb{R}^m \to \mathbb{R}$, then

$$T_1(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

dot product

- The 2nd-order Taylor polynomial is a little more complicated. What is the second derivative of f?

$$D(Df) = D(\nabla f)$$
$$= D\left(\left[\frac{\partial f}{\partial x} \ \frac{\partial f}{\partial y}\right]\right)$$
$$= \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

- This matrix is called the Hessian matrix of f
- So how does the term $\frac{1}{2} f''(x_0)(x-x_0)^2$ generalize?

$$T_2(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$
$$+ \frac{1}{2} \overbrace{\left(\underbrace{Hf(\vec{x}_0)}_{\text{matrix}} \underbrace{(\vec{x} - \vec{x}_0)}_{\text{vector}}\right)}^{\text{vector}} \overset{\text{dot product}}{\cdot} (\vec{x} - \vec{x}_0)$$

where $Hf(x_0)$ is the Hessian matrix of f at x_0.

Ex: Find $T_2(x,y)$ for $f(x,y) = ye^{-x^2} + 2$ at $(x_0, y_0) = (1, 0)$.

- First: $f(1,0) = 2$

- Then: $\nabla f(x,y) = \begin{bmatrix} -2xye^{-x^2} & e^{-x^2} \end{bmatrix}$

$$\Rightarrow \nabla f(1,0) = \begin{bmatrix} 0 & e^{-1} \end{bmatrix}$$

- Furthermore:

$$Hf(x,y) = \begin{bmatrix} -2ye^{-x^2} + 4x^2ye^{-x^2} & -2xe^{-x^2} \\ -2xe^{-x^2} & 0 \end{bmatrix}$$

$$\Rightarrow Hf(1,0) = \begin{bmatrix} 0 & -2e^{-1} \\ -2e^{-1} & 0 \end{bmatrix}$$

- Lastly, $\vec{x} - \vec{x}_0 = (x,y) - (1,0)$

$$= \begin{bmatrix} x-1 & y \end{bmatrix}$$

- Thus:

$$T_2(x,y) = f(1,0) + \nabla f(1,0) \cdot \begin{bmatrix} x-1 & y \end{bmatrix}$$

$$+ \frac{1}{2}\left(Hf(1,0)\begin{bmatrix} x-1 & y \end{bmatrix}\right) \cdot \begin{bmatrix} x-1 & y \end{bmatrix}$$

$$= 2 + \begin{bmatrix} 0 & e^{-1} \end{bmatrix} \cdot \begin{bmatrix} x-1 & y \end{bmatrix}$$

$$+ \frac{1}{2}\left(\begin{bmatrix} 0 & -2e^{-1} \\ -2e^{-1} & 0 \end{bmatrix}\begin{bmatrix} x-1 \\ y \end{bmatrix}\right) \cdot \begin{bmatrix} x-1 & y \end{bmatrix}$$

$$= 2 + e^{-1}y \cancel{+} - 2e^{-1}(x-1)y$$

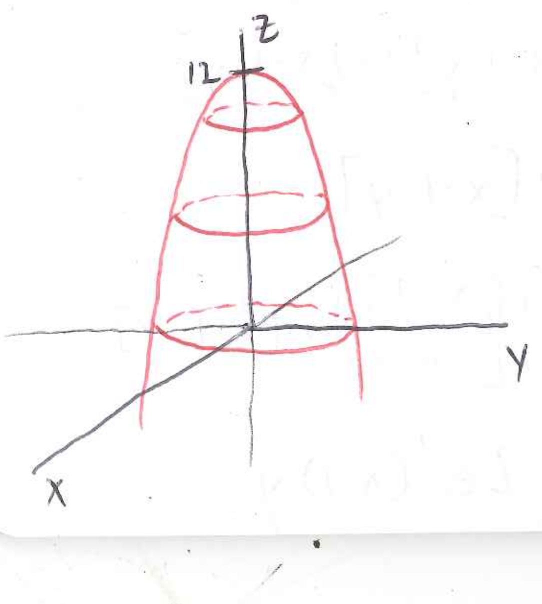
4.3 Extreme Values of Real-Valued Functions

- Recall: if $f: \mathbb{R} \to \mathbb{R}$ has a local max (or min) at x_0, then $f'(x_0) = 0$.
  - Second derivative test: If $f'(x_0) = 0$ and $f''(x_0) > 0$, then f has a local min at x_0. If $f'(x_0) = 0$ and $f''(x_0) < 0$, then f has a local max at x_0.

- How do we generalize to functions of more than 1 variable?

- If $f: \mathbb{R}^m \to \mathbb{R}$ has a local max or local min at $\vec{x}_0$, then $\nabla f(\vec{x}_0) = \vec{0}$

Ex: Consider $f(x,y) = 12 - x^2 - y^2$. It's clear that f has a local max at $\vec{x}_0 = (0,0)$.

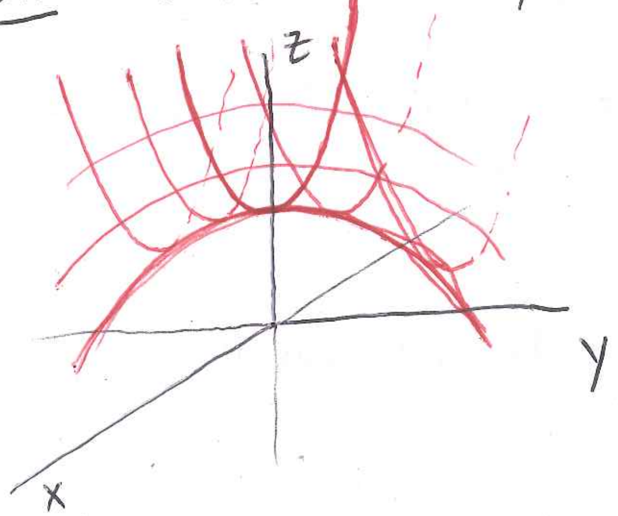


$\nabla f(x,y) = (-2x, -2y)$

$\Rightarrow \nabla f(0,0) = (0,0)$ ✓

- A point $\vec{x}_0$ such that $\boxed{\nabla f(\vec{x}_0) = \vec{0}}$ is called a <u>critical point</u> of f.
  - A critical point that is neither a local min nor local max is called a <u>saddle point</u>.

<u>Ex</u>: Consider $f(x,y) = 2x^2 - y^2 + 3$. Then



$\nabla f(x,y) = (4x, 2y)$

$\Rightarrow \vec{x}_0 = (0,0)$ is a critical point. We can see that $\vec{x}_0$ is neither a max nor a min, so it is a saddle point.

- For functions $f: \mathbb{R}^2 \to \mathbb{R}$, the second derivative test involves the Hessian matrix. Recall:

$$Hf(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

$$\Rightarrow |Hf(x,y)| = f_{xx} f_{yy} - (f_{xy})^2$$

- Suppose $\vec{x}_0$ is a critical point of f. If
  a) $|Hf| > 0$ and $f_{xx} > 0$, then $\vec{x}_0$ is a local min
  b) $|Hf| > 0$ and $f_{xx} < 0$, then $\vec{x}_0$ is a local max
  c) $|Hf| < 0$, then $\vec{x}_0$ is a saddle point
  - Note: If $|Hf| = 0$, then the test is inconclusive.

Ex: Find the local extreme points and saddle points of $f(x,y) = xye^{-x^2-y^2}$.

- First, find the critical points:

$$\nabla f(x,y) = \left(ye^{-x^2-y^2} - 2x^2ye^{-x^2-y^2},\ xe^{-x^2-y^2} - 2xy^2e^{-x^2-y^2}\right)$$

$$= e^{-x^2-y^2}\left(y - 2x^2y,\ x - 2xy^2\right)$$

$$\Rightarrow \nabla f(x,y) = (0,0) \Rightarrow \left(y - 2x^2y,\ x - 2xy^2\right) = (0,0)$$

$$\Rightarrow y - 2x^2y = 0,\quad x - 2xy^2 = 0$$

$$\Rightarrow y(1-2x^2)=0 \rightarrow y=0 \text{ or } x=\pm\frac{1}{\sqrt{2}}$$

If $y=0 \rightarrow x-2x(0)^2=0 \rightarrow x=0$

$$\Rightarrow \boxed{(0,0)} \text{ is a critical point}$$

If $x=\frac{1}{\sqrt{2}}$: $\left(\frac{1}{\sqrt{2}}\right)-2\left(\frac{1}{\sqrt{2}}\right)y^2=0$

$$\Rightarrow y^2=\frac{1}{2} \Rightarrow y=\pm\frac{1}{\sqrt{2}}$$

$$\Rightarrow \boxed{\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)} \text{ are crit. pts}$$

If $x=-\frac{1}{\sqrt{2}}$: $\left(-\frac{1}{\sqrt{2}}\right)-2\left(-\frac{1}{\sqrt{2}}\right)y^2=0$

$$\Rightarrow y^2=\frac{1}{2} \Rightarrow y=\pm\frac{1}{\sqrt{2}}$$

$$\Rightarrow \boxed{\left(\frac{-1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)} \text{ are crit. pts}$$

- Now let's calculate the Hessian matrix:

$$Hf(x,y)=\begin{bmatrix} +2x(2x^2-3)ye^{-x^2-y^2} & (2x^2-1)(2y^2-1)e^{-x^2-y^2} \\ (2x^2-1)(2y^2-1)e^{-x^2-y^2} & 2x(2y^2-3)ye^{-x^2-y^2} \end{bmatrix}$$

at $(0,0)$: $Hf(0,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow |Hf| = -1 < 0$

$\rightarrow$ saddle point

at $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$: $Hf\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \begin{bmatrix} -2e^{-1} & 0 \\ 0 & -2e^{-1} \end{bmatrix} \rightarrow |Hf| = 4e^{-2} > 0$

$\rightarrow f_{xx} > 0 \rightarrow$ local min

etc., etc.

$4e^{-2} > 0$

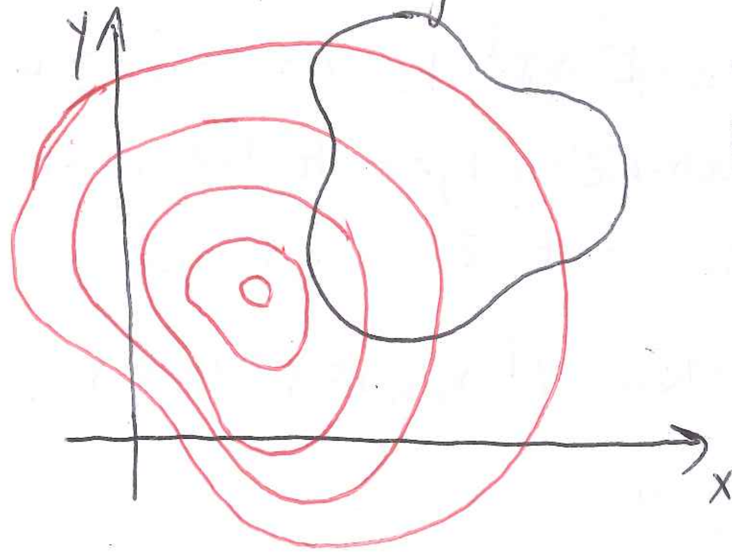
4.4 Optimization w/ Constraints and Lagrange Multipliers

- We want to optimize functions (i.e. find maxes and mins) subject to contraints.

Ex: Let $T(x,y) = x^2 - y + 200$ be the temperature of a metal disk $D = \{(x,y) \mid x^2+y^2 \le 1\}$. Where is the maximum temperature along the edge of the disk?

- What this is asking: maximize T subject to the constraint $x^2+y^2=1$.

- Let's look at the general 2D case: Consider $f: \mathbb{R}^2 \to \mathbb{R}$, subject to some constraint $g(x,y) = k$. We can parametrize the constraint w/ a path $\vec{r}(t)$.



contour curves of f (in red)

$g(x,y) = k$ (in black)