Math 202 Chapter 25 - Lecture notes 1-30

Intermediate Calculus for CSC and EOS (University of Victoria)

Chapter 2

Multivariable Calculus

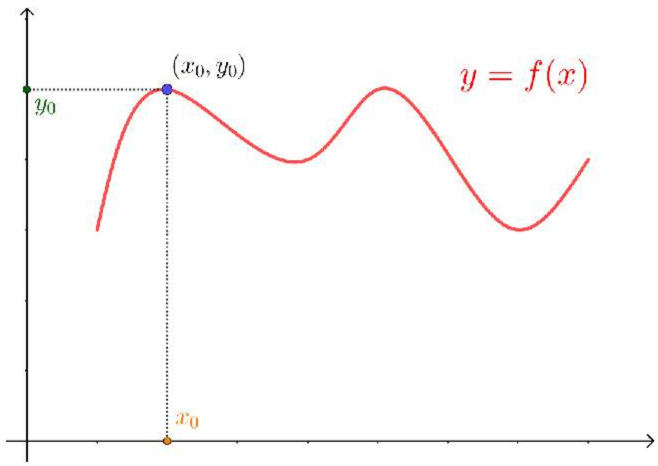
2.1 (Thomas 14.1) Multivariable Functions

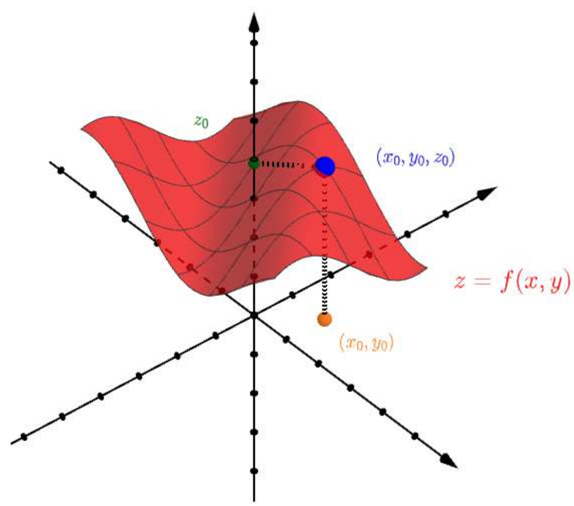
2.1.1 Function Terminology

Definition

Let $U \subset \mathbb{R}^n$. A **scalar function** is a function $f : U \to \mathbb{R}$. In other words, it's a function whose output is a number (scalar).

We have a lot of experience with scalar functions in previous courses. In $\mathbb{R}^2$, equations of the form $y = f(x)$ represent curves. In $\mathbb{R}^3$, equations of the form $z = f(x, y)$ represent surfaces. Such graphs are obtained by pairing a point in the domain with the output. The set of all such pairings forms a curve (in $\mathbb{R}^2$), surface (in $\mathbb{R}^3$), or a hypersurface (in $\mathbb{R}^n$ where $n \geq 4$).





Higher dimensional surfaces like that of the graph of the function

$$T(x, y, z) = e^{-x^2 - y^2 - z^2}$$

can't be visualized in the standard way above. New techniques are required to visualize the graphs of these functions.

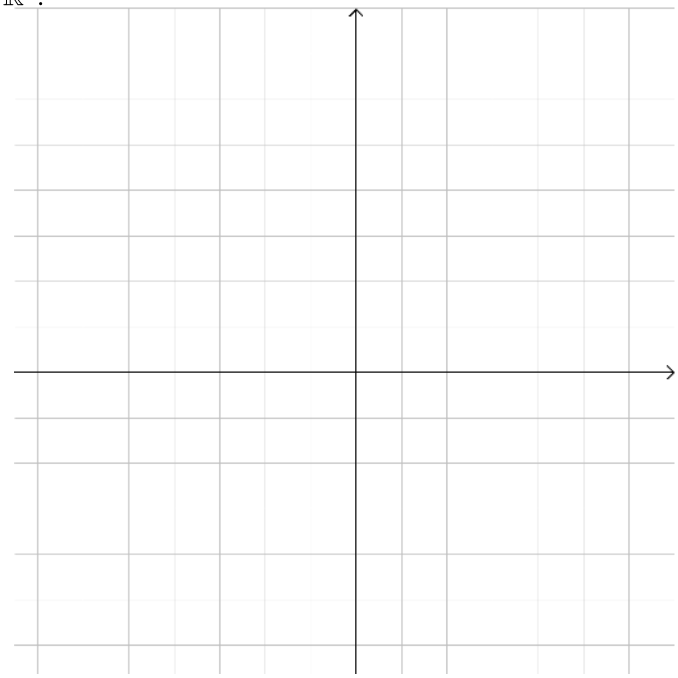
Note

Geogebra's 3D Calculator gives you a way to graph functions of the form $z = f(x, y)$! Simply bring up their download page on the course page and click on START to use the App that you want in your browser.

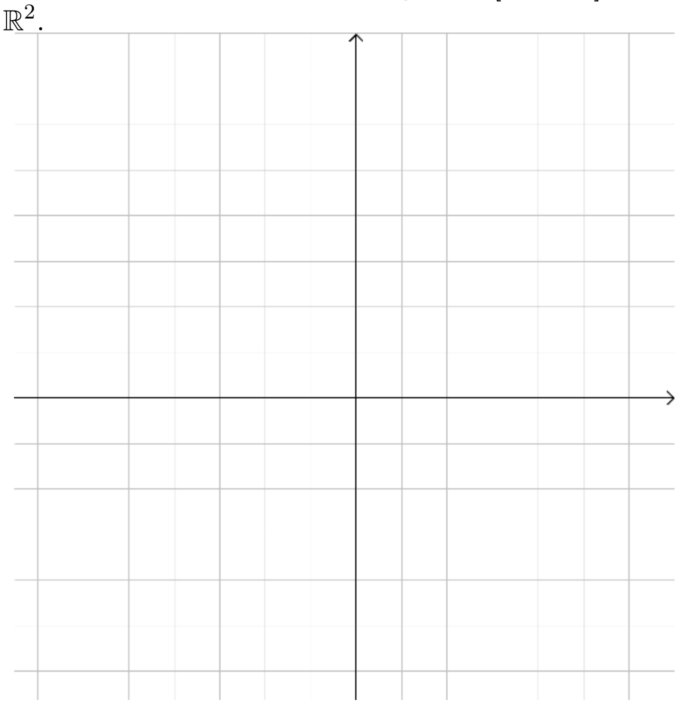
Definition

Let $f : X \to Y$ be a function. The set X is called the **domain** of f and the subset $R = \{y \in Y \mid y = f(x) \text{ for some } x \in X\} \subset Y$ is called the **range**. Simply put, the domain is the collection of all points we allow ourselves to plug into the function and the range is the collection of all obtained points of output. If f is defined on $\mathbb{R}^n$ we say that the **natural domain** of f over $\mathbb{R}$ is the collection of all points we may evaluate the function at to still obtain a real number.

Example: Determine the range and [natural] domain of the function $f(x,y) = \sqrt{y - x^2}$. Graph this domain in $\mathbb{R}^2$.



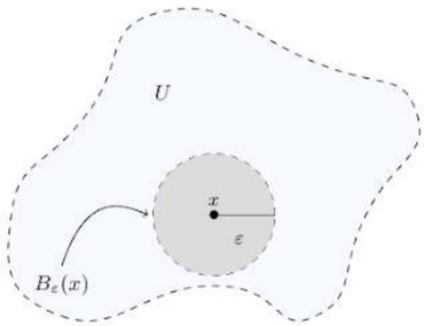
Example: Determine the range and [natural] domain of the function $f(x,y) = \sin(x/y)$. Graph this domain in $\mathbb{R}^2$.



2.1.2 Classifying Regions in $\mathbb{R}^n$

Definition

An ϵ-**neighbourhood** of a point $P \in \mathbb{R}^n$ is given by the set $B_\epsilon(P) = \{Q \in \mathbb{R}^n : d(P, Q) < \epsilon\}$. Simply, it is the collection of points lying within a distance ϵ of P.



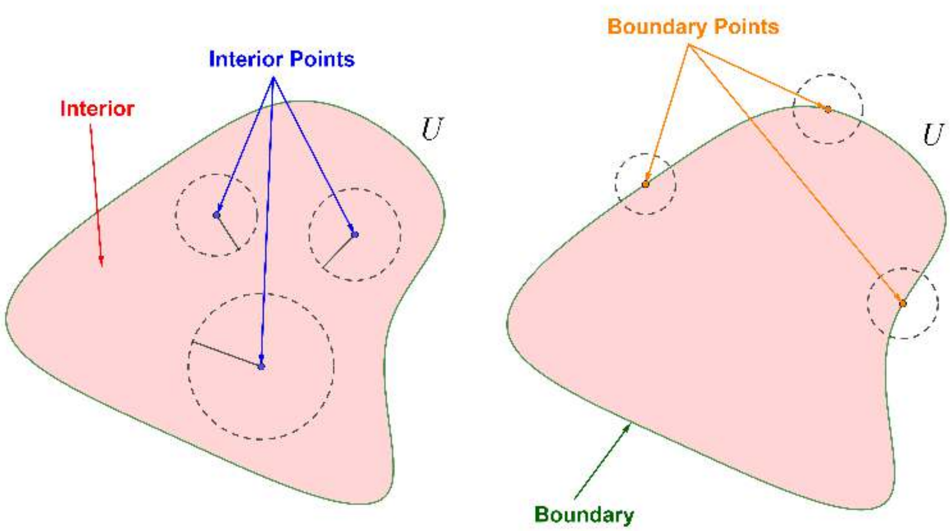
Notation

If a point P lies in a region U we denote this $P \in U$. If U is a subregion of $\mathbb{R}^n$ we denote this $U \subset \mathbb{R}^n$.

Definition

Let $U \subset \mathbb{R}^n$. Then...

- A point $P \in U$ is called an **interior point** of U if *there exists* an $\epsilon > 0$ such that $B_\epsilon(P) \subset U$. The collection of all interior points of U is called the **interior** of U. It is denoted U°.
- A point $Q \in U$ is called a **boundary point** of U if for *every* $\epsilon > 0$ we have that $B_\epsilon(Q)$ contains both interior points of U and points not in U. The collection of all boundary points is called the **boundary** of U. It is denoted ∂U.



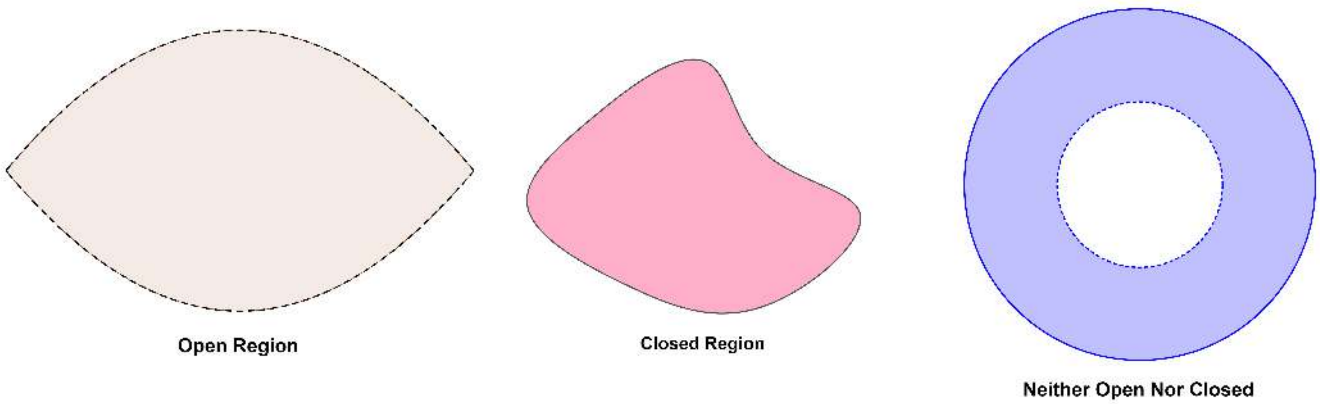
Definition

Let $U \subset \mathbb{R}^n$. Then U is called...

- **open** if every point in U is an interior point.
- **closed** if U contains all of its boundary points.

Theorem

Let $U \subset \mathbb{R}^n$, then U is open if and only if it contains none of its boundary.

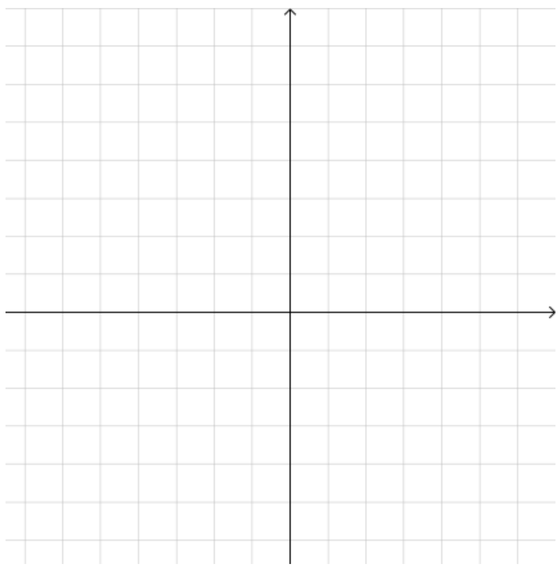


Note

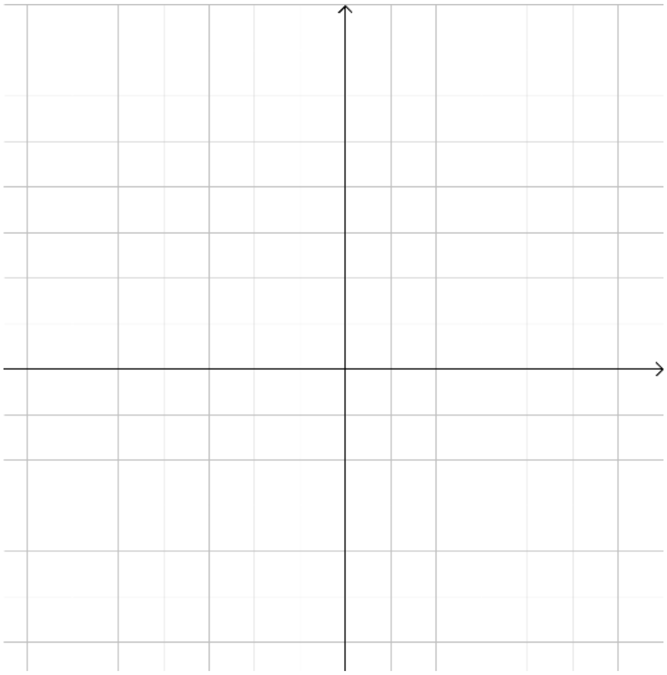
The only sets that are considered to be both open and closed are $\mathbb{R}^n$ (of all points) and the empty collection $\{\}$ (of no points). Every other set is exclusively either open, closed, or neither.

Example: Graph the domain of the following functions and determine if they are open, closed, or neither.

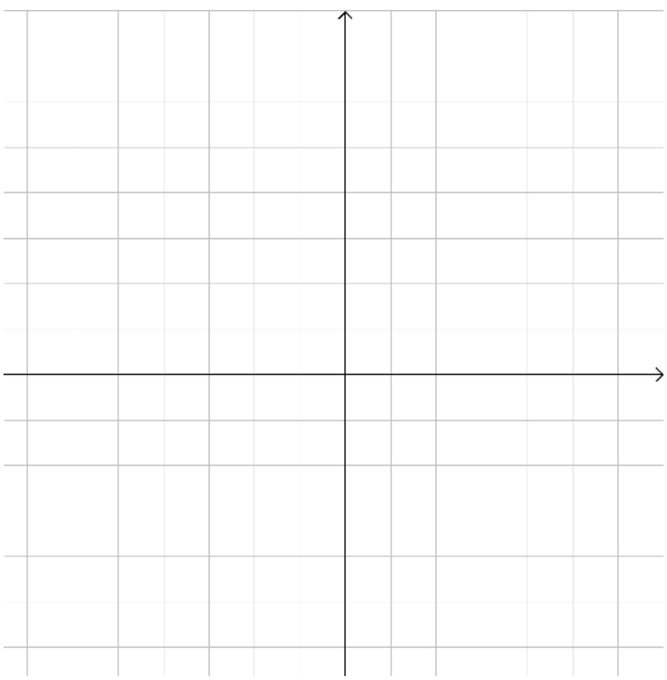
(a) $f(x, y) = \ln(y - x^2)$



(b) $f(x, y) = \ln(4 - x^2 - y^2) + \sqrt{x^2 + y^2 - 1}$

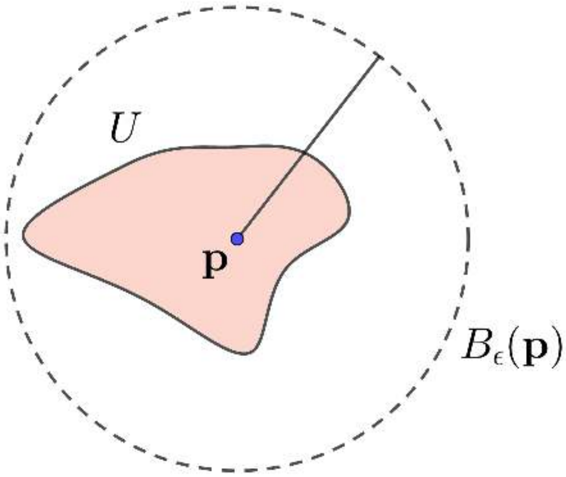


(c) $f(x, y) = \arcsin(x^2 + y^2 - 3)$



Definition

Let $U \subset \mathbb{R}^n$ and $P \in U$. If there exists an $\epsilon > 0$ such that $U \subset B_\epsilon(P)$ then U is called **bounded**. If U is not bounded we say it is **unbounded**.



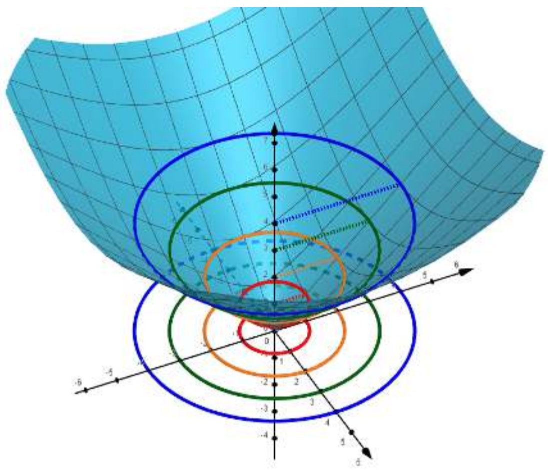
Example: Determine which of the regions in the previous example are bounded and which are unbounded.

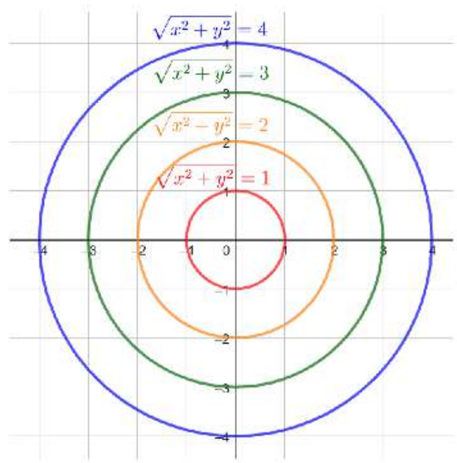
2.1.3 Level Curves and Sets

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a function and let C be a value in the range of f. The collection of all $P \in U$ such that $f(P) = C$ is called the **level set** of f associated to the output C. If $U \subset \mathbb{R}^2$ then it is instead called a **level curve**.

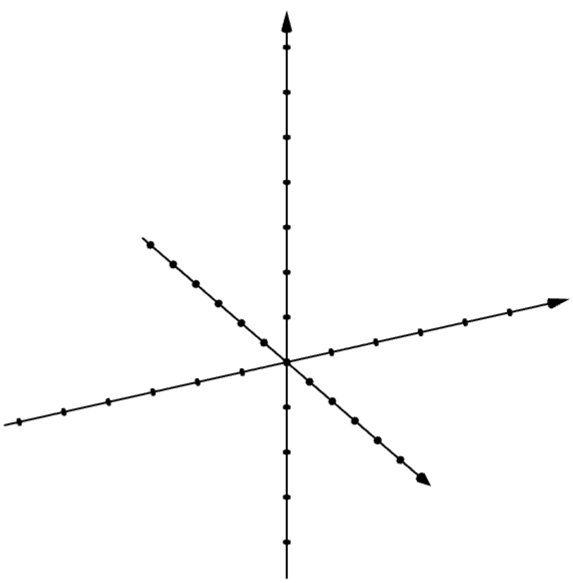
Example: Below are a few level curves of $z = \sqrt{x^2 + y^2}$. We can see visually that they are created by holding the output of the function fixed and finding all points in the domain that correspond to that fixed output.





This gives you a way to graph higher dimensional regions. For example, you can view an onion as a layering of the rings. For two-dimensional beings, this gives them a way to visualize the three-dimensional onion that is typically beyond their grasp.

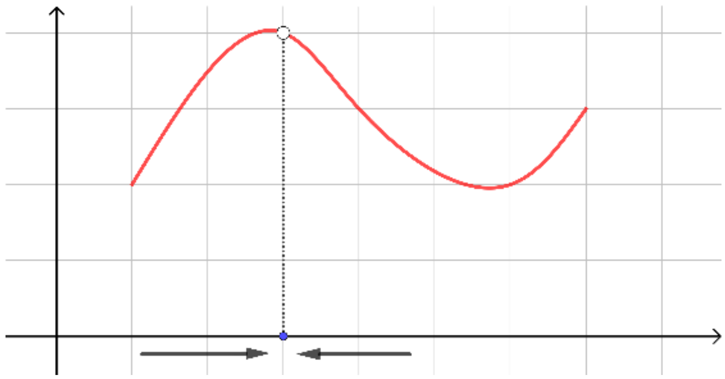
Example: Sketch the level set of $T(x, y, z) = e^{-x^2-y^2-z^2}$ associated to $T(x, y, z) = 1/2$.



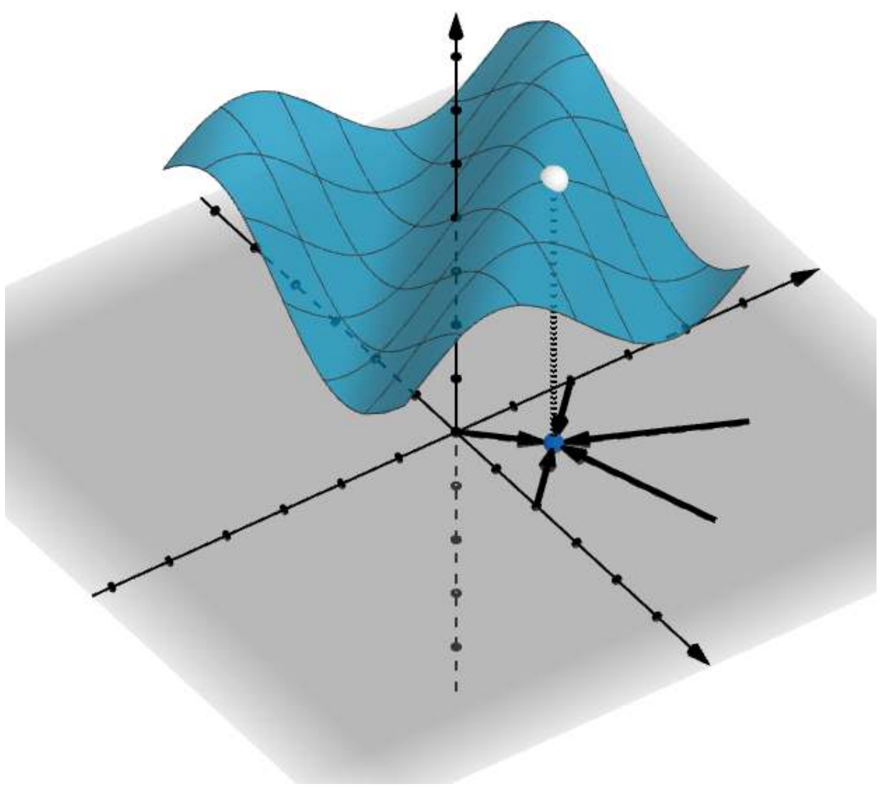
2.2 (Thomas 14.2) Limits

2.2.1 Casual Description of Limits

In Calculus I we only cared about approaching a point from two sides. That is because when we approach a point x_0 by $x \to x_0$ we only have two available options: either to approach it from values less than it or greater than it.



We lose this when we extend to $\mathbb{R}^3$ or any greater dimension. The number of paths to check when we approach a value becomes infinitely uncountable.



Casual Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$. We say that the **limit** f as P approaches P_0 is L, and is denoted $\lim\limits_{P \to P_0} f(P) = L$, provided that every path in U that approaches P_0 results in f approaching the value L.

2.2.2 Proving a Limit

There are three main ways that we will introduce.

Conjugates

Example: Compute $\lim_{(x,y)\to(0,0)} \frac{x^2 - xy}{\sqrt{x} - \sqrt{y}}$.

Substitution

Example: Compute $\lim_{(x,y)\to(0,0)} \frac{1 - \cos(xy)}{xy}$.

Using the Squeeze Lemma

Squeeze Lemma

Let $f, g, h : U \subset \mathbb{R}^n \to \mathbb{R}$ be functions such that

- $g(P) \leq f(P) \leq h(P)$ for all $P \in U$
- $\lim_{P \to P_0} g(P) = \lim_{P \to P_0} h(P) = L$

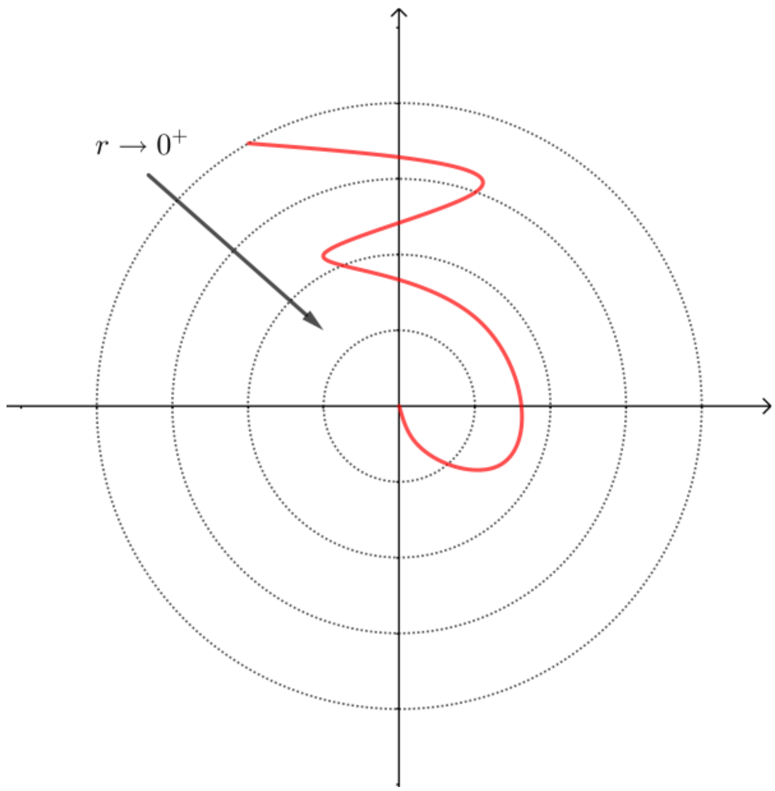
then $\lim_{P \to P_0} f(P) = L$.

Example: Determine (with proof) $\lim_{(x,y) \to (0,0)} \frac{x^2 y}{x^2 + y^2}$.

Example: Determine (with proof) $\lim_{(x,y) \to (0,0)} \frac{3x^2 y^3}{x^2 + y^4}$.

2.2.3 Using Polar Coordinates for Limits at the Origin

If we are taking a limit in $\mathbb{R}^3$ towards the origin, then we may convert the Cartesian variables (x, y) to polar coordinates by $x = r\cos(\theta)$ and $y = r\sin(\theta)$. The effect this has on the resulting variables is that $(x, y) \to (0, 0)$ gives $(r, \theta) \to (0, \theta)$ where the values of θ are dependent entirely on the path of approach.



In some niche cases this may be used in collaboration with the squeeze lemma to prove the existence of a limit. In other cases, its dependency on θ disproves a limit by demonstrating that the value of the limit depends on the path of approach.

Example: Determine $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2}$.

2.2.4 Disproving a Limit

To show a limit doesn't exist is significantly easier than proving one does exist. Merely choose two paths in the domain that both approach the same point but result in a different limit value. This is called the **Two-Path Test**.

Example: Show that the limit $\lim_{(x,y)\to(0,0)} \dfrac{2x^2y}{x^4+y^2}$ does not exist by consider the paths $x(t) = t$ and $y(t) = kt^2$ for various values of k.

Example: Show that the limit $\lim_{(x,y)\to(1,-1)} \dfrac{xy+1}{x^2-y^2}$ does not exist by considering the horizontal and vertical paths in the domain.

2.2.5 Continuity

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a function. The function f is said to be continuous at $P_0 \in U$ provided that

1. $f(P_0)$ exists;
2. $\lim\limits_{P \to P_0} f(P)$ exists; and
3. $\lim\limits_{P \to P_0} f(P) = f(P_0)$

Theorem

Provided that the operations are defined, without singularities, and are compatible: the sum, difference, product, quotient, exponentiation and composition of continuous functions results in a continuous function.

Example: Determine the region on which the function $f(x, y) = \ln(x^2 + y)e^{x^2/y} + \sin(x + y)$ is continuous.

Example: Suppose that $G(x, y) = \dfrac{y - x^2 \cos(y - x^2)}{y - x^2}$ is a function defined on all points where $y \neq x^2$. Suppose you are told that the limit of $G(x, y)$ exists along all points where $y = x^2$. Determine this limit and *extend* $G(x, y)$ so that it becomes a continuous function defined on all of $\mathbb{R}^2$.

2.3 (Thomas 14.3) Partial Derivatives

2.3.1 Interpretation and Definition of the Partial Derivative

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$. The **partial derivative** of f with respect to x_i at a point $P_0(a_1, ..., a_n)$ is defined to be the limit

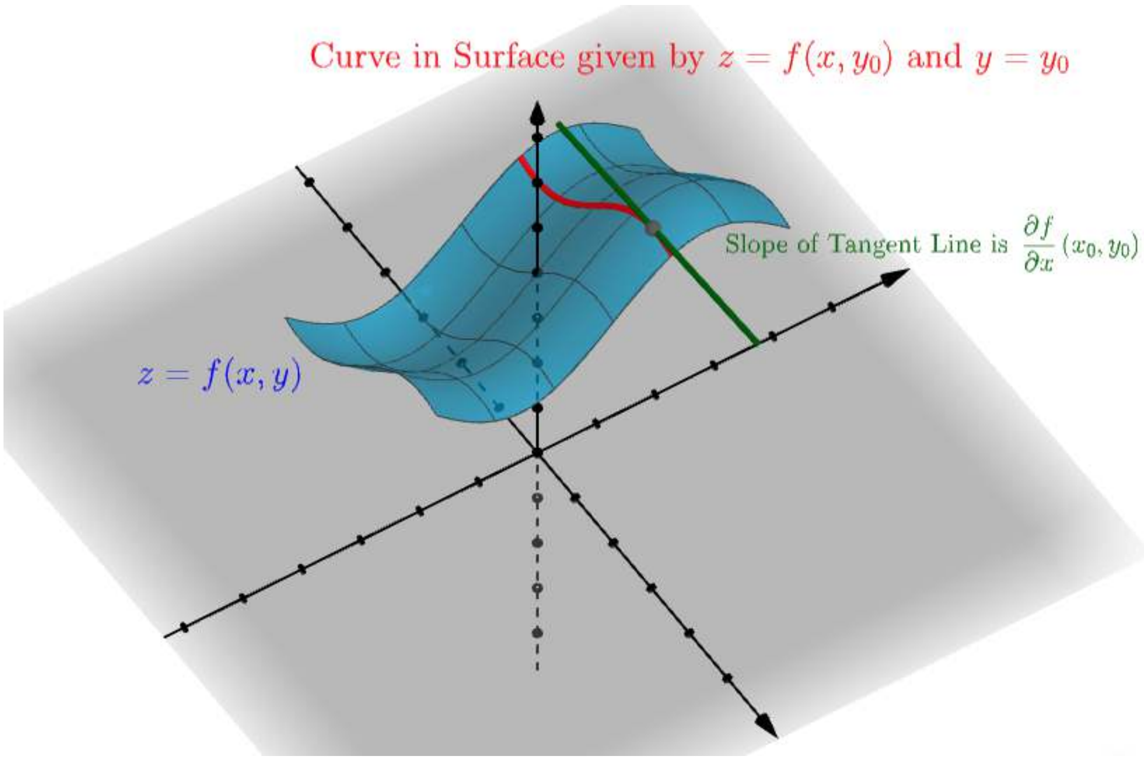
$$\frac{\partial f}{\partial x_i}(P_0) = \lim_{h \to 0} \frac{f(a_1, ..., a_i + h, ..., a_n) - f(a_1, ..., a_n)}{h}$$

Example: Let $f(x, y) = xy + y^2$. Compute $\partial f / \partial y$.

This is geometrically interpreted in the following manner. Consider the surface $z = f(x_1, ..., x_n)$ and the partial derivative $\partial f / \partial x_i$ at $P_0(a_1, ..., a_n)$. If we set all other variables other than x_i to be the constants given by

$$x_1 = a_1 \quad x_2 = a_2 \quad \cdots \quad x_{i-1} = a_{i-1} \quad x_{i+1} = a_{i+1} \quad \cdots \quad x_n = a_n$$

we obtain a curve in the surface known as a **cross section**. The partial derivative $\partial f / \partial x_i$ at P_0 represents the slope of the tangent line to that curve at the point P_0.



Notation

Consider the partial derivative of f with respect to x_i. All of the notation below is commonly used and represent the same thing.

$$\frac{\partial f}{\partial x_i} = \partial_{x_i} f = \partial_i f = f_{x_i} = f_i$$

Never represent $\dfrac{\partial f}{\partial x_i}$ as $\dfrac{df}{dx_i}$ if f is a function of more than one variable! The notation df/dx is specifically used when it is a function of one variable. Using incorrect notation will result losing marks on tests and assignments!

2.3.2 Computing Partial Derivatives by Derivative Rules

Theorem

You may obtain the partial derivative $\partial f/\partial x_i$ by assuming all other variables $x_1, ..., x_{i-1}, x_{i+1}, ..., x_n$ are constant and differentiating in the traditional sense with respect to one variable using derivative rules.

Example: Let $f(x, y) = \sin(x^2 y) + e^{x^2/y}$. Compute f_x and f_y.

Example: Let $f(x, y) = x^2 y + 3xy^2 + e^y$. Compute f_x and f_y.

2.3.3 Higher Order Derivatives

Definition

A **higher order derivative** of $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is defined to be a derivative of a derivative. That is,

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left[\frac{\partial f}{\partial x_i}\right]$$

If $x_i \neq x_j$ we say the above is a **mixed partial derivative**. Higher order derivative are defined recursively in the same manner.

Notation

All of the following are different notation that represent the same thing:

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \partial^2_{x_j x_i} f = \partial^2_{ji} f = f_{x_i x_j} = f_{ij}$$

Notice the order of symbols changes in the subscript notation. This notation generalizes in the same manner to higher order derivatives.

Example: Compute all the second order partial derivatives of the function $f(x, y) = \arctan(y/x)$.

(Continued...)

Mixed Partial Theorem

If $f : U \subset \mathbb{R}^n \to \mathbb{R}$ and its partial derivatives f_i, f_j, f_{ij} and f_{ji} are defined in a neighbourhood of P_0 and all are continuous at P_0 then $f_{ij}(P_0) = f_{ji}(P_0)$.

Example: Show that $f(x, y) = \ln(x^2 + y^2)$ show that this function satisfies the Laplace equation $f_{xx} + f_{yy} = 0$.

$$f_x = \frac{1}{x^2+y^2} \rightsquigarrow (2x+0) = \frac{2x}{x^2+y^2}$$

$$f_{xx} = \frac{(x^2+y^2)(2) - (2x)(2x+0)}{(x^2+y^2)^2}$$

$$= \frac{2x^2+2y^2-4x^2}{(x^2+y^2)^2} = \frac{2y^2-2x^2}{(x^2+y^2)^2}$$

By symmetry

$$f_{yy} = \frac{2x^2-2y^2}{(x^2+y^2)^2}$$

and so $$f_{xx}+f_{yy} = \frac{2y^2-2x^2}{(x^2+y^2)^2} - \frac{2x^2-2y^2}{(x^2+y^2)^2}$$

$$= \frac{0}{(x^2+y^2)^2} = 0 \,//$$

2.4 (Thomas 14.4) The Multivariable Chain Rule

2.4.1 Branch Diagrams and the Chain Rule

Multivariable Chain Rule

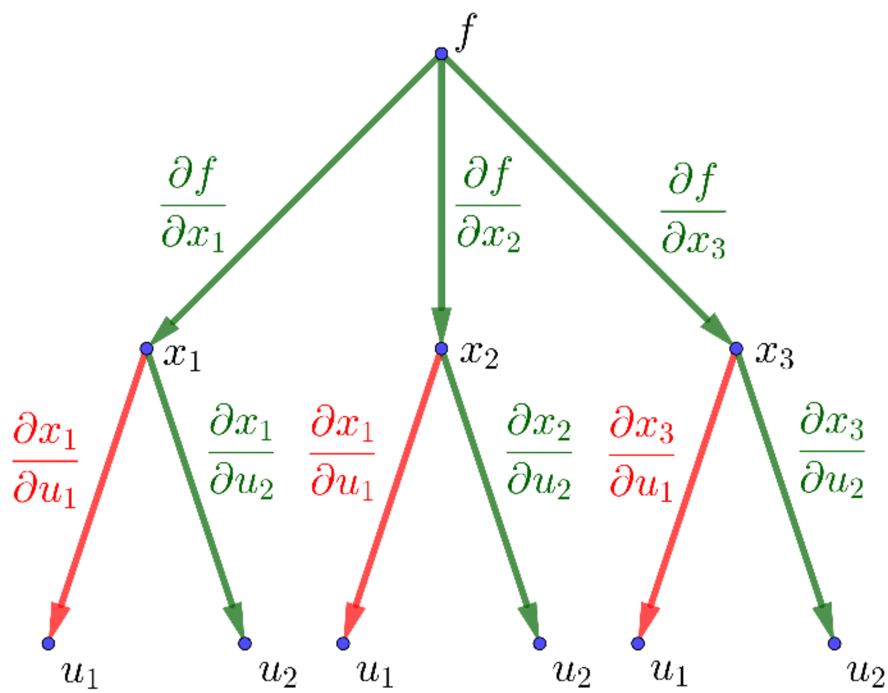
Let $f(X)$ where $X(x_1, ..., x_n) \in \mathbb{R}^n$ be a differentiable function and let $X(U) = (x_1(U), ..., x_n(U))$ where $U(u_1, ..., u_m) \in \mathbb{R}^m$ be differentiable, then the derivative of the function $g(U) = f(X(U))$ with respect to u_i is

$$\frac{\partial g}{\partial u_i} = \frac{\partial f}{\partial x_1}(P)\frac{\partial x_1}{\partial u_i}(U) + \frac{\partial f}{\partial x_2}(P)\frac{\partial x_2}{\partial u_i}(U) + \cdots + \frac{\partial f}{\partial x_n}(P)\frac{\partial x_n}{\partial u_i}(U)$$

Note

There is a nice visual aid that can be used to derive the multivariable chain rule formula given any function. This is given by a **branch diagram** below.

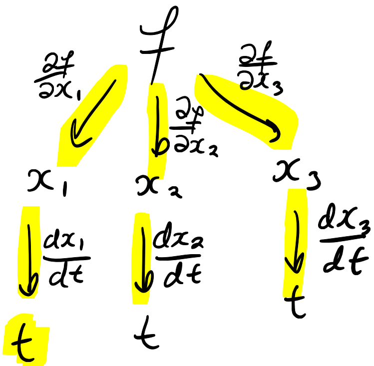
$$\frac{\partial g}{\partial u_2} = \frac{\partial f}{\partial x_1}(\mathbf{x})\frac{\partial x_1}{\partial u_2}(\mathbf{u}) + \frac{\partial f}{\partial x_2}(\mathbf{x})\frac{\partial x_2}{\partial u_2}(\mathbf{u}) + \frac{\partial f}{\partial x_3}(\mathbf{x})\frac{\partial x_3}{\partial u_2}(\mathbf{u})$$



Conditions:

$f(\mathbf{x}) = f(x_1, x_2, x_3)$ **and** $\mathbf{x}(\mathbf{u}) = \mathbf{x}(u_1, u_2)$

Example: Let $g(t) = f(x_1(t), x_2(t), x_3(t))$ where $f(x_1, x_2, x_3) = x_1^2 - x_2^2 + e^{2x_3}$ and $x_1(t) = e^{2t}$, $x_2(t) = e^{-t}$, $x_3 = 2t$. Compute $\frac{dg}{dt}$.



$$\frac{dg}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial f}{\partial x_3}\frac{dx_3}{dt}$$

$$= (2x_1+0+0)(2e^{2t}) + (0-2x_2+0)(-e^{-t})$$

$$+ (0-0+2e^{2x_3})(2)$$

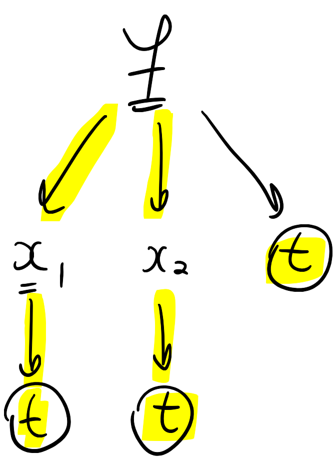
$$= 4x_1e^{2t} + 2x_2e^{-t} + 4e^{2x_3}$$

use $x_1 = e^{2t}$, $x_2 = e^{-t}$, $x_3 = 2t$

$$= 4(e^{2t})e^{2t} + 2(e^{-t})e^{-t} + 4e^{2(2t)}$$

$$= 4e^{4t} + 2e^{-2t} + 4e^{4t}$$

Example: Let $g(t) = f(x_1(t), x_2(t),$ [redacted]$, t)$ where $f(t, x_1, x_2,$ [redacted]$) = \frac{x_1+x_2}{x_2+t}$ and $x_1(t) = t^2$, $x_2(t) = t^3$. Compute $\frac{dg}{dt}$ at $t = 3$.



$$\frac{dg}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial f}{\partial t}$$

$$= \left(\frac{1}{x_2+t}\right)(2t) + \frac{t-x_1}{(x_2+t)^2}(3t^2) + \left(-\frac{x_1+x_2}{(x_2+t)^2}\right)$$

when $t=3 \Rightarrow x_1 = 3^2 = 9$ and $x_2 = 3^3 = 27$

so $g'(3) = \frac{1}{27+3}(6) + \frac{3-9}{(27+3)^2}(27) - \frac{9+27}{(27+3)^2} = A$ Number

$$f_1 = \frac{\partial}{\partial x_1}\left[\frac{x_1+x_2}{x_2+t}\right]$$

$$= \frac{1}{x_2+t}\frac{\partial}{\partial x_1}[x_1+x_2]$$

$$= \frac{1}{x_2+t}(1+0)$$

$$f_2 = \frac{\partial}{\partial x_2}\left[\frac{x_1+x_2}{x_2+t}\right]$$

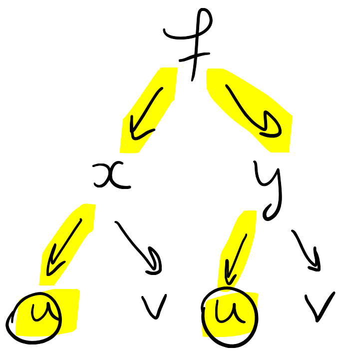
$$= \frac{(x_2+t)\frac{\partial}{\partial x_2}[x_1+x_2] - (x_1+x_2)\frac{\partial}{\partial x_2}[x_2+t]}{(x_2+t)^2}$$

$$\frac{(x_2+t)(0+1) - (x_1+x_2)(1+0)}{(x_2+t)^2}$$

$$f_t = \frac{\partial}{\partial t}\left[\frac{x_1+x_2}{x_2+t}\right] = (x_1+x_2)\frac{\partial}{\partial t}\left[(x_2+t)^{-1}\right] = \frac{\cancel{x_2}+t-x_1-\cancel{x_2}}{(x_2+t)^2} = \frac{t-x_1}{(x_2+t)^2}$$

$$= -(x_1+x_2)\cdot\frac{1}{(x_2+t)^2}$$

Example: Let $g(u,v) = f(x(u,v), y(u,v))$ where $f(x,y) = 3x^2 - 2xy + y^2$, $x(u,v) = 3u + 2v$ and $y(u,v) = 4u - v$. Compute $\dfrac{\partial g}{\partial u}$.



$$\frac{\partial g}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}$$

$$= (6x - 2y + 0)(3+0) + (0 - 2x + 2y)(4-0)$$

$$= 3\big(6(3u+2v) - 2(4u-v)\big) + 4\big(-2(3u+2v) + 2(4u-v)\big)$$

$$= 54u + 36v - 24u + 6v - 24u - 16v + 32u - 8v$$

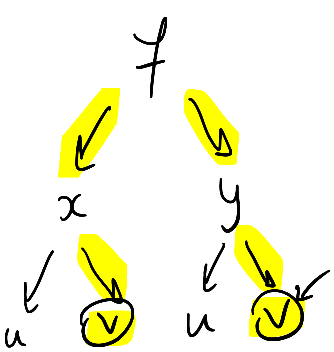
$$= 38u + 18v$$

Example: Let $g(u,v) = f(x,y)$ where $f(x,y)$ is a differentiable function with $x(u,v) = 3u + 2v$ and $y = uv$. Given the following table below:

	$(3,-1)$	$(2,2)$	$(4,-2)$	$(2,-1)$	$(-2,1)$
$f_x(x,y)$	7	2	-3	5	-4
$f_y(x,y)$	-6	3	-2	-2	8

use the multivariable chain rule to compute $\dfrac{\partial g}{\partial v}$ when $(u,v) = \boxed{(2,-1)}$

$$\frac{\partial f}{\partial x} \Leftrightarrow f_x$$



$$\frac{\partial g}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}$$

$$= f_x(x,y)\,x_v(u,v) + f_y(x,y)\,y_v(u,v)$$

$u = 2$ and $v = -1 \Rightarrow x = 3u + 2v = 3(2) + 2(-1)$
$= 6 - 2 = 4$

$y = uv = (2)(-1) = -2$

at $(u,v) = (2,-1)$

$y_v = u$

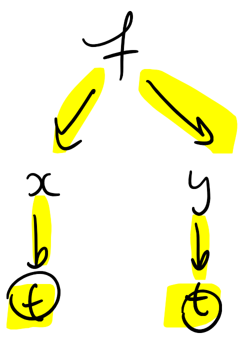
$$\Rightarrow \frac{\partial g}{\partial v}(2,-1) = f_x(4,-2)\cdot(2) + f_y(4,-2)\cdot(2)$$

$$= (-3)(2) + (-2)(2)$$

$$= -6 - 4 = -10 \,//$$

You may be required to compute higher order derivatives by repeated use of the multivariable chain rule.

Example: If $g(t) = f(x,y)$ where $x = t^2$ and $y = 2t+1$ compute $\dfrac{d^2g}{dt^2}$ in terms of the partial derivatives of $f(x,y)$.



$$\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

$$= f_x \cdot (2t) + f_y \cdot (2)$$

$$= 2t f_x + 2 f_y$$

$$\frac{d^2g}{dt^2} = \frac{d}{dt}\left[\frac{dg}{dt}\right] = \frac{d}{dt}\left[2t f_x + 2 f_y\right]$$

$$= 2\frac{d}{dt}\left[t f_x\right] + 2\frac{d}{dt}\left[f_y\right]$$

$$= 2\left\{(1) f_x + t\frac{d}{dt}\left[f_x\right]\right\} + 2\frac{d}{dt}\left[f_y\right]$$

$$= 2\left\{f_x + t\left(\frac{\partial(f_x)}{\partial x}\frac{dx}{dt} + \frac{\partial(f_x)}{\partial y}\frac{dy}{dt}\right)\right\} + 2\left\{\frac{\partial(f_y)}{\partial x}\frac{dx}{dt} + \frac{\partial(f_y)}{\partial y}\frac{dy}{dt}\right\}$$

$$= 2\left\{f_x + t\left(f_{xx}\cdot 2t + f_{xy}\cdot 2\right)\right\} + 2\left\{f_{yx}\cdot 2t + f_{yy}\cdot 2\right\}$$

$$= 2 f_x + 4t^2 f_{xx} + 8t f_{xy} + 4 f_{yy}$$

$\Rightarrow$ Page 60 $\Leftarrow$

2.4.2 Implicit Differentiation

Implicit Differentiation $x_1^2 + x_2 x_3 + x_1 x_4 + y^2 + x_1 y = 0$

Let $P(x_1, ..., x_n) \in \mathbb{R}^n$, $y \in \mathbb{R}$ and $F(P, y) = 0$ where $F : \mathbb{R}^{n+1} \to \mathbb{R}$ is a differentiable function. Then provided $F_y \neq 0$ we have

$$\frac{\partial y}{\partial x_i} = -\frac{F_{x_i}}{F_y} = -\frac{\partial F/\partial x_i}{\partial F/\partial y}$$

Example: Consider the surface given implicitly by $x^2 e^{yz} = y^2 z e^x$. Compute $\dfrac{\partial z}{\partial y}$.

$$\Rightarrow \underbrace{x^2 e^{yz} - y^2 z e^x}_{F(x,y,z)} = 0$$

$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z} = -\frac{x^2 e^{yz}(z) - 2yze^x}{x^2 e^{yz}(y) - y^2(1)e^x}$$

Note

In the particular case of an implicitly given curve $F(x, y) = 0$ then the above formula gives a shockingly fast way to find the slope of the implicit curve compared to previously known methods. That is,

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

Example: Consider the curve given implicitly by $(x^2 + y^2 - 1)^3 = x^2 y^3$. Compute $\dfrac{dy}{dx}$.

$$F(x,y) = (x^2+y^2-1)^3 - x^2y^3 = 0$$

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{3(x^2+y^2-1)^2(2x) - 2xy^3}{3(x^2+y^2-1)^2(2y) - 3x^2y^2}$$

Proof of Implicit Differentiation

The proof is simple and we demonstrate it in the case of a curve given by $F(x, y) = 0$. Assuming that y is a function of x we obtain $y = y(x)$. Then the multivariable chain rule gives, upon differentiating both sides of $F(x, y(x)) = 0$ with respect to x,

$$\frac{d}{dx}[F(x, y(x))] = \frac{d}{dx}[0] \Rightarrow \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0$$

$$\Rightarrow \frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y}$$

Like before, you can compute higher order derivatives implicitly.

Example: Consider the ellipse $x^2 + 4y^2 = 1$. Compute $\dfrac{d^2y}{dx^2}$, the concavity of the curve.

$$F(x,y) = x^2 + 4y^2 - 1 = 0$$

$$\Rightarrow \frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{2x}{8y} = -\frac{1}{4}\frac{x}{y}$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left[\frac{dy}{dx}\right] = \frac{d}{dx}\left[\frac{1}{4}\cdot\frac{x}{y}\right] = -\frac{1}{4}\cdot\frac{y(1) - x\frac{dy}{dx}}{y^2}$$

$$= -\frac{1}{4}\cdot\frac{y - x\left(-\frac{1}{4}\frac{x}{y}\right)}{y^2} \times \frac{4y}{4y}$$

$$= -\frac{1}{4}\cdot\frac{4y^2 + x^2}{4y^3}$$

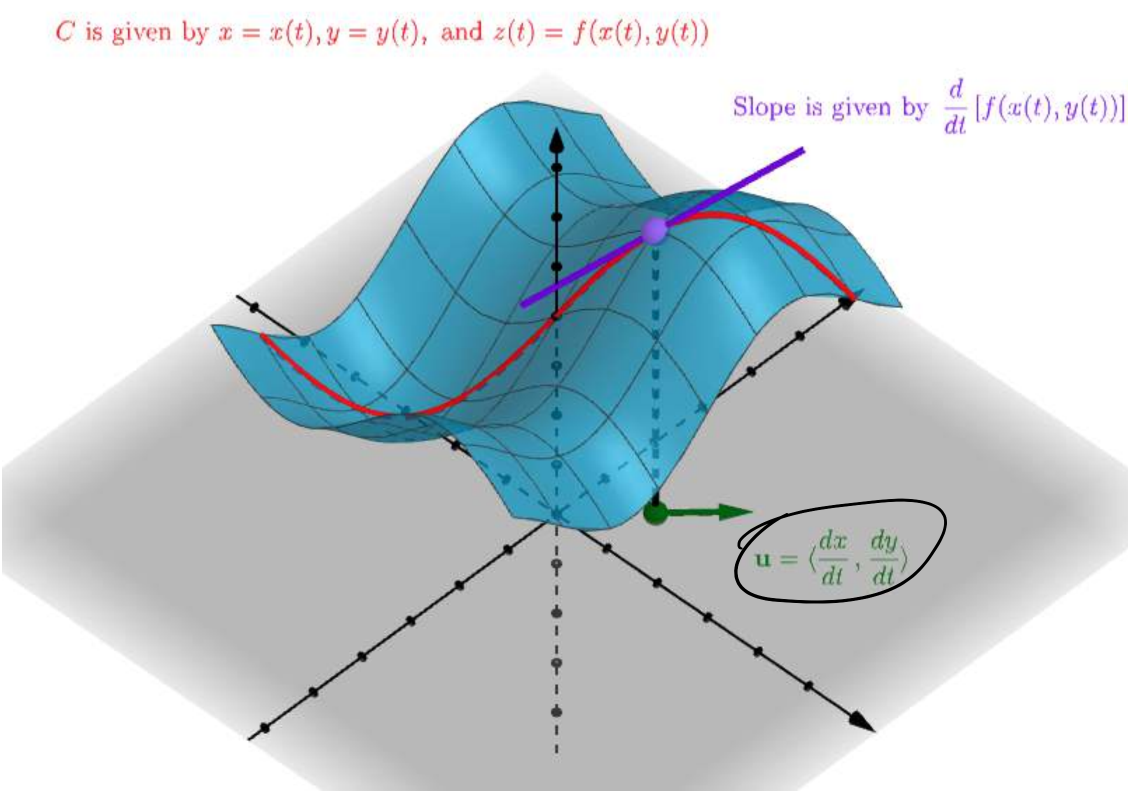
using $x^2 + 4y^2 = 1$ $\longrightarrow$ $= \dfrac{-1}{16y^3}$

2.5 (Thomas 14.5) The Directional Derivative

2.5.1 Formulating the Directional Derivative and Gradient

The partial derivatives measure the rate of change of a function $f(x_1, ..., x_n)$ along one of the components assuming all others are held constant. If we wish to simultaneously move components and determine the rate of change of the function we must construct a new formulation for this.

Keeping things simple as a demonstration, suppose we have a function $f(x, y)$ of two variables and we wish to determine the slope of the curve C at a corresponding point P in the domain moving in the direction $\mathbf{u}$.



Then by the multivariable chain rule the rate of change is given by

$$\frac{d}{dt}[f(x(t), y(t))] = \frac{\partial f}{\partial x}(P)\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}(P)\frac{dy}{dt}(t) = \frac{\partial f}{\partial x}(P)u_1 + \frac{\partial f}{\partial y}(P)u_2$$

The above of course generalizes. Before doing so, we introduce the following notation as a means to clean things up.

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then the **gradient** is defined to be the vector

$$\nabla f(P) = \left\langle \frac{\partial f}{\partial x_1}(P), ..., \frac{\partial f}{\partial x_n}(P) \right\rangle$$

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and let $\mathbf{u}$ be a unit length direction vector in $\mathbb{R}^n$. Then the **directional derivative** of f in the direction $\mathbf{u}$ at the point $P \in U$ is

$$D_{\mathbf{u}} f(P) = \nabla f(P) \cdot \mathbf{u}$$

Notation

The notation $\dfrac{\partial f}{\partial \mathbf{u}}(P)$ is very commonly used in place of $D_{\mathbf{u}} f(P)$.

Example: Find the rate of change of $f(x, y) = e^y x^2 + \arctan(xy)$ at $P(-1, 1)$ in the direction $\mathbf{v} = \langle 1, 2 \rangle$.

$$\nabla f(x,y) = \langle f_x, f_y \rangle$$

$$= \left\langle 2xe^y + \frac{1}{1+(xy)^2}(y),\ e^y x^2 + \frac{1}{1+(xy)^2}(x) \right\rangle$$

$$\nabla f(-1,1) = \left\langle 2(-1)e^1 + \frac{1}{1+(-1)^2}(1),\ e^1(-1)^2 + \frac{1}{1+(-1)^2}(-1) \right\rangle$$

$$= \left\langle -2e + \frac{1}{2},\ e - \frac{1}{2} \right\rangle$$

normalize $\vec{v}$ to unit length $\quad \hat{u} = \dfrac{1}{\|\vec{v}\|}\vec{v} = \dfrac{1}{\sqrt{1^2+2^2}}\langle 1, 2 \rangle$

$$= \frac{1}{\sqrt{5}}\langle 1, 2 \rangle$$

$$D_{\hat{u}} f(P) = \frac{\partial f}{\partial \hat{u}}(P) = \nabla f(P) \cdot \hat{u} = \left\langle -2e + \frac{1}{2},\ e - \frac{1}{2} \right\rangle \cdot \left(\frac{1}{\sqrt{5}} \langle 1, 2 \rangle \right)$$

$$= \frac{1}{\sqrt{5}}\left(\left(-2e + \frac{1}{2}\right) + \left(e - \frac{1}{2}\right)(2) \right)$$

$$= \frac{1}{\sqrt{5}}\left(-\frac{1}{2}\right) = -\frac{1}{2\sqrt{5}}$$

2.5.2 Geometry of the Directional Derivative

Since the directional derivative is defined in terms of the dot product, it inherits the same geometric interpretation and formulations.

In particular, let θ represent the acute angle between the vector $\nabla f(P)$ and unit vector $\mathbf{u}$. Then the directional derivative satisfies

$$D_{\mathbf{u}} f(P) = \nabla f(P) \cdot \mathbf{u} = \|\nabla f(P)\| \|\mathbf{u}\| \cos(\theta) = \|\nabla f(P)\| \cos(\theta)$$

Accordingly, this function is maximized when $\theta = 0$ and minimized when $\theta = \pi$.

Theorem

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function, $\mathbf{u}$ be a unit length vector in $\mathbb{R}^n$ and $P \in U$ be such that $\nabla f(P) \neq \mathbf{0}$. Then the directional derivative $D_{\mathbf{u}} f(P)$ is...

- maximized when the unit direction is given by $\mathbf{u} = \frac{1}{\|\nabla f(P)\|} \nabla f(P)$ and is given by

$$\max_{\mathbf{u}} (D_{\mathbf{u}} f(P)) = \|\nabla f(P)\|$$

- minimized when the unit direction is given by $\mathbf{u} = -\frac{1}{\|\nabla f(P)\|} \nabla f(P)$ and is given by

$$\min_{\mathbf{u}} (D_{\mathbf{u}} f(P)) = -\|\nabla f(P)\|$$

If the acute angle $\theta = \angle(\nabla f(P), \mathbf{u}) = \pi/2$, meaning the two vectors are orthogonal, then $D_{\mathbf{u}} f(P) = 0$ and there is no change.

Example: Determine the maximum rate of change of the function $f(x, y) = y \operatorname{arcsec}(x) + y^2 x$ at the point $P(\sqrt{2}, -2)$ and determine the direction in which this rate occurs.

$$\nabla f(x,y) = \left\langle y \cdot \frac{1}{x\sqrt{x^2-1}} + y^2,\ \operatorname{arcsec}(x) + 2yx \right\rangle$$

$$\nabla f(\sqrt{2},-2) = \left\langle (-2) \cdot \frac{1}{\sqrt{2}\sqrt{2-1}} + (-2)^2,\ \operatorname{arcsec}(\sqrt{2}) + 2(-2)(\sqrt{2}) \right\rangle$$

$$= \left\langle -\frac{2}{\sqrt{2}} + 4,\ \frac{\pi}{4} - 4\sqrt{2} \right\rangle$$

$$\text{Max slope} = \|\nabla f(\sqrt{2},-2)\| = \sqrt{\left(-\frac{2}{\sqrt{2}}+4\right)^2 + \left(\frac{\pi}{4} - 4\sqrt{2}\right)^2}$$

2.5.3 Orthogonality of the Gradient with Level Sets and Implicit Surfaces

Consider the simple case of a function $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ and formulate the level curve $f(x, y) = C$. The curve itself has a representation $x = x(t), y = y(t)$ and so we obtain the equation

$$f(x(t), y(t))) = C$$

Then by differentiating both sides with respect to t we obtain

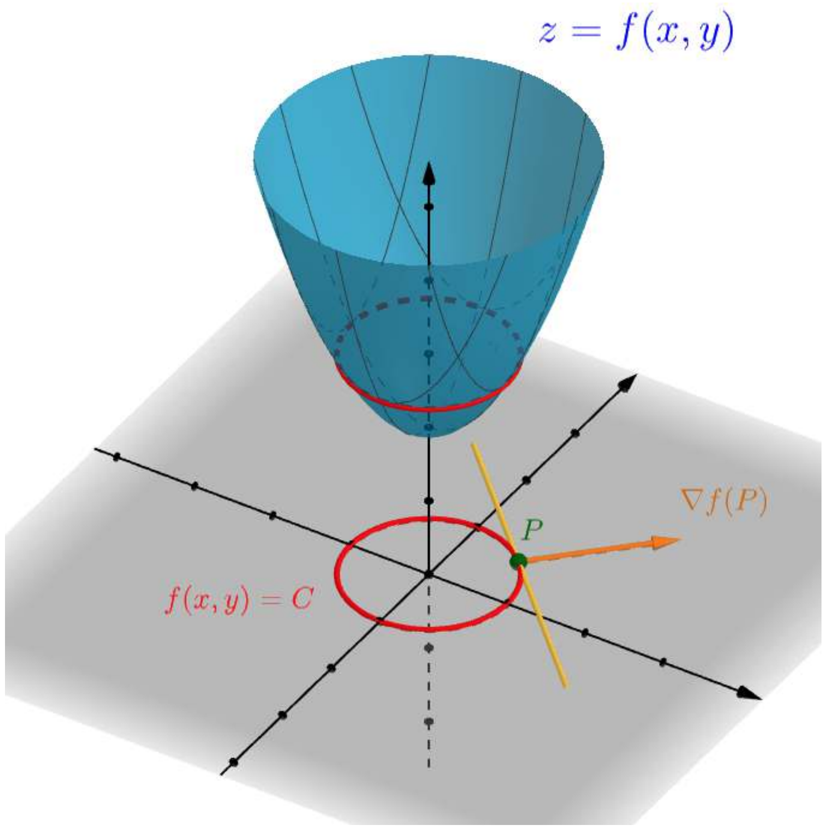
$$\frac{d}{dt}[f(x(t), y(t))] = \frac{d}{dt}[C]$$

$$\Rightarrow \nabla f(x(t), y(t)) \cdot \langle dx/dt, dy/dt \rangle = 0$$

However, as $T = \langle dx/dt, dy/dt \rangle$ is a vector representing a tangent direction to the curve then we conclude that ∇f is orthogonal to the level curve. This result extends naturally into level sets.

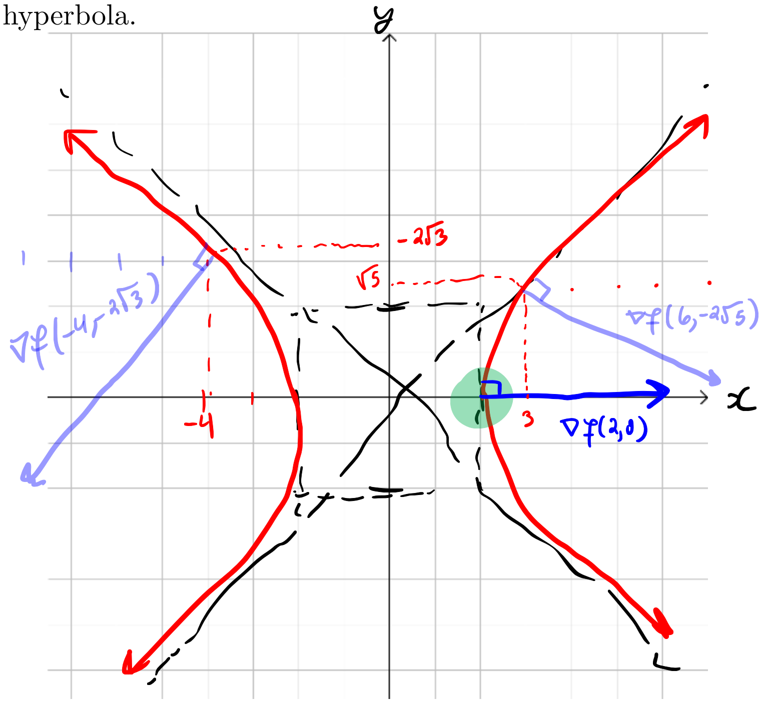
Theorem

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and let P be a point on the level set $f(X) = C$ where C is a value in the range of f. Then the gradient vector $\nabla f(P)$ is orthogonal to the level set $f(X) = C$ at P.



$\frac{x^2}{2^2} - \frac{y^2}{2^2} = 1$

Example: Using the gradient, construct vectors orthogonal to the hyperbola $x^2 - y^2 = 4$. Graph this curve in the space below and demonstrate your result is true at the points $P(2, 0)$, $Q(3, \sqrt{5})$, and $R(-4, -2\sqrt{3})$ on the hyperbola.



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

$\Rightarrow x^2 - y^2 = 4$ is a level set/curve of f.

Form

$\Rightarrow \nabla f = \langle 2x, -2y \rangle$

$\nabla f(2, 0) = \langle 4, 0 \rangle$

$\nabla f(3, \sqrt{5}) = \langle 6, -2\sqrt{5} \rangle$

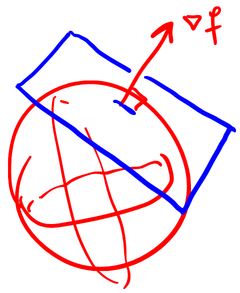
$\nabla f(-4, -2\sqrt{3}) = \langle -8, 4\sqrt{3} \rangle$

Example: Using the gradient, construct the equation of the tangent plane to the sphere $x^2 + y^2 + z^2 = 4$ at the point $P(\sqrt{2}, 1, -1)$.

Consider $f(x,y,z) = x^2 + y^2 + z^2$. Then $x^2 + y^2 + z^2 = 4$ is a level set of f when $f(x,y,z) = 4$.

Thus ∇f is orthogonal to $x^2 + y^2 + z^2 = 4$.

$\Rightarrow \nabla f = \langle 2x, 2y, 2z \rangle$



Thus $\nabla f(\sqrt{2}, 1, -1)$ is orthogonal to the plane

$= \langle 2\sqrt{2}, 2, -2 \rangle$

So $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$

$\Rightarrow \langle 2\sqrt{2}, 2, -2 \rangle \cdot \langle x - \sqrt{2}, y - 1, z + 1 \rangle = 0$

$\Rightarrow 2\sqrt{2}x - 4 + 2y - 2 - 2z - 2 = 0$

$\Rightarrow 2\sqrt{2}x + 2y - 2z = 6$

$\Rightarrow \sqrt{2}x + y - z = 3$

2.6 (Thomas 14.6) Tangent Plane and Differentials

2.6.1 Introducing the Tangent Plane and Normal Line

Tangent Plane and Normal Line

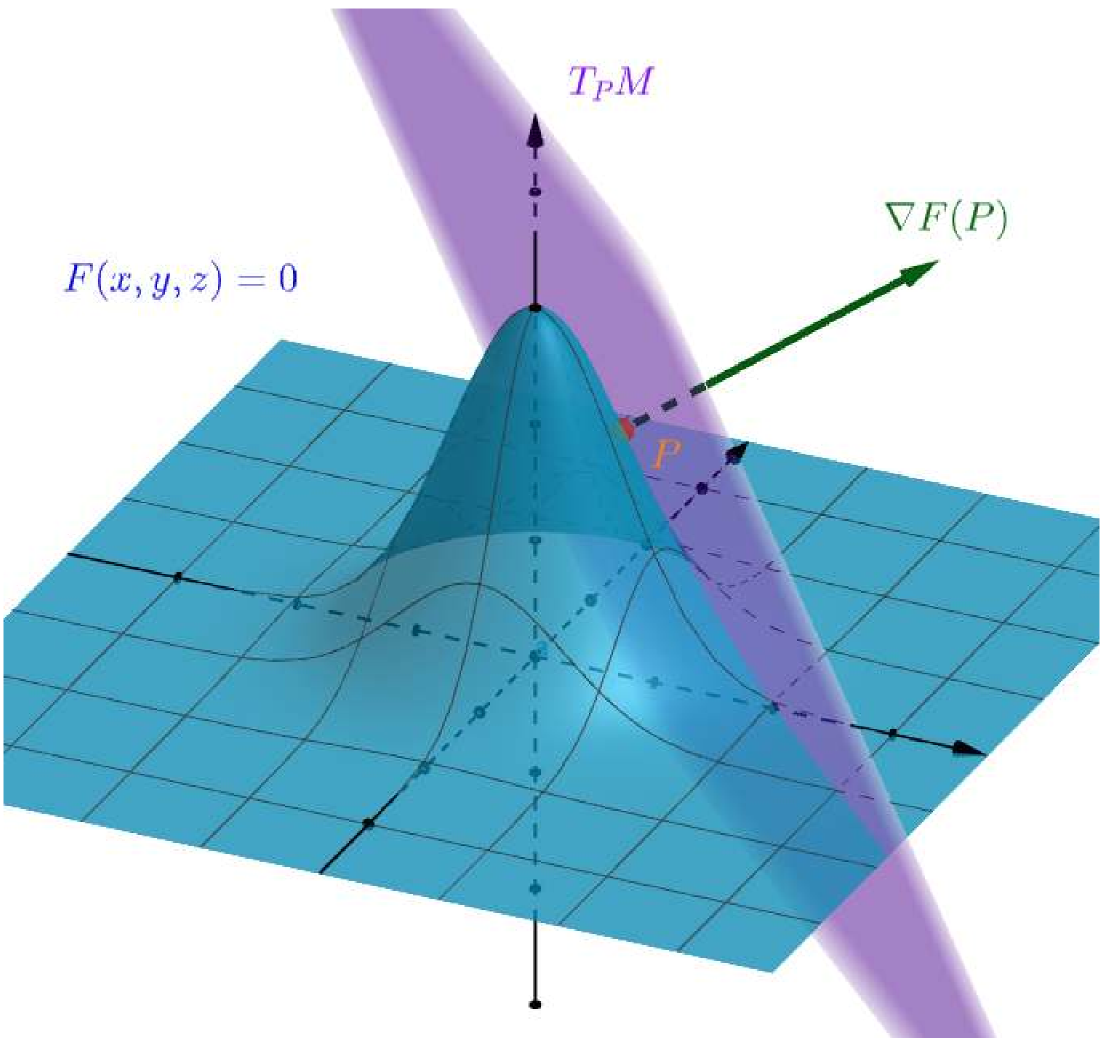
Let $F : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and consider the implicitly given surface $F(X) = 0$. Let P be a point on this surface. Then the **Tangent Plane** to $F(X) = 0$ at P is given by

$$\nabla F(P) \cdot (X - P) = 0$$

If M represents the surface $F(X) = 0$ then the collection of all points on the tangent plane is denoted $T_P M$ and is called the **Tangent Space** of M. The **Normal Line** is given by

$$\mathbf{N}(t) = \mathbf{P} + t\nabla F(P)$$

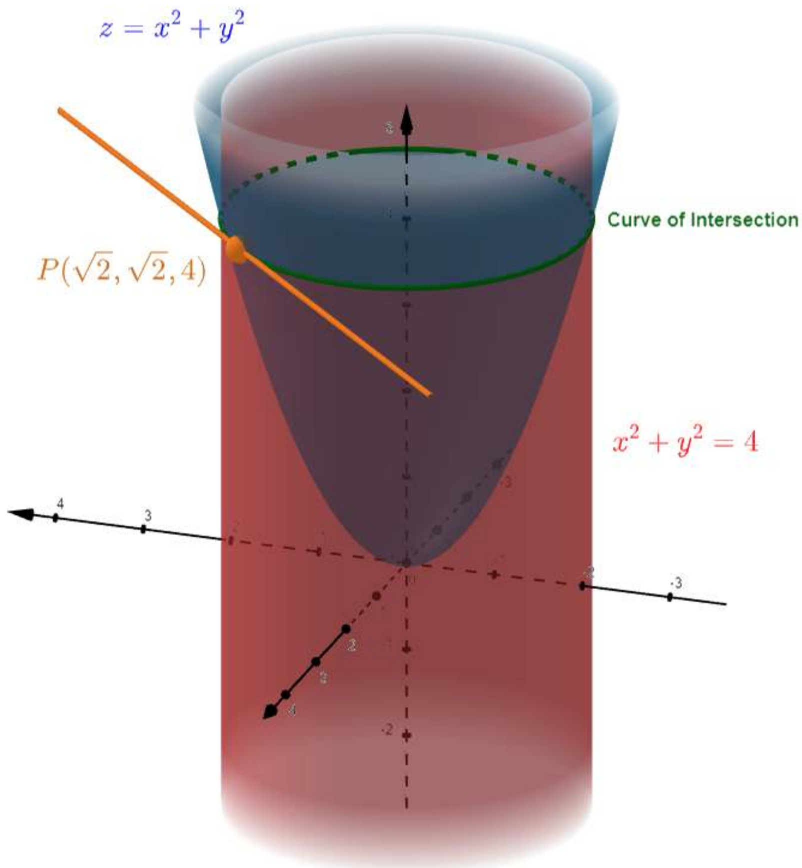
and represents the equation of the line through P and orthogonal to the surface M.



Example: Find the tangent plane and normal line to the surface $x^3 + 3x^2y^2 + y^3 + 4xy - z^2 = 0$ at the point $P(1, 1, 3)$.

Example: Find the equation of the tangent plane to $f(x, y) = \sqrt{y - x}$ at the point when $(x, y) = (1, 2)$.

Example: Find the equation of the tangent line to the surfaces $x^2 + y^2 = 4$ and $z = x^2 + y^2$ at $P(\sqrt{2}, \sqrt{2}, 4)$.



2.6.2 Differentials

Casual Description of Differentials

Let x be a variable, then the **infinitesimal** of it is denoted dx and is interpreted as the smallest change that x can have. If $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is a differentiable function then there are the following relationships between the infinitesimal of f and its variables:

- The infinitesimal change of f in the unit direction $\mathbf{u}$ in U from a point $P \in U$ a distance ds in the surface is given by

$$df = D_{\mathbf{u}} f(P) ds$$

- The infinitesimal change of f given the incremental changes in the components dx_1, ..., dx_n from a point $P \in U$ in the domain is given by

$$df = f_1(P)dx_1 + \cdots + f_n(P)dx_n$$

These expressions are called **differentials**.

Example: Estimate the change in the function $h(x, y, z) = \cos(\pi xy) + xz^2$ if we move a distance $ds = 0.1$ units from $P(-1, -1, -1)$ to the origin.

Example: Estimate the change in $z = f(x, y)$ if $z^2 x + y \ln(z) + y^2 = 0$ from the point $P(2, \sqrt{2}, 1)$ when moving with $dx = 0.1$ and $dy = 0.2$. *Hint: You will need to use implicit differentiation.*

2.6.3 Linearization

Estimation Using the Tangent Plane

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and consider the tangent plane to $z = f(X)$ at P_0 given by

$$L(X) = f(P_0) + \nabla f(P_0) \cdot (X - P_0)$$

This expression is also called the **Linearization** of $z = f(X)$ at P_0. Provided that a point Q is close to P_0 we have

$$L(Q) \approx f(Q)$$

and the approximation becomes worse the further you move away from the point P_0.

Example: Let $f(x, y) = x^2y + 3y^2x$ and use linearization to approximate the value of $f(3.2, 4.1)$. Use a calculator to compare this to a more accurate approximation.

2.7 (Thomas 14.7) Extreme Values and Optimization over $\mathbb{R}^2$ and $\mathbb{R}^3$

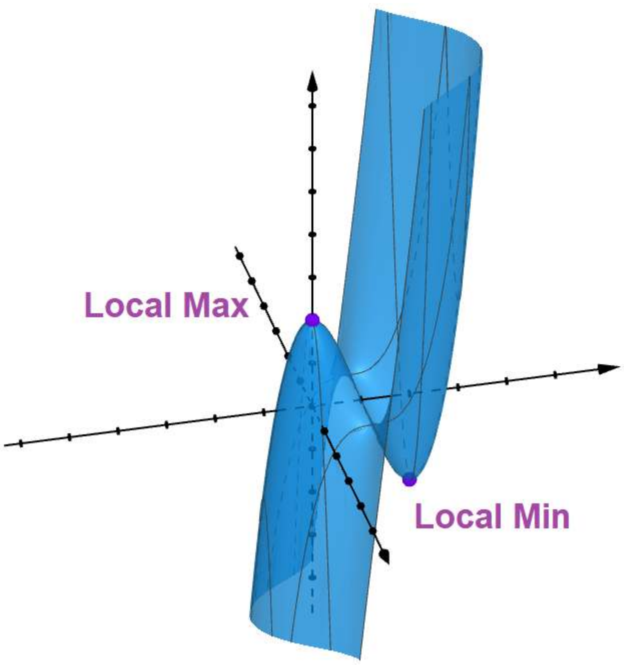
2.7.1 Extrema and the Extreme Value Theorem

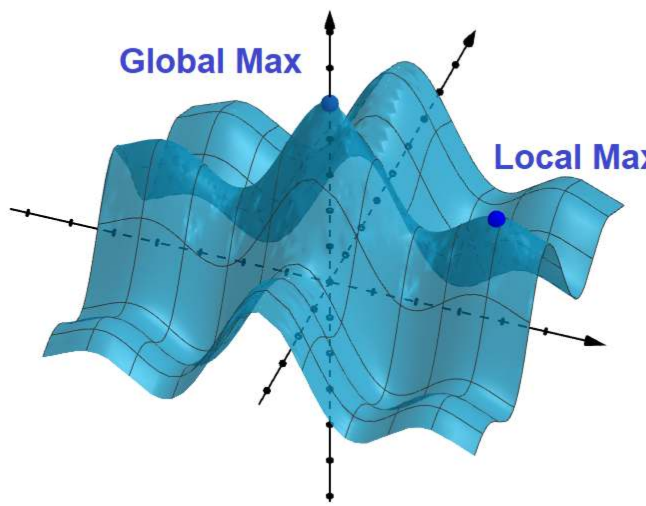
Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$. We say that a point $P_0 \in U$ is a...

- **global** or **absolute maximum** on U if $f(P_0) \geq f(P)$ for all $P \in U$.
- **global** or **absolute minimum** on U if $f(P_0) \leq f(P)$ for all $P \in U$.
- **local** or **relative maximum** on U if there exists an $\epsilon > 0$ such that $f(P_0) \geq f(P)$ for all $P \in B_\epsilon(P_0)$.
- **local** or **relative minimum** on U if there exists an $\epsilon > 0$ such that $f(P_0) \leq f(P)$ for all $P \in B_\epsilon(P_0)$.

  any of the above points are called **extrema** of the function.





Theorem

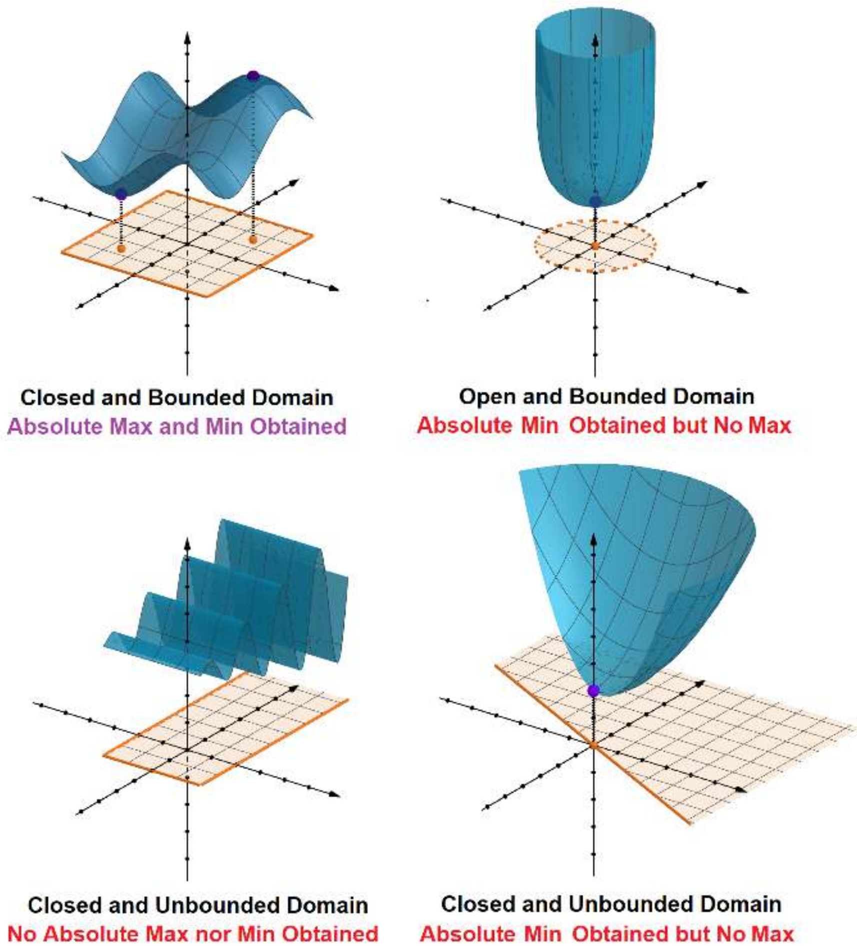
Every global extrema is also a local extrema. The reverse implication isn't necessarily true.

We will be discussing how to optimize functions of the form $w = f(x, y)$ and $w = f(x, y, z)$. Optimization of higher order functions requires knowing how to compute the determinant in arrays of sizes greater than 3×3. While not particularly difficult to introduce, we still keep it reduced to these simpler cases and alert the reader that generalization of what is to come isn't difficult to introduce.

Prior to doing so, it is beneficial for the reader to know when we can guarantee the existence of global extrema.

Extreme Value Theorem

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a continuous function. If U is closed and bounded then f obtains both an absolute maximum and an absolute minimum on U.



Example: Argue that the function $f(x_1, x_2) = \sqrt{x - y^2} + \sqrt{1 - x - y}$ has an absolute maximum and minimum on its natural domain.

2.7.2 The First Order Necessary Conditions and Critical Points

First Order Necessary Conditions

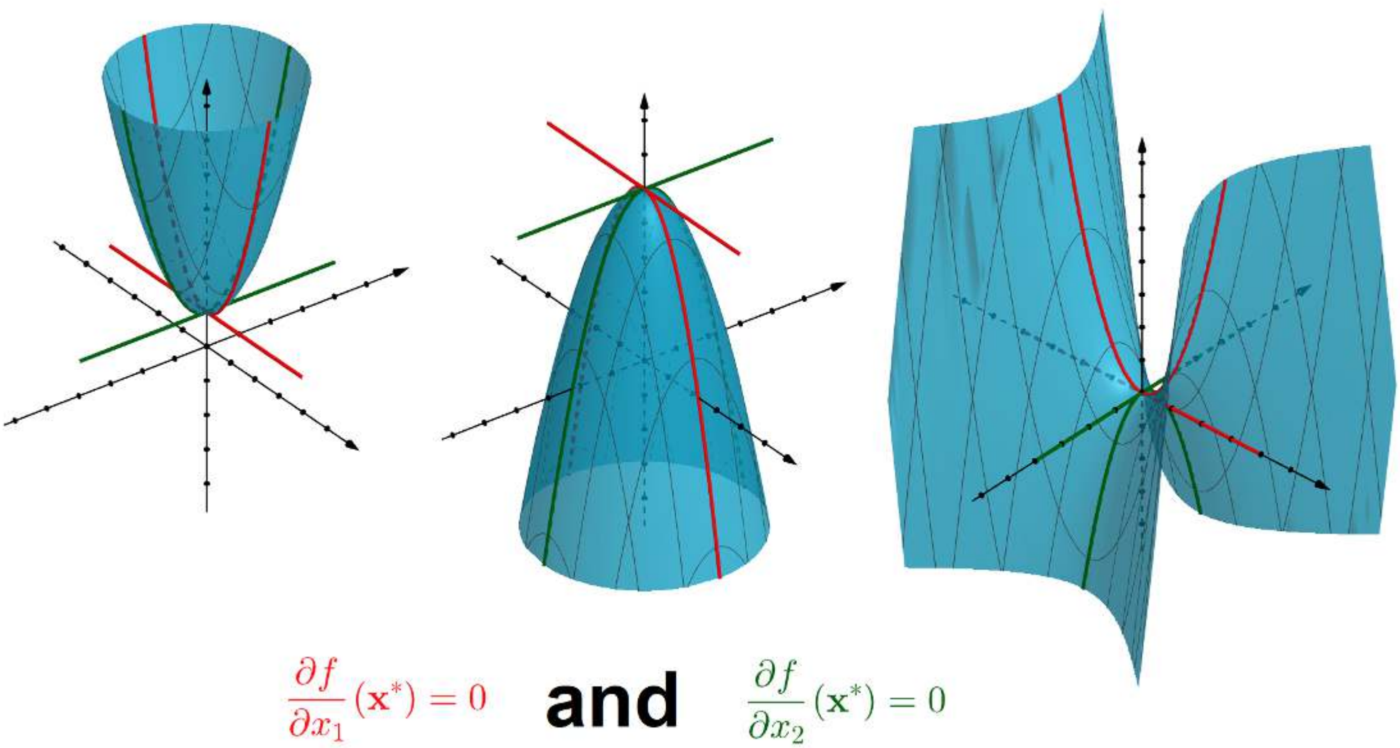
Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. If f has a local extrema at $P_0 \in U$ then

$$\nabla f(P_0) = \mathbf{0}$$

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. A point $P_0 \in U$ where either $\nabla f(P_0) = \mathbf{0}$ or $\nabla f(P_0)$ does not exist is called a **critical point** of f.

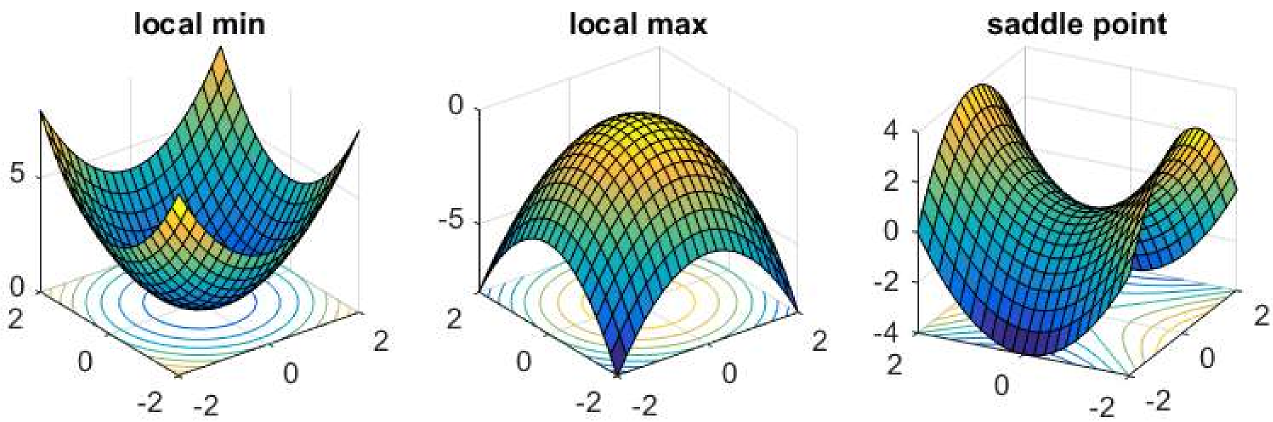
By the first order necessary conditions, critical points give the locations of potential extrema.



As the above picture illustrates, it is possible to have points where $\nabla f(P_0) = \mathbf{0}$ but are neither a local minimum or maximum. Such points correspond to regions where either the function is very flat or places where the curvature differs depending on the curve you travel on surrounding the point.

Definition

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$. We say that a point $P_0 \in U$ is a **saddle point** on U if there exist points Q and R in $B_\epsilon(P_0)$ such that $f(P_0) \leq f(Q)$ and $f(P_0) \geq f(R)$ for every $\epsilon > 0$.



Example: Find all critical points of the function

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

2.7.3 Second Order Sufficient Conditions and Classifying Points

Second Order Sufficient Conditions

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ with U an open and connected set and f twice differentiable. Suppose that $P_0 \in U$ satisfies $\nabla f(P_0) = \mathbf{0}$. Then P_0 is a...

In $\mathbb{R}^2$

- local min provided $f_{xx}(P_0) > 0$ and $\begin{vmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{vmatrix} > 0$.
- local max provided $f_{xx}(P_0) < 0$ and $\begin{vmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{vmatrix} > 0$.

In $\mathbb{R}^3$

- local min provided $f_{xx}(P_0) > 0$, $\begin{vmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{vmatrix} > 0$, and $\begin{vmatrix} f_{xx}(P_0) & f_{xy}(P_0) & f_{xz}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) & f_{yz}(P_0) \\ f_{zx}(P_0) & f_{zy}(P_0) & f_{zz}(P_0) \end{vmatrix} > 0$
- local max provided $f_{xx}(P_0) < 0$, $\begin{vmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{vmatrix} > 0$, and $\begin{vmatrix} f_{xx}(P_0) & f_{xy}(P_0) & f_{xz}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) & f_{yz}(P_0) \\ f_{zx}(P_0) & f_{zy}(P_0) & f_{zz}(P_0) \end{vmatrix} < 0$

If the sign patterns are non-zero and fall into some different pattern then the point is a saddle point. If any of the above determinants or $f_{xx}(P_0)$ are zero then the result is inconclusive and the test fails.

Example: In the previous example the critical points of $f(x, y, z) = 2x^2 + y^2 + z^4 - 4xz - 2y$ are $(0, 1, 0)$, $(-1, 1, -1)$ and $(1, 1, 1)$. Classify all three of these critical points.

(Continued...)

Example: Find and classify all extrema of the function

$$f(x, y) = 3x^2y + y^3 - 3x^2 - 3y^2 + 2$$

(Continued...)

2.7.4 Finding the Extrema of Functions over a Closed Region

If you have a closed region, then you must also check the edges of your domain as well as all critical points in the interior. By the Extreme Value Function, if the region is closed and bounded we are guaranteed that absolute extrema exist. Not only do they exist, but they are also easy to find provided your function is continuous. All potential locations of absolute extrema on closed and bounded sets are at critical points on the interior or boundary.

Example: Find the absolute maximum and minimum of the function

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

over the region bounded by the curves $x = 0$, $y = 2$ and $y = 2x$ in $\mathbb{R}^2$.

(Continued...)