Functions of Several Variables

Local Linear Approximation

Dr.A.L.Pathak

Real Variables

In our studies we have looked in depth at

- functions $f : X \longrightarrow Y$ where X and Y are arbitrary metric spaces,
- real-valued functions $f : X \longrightarrow \mathbb{R}$ where X is an arbitrary metric space,

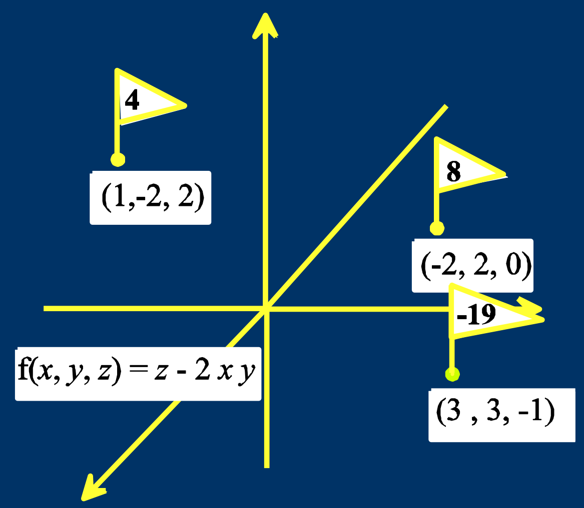
and

- functions $f : \mathbb{R} \longrightarrow \mathbb{R}$.

Now we want to look at functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

Scalar Fields

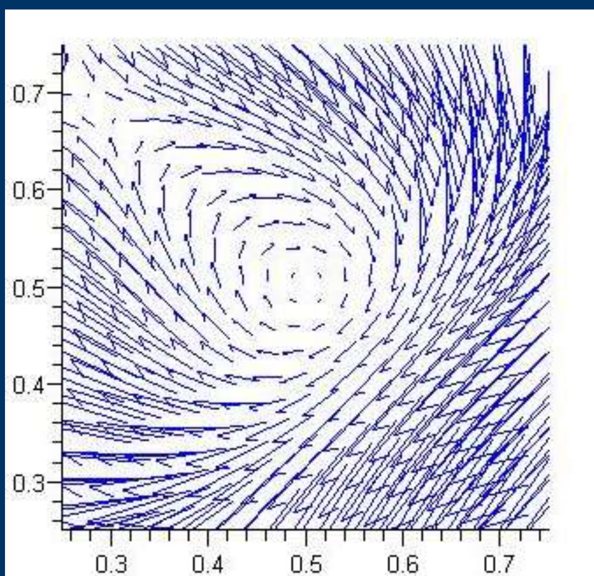
A function **F**: ☼$^n \rightarrow$ ☼ is called a Scalar Field because it assigns to each <u>vector</u> in ☼^{n} a <u>scalar</u> in ☼.



Scalar Fields: Think of the domain as a "field" in which each point is "tagged" with a number. <u>Example</u>: Each point in a room can be associated with a temperature in degrees Celsius.

Vector Fields

A function **F**: ☼^{n} → ☼^{m} (m > 1) is called a Vector Field because it assigns to each vector in ☼^{n} a vector in ☼^{m}.



Vector Fields: Think of the domain as a "field" in which each point is "tagged" with a vector. Example: domain is the surface of a river, we can associate each point with a current, which has both magnitude and direction and is therefore a vector.

Vector and Scalar Fields

- Let **F**: ☼^{n} ⟶ ☼^{m} (m > 1) be a vector field. Then there are scalar fields $F_1, F_2, \ldots, F_m$ from ☼^{n} ⟶ ☼ such that

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_m(\mathbf{x}))$$

- The functions $F_1, F_2, \ldots, F_m$ are called the <u>coordinate functions</u> of **F**.
- For example: $\mathbf{F}(x,y,z) = \left(x^2z-3,\ x+y^2z^3,\ \sin(xyz)\right)$

$$F_1(x,y,z) = x^2z-3 \qquad F_2(x,y,z) = x+y^2z^3$$

$$F_3(x,y,z) = \sin(xyz)$$

The Space ☼^{n}: Linear Algebra meets Analysis

☼^{n} is a Linear space (or vector space)---each element of ☼^{n} is a vector. Vectors can be added together and any vector multiplied by a scalar (number) is also a vector.

$$\mathbf{x}+\mathbf{y}=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$$

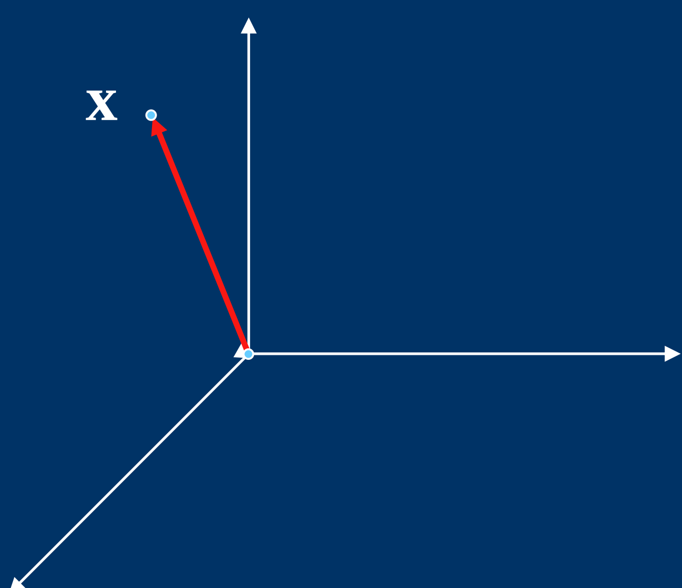
$$2\mathbf{x}=(2x_1, 2x_2, \ldots, 2x_n)$$

☼^{n} is Normed---Every element **x** in ☼^{n} has a norm $\|\mathbf{x}\|$, which is a non-negative real number and which you can think of as the "magnitude" of the vector.

The Space ☼^{n}:
Linear Algebra meets Analysis

The norm of $\|\mathbf{x}\|$ is defined to be the (usual) distance in ☼^{n} from $\mathbf{x}$ to $\mathbf{0}$.

$$\|\mathbf{x}\| = d_{\mathbf{R}^n}(\mathbf{x},\mathbf{0}) = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$



The norm in ☼ is analogous to the absolute value in ☼:

$$\|\mathbf{x}\text{-}\mathbf{y}\| = d_{\mathbf{R}^n}(x, y)$$

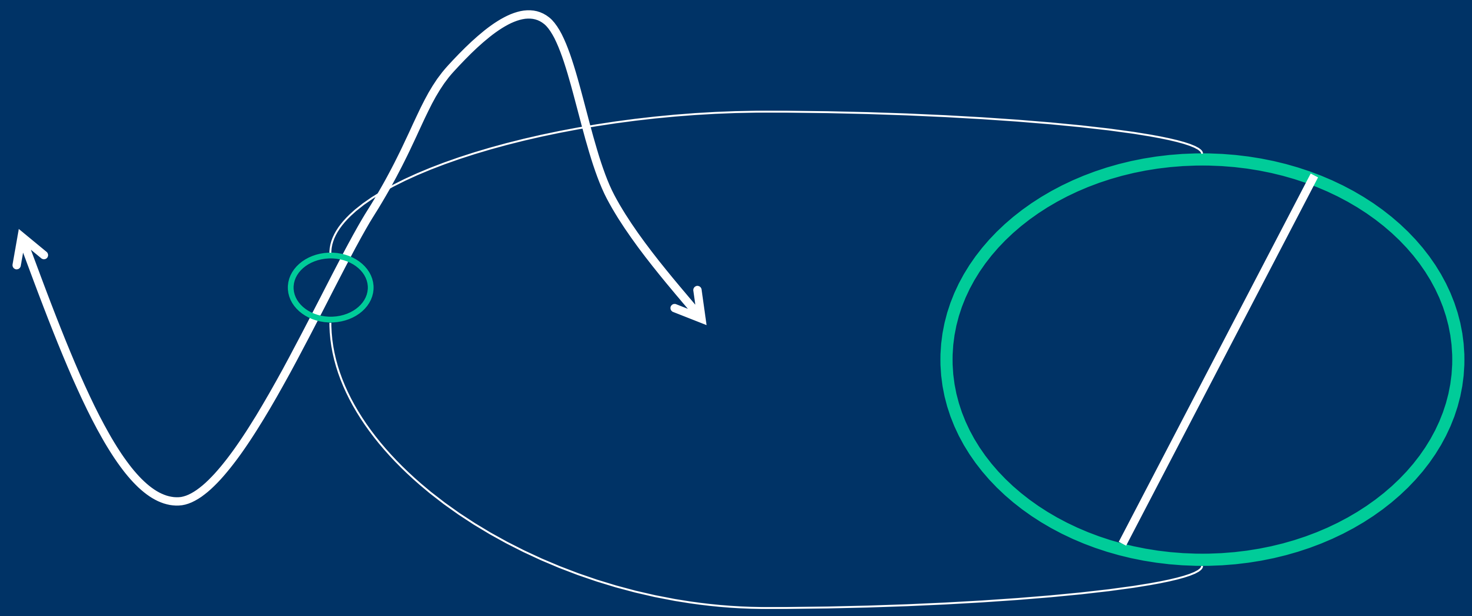
$\|\mathbf{x}\| = 0$ if and only if $\mathbf{x}=\mathbf{0}$

$$\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$$

$$\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$$

Differentiability---1 variable

- When we zoom in on a “sufficiently nice” function of one variable, we see a straight line.



Zooming and Differentiability

- We expressed this view of differentiability by saying that *f* is differentiable at *p* if there exists a real number *f* ’(*p*) such that

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

provided that *x* is “close” to *p*.

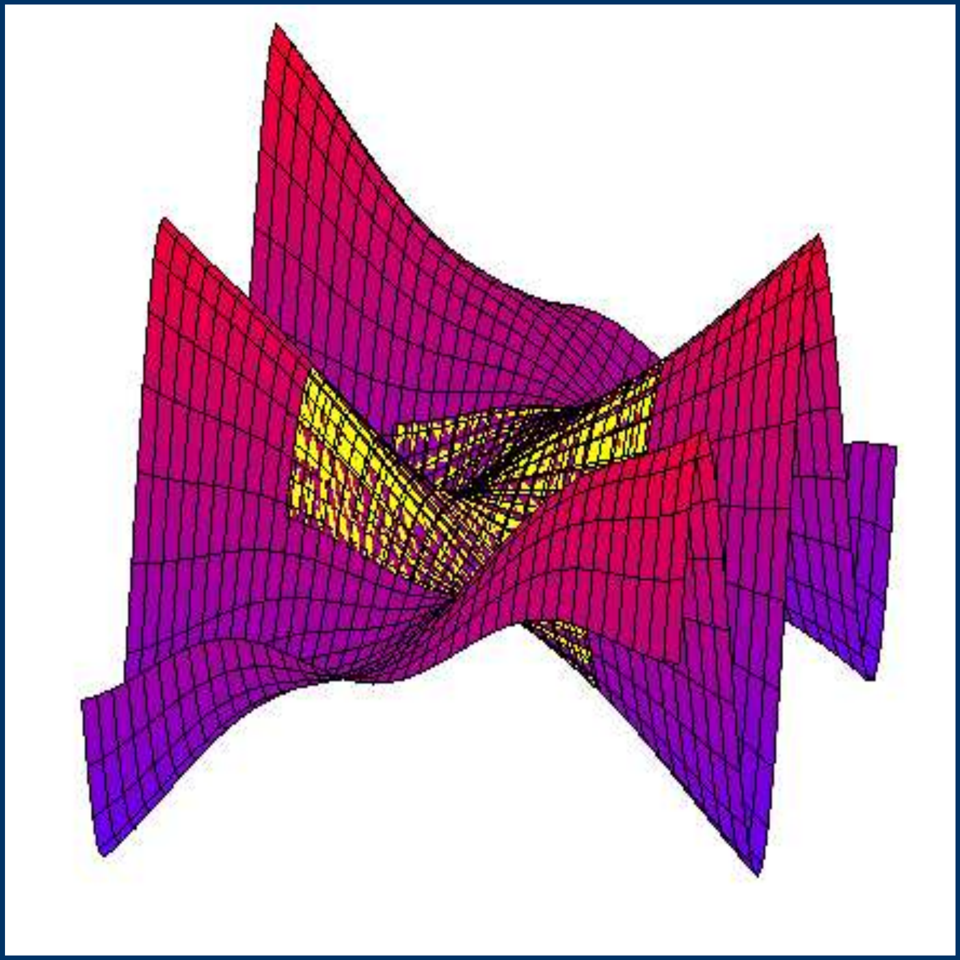
- More precisely, if for all *x*,

$$f(x) = f'(p)(x-p) + f(p) + r(x)$$

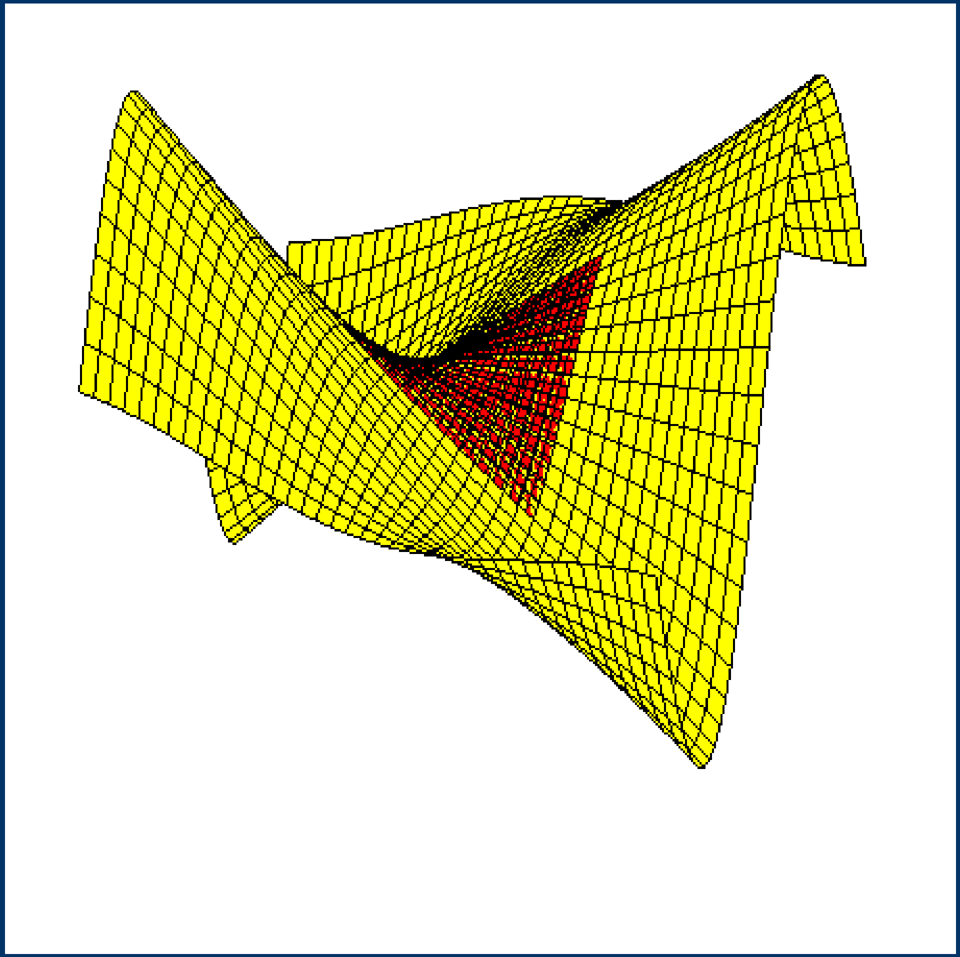
where $\dfrac{|r(x)|}{|x-p|} \to 0$ as $x \to p$.

In other words, if *f* is **locally linear** at *p***.**

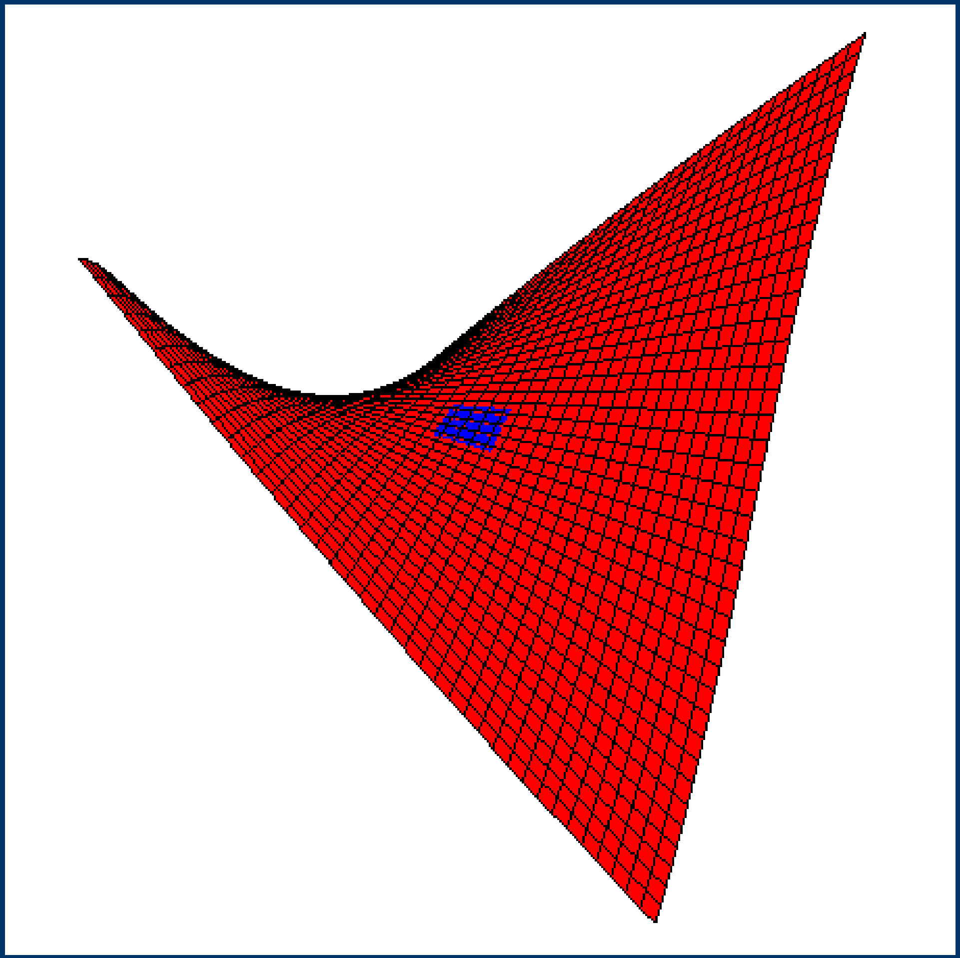
Functions of two Variables



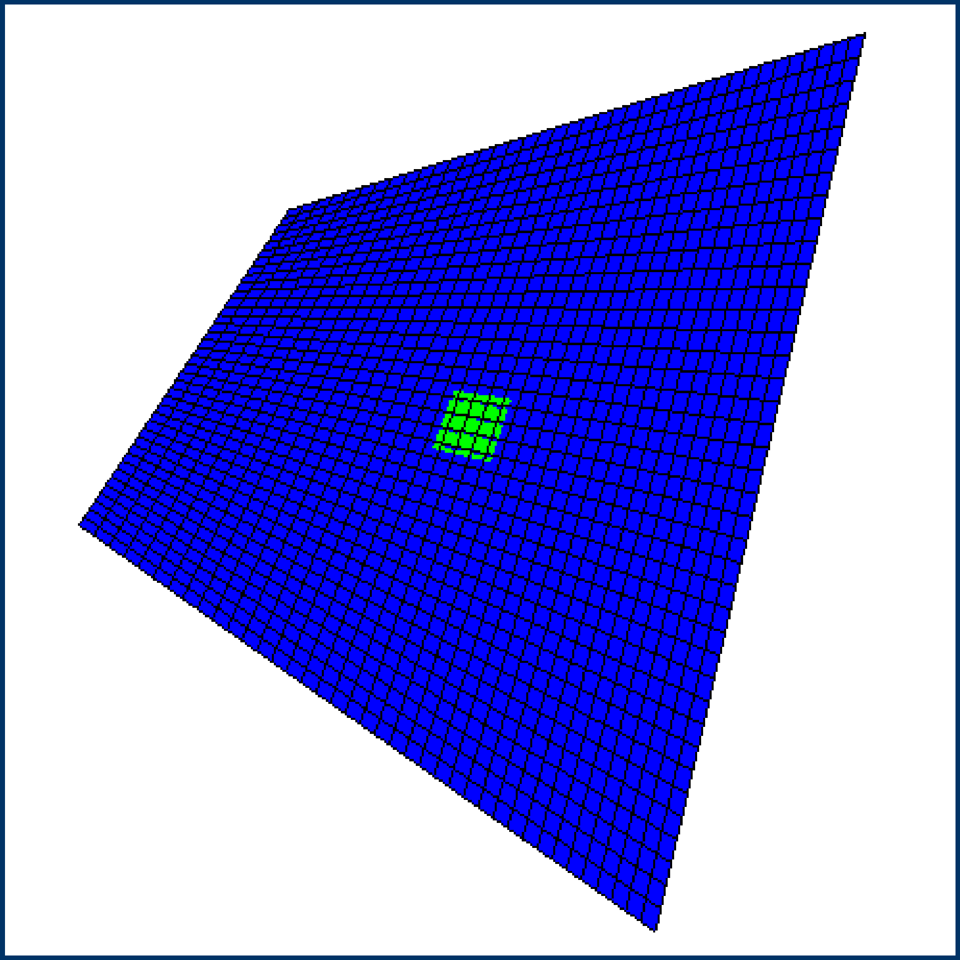
Functions of two Variables



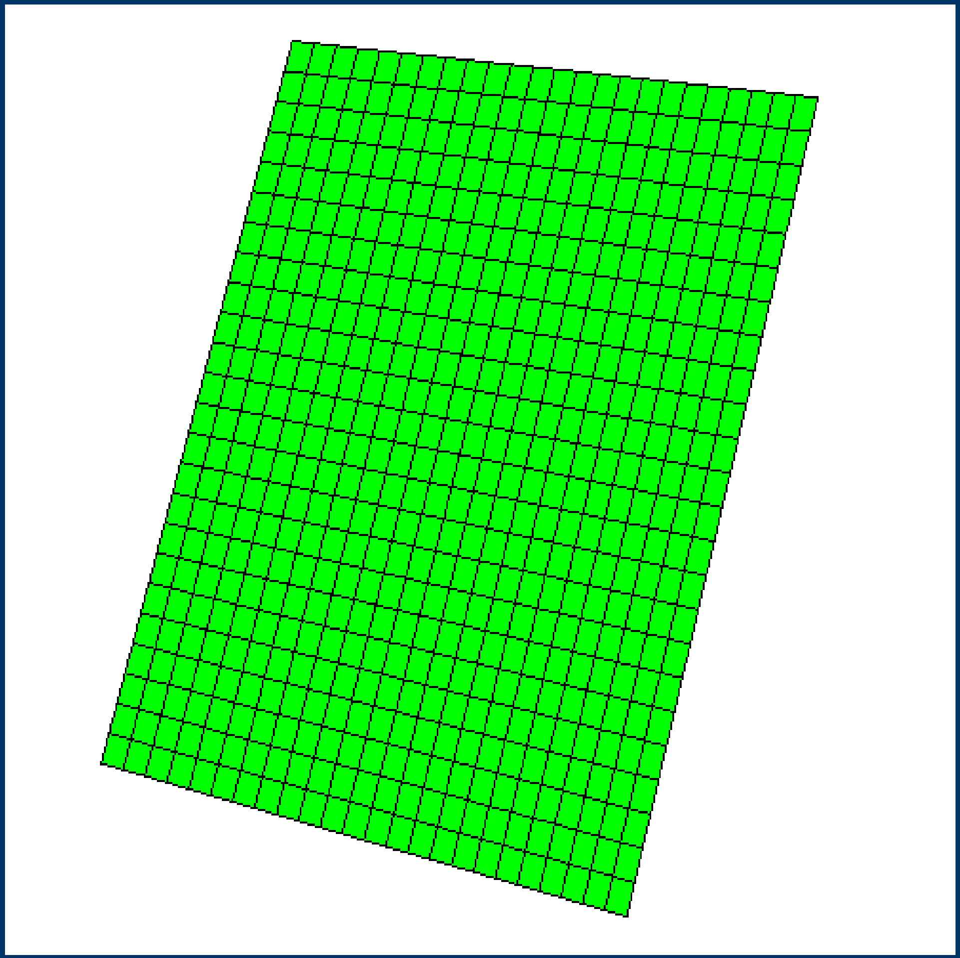
Functions of two Variables



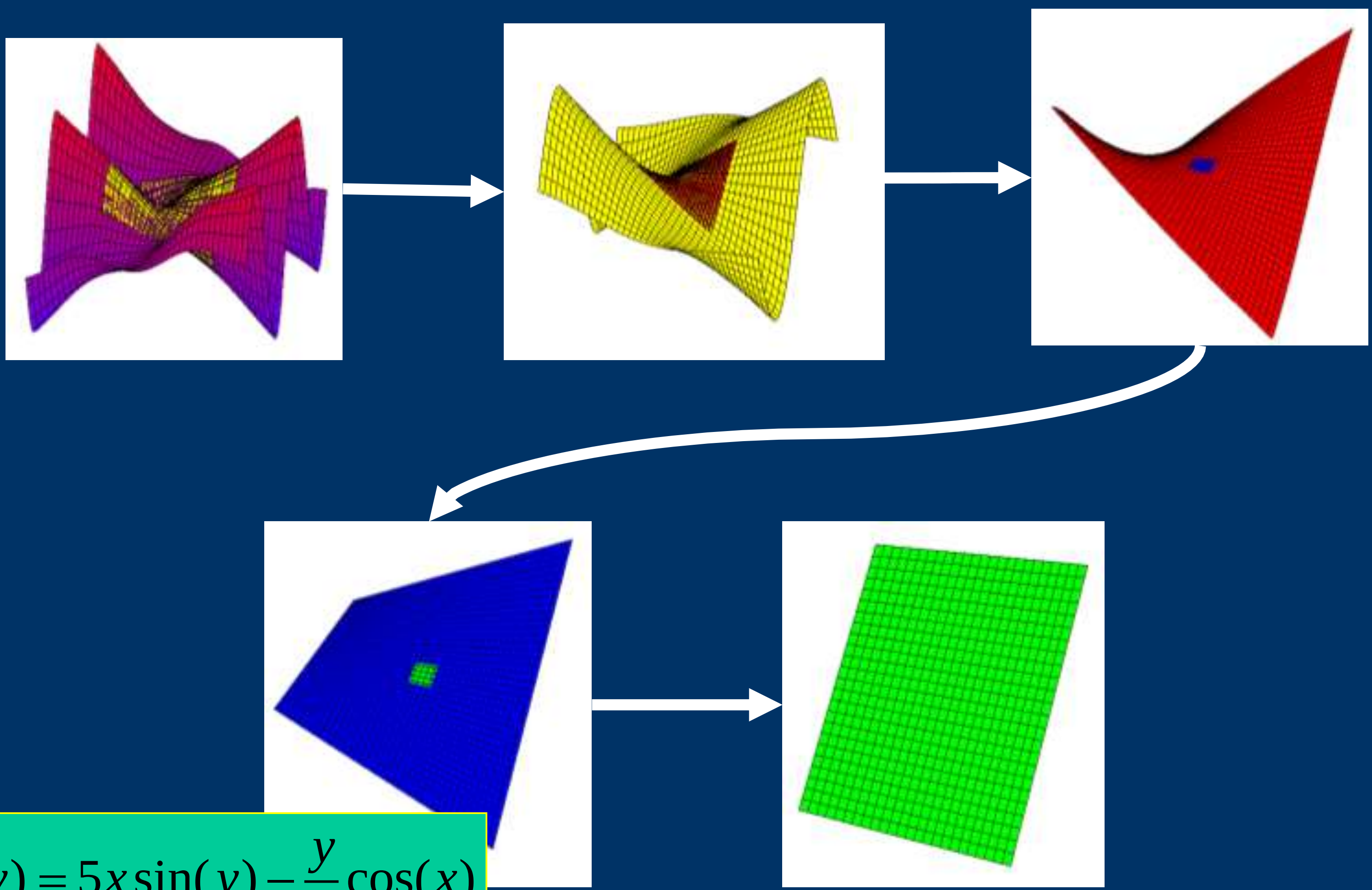
Functions of two Variables



Functions of two Variables



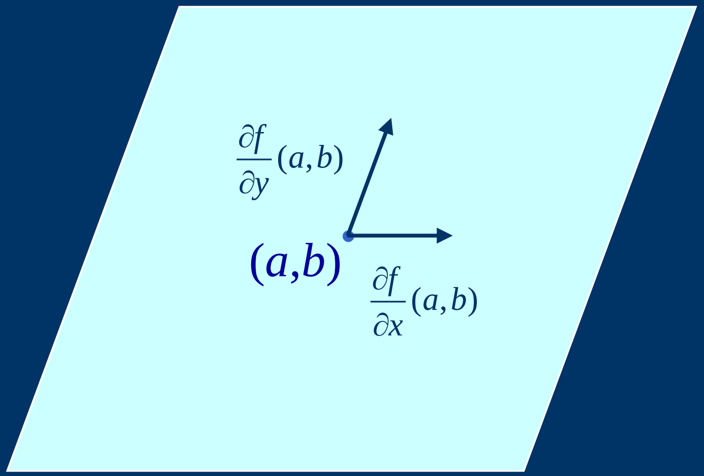
When we zoom in on a “sufficiently nice” function of two variables, we see a plane.



$$f(x,y) = 5x\sin(y) - \frac{y}{2}\cos(x)$$

Describing the Tangent Plane

- To describe a tangent <u>line</u> we need a single number---the slope.
- What information do we need to describe this plane?
- Besides the point (a,b), we need two numbers: the partials of f in the x- and y-directions.



Equation?

$$L_{(a,b)}(x,y) = \frac{\partial f(a,b)}{\partial x}(x-a) + \frac{\partial f(a,b)}{\partial y}(y-b) + f(a,b)$$

Describing the Tangent Plane

$$L_{(a,b)}(x,y) = \frac{\partial f(a,b)}{\partial x}(x-a) + \frac{\partial f(a,b)}{\partial y}(y-b) + f(a,b)$$

We can also write this equation in vector form.

Write $\mathbf{x} = (x,y)$, $\mathbf{p} = (a,b)$, and $\vec{\nabla} f(\mathbf{p}) = \left\langle \frac{\partial f(a,b)}{\partial x}, \frac{\partial f(a,b)}{\partial y} \right\rangle$

Gradient Vector!

$$L_{\mathbf{p}}(\mathbf{x}) = \vec{\nabla} f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + f(\mathbf{p})$$

Dot product!

General Linear Approximations

In the expression $L_{\mathbf{p}}(\mathbf{x}) = \vec{\nabla} f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + f(\mathbf{p})$ we can think of the gradient as a linear function on $☼^2$. (It assigns a vector to each point in $☼^2$.)

For a general function $\mathbf{F}$: $☼^n \longrightarrow ☼^m$ and for a point $\mathbf{p}$ in $☼^n$, we want to find a linear function $\mathbf{A}_{\mathbf{p}}$: $☼^n \longrightarrow ☼^m$ such that

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{A}_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + \mathbf{F}(\mathbf{p}) \qquad \text{for } \mathbf{x} \text{ "close" to } \mathbf{p}.$$

The function $\mathbf{A}_{\mathbf{p}}$ is linear in the linear algebraic sense.

General Linear Approximations

In the expression $L_{\mathbf{p}}(\mathbf{x}) = \vec{\nabla} f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + f(\mathbf{p})$ we can think of the gradient as a linear function on ☼^{2}. (It assigns a vector to each point in ☼^{2}.)

For a general function **F**: ☼$^{n} \longrightarrow$ ☼^{m} and for a point **p** in ☼^{n}, we want to find a linear function $\mathbf{A_p}$: ☼$^{n} \longrightarrow$ ☼^{m} such that

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{A_p}(\mathbf{x} - \mathbf{p}) + \mathbf{F}(\mathbf{p}) \qquad \text{for } \mathbf{x} \text{ "close" to } \mathbf{p}.$$

Note that the expression

$\mathbf{A_p}$ (**x-p**)

is not a product. It is the function $\mathbf{A_p}$ <u>acting</u> on the vector (**x-p**).

To understand

Differentiability

We need to understand

Linear Functions

Linear Functions

A function **A** is said to be linear provided that

$$\mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{A}(\mathbf{x}) + \mathbf{A}(\mathbf{y})$$

and

$$\mathbf{A}(\lambda\mathbf{x}) = \lambda\ \mathbf{A}(\mathbf{x})$$

Note that **A** (**0**) = **0**, since **A**(**x**) = **A** (**x**+**0**) = **A**(**x**)+**A**(**0**).

For a function **A**: ☼^{n} $\longrightarrow$ ☼^{m}, these requirements are very prescriptive.

Linear Functions

It is not difficult to show that if **A**: $\mathbb{R}^n \rightarrow \mathbb{R}^m$ is linear, then **A** is of the form:

$$
\begin{aligned}
\mathbf{A}(\mathbf{x}) &= \mathbf{A}(x_1, x_2, x_3, \ldots, x_n) \\
&= \begin{bmatrix}
a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n \\
a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \cdots & + & a_{2n}x_n \\
a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \cdots & + & a_{3n}x_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \cdots & + & a_{mn}x_n
\end{bmatrix}
\end{aligned}
$$

where the a_{ij}'s are real numbers for i = 1, 2, . . . m and j =1, 2, . . ., n.

Linear Functions

Or to write this another way. . .

$$
\begin{aligned}
\mathbf{A}(\mathbf{x}) &= \mathbf{A}(x_1, x_2, x_3, \ldots, x_n) \\
&= \boxed{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \\
&= \mathbf{Ax}
\end{aligned}
$$

In other words, every linear function **A** acts just like left-multiplication by a matrix. Thus we cheerfully confuse the function **A** with the matrix that represents it!

Linear Algebra and Analysis

The requirement that **A**: $\mathbb{R}^n \to \mathbb{R}^m$ be linear is very prescriptive in other ways, too.

Let **A** be an m×n matrix and the associated linear function.

- Then **A** is Lipschitz with $\|\mathbf{Ax}\| \le \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2} \|\mathbf{x}\|.$

- In particular, when $\|\mathbf{x}\| \le 1$, $\|\mathbf{Ax}\| \le \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}.$

- That is, **A** is <u>bounded</u> on the closed unit ball of $\mathbb{R}^n$.

Norm of a Linear Function

- We can thus define the norm of $\mathbf{A}$: $\mathbb{R}^n \to \mathbb{R}^m$ by

$$\|\mathbf{A}\| = \sup\{ \|\mathbf{Ax}\| : \|\mathbf{x}\| \le 1\}.$$

- Properties of this norm:

$$\|\mathbf{A}\| \le \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^{2} \right)^{1/2}.$$

$\|\mathbf{Ax}\| \le \|\mathbf{A}\|\ \|\mathbf{x}\|$ for all $x \in \mathbb{R}^n$.

$\|a\mathbf{A}\| = |a|\ \|\mathbf{A}\|$ where a is a real number.

$\|BA\| \le \|B\|\ \|A\|$ (where B is an $k\times m$ matrix and A is an $m\times n$ matrix.)

BA represents the *composition* of **B**: $\mathbb{R}^m \to \mathbb{R}^k$ and **A**: $\mathbb{R}^n \to \mathbb{R}^m$.

Norm of a Linear Function

- Further properties of this norm:

  $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ (where A and B are m×n matrices.)

  $\|a\mathbf{A}\|=|a|\ \|\mathbf{A}\|$ where a is a real number.

These two things imply that

  $\|A-B\|$ is a distance function that measures the distances between linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$.

In other words, this norm on linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ acts pretty much like the norm on $\mathbb{R}^n$. To "first order" the properties that we associate with absolute values hold for this norm.

Local Linear Approximation

$$\mathbf{A}(\mathbf{x}) \quad = \quad \mathbf{A}\mathbf{x}$$

where $\mathbf{A}$ is the $m \times n$ matrix

$$\mathbf{A} \quad = \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

For $\mathbf{x}$ "close" to $\mathbf{p}$ we want

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{A}_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + \mathbf{F}(\mathbf{p})$$

For all x, we have

$\mathbf{F}(\mathbf{x}) = \mathbf{A}_{p}(\mathbf{x}\text{-}\mathbf{p}) + \mathbf{F}(\mathbf{p}) + \mathbf{E}(\mathbf{x})$

Where $\mathbf{E}(\mathbf{x})$

is the error committed by

$\mathbf{L}_{\mathbf{p}}(\mathbf{x}) = \mathbf{A}_{\mathbf{p}}(\mathbf{x}\text{-}\mathbf{p}) + \mathbf{F}(\mathbf{p})$ in

approximating $\mathbf{F}(\mathbf{x})$

Local Linear Approximation

Fact: Suppose that **F**: $\mathbb{R}^n \to \mathbb{R}^m$ is given by coordinate functions **F**=(F_1, F_2 , . . ., F_m) and all the partial derivatives of F exist "near" $\mathbf{p} \in \mathbb{R}^n$ and are continuous at **p** , then . . .

there is some matrix $\mathbf{A}_p$ such that **F** can be approximated locally near **p** by

$$\mathbf{L}_\mathbf{p}(\mathbf{x}) = \mathbf{A}_\mathbf{p}(\mathbf{x}-\mathbf{p}) + \mathbf{F}(\mathbf{p})$$

What can we say about the relationship between the matrix $\mathbf{A}_p$ and the coordinate functions F_1, F_2, F_3, . . ., F_m?

Quite a lot, actually. . .

We Just Compute

First, I ask you to believe that if $\mathbf{A_p} = (A_1, A_2, \ldots, A_n)$ for all i and j with $1 \leq i \leq n$ and $1 \leq j \leq m$

$$\frac{\partial A_j}{\partial x_i} \equiv \frac{\partial F_j}{\partial x_i}(\mathbf{p})$$

This should not be too hard. Why?

Think about tangent lines, think about tangent planes.

Considering now the matrix formulation, what is the partial of A_j with respect to x_i?
(Note: $A_j(\mathbf{x}) = a_{j\,1}x_1 + a_{j\,2}x_2 + \ldots + a_{jn}x_n$)

The Derivative of **F** at **p**

(sometimes called the **Jacobian Matrix** of **F** at **p**)

$$\mathbf{F}'(\mathbf{p}) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{p}) & \frac{\partial F_1}{\partial x_2}(\mathbf{p}) & \frac{\partial F_1}{\partial x_3}(\mathbf{p}) & \cdots & \frac{\partial F_1}{\partial x_n}(\mathbf{p}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{p}) & \frac{\partial F_2}{\partial x_2}(\mathbf{p}) & \frac{\partial F_3}{\partial x_3}(\mathbf{p}) & \cdots & \frac{\partial F_2}{\partial x_n}(\mathbf{p}) \\ \frac{\partial F_3}{\partial x_1}(\mathbf{p}) & \frac{\partial F_3}{\partial x_2}(\mathbf{p}) & \frac{\partial F_3}{\partial x_3}(\mathbf{p}) & \cdots & \frac{\partial F_3}{\partial x_n}(\mathbf{p}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{p}) & \frac{\partial F_m}{\partial x_2}(\mathbf{p}) & \frac{\partial F_m}{\partial x_3}(\mathbf{p}) & \cdots & \frac{\partial F_m}{\partial x_n}(\mathbf{p}) \end{bmatrix}$$

Some Useful Derivatives

- Identify the derivative of each vector field at a point **p**. Guess, then verify!
- The <u>constant</u> function **F** (**x**) = **v**.
- The <u>identity</u> function **F**(**x**) = **x**.
- The <u>Linear</u> function **A**(**x**)=**A** **x**.
- A∘**F** (Where A is a linear function and F is diff'able)
- **F**+**G** (assuming both **F** and **G** are diff'able)
- a**F** (where a∈☼ and **F** is diff'able)

Continuity of the Derivative?

Theorem: Suppose that all of the partial derivatives of **F**: ☼$^{n} \longrightarrow$ ☼^{m} exist in a neighborhood around the point **p** and that they are all continuous at **p**. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that if d (**z** , **p**) $< \delta$, then

$$\|\mathbf{F}'(\mathbf{z}) - \mathbf{F}'(\mathbf{p})\| < \varepsilon.$$

In other words, if the partials exist and are continuous near **p** then the Jacobian matrix for **p** is "close" to the Jacobian matrix for any "nearby" point.

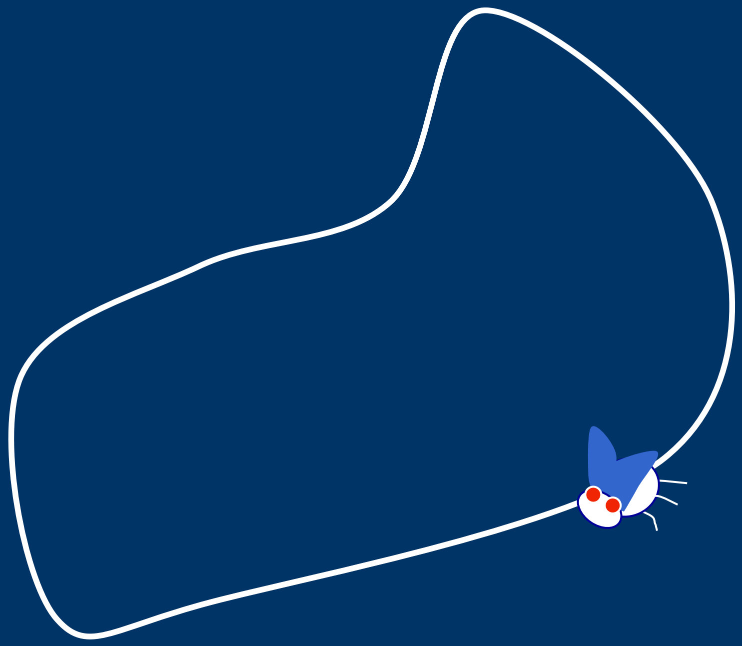
Mean Value Theorem for Vector Fields?

Theorem: Let E be an open subset of $\mathbb{R}^n$ and let $\mathbf{F}$: E $\longrightarrow \mathbb{R}^m$. Suppose that **a**, **b** and the entire line segment joining them are in E. If F is differentiable at every point on the line segment between a and b (including the endpoints) then there exists c on the segment between a and b such that

$$\left\|\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\right\| \leq \left\|\mathbf{F}'(\mathbf{c})\left(\mathbf{b}-\mathbf{a}\right)\right\| \leq \left\|\mathbf{F}'(\mathbf{c})\right\|\left\|\mathbf{b}-\mathbf{a}\right\|$$

Note: $\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a}) = \mathbf{F}'(\mathbf{c})\left(\mathbf{b}-\mathbf{a}\right)$ does not hold if m >1!
Even if n = 1 and m = 2.

Example?



Standard way to interpret **F**: ☼ → ☼^{2} is to picture a (parametric) curve in the plane. Picture a fly flying around the curve. It's velocity (a vector!) at any point is the derivative of the parametric curve at that point. What would

$$\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})=\mathbf{F}'(\mathbf{c})\left(\mathbf{b}-\mathbf{a}\right)$$

mean for a <u>closed</u> curve?

The “Take Home Message”

- The set of Linear Functions on ☼^{n} is normed and that norm behaves pretty much like the absolute value function.
- There is a multi-variable version of the Mean Value Theorem than involves inequalities in the norms.
- If the partials exist and are continuous, the Jacobian matrices corresponding to nearby points are “close” under the norm.

I will remind you of these results when we need them.