

## Lecture 1: Introduction, vectors and vector operations

Lecturer: Nikh l Srivastava

January 16

Scribe: Aditya Sengupta

**Note:** *LaTeX template courtesy of UC Berkeley EECS dept.*

## 1.1 Overview of course

Recall from 1A/1B:

- Scalar functions of scalar variables (input and output were both numbers), e.g.  $f(x) = x^2$  on  $\mathbb{R}$
- Scalar functions had domains and codomains:

$$f : I \rightarrow \mathbb{R}$$

- How to graph these functions
- Differentiation:  $\frac{dy}{dx} = f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$  = slope of tangent line of graph at  $(x, f(x))$ . Maxima and minima of functions this way.
- Integration:  $\int_a^b f(x) dx = \lim_{\Delta x \rightarrow 0} f(x_i^*) \Delta x$
- Fundamental Theorem of Calculus

$$\int_a^b f'(x) dx = f(b) - f(a)$$

In this course, we'll extend all these concepts from one dimension (scalar inputs/outputs) to two and three dimensions. This involves dealing with vector valued functions of vector variables. For example,

$$f(x, y) = x^2 + y^2 \text{ defined on } [-1, 1] \times [-1, 1]$$

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

Both the inputs and outputs can be a vector consisting of multiple scalars in these examples.

The basic idea behind multivariable calculus is the same as in single variable calculus, which is to approximate curved things/other complex shapes by flat things and to take a limit.

## 1.2 Applications of multivariable calculus

1. Classical physics/electromagnetism, because the real world has three dimensions. Great. I now have *two* courses dedicated to telling me what Maxwell's equations mean.

2. Statistics/machine learning. The backpropagation algorithm in deep learning is basically the multivariate chain rule.
3. Geometry, graphics
4. Helps develop visual intuition in three dimensions. That's a thing I can do on a computer screen by typing.

## 1.3 Organisation of content

There are four parts to the course,

1. Geometric Preliminaries
2. Differentiation
3. Integration
4. Generalisations of the Fundamental Theorem of Calculus

## 1.4 Geometric Preliminaries

A point is represented as  $(a_1, a_2) \in R^2$  in a plane, and  $(a_1, a_2, a_3) \in R^3$  in space.

$(a_1, a_2)$  are the coordinates of the point in 2D space, i.e. it is the point you reach by moving  $a_1$  along the  $x$ -axis and  $a_2$  along the  $y$ -axis. This can be generalised to three dimensions, although a representation on paper is more difficult to understand. We can use the right hand rule to do this. (Add the right hand rule in later).

Pythagoras' Theorem is a useful tool to tell us the distance between pairs of points (or more simply, the distance of a point from the origin) in  $R^2$  and  $R^3$ . This leads to the distance formula:

**Definition 1.** *The distance between two points  $(a_1, a_2, a_3)$  and  $(b_1, b_2, b_3)$  is defined to be*

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}$$

## 1.5 Vectors

A more useful concept than a point in space is a vector.

### 1.5.1 Definition

**Definition 2.** *A vector is a quantity that has a direction and a magnitude (length).*

We represent vectors as arrows. Here's an example of a vector:  $\rightarrow$

Vectors are labelled  $\vec{a}$ . They do not have a fixed position. This vector is the same as the one above:  $\rightarrow$ . However, just the use of arrows becomes inconvenient for calculations. Therefore we write vectors in 3D with respect to a coordinate system.

## 1.5.2 Basis vectors

**Definition 3.** The standard basis vectors  $\hat{i}$ ,  $\hat{j}$ , and  $\hat{k}$  are vectors of unit length along the  $x$ ,  $y$ , and  $z$  axes respectively.

(A vector of unit length is denoted  $\hat{a}$ , and the length of a vector is denoted  $|\vec{a}|$ .)

Any other vector in 3D space can be represented as a sum of these standard basis vectors. Since vectors are independent of their starting position, we can move the tail of the vector to the origin, then describe the movement along the vector from the origin as moving some amount along each of the basis vectors:

$$\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k} = \langle a_1, a_2, a_3 \rangle$$

These are the components of the vector. This seems similar to the representation of a point, with a key difference being that the vector represents the displacement from the origin to a point.

The length of a vector  $\vec{a}$  is defined as

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

by the distance formula.

## 1.5.3 Example

Consider the vector  $\vec{a} = \hat{i} + 2\hat{j} + 3\hat{k}$ . We can get the length of this vector from the 3D distance formula:

$$|\vec{a}| = \sqrt{1^2 + 2^2 + 3^2}$$

This can also be derived from the Pythagorean theorem (the 2D distance formula) by first finding the length of the components of the vector in the  $x$ - $y$  plane, then taking this as one side and the  $z$ -component as another side of a right angled triangle.

## 1.6 Basic operations on vectors

### 1.6.1 Addition

Vector addition can be visualised by moving the tail of one vector and the head of the other vector together, and completing the resultant triangle.

Algebraically, vectors  $\vec{a} = \langle a_1, a_2, a_3 \rangle$  and  $\vec{b} = \langle b_1, b_2, b_3 \rangle$  can be added by adding each corresponding pair of components:

$$\vec{a} + \vec{b} = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3 \rangle$$

## 1.6.2 Scalar Multiplication

This involves only a change in the magnitude of the vector, where the length is changed (increased, decreased, or where the scalar is negative, in the opposite direction and scaled by a corresponding positive amount).

$c\vec{a}$  means multiply  $\vec{a}$  by  $c$ . Where  $c > 0$ , the direction stays the same. Where  $c < 0$ , the direction becomes opposite. Where  $c = 0$ , it becomes the zero vector, which is the only vector not to have a direction.

Algebraically, this becomes

$$c\vec{a} = \langle ca_1, ca_2, ca_3 \rangle$$

## 1.7 Multiplication of vectors

We reject the apparently obvious notion that

$$\langle a_1, a_2, a_3 \rangle \langle b_1, b_2, b_3 \rangle = \langle a_1 b_1, a_2 b_2, a_3 b_3 \rangle$$

as it has no geometric meaning.

There are two geometrically meaningful ways to multiply vectors:

### 1.7.1 Dot Product

This takes two vectors and outputs a scalar:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

in  $R^3$ .

e.g.  $\langle 1, 2, 3 \rangle \cdot \langle 0, 1, 0 \rangle = 1 \cdot 0 + 2 \cdot 1 + 3 \cdot 0 = 2$ .

We can take the dot products of the standard basis vectors:

$$\hat{i} \cdot \hat{i} = \langle 1, 0, 0 \rangle \cdot \langle 1, 0, 0 \rangle = 1$$

$$\hat{i} \cdot \hat{j} = \langle 1, 0, 0 \rangle \cdot \langle 0, 1, 0 \rangle = 0$$

**Properties of the dot product:**

1.  $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$  (commutativity)
2.  $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$  (distributivity)

Note that this is unlike the multiplication of numbers in that  $\vec{a} \cdot \vec{b} \cdot \vec{c}$  doesn't make sense.

## 1.7.2 Geometric meaning of dot product

A vector dotted with itself gives the length squared:

$$\vec{a} \cdot \vec{a} = a_1^2 + a_2^2 + a_3^2 = |\vec{a}|^2$$

More interesting interpretation: dot product gives the angle between two vectors. Suppose  $\vec{a} = \langle a_1, a_2, a_3 \rangle$  and  $\vec{b} = \langle b_1, b_2, b_3 \rangle$ , and there is an angle  $\theta$  between them. We can work out the angle between them geometrically with the law of cosines. First define  $\vec{c} = \vec{a} - \vec{b}$ , then using the law of cosines:

$$|\vec{c}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a}||\vec{b}| \cos \theta$$

Now, from the distributive property of the dot product,

$$\begin{aligned} |\vec{c}|^2 &= \vec{c} \cdot \vec{c} = (\vec{a} - \vec{b}) \cdot (\vec{a} - \vec{b}) = \vec{a} \cdot \vec{a} - \vec{a} \cdot \vec{b} - \vec{b} \cdot \vec{a} + \vec{b} \cdot \vec{b} \\ &= |\vec{a}|^2 + |\vec{b}|^2 - 2\vec{a} \cdot \vec{b} \end{aligned}$$

On comparing this with the above, we see that

$$\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos \theta$$

Therefore, the dot product gives us an easy way to compute angles.

We can first find the vectors that define the angle using the coordinates of the relevant points,

$$\vec{a} = \vec{PQ} = \vec{OQ} - \vec{OP} = \langle 1, 1, 0 \rangle - \langle 0, 0, 1 \rangle = \langle 1, 1, -1 \rangle$$

Similarly,  $\vec{b} = \langle 0, 2, -1 \rangle$ .

Then we find the dot product:  $\vec{a} \cdot \vec{b} = 1 \cdot 0 + 1 \cdot 2 + -1 \cdot -1 = 3$

The lengths of the individual vectors can be found by taking each vector's dot product with itself,

$$\begin{aligned} |\vec{a}| &= \sqrt{1^2 + 1^2 + (-1)^2} = \sqrt{3} \\ |\vec{b}| &= \sqrt{0^2 + 2^2 + (-1)^2} = \sqrt{5} \end{aligned}$$

Therefore, we can find the cosine of the angle between them:

$$\cos \theta = \frac{2}{\sqrt{3}\sqrt{5}}$$

Special case: if  $\vec{a} \cdot \vec{b} = 0$  then the angle between them is  $\frac{\pi}{2}$ .

Math 53: Multivariable Calculus

Spring 2018

## Lecture 2: Cross product, characterising lines and planes

Lecturer: Nikhil Srivastava

January 18

Scribe: Aditya Sengupta

## 2.1 Projection operation

Projection deals with representing the components of a vector in terms of another vector, not necessarily the basis vectors.

**Definition 4.** Given  $\vec{a}, \vec{b} \neq 0$ , the vector projection of  $\vec{b}$  on  $\vec{a}$ , written as  $\text{proj}_{\vec{a}}(\vec{b})$ , is the unique vector that is parallel to  $\vec{a}$  satisfying the condition that the projection is orthogonal to  $\vec{a}$ .

We can find the magnitude and direction of the projection. The direction is the same as that of  $\vec{a}$ , which is  $\frac{\vec{a}}{|\vec{a}|}$ . The magnitude is  $|\vec{b}| \cos \theta$ , from trigonometry. Therefore the vector  $\text{proj}_{\vec{a}}(\vec{b}) = \frac{\vec{a}}{|\vec{a}|} |\vec{b}| \cos \theta$ , or more simply  $\frac{\vec{a}}{|\vec{a}|^2} (\vec{a} \cdot \vec{b})$ .

The projection gives a decomposition of any vector  $\vec{b}$ :

$$\vec{b} = \text{proj}_{\vec{a}}(\vec{b}) + (\vec{b} - \text{proj}_{\vec{a}}(\vec{b}))$$

Geometrically, we can see that the second part of this is the component perpendicular to the other vector  $\vec{a}$ .

## 2.2 Interpretation of the dot product

From the value of the dot product, we can make conclusions about the angle between the two vectors:

$$\vec{a} \cdot \vec{b} = 0 \iff \theta = \frac{\pi}{2}$$

$$\vec{a} \cdot \vec{b} > 0 \iff \theta < \frac{\pi}{2}$$

$$\vec{a} \cdot \vec{b} < 0 \iff \theta > \frac{\pi}{2}$$

## 2.3 Cross product

The cross product takes two vectors in  $\mathbb{R}^3$  and outputs a vector.

**Definition 5.** Geometrically, given  $\vec{a}, \vec{b} \in \mathbb{R}^3$ ,  $\vec{a} \times \vec{b}$  is the unique vector in  $\mathbb{R}^3$  with magnitude  $|\vec{a} \times \vec{b}|$  equal to the area of the parallelogram spanned by  $\vec{a}$  and  $\vec{b}$ , and a direction orthogonal to both  $\vec{a}$  and  $\vec{b}$  which satisfies the right hand rule.

**Remark.**  $\vec{a} \times t\vec{a} = 0$

We can simplify the magnitude using the geometric interpretation, in which the area of a parallelogram is base  $\times$  height, which works out to  $|\vec{a}||\vec{b}|\sin\theta$ .

### 2.3.1 Cross products of basis vectors

$$\hat{j} \times \hat{i}$$

The above can be worked out geometrically by drawing  $\hat{i}$  and  $\hat{j}$  in 3D space. The two define a square of side 1, therefore the vector's magnitude is 1. From the right hand rule, we see the direction is  $-\hat{k}$ . Therefore

$$\hat{j} \times \hat{i} = -\hat{k}$$

Note that if the order of parameters were switched, the right hand rule would work the other way and the answer would be  $\hat{k}$ .

From this we see that  $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ .

### 2.3.2 Algebraic definition of cross product

**Definition 6.** *The determinant of a 2x2 matrix*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The determinant gives us an algebraic definition for the cross product. Let  $\langle a_1, a_2 \rangle$  and  $\langle b_1, b_2 \rangle$  be vectors in  $\mathbb{R}^2$ . Then, the determinant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

gives the area of the parallelogram defined by  $\vec{a}$  and  $\vec{b}$ . This can be shown:

$$\begin{aligned} A &= |\vec{a}||\vec{b}|\sin\theta = |\vec{a}||\vec{b}|\sqrt{1 - \cos^2\theta} \\ &= |\vec{a}||\vec{b}|\sqrt{1 - \frac{(\vec{a} \cdot \vec{b})^2}{|\vec{a}|^2|\vec{b}|^2}} = |a_1b_2 - a_2b_1| \end{aligned}$$

This can be extended to three dimensions if we first define a 3x3 determinant in terms of 2x2 determinants:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

Geometrically, this gives us the volume of the *parallelapiped* defined by vectors  $\vec{a}, \vec{b}, \vec{c}$ .

This allows us to define the cross product algebraically:

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \hat{i} \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - \hat{j} \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \hat{k} \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

**Example**

$$\vec{a} = \langle 1, 1, 0 \rangle = \hat{i} + \hat{j}$$

$$\vec{b} = \langle 1, 0, 0 \rangle = \hat{i}$$

We apply the algebraic definition,

$$\vec{a} \times \vec{b} = \hat{i} \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} - \hat{j} \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} + \hat{k} \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = 0 + 0 - \hat{k}$$

**Remark.** The direction in  $\vec{a} \times \vec{b}$  indicates a specific side of a parallelogram.

**Remark.**  $\vec{a} \cdot (\vec{b} \times \vec{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$ .

This is the triple product of three vectors.

## 2.4 Lines and Planes

Lines and points can be described using vectors. The point  $P(x, y, z)$  is identified by its *position vector*  $\vec{r} = \langle x, y, z \rangle$ .

To specify a line in 3D, we need:

- a point on the line (specified by its position vector  $\vec{r}_o = \langle x_0, y_0, z_0 \rangle$ )
- a vector parallel to the line,  $\vec{v} = \langle a, b, c \rangle$ .

With this information, we can start at any point on the line (the position vector to the point) and add any multiple of the parallel vector  $\vec{v}$  to get any point on the line. The general expression for the line is therefore

$$\vec{r} = \vec{r}_o + t\vec{v}$$

Different values of the auxiliary variable  $t$  correspond to different points on the line. This is called a parametric equation.

**Example:** Find the vector equation of the line containing the points  $(1, 2, 3)$  and  $(0, 1, 0)$ .

Let  $\vec{r}_0 = \langle 1, 2, 3 \rangle$ . We know that the line has to go through the second point  $\vec{r}_1 = \langle 0, 1, 0 \rangle$ , therefore we can take  $\vec{v} = \vec{r}_1 - \vec{r}_0 = \langle 0, 1, 0 \rangle - \langle 1, 2, 3 \rangle = \langle -1, -1, -3 \rangle$ .



Then, we use the general expression  $\vec{r} = \vec{r}_0 + t\vec{v} = \langle 1, 2, 3 \rangle + t\langle -1, -1, -3 \rangle = \langle 1 - t, 2 - t, 3 - 3t \rangle$ .

In general, we can use this component-wise addition approach to show that

$$\langle x, y, z \rangle = \langle x_0 + at, y_0 + bt, z_0 + ct \rangle$$

This is also a parametric expression. We can convert this into three parametric scalar equations,

$$x = x_0 + at$$

$$y = y_0 + bt$$

$$z = z_0 + ct$$

We can now eliminate  $t$  by solving for  $t$  in each equation and setting them all equal to each other, which yields

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

which is valid as long as  $a, b, c \neq 0$ .

**Remark.** To describe a line segment between two points characterised by  $\vec{r}_0$  and  $\vec{r}_1$ ,

$$\vec{r} = \vec{r}_0 + t(\vec{r}_1 - \vec{r}_0), t \in [0, 1]$$

## 2.5 Planes in 3D

To specify a plane, we need

- a point in the plane (specified by its position vector  $\vec{r}_o = \langle x_0, y_0, z_0 \rangle$ )
- a normal vector perpendicular to the plane,  $\vec{n} = \langle a, b, c \rangle$ .

For any  $\vec{r}$  in the plane, the vector  $\vec{r} - \vec{r}_o$  is always perpendicular to  $\vec{n}$ . Therefore the vector equation of the plane is

$$(\vec{r} - \vec{r}_o) \cdot \vec{n} = 0$$

This is not a parametric equation.

We can expand vectors component-wise, resulting in

$$\langle x - x_0, y - y_0, z - z_0 \rangle \cdot \langle a, b, c \rangle = 0$$

$$\therefore a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \implies ax + by + cz = d$$

**Example:** Find the plane through

$$\vec{r}_0 = \langle 1, 0, 0 \rangle, \vec{r}_1 = \langle 1, 1, 0 \rangle, \vec{r}_2 = \langle 1, 1, 1 \rangle$$

We can find the normal vector to the plane by taking the cross product of the vectors between  $r_0$  and  $r_1$ , and  $r_0$  and  $r_2$ .

$$(\vec{r}_1 - \vec{r}_0) \times (\vec{r}_2 - \vec{r}_0)$$

The above vector is always orthogonal to the plane provided the points are not collinear.

**Theorem 7.** *The general equation of a plane through three points  $\vec{r}_0, \vec{r}_1, \vec{r}_2$  is*

$$(\vec{r} - \vec{r}_0) \cdot \left( (\vec{r}_1 - \vec{r}_0) \times (\vec{r}_2 - \vec{r}_0) \right) = 0$$

**Remark.** *A parametric equation for a plane can also be obtained:*

$$\vec{r} = \vec{r}_0 + t_1\vec{v}_1 + t_2\vec{v}_2$$

*as long as  $\vec{v}_1$  and  $\vec{v}_2$  are not parallel*

Math 53: Multivariable Calculus

Spring 2018

## Lecture 3: Curve parameterisation

Lecturer: Nikh l Srivastava

January 23

Scribe: Aditya Sengupta

### 3.1 Characterising the division of space by a plane

Do the points  $P(0, 1, 2)$  and  $Q(3, 4, 5)$  lie on the same side or different sides of the following plane?

$$(\vec{r} - \langle 1, 2, 3 \rangle) \cdot \langle 1, 1, 0 \rangle = 0$$

**Approaches:**

1. Find line segment PQ and decide whether it intersects the plane.
2. Find  $\vec{PA}, \vec{QA}$  for an arbitrary point  $A$  on the plane and find the normal vector defined by these (take the cross product)
3. The above, but with the specific point  $\langle 1, 2, 3 \rangle$

**Solution with the third approach:**

The vector between  $P$  and the given point is  $\vec{r}_1 = \langle -1, -1, -1 \rangle$ . The vector between  $Q$  and the given point is  $\vec{r}_2 = \langle 2, 2, 2 \rangle$ .

Now, if the angle between  $\vec{r}_1$  and  $\vec{n}$  is acute and that between  $\vec{r}_2$  and  $\vec{n}$  is obtuse (or vice versa), we know they are on opposite sides of the plane. Points on one side have acute angles, and points on the other have obtuse angles.

$$\cos \theta_1 \implies \vec{r}_1 \cdot \vec{n} = \langle -1, -1, -1 \rangle \cdot \langle 1, 1, 0 \rangle = -2 < 0 \implies \text{obtuse}$$

$$\cos \theta_2 \implies \vec{r}_2 \cdot \vec{n} = \langle 2, 2, 2 \rangle \cdot \langle 1, 1, 0 \rangle = 4 > 0 \implies \text{acute}$$

Therefore they are on opposite sides of the plane.

### 3.2 Parameterised curves

Recall that a line is characterised as  $\vec{r} = \vec{r}_0 + t\vec{v}$ . This is implicitly a function of  $t$ .

**Definition 8.** A vector valued function has real inputs (domain) and vector outputs (codomain).

A parameterisation of a curve consists of a vector valued function  $\vec{r}(t)$  which specifies position vectors of points on the curve, and a range of parameters  $a \leq t \leq b$ . Giving different inputs within the range will yield position vectors to specific points on the curve.

### 3.2.1 Example 1

$$\vec{r}(t) = \langle \cos t, \sin t \rangle, t \in [0, 2\pi]$$

Plot points for different values of  $t$ .

$t$	$\vec{r}(t) = \langle \cos t, \sin t \rangle$
0	1, 0
$\pi/2$	0, 1
$\pi/4$	$1/\sqrt{2}, 1/\sqrt{2}$

Another way of approaching this is to eliminate  $t$ .

$$x = \cos t, y = \sin t = \sqrt{1 - \cos^2 t} = \pm\sqrt{1 - x^2}$$

We see that all points satisfy  $x^2 + y^2 = 1$ , which in terms of vectors means  $|\vec{r}(t)|^2 = 1$ . This is the set of all points at a distance of 1 from the origin, which is a circle.

### 3.2.2 Example 2

$$\vec{r}(t) = \langle -\cos(2t), -\sin(2t) \rangle, t \in [0, 2\pi]$$

This is a circle that rotates twice and that starts from a different point.

### 3.2.3 Example 3

$$\vec{r}(t) = \langle \cos t, \sin t, t \rangle, t \in \mathbb{R}$$

We can visualise this by looking at its projections onto 2D planes.

In the  $x - y$  plane, this looks like the vector

$$\vec{r}(t) = \langle \cos t, \sin t \rangle$$

, which is just a circle as above.

We can visualise what happens along the  $z$ -axis because it is linear in  $t$ . We see the curve spiralling upwards helically. The curve lies in the cylinder  $x^2 + y^2 = 1$ .

### 3.2.4 Parameterised curve as moving particle

We can view a parameterised curve as describing the motion of a particle, if we think of  $t$  as time. Parameterisation can be used to compute interesting features like velocity.

**Definition 9.**

$$\lim_{t \rightarrow a} \vec{r}(t) = \lim_{t \rightarrow a} \langle x(t), y(t), z(t) \rangle = \langle \lim_{t \rightarrow a} x(t), \lim_{t \rightarrow a} y(t), \lim_{t \rightarrow a} z(t) \rangle$$

*The limit of a vector valued function is a vector with its components set to the limit of each of the corresponding components of the vector valued function.*

Conceptually, this means that as  $t$  approaches  $a$  from either direction, the vector will move towards the limit of the general vector expression as  $t$  approaches  $a$ .

$\vec{r}(t)$  is continuous at  $a$  if  $\lim_{t \rightarrow a} \vec{r}(t) = \vec{r}(a)$ .

The derivative of  $\vec{r}(t)$  is the vector

$$\frac{d\vec{r}}{dt} = \lim_{h \rightarrow 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h}$$

whenever the limit exists. This gives the vector representing an infinitesimal change in the original vector.

$$\frac{d\vec{r}}{dt} = \left\langle \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right\rangle$$

### 3.2.5 Example

$$\vec{r}(t) = \langle \cos t, \sin t \cos t \rangle, t \in [0, 2\pi]$$

We can visualise this by drawing tangent vectors to the curve at some point.

t	$\vec{r}(t)$	$\vec{r}'(t)$
0	1, 0	0, 1
$\pi/2$	0, 0	-1, -1
$\pi/4$	$1/\sqrt{2}, 1/2$	$-1/\sqrt{2}, 0$

The parameterised curve can be at the same point at different times. Velocity is a function of the parameterisation, not just the curve.

**Remark.** The unit tangent at  $t$  is defined as

$$\vec{T}(t) = \frac{\vec{r}'(t)}{|\vec{r}'(t)|}$$

This unit tangent is not defined when  $\vec{r}'(t) = 0$ .

The velocity changes if you change the parameterisation, e.g. for  $\vec{r}(t) = \langle \cos 2t, \sin 2t \rangle$  and  $\vec{s}(t) = \langle \cos t, \sin t \rangle$ ,  $\vec{r}'(t) = 2\vec{s}'(t)$ .

However, the unit tangent vector at a point does not change.

## 3.3 Mystery of the rolling cycloid

What does it look like when the marked point hits the ground, i.e. upon zooming in?

1. Makes an angle  $\theta$
2. Makes no angle (becomes infinitely sharp)
3. Becomes a smooth curve

4. Makes a loop
5. Something else?

In general, given a verbal description of some process, how do you find a parametric equation? After this has been achieved, we can use  $\vec{r}'(t)$  to answer the question.

Math 53: Multivariable Calculus

Spring 2018

## Lecture 4: Calculus with vector-valued functions

Lecturer: Nikhil Srivastava

January 25

Scribe: Aditya Sengupta

## 4.1 Mystery of the Rolling Cycloid, contd.

We start by deriving the parametric equation for the trajectory of the marked point P. We parameterise the path by the angle of the wheel (we choose the parameter that simplifies the most complicated part).

To find the vector  $\vec{OP}$  that tracks the chosen point as  $\theta$  increases, we express it as a sum of simpler vectors:

$$\vec{OP} = \vec{OA} + \vec{AB} + \vec{BP}$$

$\vec{OA}$  always has a direction  $\langle 1, 0 \rangle$  and a magnitude  $R\theta$ .

$\vec{AB}$  always has a direction  $\langle 0, 1 \rangle$  and a magnitude  $R$ .

$\vec{BP}$  always has a magnitude  $R$ , and via trigonometry we can calculate its direction as  $\langle -\sin \theta, -\cos \theta \rangle$ .

Therefore, combining all of these, we get the expression

$$\vec{r}(\theta) = \langle R\theta - R\sin \theta, R - R\cos \theta \rangle$$

This is valid on the interval  $\theta = [0, 2\pi]$ , and because  $\sin$  and  $\cos$  are periodic in  $2\pi$ , this works for all  $\theta$ .

Now that we have the equation, we can use it to examine motion of the point near  $\theta = 0$ . We do this by taking the component-wise derivative,

$$\vec{r}'(\theta) = \langle R - R\cos \theta, R\sin \theta \rangle$$

At  $\theta = 0$ ,  $\vec{r}'(0) = 0$ .

We proceed by computing velocity for small  $\theta$ ,

$$\vec{r}'(\theta) \approx \left\langle R\frac{\theta^2}{2}, R\theta \right\rangle$$

The tangent unit vector is

$$\left\langle \frac{\theta^2}{2|\theta|}, \frac{\theta}{|\theta|} \right\rangle$$

In the limit, as  $\theta \rightarrow 0^-$ ,  $\hat{T}(\theta) \approx \langle 0, 1 \rangle$ , and as  $\theta \rightarrow 0^+$ ,  $\hat{T}(\theta) \approx \langle 0, -1 \rangle$ .

Therefore, as  $\theta \rightarrow 0$ , the slope of the line becomes infinite, i.e. the correct option is 2.

## 4.2 More interesting quantities

1. Speed (scalar) =  $|\vec{r}'(t)|$  (magnitude of velocity)
2. Acceleration (vector) =  $\vec{r}''(t)$  (instantaneous rate of change of velocity)
3. Arc length (scalar)  $L$

Given a position vector-valued function  $\vec{r}(t), a \leq t \leq b$ ,

$$L := \int_a^b |\vec{r}'(t)| dt$$

### 4.2.1 Geometric visualisation of arclength

We can approximate a curve as a series of straight lines using the arclength.

Subdivide the interval  $[a, b]$  into  $a < t_1 < \dots < t_n < b$  and sum the lengths of the resultant line segments:

$$|\vec{r}(t_1) - \vec{r}(a)| + |\vec{r}(t_2) - \vec{r}(t_1)| + \dots + |\vec{r}(b) - \vec{r}(t_n)|$$

In the limit as  $n \rightarrow \infty$ , this becomes an integral, which is the same as the formula for the arclength as above.

**Remark.** *Arclength can be used as a parameterisation as well. To find this, we first define the arclength function,*

$$s(t) = \int_a^t |\vec{r}'(y)| dy$$

*Assuming the path never goes backwards, this function is invertible to get  $t(s)$ . Then the arclength parameterised path is  $\vec{r}(t) = \vec{r}(t(s)) = (\vec{r} \circ t)(s)$ .*

## 4.3 Rules of differentiation for vector-valued functions of one variable

### 4.3.1 Chain rule

Recall the scalar chain rule,

$$(f(g(x)))' = f'(g(x))g'(x)$$

This definition will not work where the domain and range of the functions to be composed do not necessarily match.

Suppose  $\vec{r}(s)$  is a vector-valued function  $\mathbb{R} \rightarrow \mathbb{R}^3$ , and  $s(t)$  is a scalar function  $\mathbb{R} \rightarrow \mathbb{R}$ . The composition here makes sense,  $\vec{r}(s(t))$  has  $\mathbb{R} \rightarrow \mathbb{R}^3$ . We can then differentiate,



$$\frac{d}{dt}\vec{r}(s(t)) = \vec{r}'(s(t))s'(t)$$

Why does this work?

$$\vec{r}(s) = \langle f(s), g(s), h(s) \rangle$$

Therefore

$$\frac{d}{dt}\langle f(s(t)), g(s(t)), h(s(t)) \rangle = \left\langle \frac{d}{dt}f(s(t)), \frac{d}{dt}g(s(t)), \frac{d}{dt}h(s(t)) \right\rangle$$

and the scalar chain rule can be applied to each of these, to yield

$$s'(t)\vec{r}'(s(t))$$

### 4.3.2 Product Rule

There are three product rules,

1.  $s(t), \vec{r}(t)$ :  $\frac{d}{dt}s(t)\vec{r}(t) = s'(t)\vec{r}(t) + s(t)\vec{r}'(t)$
2. Dot product:  $\vec{r}_1(t), \vec{r}_2(t)$

$$\frac{d}{dt}\vec{r}_1(t) \cdot \vec{r}_2(t) = \vec{r}_1'(t) \cdot \vec{r}_2(t) + \vec{r}_1(t) \cdot \vec{r}_2'(t)$$

3. Cross product: (as above)

$$\frac{d}{dt}\vec{r}_1(t) \times \vec{r}_2(t) = \vec{r}_1'(t) \times \vec{r}_2(t) + \vec{r}_1(t) \times \vec{r}_2'(t)$$

### Exercise

Suppose a curve satisfies

$$\vec{r}'(t) \times \vec{r}''(t) = 0 \forall t$$

What can you say about it?

## 4.4 Integration of vector-valued functions

This is also defined component-wise.

If  $\vec{r}(t) = \langle f(t), g(t), h(t) \rangle$ ,

$$\int_a^b \vec{r}(t) dt = \left\langle \int_a^b f(t) dt, \int_a^b g(t) dt, \int_a^b h(t) dt \right\rangle$$

This does not appear to have a clear geometric interpretation. It is not area under the curve, because the output is a vector in 3D.

The geometric interpretation is related to the vector-valued form of the Fundamental Theorem of Calculus,

$$\int_a^b \vec{r}'(t) dt = \vec{r}(b) - \vec{r}(a)$$

which is immediately apparent by vector addition on the subdivision of the arclength from section 4.2.1.

## 4.5 Warning

(our lawyers told us we had to include this)

$$\left| \int_a^b \vec{r}'(t) dt \right| \neq \int_a^b |\vec{r}'(t)| dt$$

The two are only the same when the path is a straight line. In a simplified form,

$$\left| \frac{d\vec{r}}{dt} \right| \neq \frac{d|\vec{r}|}{dt}$$

**Math 53: Multivariable Calculus**

**Spring 2018**

## Lecture 5: Functions of many variables, limits and continuity

Lecturer: Nikh l Srivastava

30 January

Scribe: Aditya Sengupta

### 5.1 Examining types of functions

In single-variable calculus, we study functions going from a domain denoted by  $x \in \mathbb{R}$  or  $x \in I$  to  $f(x) \in \mathbb{R}$ . So far in this course, we've extended this to  $t \in I \rightarrow \vec{r}(t) = \langle f(t), g(t), h(t) \rangle \in \mathbb{R}^3$ .

Now we'll extend this idea to  $(x, y) \in D \subset \mathbb{R}^2 \rightarrow f(x, y) \in \mathbb{R}$ .

**Definition 10.** If  $D$  is a region in  $\mathbb{R}^2$ , a function is a rule that assigns a unique  $f(x, y)$  to every  $(x, y) \in D$ .  $D$  is called the domain and the set  $\{f(x, y) : (x, y) \in D\}$  is called the range.

Geometrically, these functions take a point in space and return a point on the real line.

#### 5.1.1 Examples

- $f(x, y) = x + y - 1$   
The domain is not specified here, so we use the implicit definition that it is the largest subset of  $\mathbb{R}^2$  on which the definition makes sense. In this case, that is  $\mathbb{R}^2$  itself so  $D = \mathbb{R}^2$ . The range is  $\mathbb{R}$ .
- $f(x, y) = 1 - x^2 - y^2$   
The domain is once again  $D = \mathbb{R}^2$  as any  $x$  and  $y$  can be squared and subtracted. However the range is  $R = (-\infty, 1]$  as no value greater than 1 can be reached.
- $f(x, y) = \sqrt{9 - x^2 - y^2}$   
The function only makes sense if  $9 - x^2 - y^2 \geq 0$ , so the domain is  $D = \{(x, y) : 9 - x^2 - y^2 \geq 0\}$ . Geometrically this is the inside of a circle of radius 3. The range is  $[0, 3]$ .

**Remark.** We can have functions defined by physical phenomena, but we will focus on explicit rules.

#### 5.1.2 Visualising functions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

In one dimension, we visualise functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  as a graph in which we plot all points  $(x, y)$  such that  $y = f(x)$ . This can be readily extended to two variables.

The graph of  $f(x, y)$  is now a surface  $S = \{(x, y, z) : z = f(x, y) \text{ and } (x, y) \in D\}$ . The domain is a subset of the  $x - y$  plane, and each point on the surface representing the graph of the function is above or below a corresponding point in the domain.

This is a two-dimensional representation because we can parameterise the surface by two free variables. Alternatively, we can think of the function being characterised by three variables minus one constraint equation, which yields two dimensions.

## 5.2 Techniques for drawing graphs of $f(x, y)$

1. Draw traces, which are functions of one variable obtained by fixing one of the input variables.

### Example 1 - Traces

$$z = f(x, y) = -y$$

To draw the graph of this, we fix  $x = 0$ , i.e. we're only looking for solutions in the  $y - z$  plane. Then we can draw it as a line  $z = -y$  passing through the point  $(0, 0, 0)$ . Then we can draw another trace, by fixing  $x = 1$ , and once again drawing a line  $z = -y$  through the point  $(1, 0, 0)$ . In this way we can get a picture of what the entire plot would look like.

We can alternatively do this by fixing  $y = 0$ . The equation then yields  $z = 0$ . But this is a bad idea because of a heuristic we *definitely* didn't just come up with, that traces are useful when you actually get a function of one variable. Yes. That was planned.

2. A *level curve* of  $f(x, y)$  is the set of points  $(x, y)$  where  $f(x, y) = k$ . That is, we fix the output. This is also called the contour of a function.

### Example 2 - Level Curves

$$z = f(x, y) = -y$$

We can start by fixing the output as  $z = 0$ . This implies  $-y = 0$ , which gives us a line on the  $x$ -axis.  $z = 1 \implies y = -1$ , which gives us a line parallel to the  $x$ -axis. We can continue doing this on the  $x - y$  plane to get a view from above of the function. Each level curve indicates a fixed height, i.e. the intersection with a horizontal plane  $z = k$ .

A plot of many level curves with evenly spaced values of  $z$  is called a contour plot.

### Example 3 - Both

$$f(x, y) = \sqrt{1 - x^2 - y^2}, D = (x, y) : x^2 + y^2 \leq 1$$

We make a trace at  $x = 0$ :

$$z = \sqrt{1 - y^2} \implies z^2 = 1 - y^2, z \geq 0$$

which is a semicircle centered at the origin in the  $y - z$  plane.

Another trace at  $y = 0$ :

$$z^2 + x^2 = 1, z \geq 0$$

We can continue doing this to build up a full picture of the surface, but it would be easier to do a contour plot. We fix  $z = 0$ , then  $x^2 + y^2 = 1$  which is the unit circle on the  $x - y$  plane. Then,  $z = 1$  yields  $x^2 + y^2 = 0$  which is the single point  $(0, 0, 1)$ . So the contour plot is only the origin.  $z = \frac{1}{2}$  gives a smaller circle centered at the origin with radius  $\frac{\sqrt{3}}{2}$ . This gives us a clearer idea of what this looks like: the top half of a sphere.

### Example 4

$$f(x, y) = x^2 - y^2, (x, y) \in \mathbb{R}^2$$

Traces:

$x = 0 \implies z = -y^2$ , an upside-down parabola.  $y = 0 \implies z = x^2$ , a regular parabola.

Contour plot:

$$z = 0 \implies x = \pm y$$

$$z = 1 \implies x^2 - y^2 = 1$$

$$z = -1 \implies x^2 - y^2 = -1$$

So it looks like a Pringle. If Pringles were infinite. I wish Pringles were infinite.

**Remark.** All graphs are surfaces, but not all surfaces are graphs of  $f(x, y)$ . This is because graphs have the property that  $\forall(x, y) \exists \leq 1(x, y, z)$  i.e. every vertical line has  $\leq 1$  point.

## 5.3 Three Variables

A graph of  $f(x, y, z)$  defined on a domain  $D \subset \mathbb{R}^3$  is  $S = \{(x, y, z, w) : w = f(x, y, z), (x, y, z) \in D\}$

We visualise this using level surfaces, obtained by fixing the output to some value.

**Example**  $w = f(x, y, z) = x^2 + y^2 + z^2$ .

Fix  $w = 0$ , then the level surface is the origin. For every value greater than 0, we get, in this case, a sphere. So the contour plot is essentially a series of nested spheres, representing a true graph in 4D space that we cannot draw. that's wild. or according to this guy next to me, that's disgusting.

## 5.4 Limits

### 5.4.1 Review: One variable

Recall the definition of the limit of a function of one variable:

$$\lim_{x \rightarrow a} f(x) = L$$

if, as you approach  $x = a$  on either side, the value of  $f(x)$  approaches  $L$ .

**Definition 11. Epsilon-delta definition:** For every  $\varepsilon > 0, \exists \delta > 0$  such that  $|f(x) - L| < \varepsilon$  whenever  $|x - a| < \delta$ .

**Sequence definition:** for every sequence approaching  $a$ ,  $f(x) \rightarrow L$  along the sequence  $\iff \lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = L$ .

**Definition 12.**  $f$  is continuous at  $a$  if  $\lim_{x \rightarrow a} f(x) = f(a)$ .

### 5.4.2 Two variables

**Definition 13.**

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = L$$

if for every path  $C$  approaching  $(a, b)$ ,  $f(x, y) \rightarrow L$  along  $C$ .

Formally, a path is a parameterised curve  $C(t)$  on  $[a, \infty)$  such that

$$\lim_{t \rightarrow \infty} \vec{c}(t) = (a, b)$$

The main difference between this and the one-variable version of a limit is that there are now infinitely many paths along which to approach the point. This means this definition is not useful for showing a limit exists. However, we can use it to show a limit does *not* exist.

**Example**

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

We want to find the limit as  $(x, y) \rightarrow (0, 0)$ . Set  $c_1$  to be along the  $x$ -axis from some point  $(x > 0, 0)$ .

$$\lim_{(x,0) \rightarrow (0,0)} f(x, y) = \lim_{x \rightarrow 0} \frac{x^2 - 0}{x^2 + 0} = 1$$

Now, define  $c_2$  along the  $y$ - axis from some point  $(0, y > 0)$ .

$$\lim_{(0,y) \rightarrow (0,0)} f(x, y) = \lim_{y \rightarrow 0} \frac{0 - y^2}{0 + y^2} = -1$$

Therefore the limit does not exist.

Math 53: Multivariable Calculus

Spring 2018

## Lecture 6: Partial derivatives, tangent planes, linear approximation

Lecturer: Nikhil Srivastava

1 February

Scribe: Aditya Sengupta

## 6.1 How to show a function is continuous at a point

1. Understand simple functions that are continuous, or verify that they are. e.g.  $f(x, y) = h(x)$  or  $g(y)$  for some continuous functions  $g, h$ . This means all continuous functions of one variable are still continuous here.
2. Use properties of limits to conclude that if  $f(x, y)$  and  $g(x, y)$  are continuous at  $(x_0, y_0)$  then

$$\begin{array}{c}
 f + g \\
 cf, c \in \mathbb{R} \\
 fg \\
 f/g \text{ for } (g(x_0, y_0) \neq 0) \\
 h(f(x, y)), \text{ for } h(t) \text{ continuous}
 \end{array}$$

are all also continuous.

**Example:**  $f(x, y) = \sin(xy)e^{y/x} + x$  is continuous because it is the composition of several continuous functions.

## 6.2 Differentiation

In one variable,  $\frac{df}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$  whenever the limit exists. It is called *differentiable* when the limit exists.

This has the geometric interpretation of the slope of the tangent line at a point. The tangent line has the equation

$$y - y_0 = (x - x_0)f'(x_0)$$

which we can use to linearly approximate the function:

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x$$

**Definition 14.**  $f$  is differentiable at  $x_0$  if

$$f(x_0 + \Delta x) = f(x_0) + f'(x_0)\Delta x + \varepsilon(\Delta x)\Delta x$$

where  $\varepsilon(\Delta x) \rightarrow 0$  as  $\Delta x \rightarrow 0$

## 6.3 Partial Derivatives

We can use these concepts to define differentiation in two variables. Let  $f(x, y)$  be a function of two variables. Then

**Definition 15.**

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x} = \left. \frac{d}{dx} f(x, y_0) \right|_{x=x_0}$$

Graphically, this appears as the limit as we vary  $x$  for a fixed value of  $y$  (just like the trace of a 3D plot in Lecture 5).

We can similarly define differentiation with respect to  $y$ :

**Definition 16.**

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{\Delta y \rightarrow 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y} = \left. \frac{d}{dy} f(x_0, y) \right|_{y=y_0}$$

This definition, using partial derivatives, is not perfect, as we do not have an analogous idea for a linear approximation. Nonetheless it is useful because it is easy to compute - to take the derivative with respect to one of the variables, we simply treat the other as a constant and differentiate.

**Example:**

$$f(x, y) = xe^{x+y}$$

$$\frac{\partial f}{\partial x} = e^y \frac{d}{dx}(xe^x) = e^y(xe^x + e^x)$$

$$\frac{\partial f}{\partial y} = \frac{d}{dy}(e^y xe^x) = xe^x e^y$$

We can differentiate this again, but who would want to? Lots of people, not including the prof. This gives rise to the idea of a second partial derivative,

$$\frac{\partial}{\partial x} \frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2}$$

and similarly with  $y$ .

More interestingly, we can differentiate the function with respect to both, one at a time. With the above example,

$$\frac{\partial}{\partial y} \frac{\partial f}{\partial x} = f_{xy} = e^y(xe^x + e^x)$$

$$\frac{\partial}{\partial x} \frac{\partial f}{\partial y} = f_{yx} = e^y \frac{d}{dx}(xe^x) = e^y(xe^x + e^x)$$

We see that the partial derivatives are independent of the order in which they are applied. This is Clairaut's theorem, that if  $f_{xy}$  and  $f_{yx}$  are continuous then they are equal.



### 6.3.1 Tangent Plane to a Two-Variable Function

**Definition 17.** If  $f(x, y)$  is differentiable, the tangent plane to the graph of  $f$  at  $(x_0, y_0)$  is the plane containing the two tangent lines defined by

$$\frac{\partial f}{\partial x}(x_0, y_0)$$

and

$$\frac{\partial f}{\partial y}(x_0, y_0)$$

(here we denote a partial derivative wrt  $x$  as  $f_x$  and the same with  $y$ )

Call  $T_1$  the first tangent line, going through the point  $\vec{r}_0 = \langle x_0, y_0, f(x_0, y_0) \rangle$  and with direction  $\vec{v}_1 = \langle 1, 0, \frac{\partial}{\partial x} f(x_0, y_0) \rangle$ .

Similarly,  $T_2$  has  $\vec{r}_0 = \langle x_0, y_0, f(x_0, y_0) \rangle$  and  $\vec{v}_2 = \langle 0, 1, \frac{\partial}{\partial y} f(x_0, y_0) \rangle$ .

We can define a plane with these two tangent vectors, which has the equation

$$\vec{r}(t) = \vec{r}_0 + t_1 \vec{v}_1 + t_2 \vec{v}_2$$

$$\langle x, y, z \rangle = \langle x_0, y_0, f(x_0, y_0) \rangle + t_1 \langle 1, 0, f_x(x_0, y_0) \rangle + t_2 \langle 0, 1, f_y(x_0, y_0) \rangle$$

Eliminating parameters,

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

### 6.3.2 Linear Approximation

**Definition 18.**  $f(x, y)$  is differentiable at  $(x_0, y_0)$  if

$$f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + \varepsilon_1(\Delta x)\Delta x + \varepsilon_2(\Delta y)\Delta y$$

for  $\varepsilon_1, \varepsilon_2 \rightarrow 0$  as  $\Delta x, \Delta y \rightarrow 0$ .

We now have a full picture of differentiation in two-variable functions.

**Theorem 19.** If  $f_x$  and  $f_y$  exist and are continuous at  $(x_0, y_0)$  then  $f$  is differentiable at  $(x_0, y_0)$ .

### 6.3.3 Example of an application

Find  $\sqrt{(3.01)^2 + (3.98)^2}$ .

We do this by defining a function  $f(x, y) = \sqrt{x^2 + y^2}$ . Given that it is differentiable at  $(3, 4)$  (which can be easily verified) we apply the linear approximation,

$$f(3.01, 3.98) \approx f(3, 4) + (0.01)f_x(3, 4) + (-0.02)f_y(3, 4)$$

We evaluate the partial derivatives,

$$f_x = \frac{2x}{2\sqrt{x^2+y^2}}, f_y = \frac{2y}{\sqrt{x^2+y^2}}$$

which at the given point evaluate to

$$f_x(3, 4) = \frac{3}{5}, f_y = \frac{4}{5}$$

Then we plug these values into the above equation to get

$$f(3.01, 3.98) \approx 4.99$$

For comparison, the actual value  $\approx 4.99004\dots$

## 6.4 Total Derivative

The total derivative of a function  $f(x, y)$  is

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

which is a shorthand notation for the linear approximation. In the total differential, the base point  $(x_0, y_0, z_0)$  is implicit. It only expresses the relationship between rates of change.

## 6.5 Three Variables

We can extend this idea of a partial derivative to three dimensions,

$$\frac{\partial f(x_0, y_0, z_0)}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x, y_0, z_0) - f(x_0, y_0, z_0)}{\Delta x} = \frac{d}{dx} f(x, y_0, z_0)$$

and the linear approximation still holds,

$$f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) \approx f(x_0, y_0, z_0) + f_x \Delta x + f_y \Delta y + f_z \Delta z$$

## 6.6 Timely Warning from UCPD

Existence of  $f_x, f_y$  does not imply differentiability by itself. The partial derivatives must also be continuous.

**Math 53: Multivariable Calculus**

**Spring 2018**

## Lecture 7: The Chain Rule

*Lecturer: Nikhil Srivastava*

*6 February*

*Scribe: Aditya Sengupta*

there's a staaaaaarman waiting in the sky

### 7.1 Tangent Plane

The tangent plane to the graph of  $z = f(x, y)$  at  $(x_0, y_0, z_0)$  is given by

$$(z - z_0) = f_x \cdot (x - x_0) + f_y \cdot (y - y_0) \quad (7.1)$$

The tangent plane is the graph of the linear approximation. The partial derivatives are the coefficients of the linear approximation.

### 7.2 $d$ vs. $\Delta$

The book says things like  $dx = \Delta x$ . This is garbage. I have it on good authority.

The way we interpret the difference between them is that  $\frac{dy}{dx}$  is the slope of the line, which means that for every change  $\Delta x$  there is a corresponding change  $\Delta y = \frac{dy}{dx} \Delta x$ .  $\Delta x, \Delta y, \Delta z$  reduce to  $dx, dy, dz$  in the infinitesimally small limit.

### 7.3 Chain Rule

#### 7.3.1 Single Variable Chain Rule

Recall the single variable chain rule, in which  $f$  is a differentiable function of  $x$  and  $x$  is a differentiable function of  $t$ . Then we can view  $f(t) = f(x(t))$ , and its derivative with respect to  $t$  becomes

$$\frac{df}{dt} = \frac{df}{dx} \cdot \frac{dx}{dt} \quad (7.2)$$

or

$$f'(t) = f'(x(t)) \cdot x'(t) \quad (7.3)$$

The "answer" for why it works could be that we cancel the  $dx$ 's, but this is not rigorous. The actual answer can be found with the linear approximation.

*Proof.*

$$x(t_0 + \Delta t) \approx x(t_0) + \Delta t x'(t_0) \quad (7.4)$$

and similarly

$$f(x_0 + \Delta x) \approx f(x_0) + \Delta x f'(x_0) \quad (7.5)$$

We can compose these linear approximations to get

$$f(x_0 + \Delta x) \approx f(x_0) + \Delta t x'(t_0) f'(x_0) \quad (7.6)$$

Therefore, when we try to compute  $f(x(t_0 + \Delta t))$ , we get

$$f(x(t_0 + \Delta t)) \approx f(x_0) + \Delta t \frac{d}{dt} f(x(t)) \quad (7.7)$$

□

The shorthand proof using differentials is shown below:

*Proof.*

$$dx = x'(t_0) dt \quad (7.8)$$

$$df = f'(x_0) dx \quad (7.9)$$

$$\therefore df = f'(x_0) x'(t_0) dt \quad (7.10)$$

□

### 7.3.2 ~~A New~~ Multivariate Chain Rule

Let  $x, y$ , be functions of some other parameter  $t$ , and let  $f(x, y)$  be a differentiable function. Then

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

What is  $\frac{df}{dt}$ ? We can attempt to answer this as follows:

$$df = \frac{\partial}{\partial x} f dx + \frac{\partial}{\partial y} f dy \quad (7.11)$$

”Divide by  $dt$ ”, then:

$$\frac{df}{dt} = \frac{\partial}{\partial x} f \frac{dx}{dt} + \frac{\partial}{\partial y} f \frac{dy}{dt} \quad (7.12)$$

This is correct, but we do not have a picture of what it means to divide by  $dt$ . We can get an answer with the linear approximation.

For  $(x_0, y_0) = (x(t_0), y(t_0))$ ,

$$\Delta f \approx \frac{\partial}{\partial x} f \Delta x + \frac{\partial}{\partial y} f \Delta y \approx \frac{\partial f}{\partial x} (x'(t_0) \Delta t) + \frac{\partial f}{\partial y} (y'(t_0) \Delta t) \quad (7.13)$$

The above is the linear approximation of  $f$  in terms of partial derivatives, derivatives of the components, and  $\Delta t$ . We can take a limit as  $\Delta t \rightarrow 0$ , and set the resultant  $\frac{df}{dt}$  equal to the coefficients in the linear approximation above. Then we get

$$\frac{df}{dt} = \frac{\partial}{\partial x} f \cdot \frac{dx}{dt} + \frac{\partial}{\partial y} f \cdot \frac{dy}{dt} \quad (7.14)$$

Intuitively, for 2D motion, the rate of ascent is equal to the sum of the products of slope and velocity in the  $x$  and  $y$  directions.

### 7.3.3 Example 1

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

Consider  $x(t) = y(t) = t$ . Then the rate of change of  $z$  with respect to time is

$$\frac{dz}{dt} = \frac{\partial}{\partial x} z \cdot \frac{dx}{dt} + \frac{\partial}{\partial y} z \cdot \frac{dy}{dt} = (-2x) \cdot 1 + (2y) \cdot 1 = 2y - 2x = 2t - 2t = 0 \quad (7.15)$$

i.e. by walking along the lines  $y = \pm x$ , the  $z$  does not change at all. This is consistent with the contour plot. We plot the contour for  $z = 0$  and notice that the crossed lines  $y = x$  and  $y = -x$  appear.

### 7.3.4 Example 2

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

Consider  $x(t) = \cos t, y(t) = \sin t$ . Then the rate of change of  $z$  with respect to time is

$$\frac{dz}{dt} = \frac{\partial}{\partial x} z \frac{dx}{dt} + \frac{\partial}{\partial y} z \frac{dy}{dt} = (-2x) \sin t + (2y) \cos t = 2 \sin t \cos t + 2 \sin t \cos t = 2 \sin 2t \quad (7.16)$$

## 7.4 The Chain Rule Strikes Back

Let  $x$  and  $y$  depend on two variables  $s$  and  $t$ . Then, if we want to find  $\frac{\partial}{\partial s} f$  and  $\frac{\partial}{\partial t} f$ , we use linear approximations again, using total differentials.

$$df = \frac{\partial}{\partial x} f dx + \frac{\partial}{\partial y} f dy \quad (7.17)$$

We also know that

$$dx = \frac{\partial}{\partial s} x ds + \frac{\partial}{\partial t} x dt \quad (7.18)$$

and

$$dy = \frac{\partial}{\partial s}y ds + \frac{\partial}{\partial t}y dt \quad (7.19)$$

Therefore we can compose these approximations to get

$$df = \frac{\partial}{\partial x}f \left( \frac{\partial}{\partial s}x ds + \frac{\partial}{\partial t}x dt \right) + \frac{\partial}{\partial y}f \left( \frac{\partial}{\partial s}y ds + \frac{\partial}{\partial t}y dt \right) \quad (7.20)$$

Then we expand, and take terms in  $ds$  and  $dt$  common:

$$df = \left( \frac{\partial}{\partial x}f \frac{\partial}{\partial s}x + \frac{\partial}{\partial y}f \frac{\partial}{\partial s}y \right) ds + \left( \frac{\partial}{\partial x}f \frac{\partial}{\partial t}x + \frac{\partial}{\partial y}f \frac{\partial}{\partial t}y \right) dt \quad (7.21)$$

Therefore

$$df = \frac{\partial}{\partial s}f ds + \frac{\partial}{\partial t}f dt \quad (7.22)$$

### Example 1

Let  $f(x, y) = 3x + 4y + 10$ ,  $x(s, t) = s + t$ , and  $y(s, t) = s - t$ .

We want to compute  $\frac{\partial}{\partial s}f$ :

$$\frac{\partial}{\partial s}f = \frac{\partial}{\partial x}f \frac{\partial}{\partial s}x + \frac{\partial}{\partial y}f \frac{\partial}{\partial s}y = (3)(1) + (4)(1) = 3 + 4 = 7 \quad (7.23)$$

and similarly,

$$\frac{\partial}{\partial t}f = f_x x_t + f_y y_t = (3)(1) + (4)(-1) = 3 - 4 = -1 \quad (7.24)$$

which gives us the total differential,

$$df = 7ds - dt \quad (7.25)$$

We should get the same answer if we just use the linear approximation,

$$f = 3x + 4y + 10 = 3(s + t) + 4(s - t) + 10 = 7s - t + 10$$

Therefore with differentials the constant drops out and we get  $df = 7ds - dt$  as above.

### Example 2

Let  $f(x, y) = e^x \sin y$ ,  $x(s, t) = s^2t$ ,  $y(s, t) = st^2$ .

Then we use the chain rule,

$$f_s = f_x x_s + f_y y_s = e^x \sin y \cdot 2st + e^x \cos y t^2 = 2ste^{s^2t} \sin(st^2) + t^2 e^{s^2t} \cos(st^2) \quad (7.26)$$

## 7.5 Implicit Differentiation

We may in some cases be given  $z$  implicitly by an equation  $F(x, y, z) = 0$ , e.g.  $x^2 + y^2 + z^2 = 3$  in the neighbourhood of  $(1, 1, 1)$ .

**Math 53: Multivariable Calculus**

**Spring 2018**

**Lecture 8: Gradient and Directional Derivatives**

*Lecturer: Nikh l Srivastava*

*February 8*

*Scribe: Aditya Sengupta*

## 8.1 Chain Rules

We previously discussed the two forms of the multivariate chain rule,

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

when  $x$  and  $y$  are dependent on only one variable  $t$ , and

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial s}$$

when  $x$  and  $y$  depend on  $s$  and  $t$ .

It is not always clear which of these forms to use. For example, let  $z$  be a function of  $x$  and  $y$ , and let  $h(z)$  be differentiable. To find  $\frac{\partial h}{\partial x}$ , which do we use?

$$\frac{dh}{dz} \frac{\partial z}{\partial x}, \text{ or } \frac{dh}{dz} \frac{\partial z}{\partial x} + \frac{dh}{dz} \frac{\partial z}{\partial y}?$$

We can clearly see that the first expression seems more likely to be correct, because there is no reason for an  $x$ -partial derivative to be dependent on  $y$ . To explicitly see this, we can do the following:

$$\frac{\partial}{\partial x} h(z(x, y)) = \frac{dh}{dz} \cdot \frac{d}{dx} z(x, y) = \frac{dh}{dz} \frac{\partial z}{\partial x}$$

It is possible to more rigorously prove this using linear approximations:

*Proof.*

$$\begin{aligned} h(z(x_0 + \Delta x, y_0)) &\approx h\left(z(x_0, y_0) + \Delta x \frac{\partial z}{\partial x}(x_0, y_0)\right) \\ &\approx h(z(x_0, y_0)) + \frac{dh}{dz}\left(z(x_0, y_0)\right) \frac{\partial z}{\partial x}(x_0, y_0) \Delta x = h(z(x_0, y_0)) + \frac{\partial h}{\partial x} \Delta x \end{aligned}$$

□



## 8.2 General Chain Rule

For a function  $f(x_1, \dots, x_n)$ , in which each  $x_i = x_i(t_1, \dots, t_m)$ , the following holds for every  $j$ :

$$\frac{\partial f}{\partial t_j} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t_j} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial t_j}$$

## 8.3 Implicit Partial Differentiation

In cases where we are not given  $z(x, y)$  explicitly, but rather are given some function  $F(x, y, z) = 0$ , it is possible to implicitly define  $z = z(x, y)$  and differentiate. For example, let

$$F(x, y, z) = x^3 + y^3 + z^3 + 6xyz = 1$$

This is a level surface of the function  $F$ .

Math 53: Multivariable Calculus

Spring 2018

## Lecture 9: Max/Min, Lagrange Multipliers

Lecturer: Nikhil Srivastava

February 13

Scribe: Aditya Sengupta

## 9.1 Maxima and minima

We are interested in finding the maximum or minimum of a function  $f$  on a domain  $D$ .

This has many applications:

- Find the point on a plane closest to the origin
- (Regression) find the line that best fits a given set of data points
- Find the fastest route from one point to another
- Machine learning

Recall how we did this in one variable:

*Given a differentiable  $f(x)$  on a domain, interval, find a local or global maximum or minimum of  $f(x)$ .*

Local means to find some  $a$  such that  $f(a) \geq f(x)$  for  $x \in$  neighbourhood of  $a$ , and global means to find  $f(a) \geq f(x), x \in D$ .

I could draw what that means, but I'm a bit lazy and haven't entirely figured out tikzgraphics that aren't compositions of lines and circles yet. Eventually I'll go back through these notes and get all this stuff off his online notes. Same for 54.

We observe that if  $a$  is a local extremum,  $f'(a) = 0$ , i.e. the tangent line is horizontal.

*Proof.*

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

If  $f'(a) \neq 0$ , can increase/decrease. □

**Remark.** *Local extrema in the interior are critical points, but  $f'(a) = 0$  does not imply a local minima or maxima, for example  $f(x) = x^2$  has a critical point at  $x = 0$  but is not a min or max.*

We can figure out whether a critical point is a min or max by the second derivative test.

If  $f'(a) = 0$ , then:

- $f''(a) > 0 \implies$  local min
- $f''(a) < 0 \implies$  local max
- $f''(a) = 0 \implies$  inconclusive

This gives us a procedure for finding the max or min:

1. Find critical points (maybe use the second derivative test to refine this).
2. Evaluate the function on boundary points.

The global max or min is the largest or smallest of the points above, provided they exist.

**Definition 20. Extreme Value Theorem:** *if  $f$  is continuous, and  $D$  is closed and bounded, then  $f$  achieves its global max/min on  $D$ .*

## 9.2 Generalising 1D optimisation to two variables

Given a scalar function of two variables  $f(x, y)$  that is differentiable on some domain  $D \subseteq \mathbb{R}^2$ , we want to find

*local minima*  $f(a, b) \leq f(x, y) \forall x, y \in \text{disc of radius } \delta > 0 \text{ around } (a, b)$ , or

*global minima*  $f(a, b) \leq f(x, y) \forall x, y \in D$

or the equivalents for maxima.

The gradient will help us to do this, as it points in the direction of steepest ascent. We observe that if  $\nabla f(a, b) \neq 0$  and  $(a, b)$  is in the interior of the domain  $D$  then  $(a, b)$  cannot be a local extremum.

*Proof.*

$$f(a + \Delta x, b + \Delta y) \approx f(a, b) + \nabla f \cdot \langle \Delta x, \Delta y \rangle$$

So unless  $\nabla f = 0$ , we can increase or decrease  $f$ . □

Put another way, if  $\nabla f(a, b) = 0$ , then  $(a, b)$  is a critical point.

In terms of partial derivatives,  $\nabla f(a, b) = \langle f_x(a, b), f_y(a, b) \rangle = 0$  at a critical point so both partials are zero. If  $(a, b) \in \text{interior}(D)$  is a local extremum then  $f_x = f_y = 0$ .

This implies that the extrema of  $f(x, y)$  on  $D$  are either critical points or boundary points.

### Example 1

$$f(x, y) = x^2 + y^2 - 2x + 4y + 8, D = \mathbb{R}^2$$

We take the partial derivatives,

$$f_x = 2x - 2, f_y = 2y + 4$$

Therefore the gradient is

$$\nabla f(x, y) = \langle 2x - 2, 2y + 4 \rangle$$

For this to be a critical point, both the partials must be zero, therefore we get  $x = 1, y = -2$ .

To find out whether this is a max or a min, we complete the square:

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

. It is clear that at  $(1, -2)$  the value of this function is 3, and it is greater at every other point. Therefore it is a global minimum.

## Example 2

$$f(x, y) = x^2 y^2 e^{-x^2 - y^2} \text{ on } \mathbb{R}^2$$

We take the partial derivatives,

$$f_x = 2xye^{-x^2 - y^2}(1 - x^2)$$

$$f_y = 2yx^2e^{-x^2 - y^2}(1 - y^2)$$

We observe that if  $x = 0$  or  $y = 0$ , or if  $x, y = \pm 1$ , the partial derivatives are zero. The points  $f(x, 0) = 0$  and  $f(0, y) = 0$  are global minima, because the function is zero for these values but it is positive otherwise. It turns out that  $x, y = \pm 1, \pm 1$  are local maxima.

## 9.3 Second Derivative Test

Suppose  $\nabla f(a, b) = 0$ ,  $D = f_{xx}f_{yy} - (f_{xy})^2$ .

If  $D > 0$  and  $f_{xx} > 0$ , then  $(a, b)$  is a local minima. If  $D > 0$  and  $f_{xx} < 0$ , then  $(a, b)$  is a local maxima. If  $D < 0$ , then  $(a, b)$  is a saddle point. If  $D = 0$ , this is inconclusive.

But what is  $D$  conceptually? Let's look at some examples.

### Examples

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

In this case,  $D$  evaluates to 4 ( $\frac{\partial^2 f}{\partial x^2}$ ) which is greater than 0.

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

Similarly,  $D = (-2)^2 = 4$ .

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

In this case  $D = -4 < 0$ .

The reason why this has anything to do with the second derivative test statement above is in the second order Taylor approximation of  $f(x, y)$ .

$$f(x, y) \approx f(a, b) + \Delta x f_x + \Delta y f_y + \frac{\Delta x^2}{2} f_{xx} + \frac{\Delta y^2}{2} f_{yy} + \Delta x \Delta y f_{xy}$$

## 9.4 How do we know whether there are global maxima or minima?

We can apply the extreme value theorem in two dimensions to answer this.

**Theorem 21.** *If  $f(x, y)$  is continuous and  $D$  is closed and bounded, then  $f$  achieves its global max/min on  $D$ .*

$D$  is closed if  $D$  contains all of its limit points.

## 9.5 Algorithm to find extrema

1. Find critical points - perhaps using second derivative test
2. Check values on the boundary.

That isn't very clear.

So let's try and find the max of  $x^2 + 3xy - y$  on  $D = \{-1 \leq x \leq 1, -1 \leq y \leq 1\}$ .

We start by finding partial derivatives and setting them equal to zero. This yields  $x = \frac{1}{3}$  and  $y = -\frac{2}{9}$ . This gives us a potential local max or min. Finding the value at the point yields

$$f\left(\frac{1}{3}, -\frac{2}{9}\right) = \frac{1}{9}$$

To check the values on each of the boundaries, we look at each line segment in turn:

$$L_1 : x \in [-1, 1], y = -1$$

. The function in this case becomes  $f_{L_1}(x) = x^2 - 3x + 1$ , for which we can find the extrema using one-dimensional methods. Repeat this for  $L_2, L_3, L_4$  to get maxima  $m_1, \dots, m_4$ .

Then the max is the max of  $\frac{1}{9}, m_1, \dots, m_4$ . We can compute corners first plus the four critical points.

Math 53: Multivariable Calculus

Spring 2018

## Lecture 10: Lagrange Multipliers and PDEs

Lecturer: Nikhil Srivastava

15 February

Scribe: Aditya Sengupta

## 10.1 Finding extrema when boundaries are not line segments

What if  $D = \{x^2 + y^2 \leq 1\}$ ?

In this example, we could use the fact that  $x = \pm\sqrt{1-y^2}$  on the boundary, and substitute it into  $f(x, y)$ , but this won't work in general.

### 10.1.1 Constrained Optimisation

We want to maximise or minimise an *objective function*  $f(x, y)$  subject to the *constraint equation*  $g(x, y) = k$ . Note that  $x$  and  $y$  are no longer independent.

#### Examples

1. Minimise  $f(x, y) = x^2 + y^2$  subject to the constraint  $g(x, y) = xy = 2$ .
2. We could ask a similar question in three dimensions: minimise  $f(x, y, z) = 3x^2 + 4x^2y^2 + 5z^2$  subject to  $x + y + z = 1$ .
3. More practical examples would be to find the maximum volume of a box made with  $100 \text{ m}^2$  of cardboard. We see there is an objective function as well as a constraint here.

In the first example, we observe that if level curves of  $f$  and  $\{g = k\}$  intersect non-tangentially, we can move a bit in the same direction and reach a lower level curve of  $f$ .

**Theorem 22.** *If  $x_0, y_0$  is an optimum of max/min  $f(x, y)$  subject to  $g(x, y) = k$ , then  $\nabla f \parallel \nabla g$  at  $(x_0, y_0)$ , therefore*

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0) \quad (10.27)$$

where  $\lambda$  is the Lagrange multiplier.

*Proof.* Let  $g = k$  be a level curve of a function and let  $(x_0, y_0)$  be on that curve. Let  $\vec{r}(t)$  be a parameterised curve along  $\{g = k\}$  where  $\vec{r}(0) = (x_0, y_0)$ . Then  $f(\vec{r}(t))$  has a time derivative of zero, i.e.  $\nabla f(x_0, y_0) \cdot \vec{r}'(0) = 0$ . This means the gradient of  $f$  is perpendicular to the velocity at 0. As it is a level curve, the same argument applies to  $g$  (as  $g = k$  everywhere) and so  $\nabla g(x_0, y_0) \cdot \vec{r}'(0) = 0$ . Therefore the gradients of  $f$  and  $g$  are parallel.  $\square$

So let's use this to solve the first example. We know  $f_x = 2x$ ,  $f_y = 2y$ ,  $g_x = y$ , and  $g_y = x$ . Because  $\nabla f = \lambda \nabla g$ , we know that

$$2x = \lambda y, 2y = \lambda x, xy = 2$$

where we get the third equation from the constraint. So we can solve for  $x, y, \lambda$  to get

$$x = \pm\sqrt{2}, y = \pm\sqrt{2}, \lambda = \pm 2$$

## 10.2 who needs the real world anyway

**What is the maximum volume of an open top box with surface area 12  $m^2$ ?**

We want to maximise  $f(x, y, z) = xyz$  subject to  $g(x, y, z) = xy + 2yz + 2zx$ .

We take partial derivatives,

$$f_x = \lambda g_x \implies yz = \lambda(y + 2z)$$

$$f_y = \lambda g_y \implies xz = \lambda(x + 2z)$$

$$f_z = \lambda g_z \implies xy = \lambda(2y + 2x)$$

and  $xy + 2yz + 2zx = 12$ .

Solving this yields  $z = 0, \lambda = 0$ , or  $x = y$ . We reject the first two so that  $x = y$ . We also get  $x = 2z$  by a similar method.

Therefore we plug this into  $g = 12$  to get  $(x, y, z) = (2, 2, 1)$ .

**Remark.** *These equations only generate candidates for external points. They do not even assure that an optimum has been achieved. We need to ensure that the global max exists by other reasoning.*

## 10.3 What does $\lambda$ mean?

Let  $M(k)$  be the max value with constant  $g(x, y, z) = k$ , and let  $\vec{r}(k) = \langle x(k), y(k), z(k) \rangle$  be a parameterised curve such that  $f(\vec{r}(k)) = M(k)$ .

We claim that  $\frac{dM(k)}{dk} = \lambda$ , which gives us information about the variation in the solution to an optimisation problem upon varying one of the inputs.

*Proof.*

$$\frac{dM(k)}{dk} = \frac{d}{dk} f(\vec{r}(k)) = \nabla f \cdot \vec{r}'(k) = \lambda \nabla g \cdot \vec{r}'(k) = \lambda \frac{d}{dk} g(\vec{r}(k)) = \lambda \frac{dk}{dk} = \lambda$$

□

## 10.4 Partial Differential Equations

A partial differential equation relates an unknown function of two or more variables and its partial derivatives.

**examples:**  $u(x, y), \frac{\partial u}{\partial x} = 0, \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0, (\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2$ , etc etc etc.

All the laws of classical physics are partial differential equations. For example, the heat equation from the 1800s. We mathematically define heating and cooling as a function  $U(x, t), x \in \mathbb{R}$ .

In summary, the heat equation is simple:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$



Math 53: Multivariable Calculus

Spring 2018

## Lecture 11: Review

Lecturer: Nikh l Srivastava

February 20

Scribe: Aditya Sengupta

## 11.1 Finishing PDEs

The heat equation is still simple:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

Suppose we have a rod of infinite length, and we unevenly heat it up. At time  $t = 0$ , the distribution of heat  $u(x, 0)$  is uneven, and it evens out over time until the whole rod is at thermal equilibrium at some time.

We want to know the general expression for  $u(x, t)$  given an initial condition, e.g.  $u(x, 0)$ . Suppose the initial temperature for one rod is given by  $\sin x$ , and that for a second is  $\sin 10x$ . Intuitively, the second one will cool faster as the discontinuities in temperature happen over smaller  $x$  distances.

We can formalise this by thinking of the rod as many tiny elements of width  $\Delta x$  and constant temperature across that small width. A physical law tells us that the rate of heat flow is directly proportional to the slope of temperature across a point.

The heat in each element is equal to the temperature multiplied by  $\Delta x$ . Let the heat flowing in to an element at distance  $x$  and time  $t$  be  $U_x(x, t)$  and let that flowing out be  $-U_x(x + \Delta x, t)$  (the negative sign denotes the outflow). Then, the net heat change in time  $\Delta t$  is  $\beta \Delta t (U_x(x, t) - U_x(x + \Delta x, t))$ . We can compare this with the definition of heat above, which yields the equation

$$\frac{U(x, t + \Delta t) - U(x, t)}{\Delta t} = \frac{-\beta U_x(x + \Delta x, t) - U_x(x, t)}{\Delta x}$$

In the limit, this becomes

$$\frac{\partial u}{\partial t} = -\beta \frac{\partial^2 u}{\partial x^2}$$

Some solutions to this are any constant, or  $e^{-t} \sin x$ . These can be easily verified by computing the partial derivatives. So this solution models exactly what happens to a rod heated according to  $\sin x$ : the heat decays exponentially.

In general,  $e^{-k^2 t} \sin(kx)$  is a solution to this PDE. So  $\sin(10x)$  decays 100 times faster.

Another interesting solution to this is the formula for the bell curve,  $u(x, t) = \frac{1}{\sqrt{t}} e^{-\frac{x^2}{4t}}$ .

### 11.1.1 What does it mean to satisfy the heat equation anyway?

If you plug in a function that is a solution to the heat equation, you get the differential form of the heat law as stated above.

## 11.2 Wave Equation

Suppose you have an infinite string, and you displace the string according to a function  $g(x, t)$ . What happens when the string is released?

The answer is given by a thought experiment. Imagine that a string is a series of small springs, on which Hooke's law applies. If we follow a similar line of reasoning as with the heat equation, we get the PDE

$$\frac{\partial^2 g}{\partial t^2} = c^2 \frac{\partial^2 g}{\partial x^2}$$

We observe that  $g(x - ct)$  is always a solution to this PDE, which is easily verified via the chain rule. So easy, in fact, that I'm not going to do it here. This margin is too narrow to contain it.

## 11.3 Review

There's 50 minutes left in this lecture and my battery is on 14% (I have no idea how, I had like 60% two hours ago). And it's not like there's anything too new here. So I might update this section with this material later. But for now this is all.

Math 53: Multivariable Calculus

Spring 2018

## Lecture 12: Integration

Lecturer: Nikhil Srivastava

27 February

Scribe: Aditya Sengupta

## 12.1 with one variable

Let  $f(x) \geq 0$  be a continuous function on the interval  $a \leq x \leq b$ . Plotting a graph of this function gives us a one-dimensional curve in the two-dimensional plane, and it defines the region between the graph of the function and the interval  $[a, b]$ . The area of this region is represented by the definite integral,

$$\int_a^b f(x) dx$$

This leads to the concept of a Riemann sum, in which we approximate the area by  $n$  rectangles of width  $\Delta x$ , where  $x_{i+1} - x_i = \Delta x$ . If  $f$  is a nice function, the height won't change much in each subinterval.

Therefore

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{0 \leq i \leq n} Area_i = \lim_{n \rightarrow \infty} \sum_i f(x_i^*) \Delta x$$

This lead to ideas about sums and scalar multiples of integrals, the Fundamental Theorem of Calculus, etc.

## 12.2 with two variables

Let  $f(x, y) \geq 0$  be continuous and defined on a rectangular region:

$$R = a \leq x \leq b, c \leq y \leq d$$

Analogous to the idea of an *area under the curve*, we can compute a *volume under the surface*.

Subdivide the region into a grid, consisting of a set of small rectangles.

$$a < x_1 < x_2 < \cdots < x_n < b$$

$$c < y_1 < y_2 < \cdots < y_m < d$$

Then, we define the *double integral* over this surface:

$$\iint_R f(x, y) dA = \lim_{m \rightarrow \infty, n \rightarrow \infty} \sum_{i \leq n, j \leq m} vol_{ij} = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \sum_{i \leq n} \sum_{j \leq m} f(x_i^*, y_j^*) \Delta x \Delta y$$

## 12.3 Anonymous Method

Maybe it's an *iterated integral*.

$$\begin{aligned}\iint_R f(x, y) dA &= \lim_{n \rightarrow \infty} \sum_{i \leq n} \lim_{m \rightarrow \infty} \sum_{j \leq m} f(x_i^*, y_j^*) \Delta y \Delta x \\ &= \lim_{n \rightarrow \infty} \sum_{i \leq n} \int_c^d f(x_i^*, y) dy \Delta x\end{aligned}$$

This just simplifies to

$$\int_a^b \int_c^d f(x, y) dy dx$$

## 12.4 Example

What is the volume of the 3D region bounded by the surface  $z = x^2 + 2y^2 + 3$ , and the planes  $x = 1$ ,  $y = 3$ , and the planes  $x = 0$ ,  $y = 0$ ?

We draw traces and stuff like that

but the point is just to do the integrals:

$$\begin{aligned}\int_0^3 \int_0^1 (x^2 + 2y^2 + 3) dx dy \\ = \int_0^3 \left[ \frac{1}{3} + 2y^2 + 3y \right] dy = 28\end{aligned}$$

(just trust me (read: trust the prof.))

You can do it the other way too:

$$\begin{aligned}\int_0^1 \int_0^3 x^2 + 2y^2 + 3 dy dx \\ = \int_0^1 (3x^2 + 27) dx = 28\end{aligned}$$

## 12.5 Warning

Often, the integral one way is easy and the other way is impossible, even though

$$\iint f(x, y) dy dx = \iint f(x, y) dx dy$$

## 12.6 Integrating over any subset of $\mathbb{R}$

First, define some rectangle containing the subset of  $\mathbb{R}$  over which we want to integrate. Then the integral over the subset is equivalent to the integral over the rectangle. The proof is in Math 104.

You may think I'm distracted today based on how little detail I'm writing, and you'd be right because I'm actually trying to do my CS project in parallel with this. Sorry.

### Example

$$\iint_D$$

ok you know what i'm just doing this later

Math 53: Multivariable Calculus

Spring 2018

## Lecture 14: Polar Coordinates

Lecturer: Nikhil Srivastava

1 March

Scribe: Aditya Sengupta

## 14.1 Volume of a sphere

We know that the volume of a sphere of radius  $R$  is  $V = \frac{4}{3}\pi R^3$ . We can derive this from the formula of a sphere, which is

$$\{(x, y, z) : x^2 + y^2 + z^2 = R^2\}$$

Therefore the surface is twice the area of the graph of  $f(x, y) = \sqrt{R^2 - x^2 - y^2}$  over the region  $D = \{x^2 + y^2 \leq R^2\}$ . Therefore we can set up a double integral,

$$\frac{V}{2} = \iint f(x, y) dA$$

which is

$$\frac{V}{2} = \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \sqrt{R^2 - x^2 - y^2} dy dx$$

This isn't fun. Math should be fun. So to make math fun we use polar coordinates. Yes. That is a true statement. Everyone loves polar coordinates.

## 14.2 Everyone loves polar coordinates

Polar coordinates characterise a point by the distance from the origin and either one or two angles, depending on whether you're in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ . In  $\mathbb{R}^2$ , we can define a point by  $(r, \theta)$  where  $x = r \cos \theta$  and  $y = r \sin \theta$ .

Integrating in this coordinate system is a little less simple, but still reasonable. We can subdivide a region into small arcs subtending angle  $d\theta$ . Then we can invoke magic to get the trivial result

$$\iint_D f(r, \theta) dA = \int_{\alpha}^{\beta} \int_a^b f(r, \theta) r dr d\theta$$

## 14.3 Volume of a sphere in the coordinates everyone loves

The integral we were looking at before in polar coordinates is

$$\int_0^{2\pi} \int_0^R \sqrt{R^2 - r^2} r dr d\theta$$

which we can evaluate to get

$$\frac{V}{2} = \frac{2\pi}{3} R^3$$

## 14.4 More examples

ide

Lecture 15: Change of variables in double integrals

*Lecturer: Nikhil Srivastava*

*6 March*

*Scribe: Aditya Sengupta*

Blank for now. (Wasn't here.)



**Math 53: Multivariable Calculus**

**Spring 2018**

## Lecture 16: Triple integrals

*Lecturer: Nikh l Srivastava*

*8 March*

*Scribe: Aditya Sengupta*

### 16.1 Change of variables, continued

In one dimension, we are familiar with the concept of a variable change. Let  $x = g(u)$ , then the integral is

$$\int_a^b f(x)dx = \int_c^d f(g(u))g'(u)du$$

In two dimensions, the concept is similar. Say we have a function

$$\iint_R f(x, y)dA$$

Then, we substitute  $x = x(u, v)$  and  $y = y(u, v)$ . Analogous to changing the limits of one-dimensional integration, we change the region  $R$  over which we're integrating. A small rectangle of area  $\Delta u\Delta v$  in  $R$  is transformed to a slightly different area in the new region  $S$ . So the integral becomes

$$\iint_S f(x(u, v), y(u, v)) \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} dudv$$

(for now, just zoom in)

We denote this determinant by

$$\left| \frac{\partial(x, y)}{\partial(u, v)} \right|$$

#### 16.1.1 Derivation for the determinant

Consider a small rectangle of side lengths  $\Delta u$  and  $\Delta v$ . The sides are vectors  $\langle \Delta u, 0 \rangle$  and  $\langle 0, \Delta v \rangle$ . To find the magnitude of a small change in  $u$  in the transform, we take the partial derivative of the point mapped as  $(x, y)$  to get one side length of the new parallelogram  $\langle x_u, y_u \rangle \Delta u$ .

By symmetry the other is  $\langle x_v, y_v \rangle \Delta v$ . Therefore, the area is

$$A = \Delta u \Delta v \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}$$

### 16.1.2 Example

$$\iint_R \frac{1}{\sqrt{xy}} dA$$

where  $R$  is defined by the lines

$$x + y = 1, y = x/3, x + y = 4, y = 3x$$

We make the substitution  $x = u^2, y = v^2$ . Then, with the definition  $T(u, v) = (u^2, v^2)$ , we can find the region  $S$  such that  $T(S) = R$  is a one-to-one correspondence.

On  $y = 3x, v^2 = 3u^2$ .

On  $y = x/3, v^2 = u^2/3$ .

On  $x + y = 1, u^2 + v^2 = 1$ .

On  $x + y = 4, u^2 + v^2 = 4$ .

This defines our transformed region.

The Jacobian is:

$$\begin{vmatrix} 2u & 0 \\ 0 & 2v \end{vmatrix} = 4uv$$

Therefore, the integral becomes

$$\iint_R \frac{1}{\sqrt{xy}} dA = \iint_S \frac{1}{uv} 4uv du dv = 4 \iint_S du dv$$

which is just an area computation, which we can do in polar coordinates.

The lines make angles of  $\frac{\pi}{6}$  and  $\frac{\pi}{3}$  respectively, and  $r$  goes from 1 to 2. So the integral becomes

$$4 \int_{\pi/6}^{\pi/3} \int_1^2 r dr d\theta = \pi$$

## 16.2 Triple Integrals

Let  $R = [a, b] \times [c, d] \times [r, s]$ , and let  $f : R \rightarrow \mathbb{R}$ . Then, we subdivide the regions as follows:

$$a < x_0 < x_1 < \cdots < x_n < b$$

$$c < y_0 < y_1 < \cdots < y_n < d$$

$$r < z_0 < z_1 < \cdots < z_n < s$$

Define  $\Delta x = x_i - x_{i-1}$ , and  $\Delta y$  and  $\Delta z$  similarly. Then, take a sample point in one of the *sub-boxes* on the region  $(x_i^*, y_j^*, z_k^*)$ . We can find the volume of the entire region by summing over the volumes of each sub-box:

$$\lim_{n \rightarrow \infty} \sum_{i,j,k=1}^n f(x_i^*, y_j^*, z_k^*) \Delta x \Delta y \Delta z$$

i.e. as an integral, this is

$$\int \int \int_R f(x, y, z) dV$$

As with double integrals, we do this as an iterated integral.

### 16.2.1 Example

$$R = [0, 1] \times [2, 3] \times [4, 5]$$

and  $f(x, y, z) = xyz$ .

The integral is

$$\begin{aligned} & \int_4^5 \int_2^3 \int_0^1 xyz dx dy dz \\ &= \int_4^5 \int_2^3 \frac{yz}{2} dy dz = \int_4^5 \frac{y^2 z}{4} \Big|_2^3 dz = \frac{5}{4} \frac{z^2}{2} \Big|_4^5 = \frac{45}{8} \end{aligned}$$

### 16.2.2 yay fun

We can define this integral for any bounded solid region.

### 16.2.3 Another example

Find  $\int \int \int x dV$  where the region is bounded by  $x = 4y^2 + 4z^2$  and  $x = 4$ .

After a lot of drawing, we get the following integral, after we decide to set it up in the order  $dx dA$ .

Actually I'm not doing this

Math 53: Multivariable Calculus

Spring 2018

## Lecture 20: Conservative Fields and Green's Theorem

Lecturer: Nikh l Srivastava

22 March

Scribe: Aditya Sengupta

## 20.1 Line Integrals, Continued?

"Recall" that  $F = P\hat{i} + Q\hat{j}$  implies that the line integral is

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b F(\vec{r}(t)) \cdot \vec{r}'(t) dt = \int_C P dx + \int_C Q dy$$

$$F = \nabla f \text{ for some continuous } f \implies \int_C \nabla f \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a))$$

We can say that  $F$  is conservative if the value of its line integral depends only on the end points, or that  $\int_C \vec{F} \cdot d\vec{r} = 0$  over every closed  $C$ . This in turn suggests that  $Q_x = P_y$ .

Suppose  $F = \langle P, Q \rangle = \nabla f = \langle f_x, f_y \rangle$ . This implies  $P = f_x, Q = f_y$ . Then  $P_y = Q_x$  by Clairaut's theorem.

Therefore we can check if a field is conservative by checking whether the opposite partial derivatives are equal. For example,  $F = -y\hat{i} + x\hat{j}$  is not conservative because  $Q_x = 1$  and  $P_y = -1$ .

The  $Q_x = P_y$  condition is not an if-and-only-if relation. If  $Q_x = P_y$  and  $F$  exists on a simply connected domain, then  $F$  is conservative. So what's a simply connected domain?

**Definition 23.** A domain  $D$  is simply connected if every simple closed curve encloses a region containing only points of  $D$ .

**Definition 24.**  $C$  is simple if it does not intersect itself.

**Definition 25.** It's turtles all the way down.

**Theorem 26.** Jordan Curve Theorem: a simple closed continuous curve divides the plane into two regions, one of which is simply connected.

## 20.2 Nonconservative Field

$$F = P\hat{i} + Q\hat{j}$$

$D$  simply connected

$Q_y - P_x = 0$  for conservative fields. This is called the *curl* of the vector field.

If the curl is zero, there is no rotation of the vector field. If it's positive, the field is counterclockwise, and if it's negative, the field is clockwise.

## 20.3 Green's Theorem

**Theorem 27.** *if  $C$  is a simple closed curve enclosing a region  $D$ , and  $F = P\hat{i} + Q\hat{j}$  is a vector field, then*

$$\oint_C \vec{F} \cdot d\vec{r} = \iint \nabla \times \vec{F} dA$$

In terms of our earlier example, this means

$$\int_C Pdx + \int_C Qdy = \iint (Q_x - P_y)dA$$

Math 53: Multivariable Calculus

Spring 2018

## Lecture 21: Green's Theorem and Divergence

Lecturer: Nikhil Srivastava

3 April

Scribe: Aditya Sengupta

Sorry about the lack of images. Use your imagination a bit? :P

Wow emoticons don't typeset nicely at all

## 21.1 Review of Green's Theorem

Given a vector field  $F = P\hat{i} + Q\hat{j}$ , the line integral over a simple closed positively oriented curve  $C$  is

$$\int_C \vec{F} \cdot d\vec{r} = \iint_D \text{curl}(F) dA$$

where the curl of  $F$  is  $Q_x - P_y$ . More explicitly, this means

$$\int_C Pdx + Qdy = \iint_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)$$

## 21.2 Applications

1. Can change complicated  $\int_C$  into simple  $\iint_D$  and vice versa
2. Can calculate the area of complicated curves, such as Florida.

The area is represented as a double integral over a weird-shaped domain,  $\iint_D dA$ . By Green's theorem, this is equal to

$$\int_C Pdx + Qdy$$

where  $F = \langle P, Q \rangle$  such that  $\text{curl}(F) = 1$ . For example,  $F = x\hat{j}$ .

If we take the above example, this becomes a Riemann sum:

$$\int_C xdy \approx \sum_{i=1}^n x(t_i)y'(t_i)\Delta t$$

3. A vector field  $F$  is conservative if and only if  $\oint_C \vec{F} \cdot d\vec{r} = 0$  for closed  $C$ .  $F$  being conservative also implies  $Q_x - P_y = 0$ . Therefore, we can use Green's theorem to prove the following:

**Theorem 28.** *If  $D$  is simply connected and  $Q_x - P_y = 0$  on  $D$ , then  $F = \nabla f$  on  $D$ .*

*Proof.* Pick some simple closed curve  $C$ . We know that the curl of  $F$  is 0. This means

$$\int_C \vec{F} \cdot d\vec{r} = \int \int_D \text{curl}(F) dA = 0$$

Therefore,  $F$  is the gradient of another function. □

## 21.3 Proof of Green's Theorem

We want to show that  $\int_C P dx + Q dy = \int \int_D (Q_x - P_y) dA$  for components  $P, Q$  of a two-dimensional vector  $F$ .

One way of approaching this is to show the following:

$$\int_C P dx = \int \int_D -P_y dA$$

$$\int_C Q dy = \int \int_D Q_x dA$$

which are, respectively, the  $Q = 0$  and  $P = 0$  cases. Then we can simply add these two results to get Green's theorem.

### 21.3.1 Proving the first statement

It's not entirely clear how to go about this, because we don't know how to do double integrals over arbitrary regions. So, as a first step, we prove this statement for Type I regions:

$$\int_C P dx = \int \int_D -P_y dA, D = \{a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

We calculate the line integral, by dividing it into four parts  $C_1$  through  $C_4$ . The line integral evaluates to

$$\int_C P dx = \int_{C_1} P dx + \int_{C_2} P dx + \int_{C_3} P dx + \int_{C_4} P dx$$

Since  $x$  does not change on  $C_2$  and  $C_4$  (get pictures), the second and fourth integrals become zero.

$$= \int_a^b P(x, g_1(x)) dx + \int_b^a P(x, g_2(x)) dx = \int_a^b P(x, g_1(x)) - P(x, g_2(x)) dx$$

Now for the double integral:

$$\int \int_D -P_y dA = \int_a^b \int_{g_1(x)}^{g_2(x)} -\frac{\partial}{\partial y} P(x, y) dy dx$$

By the Fundamental Theorem of Calculus, this is equal to

$$\int_a^b -\left(P(x, g_2(x)) - P(x, g_1(x))\right) dx$$

which is the same as the above.

Therefore,

$$\int_C P dx = \int \int_D -P_y dA$$

for Type I regions.

### 21.3.2 Extending this to general regions

The idea behind this is to decompose complicated regions into simpler ones. Let an arbitrary region consist of sub-regions  $R_1$  and  $R_2$ , traced by curves  $C_1$  and  $C_2$  respectively. Suppose we know that the statement above is true for both regions, i.e.

$$\int_{C_1} P dx = \int \int_{R_1} -P_y dA$$

$$\int_{C_2} P dx = \int \int_{R_2} -P_y dA$$

On the right hand side, we can add these up to get an area integral over the entire region. On the left side, we get a line integral over the original  $C$ , plus two integrals over the interface between the regions which end up being equal and opposite, therefore they cancel.

Now, for general regions, we can subdivide them into Type I regions. We know that the statement is valid for all such regions, and for combinations of such regions, therefore the statement is valid for the larger region as well.

### 21.3.3 Repeat the above for the $Qdy$ statement

This can be done using all the same methods as above.

### 21.3.4 Add the two statements

*Proof.* of Green's theorem: See sections 21.3.1 through 21.3.3. □

## 21.4 Flux Interpretation of Green's Theorem

So far, we have considered the physical interpretation of line integrals:



$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t) dt$$

We can also express this geometrically as a tangent vector. Change  $\vec{r}'(t)dt$  to  $\hat{T}(t)|\vec{r}'(t)|dt$ . Since  $|\vec{r}'(t)|dt = ds$ , the line integral becomes

$$\int_C \vec{F} \cdot \hat{T} ds$$

We could consider the integral of  $\vec{F} \cdot \vec{n}$ , where  $\vec{n}(t)$  is the normal vector, which is  $90^\circ$  clockwise relative to  $\hat{T}$  and perpendicular to it.

Physically, if  $F$  is a velocity field, this integral of  $\vec{F} \cdot \vec{n}$  is the flux, or the net outflow across  $C$  per unit time.

Suppose  $F = x\hat{i} + y\hat{j}$ , and  $C$  is the circle of radius 1 centered at  $(1, 1)$ . As  $\vec{F}$  is not constant, it isn't exactly clear what net flux is. To make this notion mathematically precise, let's derive an expression for  $\vec{n}$ .

We know that  $\hat{T}(t) = \frac{\vec{r}'(t)}{|\vec{r}'(t)|} = \frac{\langle x'(t), y'(t) \rangle}{|\vec{r}'(t)|}$ . We can easily verify by taking a dot product that  $\langle y'(t), -x'(t) \rangle$  is perpendicular to  $\hat{T}$ . Therefore, we normalise this to find that

$$\vec{n}(t) = \frac{\langle y'(t), -x'(t) \rangle}{|\vec{r}'(t)|}$$

We can now give an explicit description of the flux:

$$\begin{aligned} Flux &= \int_C \vec{F} \cdot \vec{n} ds = \int \langle P, Q \rangle \cdot \frac{\langle y'(t), -x'(t) \rangle}{|\vec{r}'(t)|} |\vec{r}'(t)| dt \\ &= \int P(x, y)y'(t) - Q(x, y)x'(t) dt = \int_C Pdy - Qdx \end{aligned}$$

to which we can apply Green's theorem. This is slightly modified, however: we take  $P = -Q$  and  $Q = P$ . This means the area integral becomes

$$\int \int_D P_x + Q_y dA$$

This quantity is called the *divergence*.

## 21.5 Divergence Form of Green's Theorem

Define  $div(F) = P_x + Q_y$ . Then,

$$\int_C \vec{F} \cdot \vec{n} ds = \int \int_D div(F) dA$$

for a simply closed positively oriented  $C$ .

We can use this to solve the example in 21.4, in which  $F$  was not constant. We can take the divergence:

$$\operatorname{div} F = \frac{\partial}{\partial x} x + \frac{\partial}{\partial y} y = 2$$

Therefore, by this form of Green's theorem, the flux becomes

$$\int_C \vec{F} \cdot \vec{n} ds = \int \int_D \operatorname{div}(F) dA = \int \int_D 2 dA = 2\pi$$

Conclusions:

1. This was easy
2. Flux only depends on area, not on location. (*my addition:* As long as divergence is a constant and does not depend on  $x$  or  $y$ . Which is probably not always the case.)

## 21.6 Physical Meaning of Divergence

Imagine the graph of a vector field  $\langle x, y \rangle$ . There is no net inflow, and there is a net outflow. This is reflected in that the divergence has a positive value. For the vector field  $-\langle x, y \rangle$ , divergence is  $-2$ , and there is a net inflow, or a net *negative* outflow. The divergence of a constant vector field is zero, because there is no net inflow or outflow in a constant field such as  $\langle 1, 1 \rangle$ .

Therefore, the divergence represents the flux over a small area. (Rate of net outflow over unit time and area)

**Math 53: Multivariable Calculus**

**Spring 2018**

**Lecture 22: Curl and Divergence, Parameterized Surfaces, Surface Area**

*Lecturer: Nikhil Srivastava*

*5 April*

*Scribe: Aditya Sengupta*

The goal for the rest of the course is to generalise the ideas of curl and divergence to three dimensions. In order to do this, we need to formalize a two-dimensional analogue of a line integral, which turns out to be a *surface integral*.

## 22.1 Curl in 3D

Given a vector field  $F = \langle P, Q, R \rangle$ ,

$$\operatorname{curl}(F) = \nabla \times F = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \hat{i}(R_y - Q_z) - \hat{j}(R_x - P_z) + \hat{k}(Q_x - P_y) = \langle R_y - Q_z, P_z - R_x, Q_x - P_y \rangle$$

Why does this make sense?

### 22.1.1 The 2D curl is a special case

Let  $R = 0$  in the above. Then,  $P_z = Q_z = 0$  and  $R_x = R_y = R_z = 0$ , and the curl simplifies to

$$\operatorname{curl}(F) = \langle 0, 0, Q_x - P_y \rangle$$

### 22.1.2 It appears to measure local rotation

Take a constant  $F$ . The curl is zero both intuitively and by computing partial derivatives (which will all individually be 0).

Now, let  $F = \langle x, y, z \rangle$ . The curl is 0 again, because curl is meant to measure rotation, and plotting the corresponding vector field shows us that  $F$  only causes displacement, not rotation.

Let  $F = \langle 0, z, -y \rangle$ . The curl of  $F$  is  $\langle -2, 0, 0 \rangle$ , which is consistent with the corresponding vector field plot, which shows clockwise rotation in the  $x$ -direction.

### 22.1.3 It lets us check whether a vector field is conservative

**Theorem 29.** *If  $F = \nabla f$ ,  $\operatorname{curl}(F) = 0$ .*

*Proof.* Let  $F = \langle F_x, F_y, F_z \rangle$ . The curl of  $F$  is

$$\langle f_{zy} - f_{yz}, f_{xz} - f_{zx}, f_{yx} - f_{xy} \rangle$$

which is the zero vector by Clairaut's theorem. □

**Theorem 30.** *If  $F$  is defined on  $\mathbb{R}^3$  and  $\text{curl}(F) = 0$ , then  $F = \nabla f$  for some  $f$ .*

### Example

**Is the vector field  $F = \langle e^z, 1, xe^z \rangle$  conservative?**

We compute the determinant,

$$\begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ e^z & 1 & xe^z \end{vmatrix} = \langle 0, e^z - e^z, 0 \rangle = \vec{0}$$

Therefore the vector field is conservative. This means we know that

$$\langle e^z, 1, xe^z \rangle = \langle f_x, f_y, f_z \rangle$$

for some  $f$ . From this, we know

$$f_x = e^z, f_y = 1, f_z = xe^z$$

We integrate each of these to get:

$$\begin{aligned} f &= \int e^z dx + g_1(y, z) = \int dy + g_2(x, z) = \int xe^z dz + g_3(x, y) \\ &= xe^z + y + C \end{aligned}$$

We can show that all possible gradients of a function are the same up to a constant. Let  $f_1$  and  $f_2$  both satisfy  $F = \nabla f$ . Then,

$$\nabla f_1 - \nabla f_2 = F - F = 0$$

i.e.,  $\nabla(f_1 - f_2) = 0$  which implies they are the same up to a constant.

## 22.2 Divergence in 3D

**Definition 31.**

$$\text{div}(F) = P_x + Q_y + R_z$$

We can do the same sanity checks as with the generalization of curl, with the physical intuition that the divergence of a vector field measures the outflow of a vector field at a point.

## 22.3 Relationship between gradient, curl, and divergence

$$\text{functions} \xrightarrow{\nabla} \text{vector fields} \xrightarrow{\text{curl}} \text{vector fields} \xrightarrow{\text{div}} \text{functions}$$

Note that the curl of the divergence, or vice versa, is always 0. We can also take the divergence of the divergence of a function, which gives us the Laplacian  $\nabla^2$  of a function.

## 22.4 Surface Integration

In order to do line integrals, we needed to parameterize the curve we were integrating over. We need to parameterize a surface in much the same way to do a surface integral.

Let  $D \subseteq \mathbb{R}^2$  be the domain over which we want to do a surface integral. We parameterise it by a vector  $\vec{r}(u, v) = x(u, v)\hat{i} + y(u, v)\hat{j} + z(u, v)\hat{k}$ , where  $(u, v) \in D$ .

Suppose the surface is a graph of a function  $f(x, y)$ . Then, we can parameterise the surface as  $(x, y, z) \rightarrow (x, y, f(x, y))$ .

In a plane, where  $\vec{r}_a$  and  $\vec{r}_b$  are nonparallel vectors in the plane and  $\vec{r}_0$  is the position vector of any point in the plane, we can parameterise it by  $\vec{r}_0 + u\vec{r}_a + v\vec{r}_b, (u, v) \in \mathbb{R}^2$ .

To parameterise a sphere, we can say

$$x = \rho \sin u \cos v, y = \rho \sin u \sin v, z = \rho \cos u$$

$$u \in [0, \pi], v \in [0, 2\pi]$$

To complete the analogy, we need an analogue to the tangent or velocity vector in line integrals. We define a  $u$ -velocity and a  $v$ -velocity, where  $\vec{r}_u = \langle x_u, y_u, z_u \rangle$  and the same with  $v$ . Therefore, we can define the tangent plane to be

$$\vec{s}(u, v) = \vec{r}(u_0, v_0) + u\vec{r}_u(u_0, v_0) + v\vec{r}_v(u_0, v_0)$$

Now, with only three minutes remaining, we're ready to do an integral over a surface!

In the  $u - v$  plane, a rectangle of area  $\Delta u \Delta v$  is mapped onto a strange curvy thing.

Math 53: Multivariable Calculus

Spring 2018

## Lecture 23: Midterm 2 Review

Lecturer: Nikh l Srivastava

10 April

Scribe: Aditya Sengupta

**23.1 Double integrals**

The idea of a double integral is that we partition the region over which we are integrating,  $D$ , into infinitesimally small rectangles of area  $\Delta x \Delta y$ .

$$\int_D f(x, y) dx dy = \int_D f(x, y) dA = \lim_{n \rightarrow \infty} \sum_{i, j \leq n} f(x_i^*, y_j^*) \Delta x \Delta y$$

If we take  $f = 1$ , then we get  $\int_D dA$  which is the area of the region.

**23.1.1 Computing double integrals****23.1.1.1 Type 1**

$$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) dy dx$$

**23.1.1.2 Type 2**

$$\int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) dx dy$$

**23.2 Polar coordinates and change of variables**

$$x = r \cos \theta, y = r \sin \theta$$

This allows us to more easily describe regions more naturally defined by circles and arcs, making integration easier:

$$\int f(x, y) dA = \int f(r \cos \theta, r \sin \theta) r dr d\theta$$

The extra  $r$  term comes from the Jacobian of the associated Cartesian-polar transformation. In general, a change of variables allows us to transform an arbitrarily-shaped region into a rectangle that we can easily integrate over.

The multiplier for a change of variables is *the Jacobian*:

$$J = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}$$

If we change an integral in  $dx dy$  into one in  $du dv$  by a transformation  $T$ , we get the transformed integral

$$\int \int_D f(x, y) dx dy = \int \int_S f(T(u, v)) \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} du dv$$

## 23.3 Triple Integrals

These are just double integrals with one more integral, with the physical interpretation being that they are a volume instead of an area. This allows us to define more coordinate systems:

### 23.3.1 Cylindrical Coordinates

A point in cylindrical coordinates is characterised by  $r, \theta, z$  where  $x = r \cos \theta, y = r \sin \theta$ .

$$\int \int \int f dV = \int \int \int f r dz dr d\theta$$

### 23.3.2 Spherical Coordinates

We can also generalize polar coordinates in 3D with a sphere.

$$x = \rho \sin \varphi \cos \theta, y = \rho \sin \varphi \sin \theta, z = \rho \cos \varphi$$

and the integral becomes

$$\int \int \int f dV = \int \int \int f \rho^2 \sin \phi d\rho d\varphi d\theta$$

## 23.4 Line Integrals

The line integral of a function (scalar field)  $\int_C f ds$ , where  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is found by by defining a position vector  $\vec{r}(t), t \in [a, b]$ . Then the line integral becomes

$$\int_C f(\vec{r}) ds = \int_a^b f(\vec{r}(t)) |\vec{r}'(t)| dt$$

$\vec{r}(t)$  and  $\vec{r}'(t)$  are both vector-valued functions from  $\mathbb{R}$  to  $\mathbb{R}^2$ , whose output consists of components  $\langle x(t), y(t) \rangle$  or their time derivatives. It follows that  $|\vec{r}'(t)| = \sqrt{(x'(t))^2 + (y'(t))^2}$ .

An easier way of going about this is to consider the line integral equal to  $\int_C \vec{F} \cdot d\vec{r}$ , which is

$$\int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t) dt = \int_a^b \left[ P(\vec{r}(t)) \frac{dx(t)}{dt} + Q(\vec{r}(t)) \frac{dy(t)}{dt} \right] dt$$

This depends on the orientation of  $C$ , but is much simpler to compute.

## 23.5 Conservative Fields/FTLI

If  $F = \nabla f$  for some  $f$ , then we can change the vector field  $\vec{F}$  into a scalar field  $f$ . So if  $C$  is still parameterised by  $t \in [a, b]$ , then we can say

$$\int_C \vec{F} \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a))$$

In one dimension, this is simply the fundamental theorem of calculus. The statement that this is true means that the line integral is path independent, which is the statement that  $\vec{F}$  is a conservative field. If  $\vec{F} = \langle P, Q \rangle$ ,  $\vec{F}$  is conservative if and only if  $Q_x = P_y$ .

## 23.6 Green's Theorem

This is the generalization of integration by parts in two dimensions. If  $C$  is a simple, closed, positively oriented curve enclosing a region  $D$ , we can say that

$$\int_C \vec{F} \cdot d\vec{r} = \int_C P dx + Q dy = \int \int_D (Q_x - P_y) dA$$



Math 53: Multivariable Calculus

Spring 2018

## Lecture 27: Stokes' Theorem

Lecturer: Nikhil Srivastava

24 April

Scribe: Aditya Sengupta

## 27.1 The Story So Far and Stokes' Theorem

1. The fundamental theorem of line integrals:

$$\int_C \vec{F} \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a))$$

2. Green's Theorem:

$$\int \int_D \text{curl}(F) dA = \oint \vec{F} \cdot d\vec{r}$$

3. Gauss's Theorem:

$$\int \int \int \nabla \cdot \vec{F} dV = \iint_S \vec{F} \cdot \vec{n} dS$$

These have a common pattern of changing an integral over a region into a lower-dimensional integral over the boundary of that region. So, we can extend this to:

4. Stokes' Theorem

2D  $\rightarrow$  S, enclosed surface in  $\mathbb{R}^3$ . 1D  $\rightarrow$  C, boundary of S, positively oriented.

Let  $\vec{F} = \langle P, Q, R \rangle$  be a vector field defined on S with continuous partial derivatives. Then, we can state Stokes' Theorem as follows:

**Theorem 32.**

$$\int \int_S \text{curl}(\vec{F}) \cdot \vec{n} dS = \oint_C \vec{F} \cdot d\vec{r}$$

This requires us to define *positively oriented*. We say that if you are walking on a positively oriented curve, with the interior to the left, then the normal is pointing upward. Alternatively, we can use a right hand rule, in which if the index finger of your right hand is tangent to C, and your middle finger points to the interior of S, then your thumb is the normal vector.

Verbally, Stokes' Theorem tells us that the net circulation (work) done by a field over a path is equal to the local circulation (curl) of that field along the surface.



### Example 1

Let  $\vec{F} = \langle P, Q, 0 \rangle$ , let  $S$  be a region in the  $x - y$  plane oriented upwards, and let  $C$  be the positively oriented boundary of  $S$ . To verify Stokes' Theorem, we compute the surface integral of the curl:

$$\text{curl}(\vec{F}) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & 0 \end{vmatrix}$$

the  $z$  component of which is  $Q_x - P_y$ . We disregard the other two because the normal vector is  $\langle 0, 0, 1 \rangle$ . Therefore, on the left hand side we get

$$\int \int_S (Q_x - P_y) dA = \oint_a^b \langle P, Q, 0 \rangle \cdot \langle x(t), y(t), 0 \rangle dt = \int_C P dx + Q dy$$

which is true by Green's Theorem.

## Example 2

Let  $S$  be the region of the plane  $3x + 2y + z = 1$  in the first octant. We want to find

$$\oint_C \vec{F} \cdot d\vec{r}$$

where  $F = \langle 1, x + yz, xy - \sqrt{z} \rangle$  and  $C$  is the positively oriented boundary of  $S$ .

We apply Stokes' Theorem. The normal vector is  $\frac{\langle 3, 2, 1 \rangle}{\sqrt{14}}$ , and the surface is the graph of  $z = 1 - 3x - 2y$ . Since the graph lies in the first octant, the domain is equal to

$$D = \{x \geq 0, y \geq 0, 3x + 2y \leq 1\}$$

Then, the curl is

$$\text{curl}(F) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 1 & x + yz & xy - \sqrt{z} \end{vmatrix} = \langle x - y, -y, 1 \rangle$$

Therefore, the integral becomes

$$\frac{1}{\sqrt{14}} \int \int_S \langle x - y, -y, 1 \rangle \cdot \langle 3, 2, 1 \rangle dS = \frac{1}{\sqrt{14}} \int \int_D (3(x - y) - 2y + 1) \sqrt{14} dA$$

where we add in a magnification factor of  $|\vec{r}_x \times \vec{r}_y| = \sqrt{14}$