Multivariable Calculus Math S21A

Oliver Knill **Harvard University**

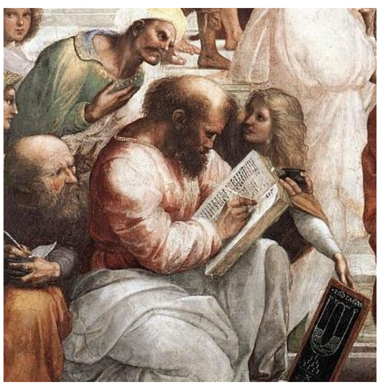
Chapter 1. Geometry and Space

Section 1.1: Space and Distance

René Descartes

Points P in space are described by **Coordinates** like $P = (3, 4, 5)$ As promoted by **René Descartes** in the 16'th century, geometry can be treated algebraically using **coordinate systems**. The **distance** between $P(x, y, z)$ and $Q = (a, b, c)$ is defined as $d(P, Q) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$. This is motivated by **Pythagoras theorem** which we will prove. We explore geometric objects in the plane and in space. We focus **cylinders**, **planes** or **spheres** and learn how to find the **center** and **radius** of a sphere. This is the **completion of the square**.

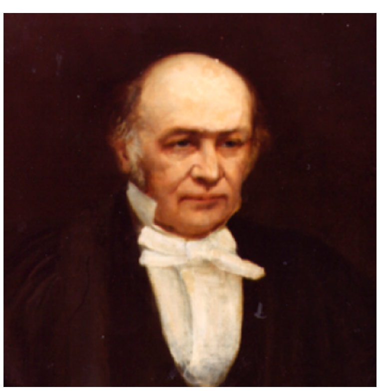
Section 1.2: Vectors and Dot product



Pythagoras

Two points P, Q define a **vector** $\vec{PQ} = -\vec{QP}$. Vectors describe **velocities**, **forces**, **color** or **data**. The **components** of $\vec{PQ}$ connecting $P = (a, b, c)$ with $Q = (x, y, z)$ are the entries of the **column vector** $[x-a, y-b, z-c]^T$. The **zero vector** is $\vec{0} = [0, 0, 0]^T$. The **standard basis vectors** are $\vec{i} = [1, 0, 0]^T, \vec{j} = [0, 1, 0]^T, \vec{k} = [0, 0, 1]^T$. **Addition**, **subtraction** and **scalar multiplication** work geometrically and algebraically. The **dot product** $\vec{v} \cdot \vec{w}$ is a **scalar** giving **length** $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$ and **direction** $\vec{v}/|\vec{v}|$ for $|\vec{v}| \neq 0$. The **angle** is defined by $\vec{v} \cdot \vec{w} = |\vec{v}||\vec{w}| \cos \alpha$ as justified by **Cauchy-Schwarz** $|\vec{u} \cdot \vec{v}| \leq |\vec{u}||\vec{v}|$. The cos-formula follows. If $\vec{v} \cdot \vec{w} = 0$, we say $\vec{v}, \vec{w}$ are **perpendicular**, giving **Pythagoras** $|\vec{v} + \vec{w}|^2 = |\vec{v}|^2 + |\vec{w}|^2$.

Section 1.3: Cross and Triple Product



Rowan Hamilton

The **cross product** $\vec{v} \times \vec{w}$ of $\vec{v} = [a, b, c]^T$ and $\vec{w} = [p, q, r]^T$ is defined as $[br - cq, cp - ar, aq - bp]^T$. It is perpendicular to $\vec{v}$ and $\vec{w}$. In two dimensions, the cross product is a scalar $[a, b]^T \times [p, q]^T = aq - bp$. The product is useful to compute **areas** of parallelograms, the **distance** between a point and a line, or to **construct** a plane through three points or to **intersect** two planes. We prove a formula $|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}| \sin(\alpha)$ which allows us to define the **area** of the parallelepiped spanned by $\vec{v}$ and $\vec{w}$. The **triple scalar product** $(\vec{u} \times \vec{v}) \cdot \vec{w}$ is a scalar and defines the **signed volume** of the parallelepiped spanned by $\vec{u}, \vec{v}$ and $\vec{w}$. Its sign gives the **orientation** of the coordinate system defined by the three vectors. The triple scalar product is 0 if and only if the three vectors are in a common plane.

Section 1.4: Lines and Planes

Arthur Cayley

Because $[a,b,c]^T = \vec{n} = \vec{u} \times \vec{v}$ is perpendicular to $\vec{x} - \vec{w}$, if $\vec{x}, \vec{w}$ are in the plane spanned by $\vec{u}$ and $\vec{v}$, points on a plane satisfy $ax + by + cz = d$. We often know the normal vector $\vec{n} = [a,b,c]^T$ to a plane and can determine the constant d by plugging in a known point (x,y,z) on equation $ax + by + cz = d$. The parametrization $\vec{x}(t,s) = \vec{w} + t\vec{u} + s\vec{v}$ is an other way to represent surfaces. We introduce **lines** by the parameterization $\vec{r}(t) = \vec{OP} + t\vec{v}$, where P is a point on the line and $\vec{v} = [a,b,c]^T$ is a vector telling the direction of the line. If $P = (o,p,q)$, and a,b,c are all non-zero then $(x-o)/a = (y-p)/b = (z-q)/c$ is called the **symmetric equation** of a line. It can be interpreted as the intersection of two planes. As an application of the dot and cross products, we look at various **distance formulas**.

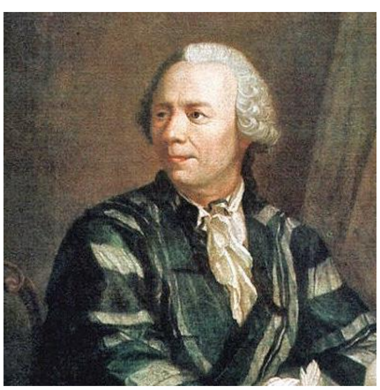
Chapter 2. Curves and Surfaces

Section 2.1: Level Curves and Surfaces

Claudius Ptolemy

The **graph** of a function $f(x,y)$ of two variables is defined as the set of points (x,y,z) for which $g(x,y,z) = z - f(x,y) = 0$. We look at examples and match some graphs with functions $f(x,y)$. **Generalized traces** like $f(x,y) = c$ are called **level curves** of f and help to visualize surfaces. The set of all level curves forms a **contour map**. After a short review of **conic sections** like **ellipses**, **parabola** and **hyperbola** in two dimensions, we look at more general surfaces of the form $g(x,y,z) = 0$. We start with the **sphere** and the **plane**. If $g(x,y,z)$ is a function which only involves linear and quadratic terms, the level surface is called a **quadric**. Important quadrics are **spheres**, **ellipsoids**, **cones**, **paraboloids**, **cylinders** as well as **hyperboloids**.

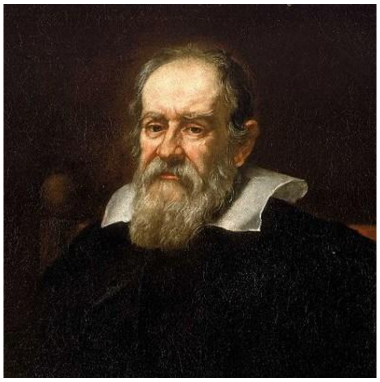
Section 2.2: Parametric Surfaces



Leonhard Euler

Surfaces are described implicitly or parametrically. Examples of implicit descriptions $g(x,y,z) = 0$ are $x^2 + y^2 + z^2 - 1 = 0$. Examples of **parametrizations** $\vec{r}(u,v) = [x(u,v), y(u,v), z(u,v)]^T$ are the **sphere** $\vec{r}(\theta,\phi) = [\rho\cos(\theta)\sin(\phi), \rho\sin(\theta)\sin(\phi), \rho\cos(\phi)]^T$, where ρ is fixed and ϕ, θ are the **Euler angles**. Using computers, one can **visualize** also complicated surfaces. Parametrization of surfaces is important in **geodesy**, where they appear as maps or in **computer generated imaging**, where the parameterization $\vec{r}(u,v)$ is called the **"uv-map"**. Parametrizations of surfaces make use of **cylindrical coordinates** (r,θ,z), where $r \geq 0$ is the distance to the z-axes and $0 \leq \theta < 2\pi$ is an angle. **spherical coordinates** (ρ,θ,ϕ) use ρ, the distance to $(0,0,0)$ and θ, ϕ, the Euler angles.

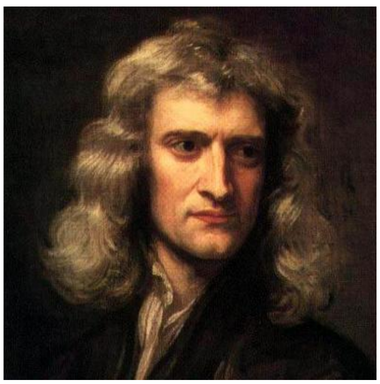
Section 2.3: Parametric Curves



Johannes Kepler

The parametrization $\vec{r}(t) = [x(t), y(t), z(t)]^T$, of a curve using **parameter** t given in the interval $I = [a, b]$ contains more information than the curve itself. It tells also, how the curve is traced if t is interpreted as **time**. Differentiation of a parametrization $\vec{r}(t)$ leads to the **velocity** $\vec{r}\,'(t)$, a vector which is **tangent** to the curve at $\vec{r}(t)$. A second differentiation with respect to t gives the **acceleration** vector $\vec{r}\,''(t)$. The **speed** $|r'(t)|$ is a scalar. We also learn how to get from $\vec{r}''(t)$ and $\vec{r}\,'(0)$ and $\vec{r}(0)$ the position $\vec{r}(t)$ by integration. A special case is the **free fall**, where the acceleration vector is constant.

Section 2.4: Arc length and Curvature



Isaac Newton

The **arc length** of a curve is defined as a limiting length of polygons and leads to the **arc length** integral $\int_a^b |r'(t)|\, dt$. A re-parametrization of a curve does not change the arc length. The **curvature** $\kappa(t)$ of a curve measures how much a curve is bent. Acceleration and curvature involve second derivatives. Curvature is a quantity which does not depend on parameterizations. One "feels" acceleration and "sees" curvature $\kappa(t) = |T''(t)|/|T'(t)| = |\vec{r}\ '(t) \times \vec{r}\ ''(t)|/|r\ '(t)|$, where $\vec{T}(t) = \vec{r}'(t)$ is the **unit normal vector** $\vec{T}$. Together with **normal vector** $\vec{N}$ and **bi-normal vector** $\vec{B}$ the 3 vectors form an orthonormal frame.

Chapter 3. Linearization and Gradient

Section 3.1: Partial Derivatives

Alexis Clairot

Continuity questions in multi variables can be more interesting than in one dimension. It can happen for example that $t \to f(t\vec{v})$ is continuous for every $\vec{v}$ but that f is not still continuous. Discontinuities naturally appear with **catastrophes**, changes of the minimum of a critical point. **Partial derivatives** $f_x = \partial_x f = \frac{\partial f}{\partial x}$ satisfy **Clairot's theorem** $f_{xy} = f_{yx}$ for **smooth** functions (functions one can differentiate arbitrarily). We look then at some **partial differential equations** (PDE). Examples are the **transport** $f_x(t, x) = f_t(t, x)$, the **wave** $f_{tt}(t, x) = f_{xx}(t, x)$ and the **heat equation** $f_t(t, x) = f_{xx}(t, x)$.

Section 3.2 Linear Approximation

Brook Taylor

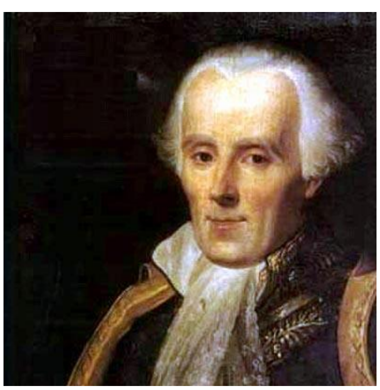
Linearization is an important concept in science because many physical laws are linearization of more complicated laws. Linearization is also useful to estimate quantities. After a review of linearization of functions of one variables, we introduce the **linearization** of a function $f(x,y)$ of two variables at a point (p,q). It is defined as the function $L(x,y) = f(p,q) + f_x(p,q)(x-p) + f_y(p,q)(y-q)$. The **tangent line** $ax + by = d$ at a point (p,q) is a level curve of L and $a = f_x, b = f_y$. Linearization works similarly in three dimensions, where it allows to compute the **tangent plane** $ax + by + cz = d$. The key is the **gradient** $f' = \nabla f = [f_x, f_y, f_z]^T$. We don't cover higher order approximations but they could be done. For nice functions of several variables there is **Taylor theorem** $f(\vec{x}) = f(\vec{p}) + f'(\vec{p}) \cdot (\vec{x} - \vec{p}) + f''(0)(\vec{x} - \vec{p}) \cdot (\vec{x} - \vec{p})/2 + \cdots$ as in one dimensions. The second term $f''(0)$ is a 3×3 matrix which contains in its entries all the mixed derivatives of f at $\vec{p}$.

Section 3.3: Implicit Differentiation

Gottfried Wilhelm Leibniz

The **chain rule** $d/dt f(g(t)) = f'(g(t))g'(t)$ in one dimension can be generalized to higher dimensions. It becomes $d/dt f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$, where $\nabla f = [f_x, f_y, f_z]^T$ is the gradient. Written out, this formula is $d/dt f(x(t), y(t), z(t)) = f_x(x(t), y(t), z(t))x'(t) + f_y(x(t), y(t), z(t))y'(t) + f_z(x(t), y(t), z(t))z'(t)$. All other chain rule versions can be derived from this like if you have a function of several variables or vector-valued functions. A nice application of the chain rule is **implicit differentiation**: if $f(x,y,z) = 0$ defines a surface which looks locally like $z = g(x,y)$ and because $f_x + f_z z' = 0$ we can compute the partial derivatives $g_x = -f_x/f_z$ and $g_y = -f_y/f_z$ of g without knowing g.

Section 3.4: Steepest Ascent



Pierre-Simon Laplace

The **gradient** helps to understand the geometry of surfaces $g(x,y,z) - 0$ because it is perpendicular to the **level surface** $f(x,y,z) = c$. One can see this by linearization or by using the chain rule for a curve $\vec{r}(t)$ on the surface $f(\vec{r}(t)) = 0$. A special case is the plane $g(x,y,z) = ax + by + cz = d$, where $\nabla g = [a,b,c]^T$. The gradient helps to find tangent planes and tangent lines. We introduce the **directional derivative** $D_{\vec{v}}(f)$ as $D_{\vec{v}} f = \nabla f \cdot \vec{v}$ for unit vectors $\vec{v}$. Partial derivatives are special directional derivatives. The direction of the normal vector gives a non-negative partial derivative. Moving into the direction of the normal vector, increases f because $D_{\nabla f/|\nabla f|} f = |\nabla f|$. In other words, the gradient vector points in the direction of **steepest ascent**.

Chapter 4. Extrema and Double integrals

Section 4.1: Maxima and Minima

Pierre de Fermat

To **maximize** $f(x, y)$, first identify **critical points**, points where the gradient vanishes: $\nabla f(x, y) = [0, 0]^T$. The nature of critical points can be established using the **second derivative test**. Let (p, q) be a critical point and let $D = f_{xx}f_{yy} - f_{xy}^2$ denote the **discriminant** of f at this critical point. There are three fundamentally different cases: **local maxima**, **local minima** as well as **saddle points**. If $D < 0$, then (p, q) is a saddle point, if $D > 0$ and $f_{xx} < 0$ then we have a local maximum, if $D > 0$ and $f_{xx} > 0$ then we have a local minimum. If $D = 0$, we can not determine the nature of the critical point from the second derivatives alone. **Global maxima** are places where the $f(x_0, y_0) \geq f(x, y)$ for all (x, y).

Section 4.2: Lagrange Multipliers

Joseph Louis Lagrange

We can maximize $f(x, y)$ in the presence of a **constraint** $g(x, y) = 0$. A necessary condition for a maximum is ∇f and ∇g are parallel. The corresponding system of equations are called the **Lagrange equations**. They are a system of nonlinear equations $\nabla f = \lambda \nabla g, g = 0$. Extrema solve this equation of $\nabla g = 0$. When we maximize or minimize functions on a domain bounded by a curve $g(x, y) = 0$, we have to solve two problems: find the extrema in the interior and the extrema on the boundary. The second problem is a **Lagrange problem**. With the same method we can also maximize or minimize functions $f(x, y, z)$ of three variables, under the constraint $g(x, y) = 0, h(x, y) = 0$. In two or three dimensions, extrema could also be obtained without Lagrange by looking at the equation $\nabla f \times \nabla g = \vec{0}$. Still, the Lagrange framework is very general and works in any dimension.

Section 4.3: Double integrals

Guido Fubini

Integration in two dimensions is first done on **rectangles**, then on regions G bounded by graphs of functions. Depending on whether curves $y = c(x), y = d(x)$ or curves $x = a(y), y = b(y)$ are the boundaries, we call the region **left-to-right region** or **bottum-to-top region**. As in one dimension, there is a **Archimedian sum** or **Riemann sum approximation** of the integral. This allows us to derive results like **Fubini's theorem** on a rectangular region or the change of the order of integration which often enables the integration. The double integral $\int\int_G f(x, y) \; dxdy$ is the signed volume under the graph of the function of two variables. Double integrals define **area** if $f(x, y) = 1$. By **changing of order of integration** in regions which are of both times, we sometimes can integrate integrals which are impossible.

Section 4.4: Polar Integration

Bonaventura Cavalieri

Some regions can be described better in **polar coordinates** (r, θ), where $r \geq 0$ is the distance to the origin and θ is the **polar angle** with the positive x-axes. Examples of regions which can be treated like that are **polar region** is $0 \leq r \leq g(\theta)$ which trace flower-like shapes in the plane. An other application of double integrals is **surface area**. We derive the formula $\int \int_R |r_u \times r_v| \; dudv$ and give examples like graphs, surfaces of revolution and especially the sphere. Similar as for arc length, it is easy to give examples, where the surface area can be computed in closed form, like triangles, parts of the sphere or cylinder or paraboloid. Polar integration also helps to find one-dimensional integrals which otherwise would be difficult to obtain.

Chapter 5. Line integrals

Section 5.1: Triple Integrals

Archimedes of Syracuse

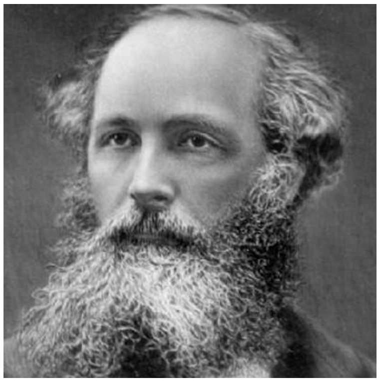
Triple integrals can measure volume, moment of inertia or the center of mass of a solid. First introduced for **cuboids**, then to more general regions like solids, sandwiched between the graphs of two functions $g(x, y)$ and $h(x, y)$. Applications are computations of **mass** $\int \int \int_E \delta(x, y, z) \; dxdydz$, **moment of inertia** $\int \int \int_E (x^2 + y^2 + z^2) \; dxdydz$, **center of mass**, $\int \int \int [x, y, z]^T \; dV$ the **expectation** $\mathrm{E}[X] = \int \int \int X(x, y, z) \; dV / \int \int \int \; dV$ of a random variable $X(x, y, z)$ on a region Ω.

Section 5.2: Spherical Integration

Bernhard Riemann

Some objects can be described better in **cylindrical coordinates** (r, θ, z), which are just polar coordinates for the x, y variables in space, with an additional z coordinate. Examples of such regions are parts of cylinders or solids of revolution. The important factor to include when changing to cylindrical coordinates is r. Other regions are integrated over better in **spherical coordinates** (ρ, ϕ, θ) with $\rho \geq 0$, the distance to the origin, the angles $\phi \in [0, \pi]$ and $\theta \in [0, 2\pi)$. Example of such regions are parts of cones or spheres. The important factor to include when changing to spherical coordinates is $\rho^2 \sin(\phi)$. As an application, we can compute **moments of inertia** of some bodies.

Section 5.3: Vector Fields



James Maxwell

Vector fields occur as force fields or velocity fields or in phase portraits of mechanics or in population dynamics. An important class are **gradient fields**. We look at examples in two or three dimensions. We learn how to **match vector fields** with formulas and introduce **flow lines**, parametrized curves $\vec{r}(t)$ for which the vector $\vec{F}(\vec{r}(t))$ is parallel to $\vec{r}\,'(t)$ at all times. Given a parametrized curve $\vec{r}(t)$ and a vector field $\vec{F}$, we can define the **line integral** $\int_C F(\vec{r}(t))\ \vec{r}\,'(t)\ dt$ along a curve in the presence of a vector field. An important example is the case if $\vec{F}$ is a **force field**. The line integral is then **work**.

Section 5.4: Line Integrals

André-Marie Ampère

For **conservative vector fields** one can evaluate a line integral using the **fundamental theorem of line integrals**. The property **conservative** is also called **path independence** or **conservative** or being a **gradient field** $\vec{F} = \nabla f$. It is equivalent to being **irrotational** $\text{curl}(F) = Q_x - P_y = 0$ if the topological condition of **simply connected** is satisfied: any closed curve can be contracted continuously to a point within the region. The region $\{(x,y) \mid x^2 + y^2 > 1\}$ for example is not simply connected because the path $[2\cos(t), 2\sin(t)]^T$ can not be pulled together to a point. In two dimensions, the curl of a field $\text{curl}([P,Q]^T) = Q_x - P_y$ measures the **vorticity** of the field and if this is zero, the line integral along a simply connected region is zero.

Chapter 6. Integral theorems

Section 6.1: Green's Theorem

Mikhail Ostrogradsky

Greens theorem equates the line integral along a boundary curve C with a double integral of the curl inside the region G: $\int\int_G \text{curl}(\vec{F})(x,y)\ dxdy = \int_C \vec{F}(\vec{r}(t))\ \vec{r}\,'(t)\ dt$. The theorem is useful to **compute areas**: take a field $\vec{F} = [0,x]^T$ which has constant curl 1. It also allows to compute complicated line integrals. Greens theorem implies that if $\text{curl}(F) = 0$ everywhere in the plane, then the field has the **closed loop property** and is therefore conservative. The **curl** of a vector field $\vec{F} = [P,Q,R]^T$ in three dimensions is a new vector field which can be computed as $\nabla \times \vec{F}$. The three components of $\text{curl}(F)$ are the vorticity of the vector field in the x,y and z direction.

Section 6.2: Flux Integrals

Siméon Denis Poisson

Given a surface S and a fluid moving with velocity field $\vec{F}(x,y,z)$ at (x,y,z). the amount of fluid which passes through the membrane S in unit time is the **flux**. This integral is $\int\int_S \vec{F} \cdot (\vec{r}_u \times \vec{r}_v)\ dudv$. The angle between $\vec{F}$ and the normal vector $\vec{n} = \vec{r}_u \times \vec{r}_v$ determines the sign of $\vec{dS} = \vec{F} \cdot \vec{n}\ dudv$. Many concepts are used in this definition: the parametrization of surfaces, the dot and cross product, as well as double integrals. We discuss how the derivatives **div**, **grad** and **curl** fit together. In one dimensions, there is only one derivative, in two dimensions, there are two derivatives grad and curl and in three dimensions, there are three derivatives grad, , curl and div.

Section 6.3: Stokes Theorem

George Gabriel Stokes

Stokes theorem equates the line integral along the boundary C of the surface with the flux of the curl of the field through the surface: $\int_C \vec{F}\ dr = \int\int_S \text{curl}(F)\dot{d}S$. The correct orientations of the surface is important. The theorem allows to illustrate the **Maxwell equations** in electromagnetism and explains why the line integral of an irrotational field along a closed curve in space is zero if the region, where $\vec{F}$ is defined in a simply connected region. It is the flux of the curl of $\vec{F}$ through the surface S bound by the curve C. At this moment, the Mathematica project is due. The project is creative, and illustrates the strong connections of mathematics with art.

Section 6.4 Divergence Theorem

Carl Friedrich Gauss

The total **divergence** of a vector field $\vec{F} = [P,Q,R]^T$ inside **solid** E is the flux of $\vec{F}$ through the boundary S. This **divergence theorem** equates the "local expansion rate" integrated over the solid $\int\int\int_E \text{div}(\vec{F})\ dV$ of a vector field $\vec{F}$ with the flux $\int\int_S \vec{F} \cdot \vec{dS}$ through the boundary surface S of E. Overview: In one dimension, there is one integral theorem, the **fundamental theorem of calculus**. In two dimensions, we have the **the fundamental theorem of line integrals in the plane** as well as **Greens theorem**. In three dimensions we have the **fundamental theorem of line integrals in space**, **Stokes theorem** and the **divergence theorem**. These integral theorems are all of the form $\int_{\delta G} F = \int_G\ dF$, where δG is the boundary of G and dF is the derivative of F.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 1: Geometry and Distance

Lecture

1.1. A point on the **real line** $\mathbb{R}$ is given by a single coordinate x. If x is positive, it is located on the **positive axis** which is divided by zero 0 from the **negative axis**, places where the coordinates are negative. A point P in the **plane** $\mathbb{R}^2$ has two **coordinates** and is written as $P = (x, y)$. A point in space $\mathbb{R}^3$ is determined by three coordinates and written as $P = (x, y, z)$. The signs of the coordinates define four **quadrants** in $\mathbb{R}^2$ or eight **octants** in $\mathbb{R}^3$. These regions all intersect at the **origin** $O = (0, 0)$ or $O = (0, 0, 0)$ and are bound by **coordinate axes** $\{y = 0\}$ and $\{x = 0\}$ or **coordinate planes** $\{x = 0\}, \{y = 0\}, \{z = 0\}$.

1.2. In $\mathbb{R}^2$ it is custom to orient the x-axis to the "east" and the y-axis to the "north". In $\mathbb{R}^3$, the most common coordinate system is to see the xy-plane as the "ground" and imagine the z-coordinate axes pointing "up". In computer graphics or photography, the xy-plane represents the **retina** or film plate and the z-coordinate measures the distance towards the viewer. In this **photographic coordinate** system, your eyes and chin define the plane $z = 0$ and the nose points in the positive z direction. If the midpoint of your eyes is the origin of the coordinate system and your eyes have the coordinates $(1, 0, 0)$ for the right eye, $(-1, 0, 0)$ for the left eye, then the tip of your nose might have the coordinates $(0, -1, 1)$.

1.3. The **Euclidean distance** between two points $P = (x, y, z)$ and $Q = (a, b, c)$ in space is defined as

Definition: $d(P, Q) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$.

Note that this is a **definition** and not a result. It is motivated by the **theorem of Pythagoras**, but we will **prove** the later result in a moment. This distance is defined in any dimension. In the plane for example the distance of the point (x, y) to (a, b) is $\sqrt{(x-a)^2 + (y-b)^2}$. If we work in $\mathbb{R}^2$, we do not think of it as part of $\mathbb{R}^3$. Coordinates work in arbitrary dimensions. A collection of n data points defines a vector in $\mathbb{R}^n$. Working in Euclidean space $\mathbb{R}^n$ makes sense from a **data scientist point of view**. One can define the Euclidean distance between $x = (x_1, \ldots, x_n)$ and $a = (a_1, \ldots, a_n)$ as $d(x, a)^2 = \sum_{k=1}^n (x_i - a_i)^2$. Having the sum of the squares appears in statistics in **least square problems**.

1.4. **Points**, **curves**, **surfaces** and **solids** are geometric objects which can be described with **functions of several variables**. An example of a curve is a **line**, an example of a surface is a **plane**, an example of a solid is the **ball**, the interior of a **sphere**.

Definition: A **circle** of radius $r \geq 0$ centered at $P = (a, b)$ is the collection of points in $\mathbb{R}^2$ which have distance r from P. A **sphere** of radius ρ centered at $P = (a, b, c)$ is the collection of points in $\mathbb{R}^3$ which have distance $\rho \geq 0$ from P. The equation of a sphere is $(x - a)^2 + (y - b)^2 + (z - c)^2 = \rho^2$.

1.5. When **completing the square** of an equation $x^2 + bx + c = 0$, we add $(b/2)^2 - c$ on both sides of the equation in order to get $(x + b/2)^2 = (b/2)^2 - c$. Solving for x gives $x = -b/2 \pm \sqrt{(b/2)^2 - c}$. This is the **quadratic equation**. Know this equation. You don't want to waste your creative power having to re-derive this again and again.

Examples

1.6. $P = (-2, -3)$ is in the third quadrant of the plane and $P = (1, 2, 3)$ is in the positive octant of space. The point $(0, 0, -8)$ is located on the negative z axis. The point $P = (1, 2, -3)$ is below the xy-plane. Can you spot the point Q on the xy-plane which is closest to P?

1.7. **Problem:** Find the distance midpoint M of $P = (1, 2, 5)$ and $Q = (-3, 4, 7)$ and verify that $d(P, M) + d(Q, M) = d(P, Q)$. **Answer:** The distance is $d(P, Q) = \sqrt{4^2 + 2^2 + 2^2} = \sqrt{24}$. The distance $d(P, M)$ is $\sqrt{2^2 + 1^2 + 1^2} = \sqrt{6}$. The distance $d(Q, M)$ is $\sqrt{2^2 + 1^2 + 1^2} = \sqrt{6}$. Indeed $d(P, M) + d(M, Q) = d(P, Q)$.

1.8. The equation $x^2 + 5x + y^2 - 2y + z^2 = -1$ is after a **completion of the square** $(x + 5/2)^2 - 25/4 + (y - 1)^2 - 1 + z^2 = -1$ or $(x - 5/2)^2 + (y - 1)^2 + z^2 = (5/2)^2$. We see a sphere **center** $(5/2, 1, 0)$ and **radius** $5/2$.

1.9. The distance $d(P, Q) = |x - a| + |y - b|$ in the plane $\mathbb{R}^2$ is called the **taxi metric** or **Manhattan distance**. **Problem**: draw a circle of radius 2. More challenging is to draw an ellipse: the set of points whose sum of the distances from $(-2, 0)$ and $(2, 0)$ is equal to 6. You can do that with a neat geometric construction.

1.10. Draw the unit circle of the **quartic distance** $d(x, y) = (x - a)^4 + (y - b)^4$. More generally, for any $p > 1$, we get a distance $d(x, y) = (x - a)^p + (y - b)^p$. For $p = 1$, it is the **taxi metric**, for $t = 2$ it is the **Euclidean metric**, for $t \to \infty$ it goes to the distance $\max(|x - a|, |y - b|)$ which is the l^∞ metric. **Problem**: is $d(P, Q) = \sqrt{|x - a|} + \sqrt{|y - b|}$ a distance? **Answer**: no, while it satisfies $d(P, Q) = d(Q, P)$ and is zero if and only if $P = Q$, it does not satisfy the **triangle inequality** $d(A, B) + d(B, C) \geq d(A, C)$. We call a space (X, d) for which d is a distance formula satisfying $d(P, Q) = d(Q, P)$, $d(P, Q) = 0 \leftrightarrow P = Q$ and $d(A, B) + d(B, C) \geq d(A, C)$ a **metric space**.

1.11. Problem: Find an algebraic expression for the set of all points for which the sum of the distances to $A = (1,0)$ and $B = (-1,0)$ is equal to 3. **Answer:** Square the equation $\sqrt{(x-1)^2+y^2} + \sqrt{(x+1)^2+y^2} = 3$, separate the remaining single square root on one side and square again. Simplification gives $20x^2 + 36y^2 = 45$ which is equivalent to $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where a, b can be computed as follows: because $P = (a,0)$ satisfies this equation, $d(P,A) + d(P,B) = (a-1) + (a+1) = 3$ so that $a = 3/2$. Similarly, the point $Q = (0,b)$ satisfying it gives $d(Q,A) + d(P,B) = 2\sqrt{b^2+1} = 3$ or $b = \sqrt{5}/2$.

1.12. In an appendix to "Geometry" of his "Discours de la méthode" which appeared in 1637, **René Descartes** promoted the idea to use algebra to solve geometric problems. Even so Descartes mostly dealt with ruler-and compass constructions, the rectangular coordinate system is now called the **Cartesian coordinate system**. His ideas profoundly changed mathematics. But ideas do not grow in a vacuum; Davis and Hersh write that in its current form, Cartesian geometry is due as much to Descartes own contemporaries and successors as to himself. One of the first to explore higher dimensional Euclidean space was Ludwig Schläfli. [1]

1.13. The method of completion of squares is due to **Al-Khwarizmi** who lived from 780-850 and used it as a method to solve quadratic equations. Even so Al-Khwarizmi worked with numerical examples, it is one of the first important steps of algebra. His work *"Compendium on Calculation by Completion and Reduction"* was dedicated to the Caliph **al Ma'mun**, who had established research center called "House of Wisdom" in Baghdad. [2]

1.14. The Euclidean geometry described is only one of many geometries. One can work with more general **metric spaces**. An important class of metric spaces are studied in Riemannian geometry, where the distance between two points can become dependent on where we are. Space becomes curved. This is the frame work of general relativity. Formally, this can happen by changing the coefficients E, G of the metric $d(P,Q)^2 = E(x-a)^2 + G(y-b)^2$. One a sphere, where $x = \theta \in [0, 2\pi]$ is longitude and $y = \phi \in [0,\pi]$ is latitude, one would take $E = \sin^2(y), G = 1$. Two points on the arctic circle with fixed longitude have shorter distance than two points on the equator with the same fixed longitudes. It is important to think now of the surface of the sphere as a space itself, without its embedding in the ambient space. This space is curved. Our four dimensional space-time universe is curved depending on the matter distribution.

Homework

This homework is due on Tuesday, 7/2/2019.

Problem 1.1: Describe in words and draw the objects in $\mathbb{R}^3$.
a) $(z-1)^2 + (x-2)^2 = 1$.
b) $|x-1| + |y-2| + |z-3| = 4$.
c) $x^2y^2z^2 = 0$,
d) $x + 2y + 3z = 6$.
e) $|(x-1,y,z)| - |(x,y-1,z)| = 1$.
f) $x^2 - z^2 = 1$.

[1] An entertaining read is "Descartes secret notebook" by Amir Aczel which deals with an other discovery of Descartes.

[2] The book "The mathematics of Egypt, Mesopotamia, China, India and Islam, by Ed Victor Katz, page 542 contains translations of some of this work.

Problem 1.2: A data point $P = (5, 6, 7)$ gives temperature, rain and wind velocity. We can visualize this in the (t, r, v) space. a) Find the distance of P to the t-axes, where r, v are zero.
b) Find the center of the sphere $t^2 + r^2 - 6r + v^2 = 10$.

Problem 1.3: Verify that the radius of the inscribed circle in a $3 : 4 : 5$ triangle is 1. Here is a possible hint: make a picture of the triangle ABC given by $A = (0, 0), B = (4, 0), C = (0, 3)$, introduce $M = (1, 1)$ then get the coordinates of the points X, Y, Z then compute the distances to verify that the inscribed circle touches the triangle at X, Y, Z.

Problem 1.4: The figure shows rectangles of area 64 and 65 made up of matching pieces. What is going on? It is a famous and classical problem. Try first on your own!

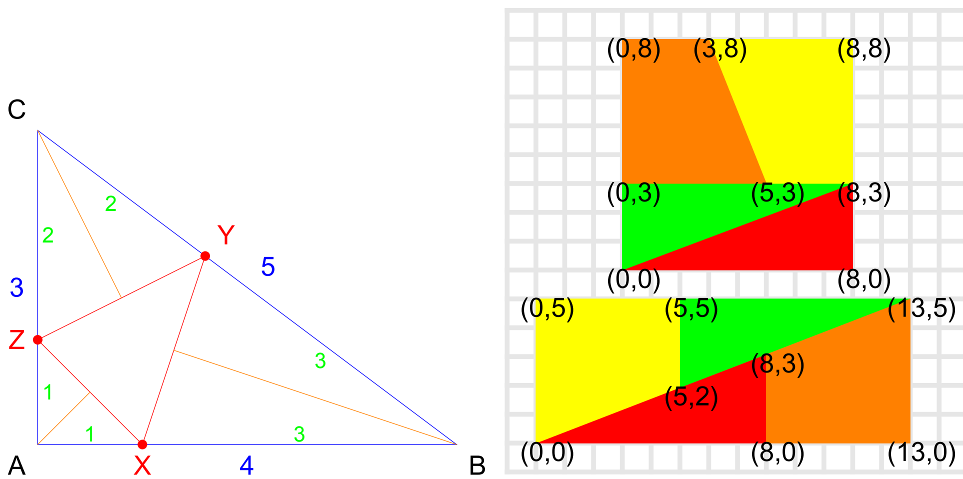


FIGURE 1. The 3-4-5 triangle and the missing square riddle.

Problem 1.5: You play billiard in the table $\{(x, y) \mid 0 \le x \le 4, 0 \le y \le 8\}$. a) Hit the ball at $(3, 2)$ to reach the hole $(4, 8)$ bouncing 3 times at the left wall and three times at the right wall and no other walls. Find the length of the shot.
b) Hit from $(3, 2)$ to reach the hole $(4, 0)$ after hitting twice the left and twice the right wall as well as the top wall $y = 8$ once. What is the length of the trajectory?

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 2: Vectors and dot product

Lecture

2.1. Two points $P = (a, b, c)$ and $Q = (x, y, z)$ in $\mathbf{R}^3$ define a **vector** $\vec{v} = \begin{bmatrix} x-a \\ y-b \\ z-c \end{bmatrix}$.
We simply write also $[x-a, y-b, z-c]^T$ or $\langle x-a, y-b, z-c\rangle$ in order to save space. It is custom in linear algebra to call $[3, 4, 5]$ a row vector and its transpose $[3, 4, 5]^T = \langle 3, 4, 5\rangle$ a **column vector**. As the vector starts at P to Q we write $\vec{v} = \vec{PQ}$. The real numbers numbers p, q, r in $\vec{v} = [p, q, r]^T$ are called the **components** of $\vec{v}$.

2.2. Vectors can be placed **anywhere** in space. [1] Two vectors with the same components are considered **equal**. Vectors can be translated into each other if their components are the same. If a vector $\vec{v}$ starts at the origin $O = (0, 0, 0)$, then $\vec{v} = [p, q, r]^T$ heads to the point (p, q, r). One can therefore identify **points** $P = (a, b, c)$ with **vectors** $\vec{v} = [a, b, c]^T$ attached to the origin. For clarity, we often draw an arrow $\vec{}$ on top of a vector variable and if $\vec{v} = \vec{PQ}$ then P is the "tail" and Q is the "head" of the vector. To distinguish vectors from points, it is custom to write $[2, 3, 4]^T$ or $\langle 2, 3, 4\rangle$ for vectors and $(2, 3, 4)$ for points.

2.3.

> **Definition:** The **sum** of two vectors is $\vec{u} + \vec{v} = [u_1, u_2]^T + [v_1, v_2]^T = [u_1 + v_1, u_2 + v_2]^T$. The **scalar multiple** is $\lambda\vec{u} = \lambda[u_1, u_2]^T = [\lambda u_1, \lambda u_2]^T$. The **difference** $\vec{u} - \vec{v}$ can best be seen as the addition of $\vec{u}$ and $(-1)\cdot\vec{v}$.

Commutativity, associativity, or distributivity rules for vectors are inherited directly from the corresponding rules for numbers.

2.4. The vectors $\vec{i} = [1, 0, 0]^T$, $\vec{j} = [0, 1, 0]^T$, $\vec{k} = [0, 0, 1]^T$ are called **standard basis vectors**. We will avoid this notation mostly but it has historically grown as some notions like the dot and cross product have grown from **quaternions** which are points (t, x, y, z) in $\mathbb{R}^4$ usually written as $t + ix + jy + kz$.

[1] In differential geometry a vector $\vec{v}$ is seen in the so called tangent space to a point P.

2.5.

Definition: The **length** $|\vec{v}|$ of a vector $\vec{v} = \vec{PQ}$ is defined as the distance $d(P,Q)$ from P to Q. A vector of length 1 is called a **unit vector**. If $\vec{v} \neq \vec{0}$, then $\vec{v}/|\vec{v}|$ is called a **direction** of $\vec{v}$. The only vector of length 0 is the 0 vector $[0,0,0]^T$.

2.6.

Definition: The **dot product** of two vectors $\vec{v} = [a,b,c]^T$ and $\vec{w} = [p,q,r]^T$ is defined as $\vec{v} \cdot \vec{w} = ap + bq + cr$.

2.7. Different notations for the dot product are used in different mathematical fields. While mathematicians write $\vec{v} \cdot \vec{w} = (\vec{v}, \vec{w})$, the **Dirac notation** $\langle \vec{v} | \vec{w} \rangle$ is used in quantum mechanics or the **Einstein notation** $v_i w^i$ or more generally $g_{ij} v^i w^j$ in general relativity is used. In statistics, it is called the **covariance** $\mathrm{Cov}[v,w]$ of centered data points. The dot product is also called **scalar product** or **inner product**. It could be generalized. Any product $g(v,w)$ which is linear in v and w and satisfies the symmetry $g(v,w) = g(w,v)$ and $g(v,v) \geq 0$ and $g(v,v) = 0$ if and only if $v = 0$ can be used as a dot product. An example is $g(v,w) = 2v_1w_1 + 3v_2w_2 + 5v_3w_3$.

2.8. The dot product determines distances and distances determines the dot product. **Proof:** Write $v = \vec{v}$. Using the dot product one can express the length of v as $|v| = \sqrt{v \cdot v}$. On the other hand, from $(v+w) \cdot (v+w) = v \cdot v + w \cdot w + 2(v \cdot w)$ can be solved for $v \cdot w$:

$$v \cdot w = (|v+w|^2 - |v|^2 - |w|^2)/2 \ .$$

2.9. The **Cauchy-Schwarz inequality** is

Theorem: $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$.

Proof. If $|w| = 0$, the statement holds as both sides are zero. Otherwise, assume $|w| = 1$ by dividing the equation by $|w|$. Now plug in $a = v \cdot w$ into the equation $0 \leq (v - aw) \cdot (v - aw)$ to get $0 \leq (v - (v \cdot w)w) \cdot (v - (v \cdot w)w) = |v|^2 + (v \cdot w)^2 - 2(v \cdot w)^2 = |v|^2 - (v \cdot w)^2$ which means $(v \cdot w)^2 \leq |v|^2$. □

2.10. Having established this, it is possible to give a definition of what an **angle** is, without referring to any geometric pictures:

Definition: The **angle** between two nonzero vectors $\vec{v}, \vec{w}$ is defined as the unique $\alpha \in [0, \pi]$ which satisfies $\vec{v} \cdot \vec{w} = |\vec{v}| \cdot |\vec{w}| \cos(\alpha)$. Since cos maps $[0, \pi]$ in a 1:1 manner to $[-1, 1]$, this is well defined.

2.11. The **Al Kashi's theorem** gives the third side length c of a triangle ABC in terms of the sides $a = d(B,C), b = d(A,C)$ and α, the angle at the vertex C

Theorem: $a^2 + b^2 = c^2 - 2ab\cos(\alpha)$.

Proof. Define $\vec{v} = \vec{AB}, \vec{w} = \vec{AC}$. Because $c^2 = |\vec{v} - \vec{w}|^2 = (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = |\vec{v}|^2 + |\vec{w}|^2 - 2\vec{v} \cdot \vec{w}$, We know $\vec{v} \cdot \vec{w} = |\vec{v}| \cdot |\vec{w}| \cos(\alpha)$ so that $c^2 = |\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}| \cdot |\vec{w}| \cos(\alpha) = a^2 + b^2 - 2ab\cos(\alpha)$. □

2.12. The **triangle inequality** tells

Theorem: $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$

Proof. $|\vec{u}+\vec{v}|^2 = (\vec{u}+\vec{v}) \cdot (\vec{u}+\vec{v}) = \vec{u}^2 + \vec{v}^2 + 2\vec{u} \cdot \vec{v} \leq \vec{u}^2 + \vec{v}^2 + 2|\vec{u} \cdot \vec{v}| \leq \vec{u}^2 + \vec{v}^2 + 2|\vec{u}| \cdot |\vec{v}| = (|\vec{u}| + |\vec{v}|)^2$. □

Definition: Two vectors are called **orthogonal** or **perpendicular** if $\vec{v} \cdot \vec{w} = 0$. The zero vector $\vec{0}$ is orthogonal to any vector. For example, $\vec{v} = [2, 3]^T$ is orthogonal to $\vec{w} = [-3, 2]^T$.

2.13. We can now prove the **Pythagoras theorem**:

Theorem: If $\vec{v}$ and $\vec{w}$ are orthogonal, then $|v - w|^2 = |v|^2 + |w|^2$.

Proof. $(\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w} + 2\vec{v} \cdot \vec{w} = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w}$. □

2.14.

Definition: The vector $\mathrm{P}(\vec{v}) = \frac{\vec{v} \cdot \vec{w}}{|\vec{w}|^2} \vec{w}$ is called the **projection** of $\vec{v}$ onto $\vec{w}$. The **scalar projection** $\frac{\vec{v} \cdot \vec{w}}{|\vec{w}|}$ is a signed length of the vector projection. Its absolute value is the length of the projection of $\vec{v}$ onto $\vec{w}$. The vector $\vec{b} = \vec{v} - P(\vec{v})$ is a vector orthogonal to the $\vec{w}$-direction.

2.15. The projection allows to **visualize** the dot product. The absolute value of the dot product is the length of the projection. The dot product is positive if v points more towards to w, it is negative if v points away from it. In the next lecture we use the projection to compute distances between various objects.

Examples

2.16. For example, with $\vec{v} = [0, -1, 1]^T$, $\vec{w} = [1, -1, 0]^T$, $\mathrm{P}(\vec{v}) = [1/2, -1/2, 0]^T$. Its length is $1/\sqrt{2}$.

2.17. The RGB **color space** consists of triples $\vec{v} = [r, g, b]^T$ describing the amount of red, green and blue of a **color**. An other coordinate system is the CMY **color space** consisting of triples $\vec{v} = [c, m, y]^T = [1-r, 1-g, 1-b]^T$, where c is **cyan**, m is **magenta** and y is **yellow**.

2.18. In physics, forces and fields $\vec{F}$ are described by vectors. The **velocity** of a curve $r(t) = [x(t), y(t), z(t)]$ is a vector attached to the point $r(t)$.

2.19. In probability theory, data are described by vectors. One calls them also **random variables**. It is in statistics, where higher dimensional spaces appear.

Homework

This homework is due on Tuesday, 7/2/2019.

Problem 2.1: Find a **unit vector** parallel to $\vec{u} + 2\vec{v} + 4\vec{w}$ if $\vec{u} = [-10, 2, 9]^T$ and $\vec{v} = [1, 1, 3]^T$ and $\vec{w} = [3, 1, 1]^T$.

Problem 2.2: An **Euler brick** is a **cuboid** with side lengths a, b, c such that all face diagonals are integers.
a) Verify that $\vec{v} = [a, b, c]^T = [275, 252, 240]^T$ is a vector which leads to an Euler brick. Halcke found the first one in 1719.
b) (*) Verify that $[a, b, c]^T = [u(4v^2 - w^2), v(4u^2 - w^2), 4uvw]^T$ leads to an Euler brick if $u^2 + v^2 = w^2$.
(Sounderson 1740) If also the space diagonal $\sqrt{a^2 + b^2 + c^2}$ is an integer, an Euler brick is called **perfect**. Nobody has found one, nor proven that it can not exist.

Problem 2.3: **Colors** are encoded by vectors $\vec{v} = [$ red , green , blue $]^T$. The red, green and blue components of $\vec{v}$ are all real numbers in the interval $[0, 1]$.
a) Determine the angle between the colors yellow and magenta.
b) What is the vector projection of the magenta-orange mixture $\vec{x} = (\vec{v} + \vec{w})/2$ onto green $\vec{y}$?

Problem 2.4: A rope is wound exactly 5 times around a stick of circumference 1 and length 12. How long is the rope?

Problem 2.5: a) Find the angle between the main diagonal of the unit cube and one of the face diagonals. Assume that both diagonals pass through a common vertex.
b) Find the vector projection of the main diagonal $\vec{v} = [1, 1, 1]^T$ onto the side diagonal $\vec{w} = [1, 1, 0]^T$.
c) Find the scalar projection of $\vec{v}$ on $\vec{w}$.

2.20. We live in a time, where **data** are increasingly important. This is good news for multi-variable calculus, as data points are usually given as points in an Euclidean space and analyzed using tools of multi-variable calculus. There are various ways how data can be stored, in a computer as lists of **bits**, in a quantum computer as a list of **qbits**, in a **relational database** as a list of tables, in a **graph data base** as a list of graphs. A sort of graph database has been designed already by the Incas in the form of **Khipu**, which are also called "talking knots". In a picture, data are color values attached in an array, a **song** is an array of amplitudes, a movie is an array of pictures, which each is an array of color vectors. It does not matter, in the end, we can store information as a point in a Euclidean space. The Harvard Khipu data base for example is a standard relational database encoding the three dimensional knots which are still available. We look here at an example, which illustrates the **story of data**:

2.21. The next table shows some data from the 2018 rankings of the top universities. The data were obtained from the website

$$https://www.topuniversities.com/university-rankings/world-university-rankings/2018$$

There are different such rankings. It is important that you are aware about how arbitrary such a "ranking" can be done. This is especially if you think about how the data were actually obtained. Which of the data can be considered objectively reproducible, which ones are more subjective?

2.22. The following concrete data provide for each institution a vector:

[Overall,Academic,Employer,Faculty/Student ratio,Citations,Int. faculty, Int. students]

MIT	100	100	100	100	99.8	100	95.5
Stanford	98.6	100	100	100	99	99.8	70.5
Harvard	98.5	100	100	99.3	99.8	92.1	75.7
Caltech	97.2	98.7	81.2	100	100	96.8	90.3
Oxford	96.8	100	100	100	83	99.6	98.8
Cambridge	95.6	100	100	100	77.2	99.4	97.9
ETH	95.3	98.2	96.2	82.4	98.7	100	98.6
Imperial	93.3	98.7	99.9	99.9	67.8	100	100
Chicago	93.2	99.6	90.7	97.4	83.6	74.2	82.5
UCL	92.9	99.3	99.2	99.2	66.2	98.7	100

2.23. Can you figure out how the first entry is computed? It is an average of the 6 entries
$\vec{x} = [Academic, Faculty - StudentRatio, Citations, International, IntStudents]$.
There is a mystery vector $\vec{v} = [a, b, c, d, e, f]$ which averages the ranking as a **dot product** of $\vec{x}$ with $\vec{v}$. Check also

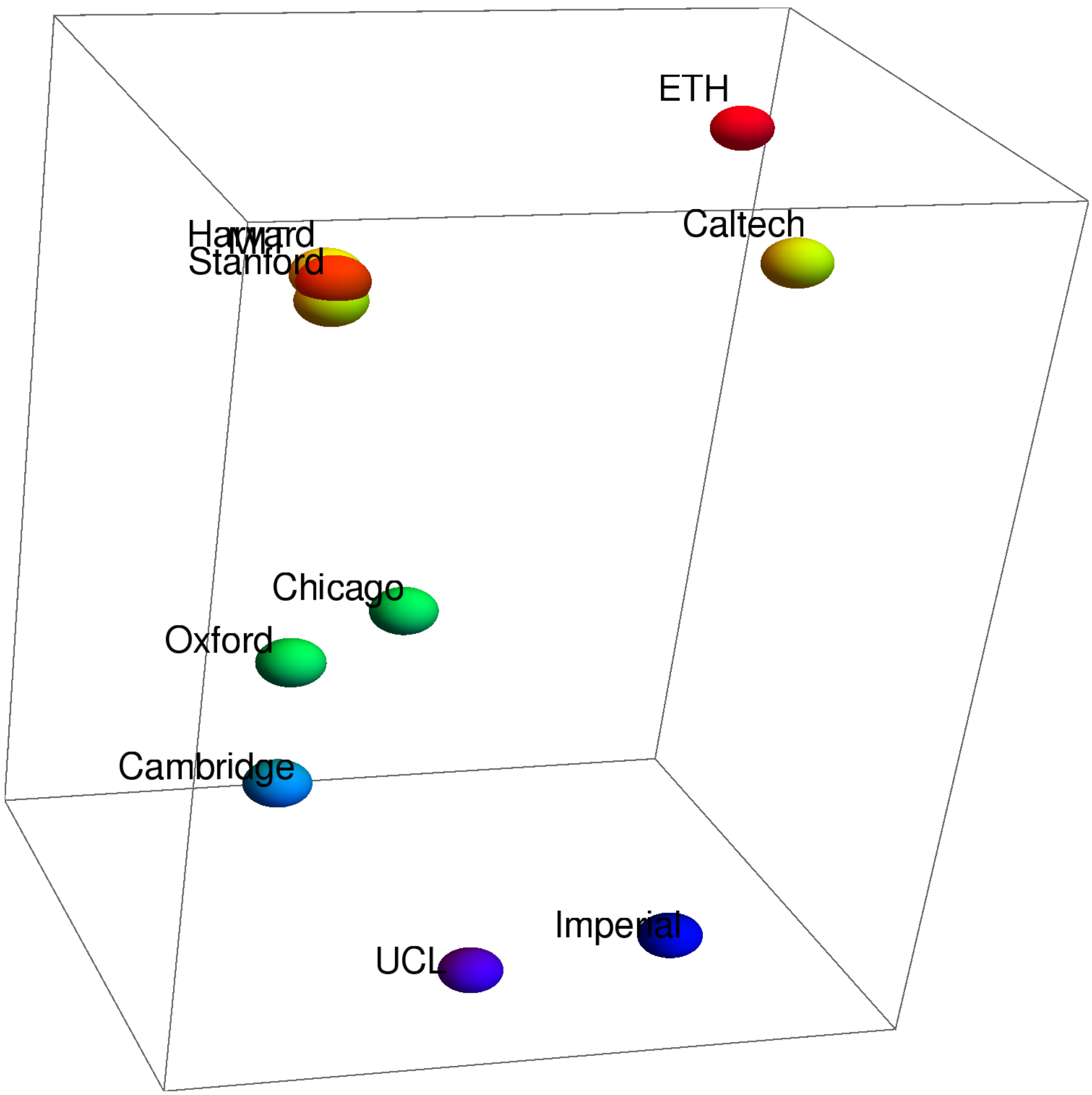
$$https://www.topuniversities.com/qs-world-university-rankings/methodology.$$

2.24. Obviously, the overall ranking depends on the key. There is a ranking key so that the university UCL is on the top for example. Can you find one? This requires to find a ranking vector $\vec{v}$ for which $\vec{x} \cdot \vec{v}$ is maximal for UCL.

2.25. How can we **visualize data**? In the next figure, we plot three relevant data points given by

$$\vec{x} = [Academic, Faculty - StudentRatio, Citations]$$

By comparing the data with the plot, can you figure out, what each of the coordinate axes is? Also this visual representation produces some kind of ranking. Being a graduate of ETH myself and being at Harvard, having been at Caltech, the picture has been turned so that these universities look particularly good ... It is important that you are aware of such manipulations. They are everywhere!



The mystery vector is $\vec{v} = [0.4, 0.1, 0.2, 0.2, 0.05, 0.05]$.

Unit 3: Cross product

Lecture

3.1. The **cross product** of two vectors $\vec{v} = [v_1, v_2]^T$ and $\vec{w} = [w_1, w_2]^T$ in the plane is the **scalar** $\vec{v} \times \vec{w} = v_1 w_2 - v_2 w_1$. To remember this, you can write it as a determinant of a 2×2 matrix $A = \begin{bmatrix} v_1 & v_2 \\ w_1 & w_2 \end{bmatrix}$, which is the product of the diagonal entries minus the product of the side diagonal entries.

3.2.

Definition: The **cross product** of two vectors $\vec{v} = [v_1, v_2, v_3]^T$ and $\vec{w} = [w_1, w_2, w_3]^T$ in space is defined as the **vector**

$$\vec{v} \times \vec{w} = [v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1]^T \,.$$

To remember this, we can write the product as a "determinant":

$$\begin{bmatrix} i & j & k \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} i & & \\ & v_2 & v_3 \\ & w_2 & w_3 \end{bmatrix} - \begin{bmatrix} & j & \\ v_1 & & v_3 \\ w_1 & & w_3 \end{bmatrix} + \begin{bmatrix} & & k \\ v_1 & v_2 & \\ w_1 & w_2 & \end{bmatrix}$$

which is $\vec{i}(v_2 w_3 - v_3 w_2) - \vec{j}(v_1 w_3 - v_3 w_1) + \vec{k}(v_1 w_2 - v_2 w_1)$ using the notation $\vec{i} = [1, 0, 0]$, $\vec{j} = [0, 1, 0]$ and $\vec{k} = [0, 0, 1]$.

3.3. Examples: the cross product of $[1, 2]^T$ and $[4, 5]^T$ is $5 - 8 = -3$. The cross product of $[1, 2, 3]^T$ and $[4, 5, 1]^T$ is $[-13, 11, -3]^T$. The cross product is clearly anti-commutative: $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$.

Theorem: In $\mathbb{R}^3$, the vector $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$ and has length $|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}| \sin(\alpha)$.

Proof. To see the orthogonality, verify for example that $\vec{v} \cdot (\vec{v} \times \vec{w}) = 0$. To check the length formula, we use the **Lagrange's identity** $|\vec{v} \times \vec{w}|^2 = |\vec{v}|^2 |\vec{w}|^2 - (\vec{v} \cdot \vec{w})^2$ which is also called **Cauchy-Binet** formula. We will do that by direct computation in class. To finish up, use $|\vec{v} \cdot \vec{w}| = |\vec{v}||\vec{w}| \cos(\alpha)$. $\square$

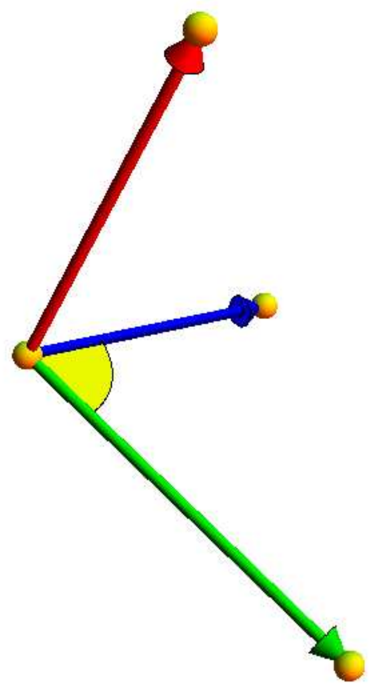


FIGURE 1. The cross product produces a vector perpendicular to two vectors. The length of the vector is the area of the parallelogram.

3.4. The statement can intuitively also to be seen by choosing a coordinate system in which the vectors are given as In that special case $\vec{v} = [a, 0, 0]^T$ and $\vec{w} = [b\cos(\alpha), b\sin(\alpha), 0]^T$, we have $\vec{v} \times \vec{w} = [0, 0, ab\sin(\alpha)]^T$ which has length $|ab\sin(\alpha)|$. This argument however assumes that the cross product does not change, if we change the coordinate system.

3.5. The absolute value respectively length $|\vec{v} \times \vec{w}|$ defines the **area of the parallelogram** spanned by $\vec{v}$ and $\vec{w}$. As stated as a **definition**, nothing needs to be proven. The definition fits with our common intuition we have about area because $|\vec{w}|\sin(\alpha)$ is the height of the parallelogram with base length $|\vec{v}|$.

3.6. The **trigonometric sin-formula** relates the side lengths a, b, c and angles α, β, γ of a general triangle:

Theorem: $\frac{a}{\sin(\alpha)} = \frac{b}{\sin(\beta)} = \frac{c}{\sin(\gamma)}$.

Proof. We can express the area of the triangle in three different ways:

$$ab\sin(\gamma) = bc\sin(\alpha) = ac\sin(\beta) .$$

Divide the first equation by $\sin(\gamma)\sin(\alpha)$ to get one identity. Divide the second equation by $\sin(\alpha)\sin(\beta)$ to get the second identity. □

3.7. It follows from the sin-formula and the fact that $\sin(\alpha) = 0$ if $\alpha = 0$ or $\alpha = \pi$ that $\vec{v} \times \vec{w}$ is zero if and only if $\vec{v}$ and $\vec{w}$ are **parallel**, that is if $\vec{v} = \lambda\vec{w}$ for some real λ. The cross product can therefore be used to check whether two vectors are parallel or not. Note that v and $-v$ are considered parallel even so sometimes the notion **anti-parallel** is used.

3.8.

Definition: The scalar $[\vec{u}, \vec{v}, \vec{w}] = \vec{u} \cdot (\vec{v} \times \vec{w})$ is called the **triple scalar product** of $\vec{u}, \vec{v}, \vec{w}$. The absolute value of $[\vec{u}, \vec{v}, \vec{w}]$ defines the **volume of the parallelepiped** spanned by $\vec{u}, \vec{v}, \vec{w}$. The **orientation** of three vectors is defined as the sign of $[\vec{u}, \vec{v}, \vec{w}]$. It is positive if the three vectors define a **right-handed** coordinate system.

3.9. Again, there was no need to prove anything because we **defined** volume and orientation. Why does this fit with our intuition? The value $h = |\vec{u} \cdot \vec{n}|/|\vec{n}|$ is the height of the parallelepiped if $\vec{n} = (\vec{v} \times \vec{w})$ is a normal vector to the ground parallelogram of area $A = |\vec{n}| = |\vec{v} \times \vec{w}|$. The volume of the parallelepiped is $hA = (\vec{u} \cdot \vec{n}/|\vec{n}|)|\vec{v} \times \vec{w}|$ which simplifies to $\vec{u} \cdot \vec{n} = |(\vec{u} \cdot (\vec{v} \times \vec{w})|$ which is the absolute value of the triple scalar product. The vectors $\vec{v}, \vec{w}$ and $\vec{v} \times \vec{w}$ form a **right handed coordinate system**. If the first vector $\vec{v}$ is your thumb, the second vector $\vec{w}$ is the pointing finger then $\vec{v} \times \vec{w}$ is the third middle finger of the right hand. For example, the vectors $\vec{i}, \vec{j}, \vec{i} \times \vec{j} = \vec{k}$ form a right handed coordinate system. Since the triple scalar product is linear with respect to each vector, we also see that volume is additive. Adding two equal parallelepipeds together for example gives a parallelepiped with twice the volume.

Examples

3.10. Problem: Find the volume of the parallelepiped which has the vertices $O = (1, 1, 0), P = (2, 3, 1), Q = (4, 3, 1), R = (1, 4, 1)$. **Answer**: We first see that the solid is spanned by the vectors $\vec{u} = [1, 2, 1]^T$, $\vec{v} = [3, 2, 1]^T$, and $\vec{w} = [0, 3, 1]^T$. We get $\vec{v} \times \vec{w} = [-1, -3, 9]^T$ and $\vec{u} \cdot (\vec{v} \times \vec{w}) = 2$. The volume is 2.

3.11. Problem: Two apples have the same shape, but one has a 3 times larger diameter. What is their weight ratio? **Answer**. For a parallelepiped spanned by $[a, 0, 0]^T$ $[0, b, 0]^T$ and $[0, 0, c]^T$, the volume is the triple scalar product abc. If a, b, c are all tripled, the volume gets multiplied by a factor 27. Now cut each apple into the same amount of parallelepipeds, the larger one with slices 3 times as large too. Since each of the pieces has 27 times the volume, also the apple is 27 times heavier!

3.12. Problem. A **3D scanner** is used to build a 3D model of a face. It detects a triangle which has its vertices at $P = (0, 1, 1), Q = (1, 1, 0)$ and $R = (1, 2, 3)$. Find the area of the triangle. **Solution.** We have to find the length of the cross product of $\vec{PQ}$ and $\vec{PR}$ which is $[1, -3, 1]^T$. The length is $\sqrt{11}$.

3.13. Problem. The scanner now detects an other point $A = (1, 1, 1)$. On which side of the triangle is it located if the cross product of $\vec{PQ}$ and $\vec{PR}$ is considered the direction "up". **Solution.** The cross product is $\vec{n} = [1, -3, 1]^T$. We have to see whether the vector $\vec{PA} = [1, 0, 0]^T$ points into the direction of $\vec{n}$ or not. To see that, we have to form the dot product. It is 1 so that indeed, A is "above" the triangle. Note that a triangle in space a priori does not have an orientation. We have to tell, what direction is "up". That is the reason that file formats for 3D printing like contain the data for three points in space as well as a vector, telling the direction.

Homework

This homework is due on Tuesday, 7/2/2019.

Problem 3.1: a) Find a unit vector perpendicular to the space diagonal $[1, 1, 1]^T$ and the face diagonal $[1, 1, 0]^T$ of the cube.
b) Find the volume of the parallelepiped for which the base parallelogram is given by the points $P = (5, 2, 2), Q = (3, 1, 2), R = (1, 4, 2), S = (-1, 3, 2)$ and which has an edge connecting P with $T = (5, 6, 8)$.
c) Find the area of the base and use b) to get the height of the parallelepiped.

Problem 3.2: a) Assume $\vec{u}+\vec{v}+\vec{w} = \vec{0}$. Verify that $\vec{u}\times\vec{v} = \vec{v}\times\vec{w} = \vec{w}\times\vec{u}$.
b) Find $(\vec{u} + \vec{v}) \cdot (\vec{v} \times \vec{w})$ if $\vec{u}, \vec{v}, \vec{w}$ are unit vectors which are orthogonal to each other and $\vec{u} \times \vec{v} = \vec{w}$.

Problem 3.3: To find the equation $ax + by + cz = d$ for the plane which contains the point $P = (1, 2, 3)$ as well as the line which passes through $Q = (3, 4, 4)$ and $R = (1, 1, 2)$, we find a vector $[a, b, c]^T$ normal to the plane and fix d so that P is in the plane.

Problem 3.4: Verify the "BAC minus CAB" formula (due to Lagrange) $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$ for general vectors $\vec{a}, \vec{b}, \vec{c}$ in space.

Problem 3.5: A product $*$ is said to satisfy the **cancellation property** if for all $x, y, z \neq 0$: $x * z = y * z$ implies that $x = y$.
a) Does the dot product satisfy the cancellation property?
b) Does the cross product satisfy the cancellation property?

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 4: Lines and Planes

Lecture

4.1. A point $P = (p, q, r)$ and a vector $\vec{v} = [a, b, c]^T$ define the **line**

$$L = \{ \begin{bmatrix} p \\ q \\ r \end{bmatrix} + t \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \ t \in \mathbb{R} \} .$$

The line consists of all points obtained by adding a multiple of the vector $\vec{v} = [a, b, c]^T$ to the vector $\vec{OP} = [p, q, r]^T$. It contains the point P as well as a copy of $\vec{v} = \vec{PQ}$ attached to P. Every vector contained in the line is necessarily parallel to $\vec{v}$. We think about the parameter t as "time". At $t = 0$, we are at the end point P of $\vec{OP}$ and at $t = 1$, we are at the end point Q of $\vec{OQ} = \vec{OP} + \vec{v}$.

4.2. If t is restricted to values in a **parameter interval** $[t_1, t_2]$, then $L = \{[p, q, r]^T + t[a, b, c]^T, \ t_1 \leq t \leq t_2 \}$ is a **line segment** which connects $\vec{r}(t_1)$ with $\vec{r}(t_2)$. For example, to get the line through $P = (1, 1, 2)$ and $Q = (2, 4, 6)$, form the vector $\vec{v} = \vec{PQ} = [1, 3, 4]^T$ and get $L = \{[x, y, z]^T = [1, 1, 2]^T + t[1, 3, 4]^T ; \}$. This can be written also as $\vec{r}(t) = [1 + t, 1 + 3t, 2 + 4t]^T$. If we write $[x, y, z]^T = [1, 1, 2]^T + t[1, 3, 4]^T$ as a collection of equations $x = 1 + 2t, y = 1 + 3t, z = 2 + 4t$ and solve the first equation for t:

$$L = \{(x, y, z) \mid (x - 1)/2 = (y - 1)/3 = (z - 2)/4 \} .$$

4.3. The line $\vec{r} = \vec{OP} + t\vec{v}$ defined by $P = (p, q, r)$ and vector $\vec{v} = [a, b, c]^T$ with nonzero a, b, c satisfies the **symmetric equations**

$$\frac{x - p}{a} = \frac{y - q}{b} = \frac{z - r}{c} .$$

The reason is that each of these expressions is equal to t. These symmetric equations have to be modified a bit one or two of the numbers a, b, c are zero. If $a = 0$, replace the first equation with $x = p$, if $b = 0$ replace the second equation with $y = q$ and if $c = 0$ replace third equation with $z = r$. The interpretation is that the line is written as an intersection of two planes.

4.4. A point P and two vectors $\vec{v}, \vec{w}$ define a **plane** $\Sigma = \{\vec{OP} + t\vec{v} + s\vec{w}$, where t, s are real numbers $\}$.
An example is $\Sigma = \{[x, y, z]^T = [1, 1, 2]^T + t[2, 4, 6]^T + s[1, 0, -1]^T \}$. This is called the **parametric description** of a plane.

4.5. If a plane contains the two vectors $\vec{v}$ and $\vec{w}$, then the vector $\vec{n} = \vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$. Because also the vector $\vec{PQ} = \vec{OQ} - \vec{OP}$ is perpendicular to $\vec{n}$, we have $(Q - P) \cdot \vec{n} = 0$. With $Q = (x_0, y_0, z_0)$, $P = (x, y, z)$, and $\vec{n} = [a, b, c]^T$, this means $ax + by + cz = ax_0 + by_0 + cz_0 = d$. The plane is therefore described by a single equation $ax + by + cz = d$. We have shown:

Theorem: The equation for a plane containing $\vec{v}$ and $\vec{w}$ and a point P is $ax + by + cz = d$, where $[a, b, c]^T = \vec{v} \times \vec{w}$ and where d is obtained by plugging in P.

4.6. Problem: Find the equation of a plane which contains the three points $P = (-1, -1, 1), Q = (0, 1, 1), R = (1, 1, 3)$.
Answer: The plane contains the two vectors $\vec{v} = \vec{PQ} = [1, 2, 0]^T$ and $\vec{w} = \vec{PR} = [2, 2, 2]^T$. The normal vector $\vec{n} = \vec{v} \times \vec{w} = [4, -2, -2]^T$ leads to the equation $4x - 2y - 2z = d$. The constant d is obtained by plugging in the coordinates of one of the points. In our case, it is $4x - 2y - 2z = -4$.

4.7. Problem: Find the angle between the planes $x + y = -1$ and $x + y + z = 2$, The **angle between the two planes** $ax + by + cz = d$ and $ex + fy + gz = h$ is defined as the angle between the two normal vectors $\vec{n} = [a, b, c]^T$ and $\vec{m} = [e, f, g]^T$.
Answer: find the angle between $\vec{n} = [1, 1, 0]^T$ and $\vec{m} = [1, 1, 1]^T$. It is $\arccos(2/\sqrt{6})$.

Examples

4.8. To practice the concepts, we look at **distance formulas**.

1) If P is a point and $\Sigma : \vec{n} \cdot \vec{x} = d$ is a plane containing a point Q, then

$$d(P, \Sigma) = \frac{|\vec{PQ} \cdot \vec{n}|}{|\vec{n}|}$$

is the distance between P and the plane. Proof: use the angle formula in the denominator. For example, to find the distance from $P = (7, 1, 4)$ to $\Sigma : 2x + 4y + 5z = 9$, we find first a a point $Q = (0, 1, 1)$ on the plane. Then compute

$$d(P, \Sigma) = \frac{|[-7, 0, -3]^T \cdot [2, 4, 5]^T|}{|[2, 4, 5]^T|} = \frac{29}{\sqrt{45}} .$$

2) If P is a point in space and L is the line $\vec{r}(t) = Q + t\vec{u}$, then

$$d(P, L) = \frac{|(\vec{PQ}) \times \vec{u}|}{|\vec{u}|}$$

is the distance between P and the line L. Proof: the area divided by base length is height of parallelogram. For example, to compute the distance from $P = (2, 3, 1)$ to the line $\vec{r}(t) = (1,1,2) + t(5,0,1)$, compute

$$d(P,L) = \frac{|[-1,-2,1]^T \times [5,0,1]^T|}{[5,0,1]^T} = \frac{|[-2,6,10]^T|}{\sqrt{26}} = \frac{\sqrt{140}}{\sqrt{26}} .$$

3) If L is the line $\vec{r}(t) = Q + t\vec{u}$ and M is the line $\vec{s}(t) = P + t\vec{v}$, then

$$d(L,M) = \frac{|(\vec{PQ}) \cdot (\vec{u} \times \vec{v})|}{|\vec{u} \times \vec{v}|}$$

is the distance between the two lines L and M. Proof: the distance is the length of the vector projection of $\vec{PQ}$ onto $\vec{u} \times \vec{v}$ which is normal to both lines. For example, to compute the distance between $\vec{r}(t) = (2,1,4) + t(-1,1,0)$ and M is the line $\vec{s}(t) = (-1,0,2) + t(5,1,2)$ form the cross product of $[-1,1,0]^T$ and $[5,1,2]^T$ is $[2,2,-6]^T$. The distance between these two lines is

$$d(L,M) = \frac{|(3,1,2) \cdot (2,2,-6)|}{|[2,2,-6]^T|} = \frac{4}{\sqrt{44}} .$$

4) To get the distance between two planes $\vec{n} \cdot \vec{x} = d$ and $\vec{n} \cdot \vec{x} = e$, then their distance is

$$d(\Sigma, \Pi) = \frac{|e-d|}{|\vec{n}|}$$

Non-parallel planes have distance 0. Proof: use the distance formula between point and plane. For example, $5x + 4y + 3z = 8$ and $10x + 8y + 6z = 2$ have the distance

$$\frac{|8-1|}{|[5,4,3]^T|} = \frac{7}{\sqrt{50}} .$$

FIGURE 1. The **global positioning system** GPS uses the fact that a receiver can get the difference of distances to two satellites.

Homework

This homework is due on Tuesday, 7/2/2019.

Problem 4.1: Given the three points $P = (7, 4, 5)$ and $Q = (1, 3, 9)$ and $R = (4, 2, 10)$. find the parametric and symmetric equation for the line perpendicular to the triangle PQR passing through its center of mass $(P + Q + R)/3 = (4, 3, 8)$.

Problem 4.2: A regular tetrahedron has vertices at the points $P_1 = (0, 0, 6)$,$P_2 = (0, \sqrt{32}, -2)$, $P_3 = (-\sqrt{24}, -\sqrt{8}, -2)$ and $P_4 = (\sqrt{24}, -\sqrt{8}, -2)$. Find the distance between two edges which do not intersect.

Problem 4.3: Find a parametric equation for the line through the point $P = (3, 1, 2)$ that is perpendicular to the line $L : x = 1+4t, y = 1-4t, z = 8t$ and intersects this line in a point Q.

Problem 4.4: Given three spheres of radius 9 centered at $A = (1, 2, 0), B = (4, 5, 0), C = (1, 3, 2)$. Find a plane $ax + by + cz = d$ which touches all of three spheres from the same side.

Problem 4.5: a) Find the distance between the point $P = (3, 3, 4)$ and the line $2x = 2y = 2z$.
b) Parametrize the line $\vec{r}(t) = [x(t), y(t), z(t)]^T$ in a) and find the minimum of the function $f(t) = d(P, \vec{r}(t))^2$. Verify that the minimal value agrees with a).

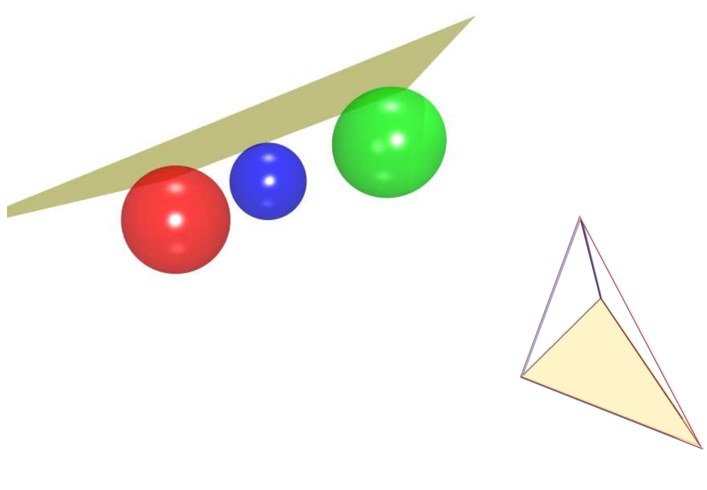


Figure 2. The sphere problem and the the tetrahedron.

Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2019

MULTIVARIABLE CALCULUS

MATH S-21A

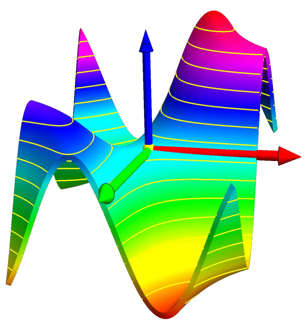
Unit 5: Functions

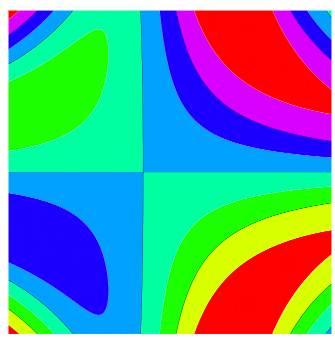
Lecture

5.1.

> **Definition:** A **function of two variables** $f(x,y)$ is a rule which assigns to two numbers x,y a third number $f(x,y)$.

The function $f(x,y) = x^5y - 2y^2$ for example assigns to $(2,3)$ the number $96 - 18 = 78$.





5.2. In general, a function is assumed to be defined for all points (x,y) in $\mathbb{R}^2$. An example is $f(x,y) = |x|^7 + e^{\sin(xy)}$. Sometimes however it is required to restrict the function to a **domain** D in the plane. For example, if $f(x,y) = \log|y| + \sqrt{x}$, then (x,y) is only defined for $x > 0$ and for $y \neq 0$. The **range** of a function f is the set of values which the function f reaches. The function $f(x,y) = 3 + x^2/(1+x^2)$ for example takes all values $3 \leq z < 4$. While $z = 3$ is reached for $x = y = 0$, the value $z = 4$ is not attained.

5.3.

> **Definition:** The set $\{(x,y,f(x,y)) \mid (x,y) \in D\} \subset \mathbb{R}^3$ is the **graph** of f.

Graphs are **surfaces** which allow to see the function visually. It is important not to mix up the graph of a function with the function itself. The function is a **rule** which assigns to (x,y) a third number, the graph is a **geometric object** in three dimensional space. The modern point of view of functions only came at the beginning of the 19th century. It might look irrelevant for you but it is like in computer science, where we distinguish different data structures, even so one can bring one into the other.

5.4. Here are some examples

Example $f(x,y)$	domain D of f	range = $f(D)$ of f
$\sqrt{9-x^2-y^2}$	closed disc $x^2+y^2 \leq 9$	$[0,3]$
$-\log(1-x^2-y^2)$	open unit disc $x^2+y^2<1$	$(0,\infty)$
$f(x,y)=x^2+y^3-xy+\cos(xy)$	plane R^2	the real line
$\sqrt{4-x^2-2y^2}$	$x^2+2y^2 \leq 4$	$[0,2]$
$1/(x^2+y^2-1)$	all except unit circle	$R \setminus (1,0]$
$1/(x^2+y^2)^2$	all except origin	positive real axis

5.5.

Definition: The set $\{(x,y) \mid f(x,y)=c=\text{const}\ \}$ is called a **contour curve** or **level curve** of f. A collection of contour curves is a **contour map**.

For example, for $f(x,y)=4x^2+3y^2$, the level curves $f=c$ are **ellipses** if $c>0$. Drawing several contour curves $\{f(x,y)=c\ \}$ simultaneously produces a **contour map** of the function f.

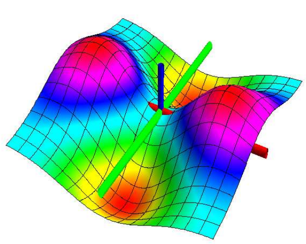
5.6. Level curves allow to visualize and analyze functions $f(x,y)$ without leaving the two dimensional space. The picture below for example shows the level curves of the function $\sin(xy)-\sin(x^2+y)$. Contour curves are everywhere: they appear as **isobars**=curves of constant pressure, or **isoclines**= curves of constant (wind) field direction, **isothermes**= curves of constant temperature or **isoheights** =curves of constant height.

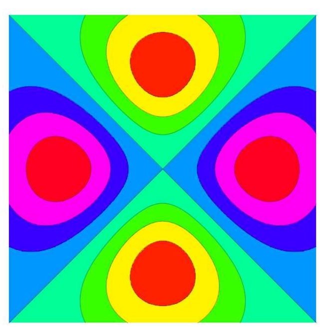
Examples

5.7. For $f(x,y)=x^2-y^2$, the set $x^2-y^2=0$ is the union of the lines $x=y$ and $x=-y$. The set $x^2-y^2=1$ consists of two hyperbola with with their "noses" at the point $(-1,0)$ and $(1,0)$. The set $x^2-y^2=-1$ consists of two hyperbola with their noses at $(0,1)$ and $(0,-1)$.

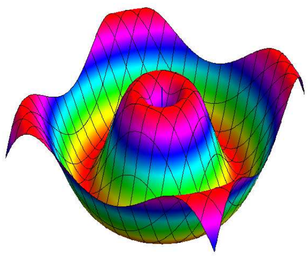
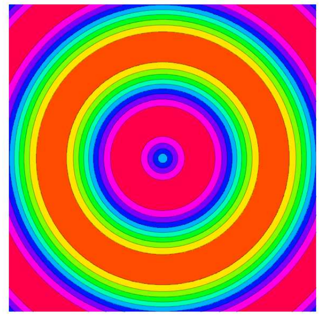
5.8. The function $f(x,y)=1-2x^2-y^2$ has contour curves $f(x,y)=1-2x^2+y^2=c$ which are ellipses $2x^2+y^2=1-c$ for $c<1$.

5.9. For the function $f(x,y)=(x^2-y^2)e^{-x^2-y^2}$, we can not find explicit expressions for the contour curves $(x^2-y^2)e^{-x^2-y^2}=c$. We can draw the traces curve $(x,0,f(x,0))$ or $(0,y,f(0,y)$ or then use a computer:





5.10. The surface $z = f(x,y) = \sin(\sqrt{x^2+y^2})$ has concentric circles as contour curves.

5.11. In applications, discontinuous functions can occur. The temperature of water in relation to pressure and volume is an example. One experiences then **phase transitions**, places where the function value can jumps. Mathematicians study singularities in a mathematical field called "catastrophe theory".

5.12.

Definition: A function $f(x,y)$ is called **continuous** at (a,b) if there is a finite value $f(a,b)$ with $\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$. This means that for any sequence (x_n,y_n) converging to (a,b), also $f(x_n,y_n) \to f(a,b)$. A function is **continuous** in $G \subset \mathbb{R}^2$ if it is continuous at every point (a,b) of G.

5.13. Continuity means that if (x,y) is close to (a,b), then $f(x,y)$ must be close to $f(a,b)$. Continuity for functions of more than two variables is defined in the same way. The bad news is that continuity is not always easy to check. The good news is that in general we do not have to worry about continuity. Lets look at some examples:

5.14. Example: For $f(x,y) = (xy)/(x^2+y^2)$, we have

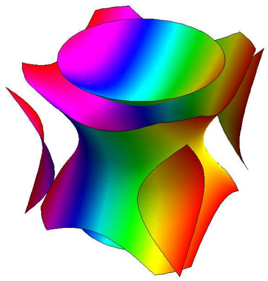
$$\lim_{(x,x)\to(0,0)} f(x,x) = \lim_{x\to 0} \frac{x^2}{2x^2} = \frac{1}{2}$$

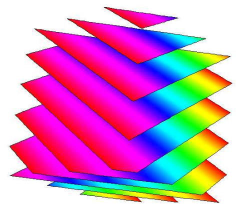
and $\lim_{(x,0)\to(0,0)} f(0,x) = \lim_{(x,0)\to(0,0)} 0 = 0$. The function is not continuous at $(0,0)$.

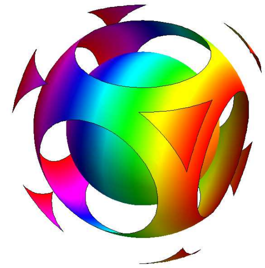
5.15. Example: The function $f(x,y) = (x^2y)/(x^2+y^2)$ is better described using polar coordinates: $f(r,\theta) = r^3\cos^2(\theta)\sin(\theta)/r^2 = r\cos^2(\theta)\sin(\theta)$. We see that $f(r,\theta) \to 0$ uniformly in θ if $r \to 0$. The function is continuous as we can extend it and extend the value to $f(0,0) = 0$. It is custom in mathematics to consider the above function **to be continuous**. The reason is that there is a **unique way** to give a function value at the undefined point.

5.16. A simpler example: the function $f(x,y) = (x^2-y^2)/(x+y)$ is continuous everywhere. Yes, the function is not defined a priori at $x+y=0$ but as it is outside this line equal to $f(x,y) = x-y$, there is a unique continuation to the entire plane and this continuation is $x-y$ whi

5.17. A function of three variables $g(x, y, z)$ assigns to three variables x, y, z a real number $g(x, y, z)$. The function $f(x, y, z) = x^2 + y - z$ for example satisfies $f(3, 2, 1) = 10$. We can visualize a function by **contour surfaces** $g(x, y, z) = c$, where c is constant. It is an **implicit description** of the surface. The contour surface of $g(x, y, z) = x^2 + y^2 + z^2 = c$ is a sphere if $c > 0$. To understand a contour surface, it is helpful to look at the **traces**, the intersections of the surfaces with the coordinate planes $x = 0, y = 0$ or $z = 0$.



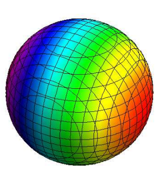




5.18. The function $g(x, y, z) = 2 + \sin(xyz)$ could define a temperature distribution in space. We can no more draw the graph of g because that would be an object in 4 dimensions. We can however draw surfaces like $g(x, y, z) = 0$.

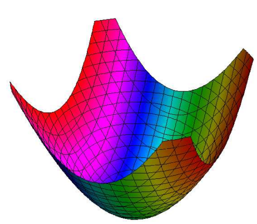
5.19. The level surfaces of $g(x, y, z) = x^2 + y^2 + z^2$ are spheres. The level surfaces of $g(x, y, z) = 2x^2 + y^2 + 3z^2$ are ellipsoids. The equation $ax + by + cz = d$ is a plane. With $\vec{n} = [a, b, c]^T$ and $\vec{x} = [x, y, z]^T$, we can rewrite the equation $\vec{n} \cdot \vec{x} = d$. If a point $\vec{x}_0$ is on the plane, then $\vec{n} \cdot \vec{x}_0 = d$. so that $\vec{n} \cdot (\vec{x} - \vec{x}_0) = 0$. This means that every vector $\vec{x} - \vec{x}_0$ in the plane is orthogonal to $\vec{n}$. For $f(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + kz + m$ the surface $f(x, y, z) = 0$ is called a **quadric**.

Sphere



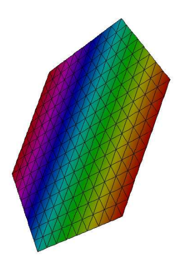
$(x-a)^2+(y-b)^2+(z-c)^2 = r^2$

Paraboloid



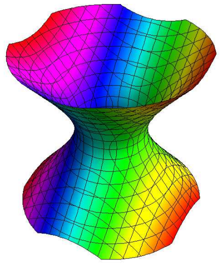
$(x - a)^2 + (y - b)^2 - c = z$

Plane



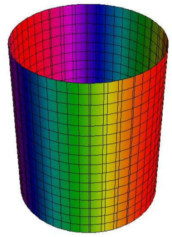
$ax + by + cz = d$

One sheeted hyperboloid



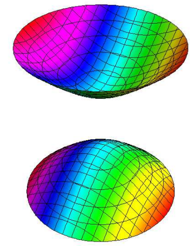
$(x-a)^2+(y-b)^2-(z-c)^2 = r^2$

Cylinder



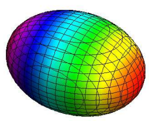
$(x-a)^2+(y-b)^2 = r^2$

Two sheeted hyperboloid



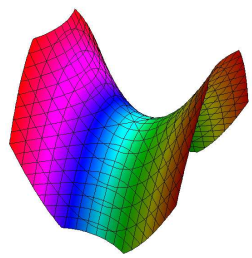
$(x-a)^2+(y-b)^2-(z-c)^2 = -r^2$

Ellipsoid



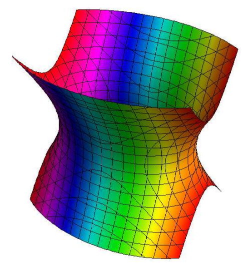
$x^2/a^2+y^2/b^2+z^2/c^2 = 1$

Hyperbolic paraboloid



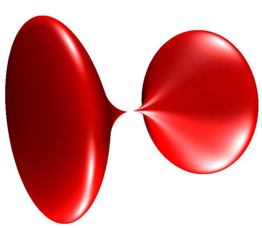
$x^2-y^2+z = 1$

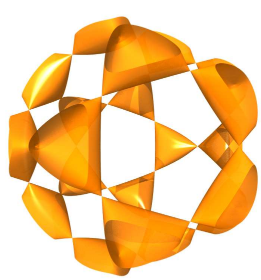
Elliptic hyperboloid

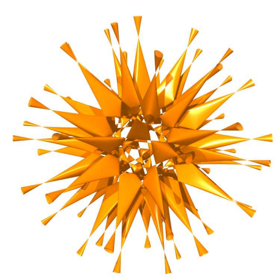


$x^2/a^2+y^2/b^2-z^2/c^2 = 1$

Problem: Higher order polynomial surfaces can be intriguingly beautiful and are sometimes difficult to describe. If f is a polynomial in several variables and $f(x,x,x)$ is a polynomial of degree d, then f is called a **degree d polynomial surface**. Degree 2 surfaces are **quadrics**, degree 3 surfaces **cubics**, degree 4 surfaces **quartics**, degree 5 surfaces **quintics**, degree 10 surfaces **decics** and so on.







Homework

This homework is due on Tuesday, 7/9/2019.

Problem 5.1: Plot the graph of the function $f(x, y) = \sin(x^2/3) - \sin(y^2/3)$ on the region $-\pi \leq x \leq \pi, -\pi \leq y \leq \pi$. How many mountain peaks do you count inside that region? How many sinks (local minima) do you count? Do not look at points on the boundary. You can do it by hand, or Wolfram alpha or Mathematica or a graphing calculator to check your picture. Do not compute the extrema analytically yet. We will do that later. This is a graphing problem.

Problem 5.2: a) Determine the domain and range of the **logarithmic mean**

$$f(x, y) = \frac{(y - x)}{\log(y) - \log(x)}$$

where log the natural logarithm.
b) The function is not defined at $x = y$ but one can define $f(x, y)$ on the diagonal $x = y$. Use Hôpital, show that the limit $\lim_{x\to 2} f(x, 2)$ exists.
c) The function is also not defined at first if $x = 0$ or $y = 0$. Show that the limit $\lim_{x\to 0} f(x, 2)$ exists.

Problem 5.3: a) Use the computer to draw the level surface $x^2 - y^2 + z^2 - x^4y^4z^4 - x^2y^2z^2 = 0$ with x, y, z all in $[-2, 2]$
b) Do the same for the $((x^2 + y^2)^2 - x^2 - y^2)^2 + z^2 = 0.05$ with with x, y, z all in $[-2, 2]$

Problem 5.4: a) Sketch the graph and contour map of
$f(x, y) = \cos(1 + 2x^2 + y^2)/(1 + 2x^2 + y^2)$.
b) Sketch the contour map of $g(x, y) = 2|x| - 5|y|$.
c) Sketch the contour map of $h(x, y) = (x^2 + y^2)^2 - x^2 + y^2$.

Problem 5.5: a) Verify that the line $\vec{r}(t) = [1, 3, 2]^T + t[1, 2, 1]^T$ is part of the hyperbolic paraboloid $z^2 - x^2 - y = 0$.
b) Verify that the line $\vec{r}(t) = [1 + t, 1 - t, t]^T$ is part of the one sheeted hyperboloid $x^2 + y^2 - 2z^2 = 2$.
c) As also the line $\vec{r}(s) = [1 - s, 1 + s, s]^T$ is part of the same hyperboloid, what is the intersection of the hyperboloid with the plane $\vec{r}(t, s) = [1 + t - s, 1 - t + s, t + s]^T$?

MULTIVARIABLE CALCULUS

MATH S-21A

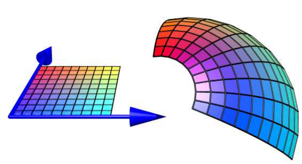
Unit 6: Parametrized Surfaces

Lecture

6.1. Surfaces can be described in two fundamentally different ways: first as **level surfaces** $g(x, y, z) = c$ and then through **parametrization**. What we have seen already for planes can be done more generally for other surfaces. Let us first look at the general setup:

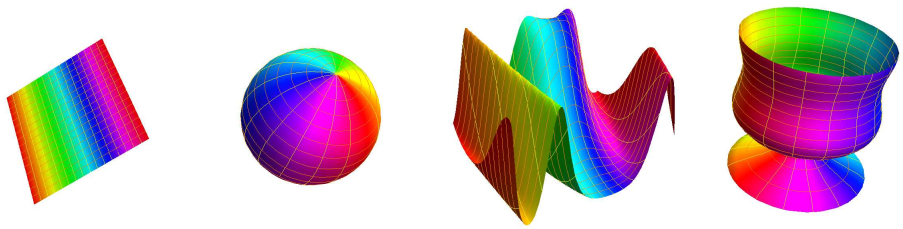
> **Definition:** A **parametrization** of a surface is a vector-valued function
> $$\vec{r}(u, v) = [x(u, v), y(u, v), z(u, v)]^T ,$$
> where $x(u, v), y(u, v), z(u, v)$ are three functions of two variables. The **parameters** u, v serve as coordinates on the surface. If we plug in concrete values like $u = 3, v = 2$ for example in a function $\vec{r}(u, v) = [u - 2, v^2, u^3 - v]^T$, we get a concrete point $\vec{r}(3, 2) = [1, 4, 25]^T$ in $\mathbb{R}^3$.

6.2. Because two parameters u and v are involved, the map $\vec{r}$ is also called uv**-map**. And like uv-light, it looks cool. If you like a fancy description, a parametrization is a map from $\mathbb{R}^2$ to $\mathbb{R}^3$. A **parametrized surface** is the image of the uv-map. The domain R of the uv-map is called the **parameter domain**. The parametrization is what you are **doing**, the surface itself is something you **see**. There are many different parametrizations of the same surface.



> **Definition:** If the first parameter u is kept constant, then $v \mapsto \vec{r}(u, v)$ is a curve on the surface. Similarly, if v is constant, then $u \mapsto \vec{r}(u, v)$ traces a curve the surface. These curves are called **grid curves**.

Parametric surfaces can become complex. In that case, it is better to explored them with the help of a computer. The following four examples are important building blocks for more general surfaces.



Definition: A point $(x, y) \neq (0, 0)$ in the plane has the **polar coordinates** $r = \sqrt{x^2 + y^2}, \theta$, where θ is the angle from the positive x-axes to the point in counter clockwise direction. For $x > 0, y > 0$ it is $\arctan(y/x)$. In general $(x, y) = (r\cos(\theta), r\sin(\theta))$. A common choice is to take $\theta \in [0, 2\pi)$. The point $((0, -1)$ has then the polar coordinates $(r, \theta) = (1, 3\pi/2)$.

6.3. Note that the formula $\theta = \arctan(y/x)$ defines the angle θ only up to an addition of an integer multiple of π. The points $(1, 2)$ and $(-1, -2)$ for example have the same θ value. In order to get the correct θ value one can could take $\arctan(y/x)$ in $(-\pi/2, \pi/2]$, where $\pi/2$ is the limit when $x \to 0^+$, then add π if $x < 0$ or if $x = 0$ and $y < 0$.

Definition: The coordinate system obtained by representing points in space as

$$(x, y, z) = (r\cos(\theta), r\sin(\theta), z)$$

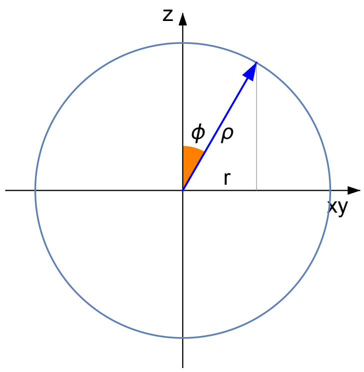
is called the **cylindrical coordinate system**.

Definition: **Spherical coordinates** use the distance ρ to the origin as well as two angles θ and ϕ called **Euler angles**. The first angle θ is the angle we have used in polar coordinates. The second angle, ϕ, is the angle between the vector $\vec{OP}$ and the z-axis. A point has the **spherical coordinate**

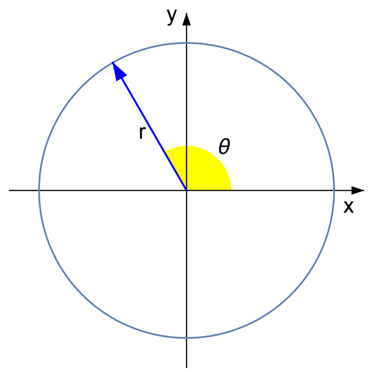
$$(x, y, z) = (\rho\cos(\theta)\sin(\phi), \rho\sin(\theta)\sin(\phi), \rho\cos(\phi)) .$$

We always use $0 \leq \theta < 2\pi, 0 \leq \phi \leq \pi, \rho \geq 0$.

The following figures allow you to derive the formulas. The distance to the z axes is $r = \rho \sin(\phi)$ and the height $z = \rho \cos(\phi)$ can be read off by the left picture, the coordinates $x = r\cos(\theta), y = r\sin(\theta)$ can be seen in the right picture.



$$x = \rho \cos(\theta) \sin(\phi),$$
$$y = \rho \sin(\theta) \sin(\phi),$$
$$z = \rho \cos(\phi)$$



Examples

6.4. A plane has the parametrization $\vec{r}(s,t) = \vec{OP} + s\vec{v} + t\vec{w}$ and the implicit equation $ax + by + cz = d$. To get from parametric to implicit, find the normal vector $\vec{n} = \vec{v} \times \vec{w}$. To get from implicit to parametric, find two vectors $\vec{v}, \vec{w}$ normal to the vector $\vec{n}$. For example, find three points P, Q, R on the surface and form $\vec{u} = \vec{PQ}, \vec{v} = \vec{PR}$.

6.5. The **sphere** $\vec{r}(u,v) = [a,b,c]^T + [\rho\cos(u)\sin(v), \rho\sin(u)\sin(v), \rho\cos(v)]^T$ can be brought into the implicit form by finding the center and radius $(x-a)^2 + (y-b)^2 + (z-c)^2 = \rho^2$.

6.6. The parametrization of a graph is $\vec{r}(u,v) = [u, v, f(u,v)]^T$. It can be written in implicit form as $z - f(x,y) = 0$.

6.7. The surface of revolution is in parametric form given as $\vec{r}(u,v) = [g(v)\cos(u), g(v)\sin(u), v]^T$. It has the implicit description $\sqrt{x^2+y^2} = r = g(z)$ which can be rewritten as $x^2 + y^2 = g(z)^2$.

6.8. Here are some level surfaces in cylindrical coordinates:

Problem: $r = 1$ is a **cylinder**, $r = |z|$ is a **double cone**, $r^2 = z$ **elliptic paraboloid**, $\theta = 0$ is a **half plane**, $r = \theta$ is a **rolled sheet of paper**.

Problem: $r = 2 + \sin(z)$ is an example of a **surface of revolution**.

6.9. Here are some level surfaces described in spherical coordinates:

Problem: $\rho = 1$ is a **sphere**, the surface $\phi = \pi/4$ is a **single cone**, $\rho = \phi$ is an **apple shaped surface** and $\rho = 2 + \cos(3\theta)\sin(\phi)$ is an example of a **bumpy sphere**.

Homework

This homework is due on Tuesday, 7/9/2019.

Problem 6.1: Find a parametrization for the plane which contains the three points $P = (4,7,1)$,$Q = (2,2,1)$ and $R = (1,3,5)$.

Problem 6.2: Plot the surface with the parametrization

$$\vec{r}(u,v) = [\cos(t)\sin(s), \sin(t)\sin(s) + (\cos(t)\sin(s))^4, \cos(s)/2]^T ,$$

where $0 \leq u \in \pi$ and $0 \leq v \leq 2\pi$. You can use technology if you like.

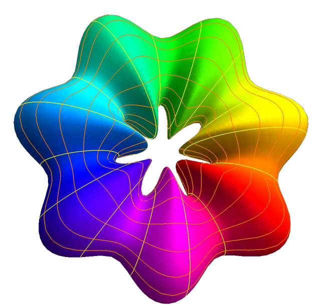
Problem 6.3: a) Find a parametrisations of the lower half of the ellipsoid $x^2/25 + y^2/16 + (z-3)^2 = 1, z \geq 3$ by using that the surface is a graph $z = f(x,y)$ on a suitable domain.
b) Find a second parametrization but use angles ϕ, θ similarly as for the sphere.

Problem 6.4: Find a parametrisation of the **torus candy** (www.math-candy.com) given as the set of points which have distance $6 + 2\cos(7\theta)$ from the circle

$$[10\cos(\theta), 10\sin(\theta), 0]^T ,$$

where θ is the angle occurring in cylindrical and spherical coordinates. We can assure you that the candy melts wonderfully on your tongue.

Hint: Use r, the distance of a point $(x,y,)$ to the z-axis. This distance is $r = (10 + (6 + 2\cos(7\theta))\cos(\psi))$ if ψ is the angle for the circle winding around the candy. You can also use that $z = (6 + 2\cos(7\theta))\sin(\psi)$. To finish the parametrization problem, translate back to Cartesian coordinates.



Problem 6.5: a) What is the equation for the surface $x^2 + y^2 = 3x + z^2$ in cylindrical coordinates?
b) Describe in words or draw a sketch of the surface whose equation is $\rho = |\sin(4\phi)|$ in spherical coordinates (ρ, θ, ϕ).

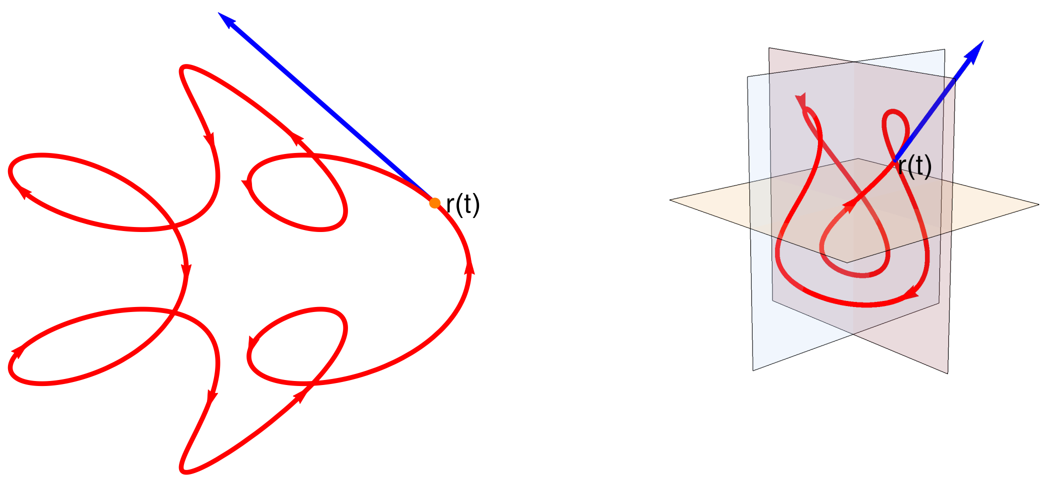
MULTIVARIABLE CALCULUS

MATH S-21A

Unit 7: Parametrized curves

Lecture

Definition: A **parametrization** of a planar curve is a map $\vec{r}(t) = [x(t), y(t)]^T$ from a **parameter interval** $R = [a, b]$ to the plane. The functions $x(t)$ and $y(t)$ are called **coordinate functions**. The image of the parametrization is called a **parametrized curve** in the plane. Similarly, the parametrization of a **space curve** is $\vec{r}(t) = [x(t), y(t), z(t)]^T$. The image of $\vec{r}$ is called a **parametrized curve** in space.



7.1. We think of the **parameter** t as **time** and the parametrization as a **drawing process**. The curve is the result what you **see**. For a fixed time t, we have a vector $[x(t), y(t), z(t)]^T$ in space. As t varies, the end point of this vector moves along the curve. The parametrization contains **more information** about the curve then the curve itself. It tells for example how fast the curve was traced.

7.2. Curves can describe the paths of particles, celestial bodies, or other quantities which change in time. Examples are the motion of a star moving in a galaxy, or economical data changing in time. Here are some more places, where curves appear:

Strings or knots	are closed curves in space.
Molecules	like RNA or proteins.
Graphics:	grid curves produce a mesh of curves.
Typography:	fonts represented by Bézier curves.
Relativity:	curve in space-time describes the motion of an object
Topology:	space filling curves, boundaries of surfaces or knots.

Definition: Any vector parallel to the velocity $\vec{r}\,'(t)$ is called **tangent** to the curve at $\vec{r}(t)$.

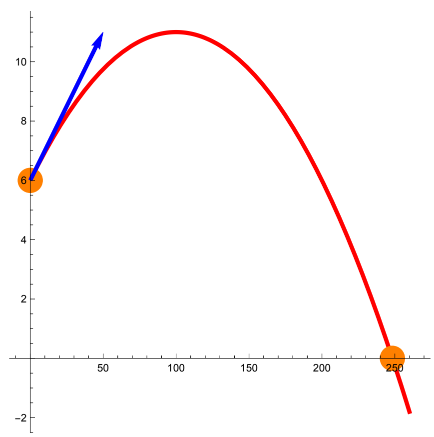
You know from single variable the **addition rule** $(f+g)' = f' + g'$, the **scalar multiplication rule** $(cf)' = cf'$ and the **Leibniz rule** $(fg)' = f'g + fg'$ as well as the **chain rule** $(f(g))' = f'(g)g'$. They generalize to vector-valued functions.

$(\vec{v}+\vec{w})' = \vec{v}' + \vec{w}'$, $(c\vec{v})' = c\vec{v}'$, $(\vec{v}\cdot\vec{w})' = \vec{v}'\cdot\vec{w} + \vec{v}\cdot\vec{w}'$ $(\vec{v}\times\vec{w})' = \vec{v}'\times\vec{w} + \vec{v}\times\vec{w}'$ $(\vec{v}(f(t)))' = \vec{v}'(f(t))f'(t)$.

The process of differentiation of a curve can be reversed using the **fundamental theorem of calculus**. If $\vec{r}'(t)$ and $\vec{r}(0)$ is known, we can figure out $\vec{r}(t)$ by **integration** $\vec{r}(t) = \vec{r}(0) + \int_0^t \vec{r}'(s)\,ds$.

Assume we know the acceleration $\vec{a}(t) = \vec{r}''(t)$ at all times as well as initial velocity and position $\vec{r}'(0)$ and $\vec{r}(0)$. Then $\vec{r}(t) = \vec{r}(0) + t\vec{r}'(0) + \vec{R}(t)$, where $\vec{R}(t) = \int_0^t \vec{v}(s)\,ds$ and $\vec{v}(t) = \int_0^t \vec{a}(s)\,ds$.

The **free fall** is the case when acceleration is constant. The direction of the constant force defines what is "down". If $\vec{r}''(t) = [0,0,-10]^T$, $\vec{r}'(0) = [0,1000,2]^T$, $\vec{r}(0) = [0,0,h]^T$, then $\vec{r}(t) = [0,1000t, h + 2t - 10t^2/2]^T$.

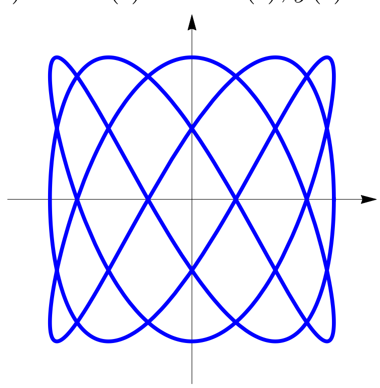


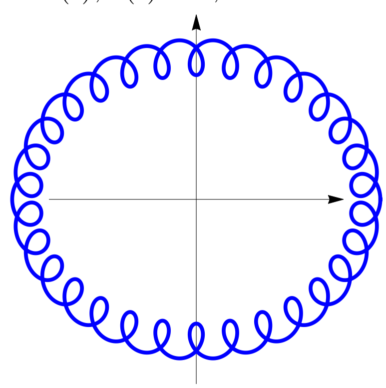
If $r''(t) = \vec{F}$ is constant, then $\vec{r}(t) = \vec{r}(0) + t\vec{r}'(0) - \vec{F}t^2/2$.

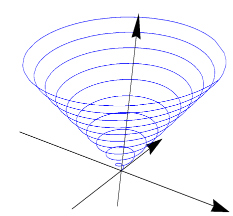
Examples

7.3. Examples:
1) The parametrization $\vec{r}(t) = [1+2\cos(t), 3+5\sin(t)]^T$ is the ellipse $(x-1)^2/4 + (y-3)^2/25 = 1$. The parametrization $\vec{r}(t) = [\cos(3t), \sin(5t)]^T$ is an example of a

Lissajous curve.
2) If $x(t) = t$, $y(t) = f(t)$, the curve $\vec{r}(t) = [t, f(t)]^T$ traces the **graph** of the function $f(x)$. For example, for $f(x) = x^2 + 1$, the graph is a parabola. 3) With $x(t) = t\cos(t), y(t) = t\sin(t), z(t) = t$ we get the parametrization of a **space curve** $\vec{r}(t) = [t\cos(t), t\sin(t), t]^T$ which traces a spiral on a cone $x^2 + y^2 = z^2$. 4) For $x(t) = 2t\cos(2t), y(t) = 2t\sin(2t), z(t) = 2t$ traces the same curve but twice as fast.
5) If $P = (a, b, c)$ and $Q = (u, v, w)$ are points in space, then $\vec{r}(t) = [a + t(u - a), b + t(v - b), c + t(w - c)]^T$ with $t \in [0, 1]$ is a **line segment** connecting P with Q. For example, $\vec{r}(t) = [1+t, 1-t, 2+3t]^T$ connects the points $P = (1, 1, 2)$ with $Q = (2, 0, 1)$.
6) For $x(t) = t\cos(t), y(t) = t\sin(t), z(t) = t$, then







The computation is done coordinate wise:

$$
\begin{array}{lll}
\text{Position} & \vec{r}\,(t) & = [\cos(3t), \sin(2t), 2\sin(t)]^T \\
\text{Velocity} & \vec{r}\,'(t) & = [-3\sin(3t), 2\cos(2t), 2\cos(t)]^T \\
\text{Acceleration} & \vec{r}\,''(t) & = [-9\cos(3t), -4\sin(2t), -2\sin(t)]^T \\
\text{Jerk} & \vec{r}\,'''(t) & = [27\sin(3t), 8\cos(2t), -2\cos(t)]^T
\end{array}
$$

7.4. Lets look at some examples of velocities and accelerations:

Example	**Velocity**
Hair growth:	0.000000005 m/s
Garden Snail	0.013 m/s
Signals in nerves:	40 m/s
Sound in air:	340 m/s
Speed of bullet:	1200-1500 m/s
Earth in solar system	30'000 m/s
Sun in galaxy:	200'000 m/s
Light in vacuum:	299'792'458 m/s

Example	**Acceleration**
Train:	0.1-0.3 m/s^2
Sprinter (100 m Dash):	3 m/s^2
Car:	3-8 m/s^2
Free fall:	1G = 9.81 m/s^2
Space X BFR:	4G m/s^2
Combat plane F35A:	9G m/s^2
Ejection from F35A:	14G m/s^2.
Electron in vacuum:	10^{15} m/s^2

Homework

This homework is due on Tuesday, 7/9/2019.

Problem 7.1: a) Sketch the plane curve

$$\vec{r}(t) = [x(t), y(t)]^T = [\cos(t) + \sin(2t), \sin(t) - \cos(2t)]^T ,$$

for $t \in [0, 2\pi]$ by plotting the points for different values of t. Calculate its velocity $\vec{r}\,'(t)$ as well as the acceleration $\vec{r}\,''(t)$ at $t = 0$.
b) Sketch the space curve

$$\vec{r}(t) = [(10 + 3\cos(17t))\cos(t), (10 + 3\cos(17t))\sin(t), 4t + 3\sin(17t)]^T$$

with $t \in [0, 5\pi]$.

Problem 7.2: Your cellphone app measures the acceleration

$$\vec{r}\,''(t) = [\cos(t), -\cos(9t), \sin(t)]^T$$

while you are riding a roller coaster. Assume you were at $(0, 0, 0)$ at time $t = 0$ with velocity $(1, 0, 0)$ at $t = 0$, what is its position $\vec{r}(t)$ at time t?

Problem 7.3: a) Two particles travel along space curves. The first is

$$\vec{r}_1(t) = [t, t^2, t^3]^T .$$

The second is

$$\vec{r}_2(t) = [1 + 2t, 1 + 6t, 1 + 14t]^T .$$

Do the particles collide? Do the particle paths intersect?
b) If $\vec{r}(t) = [\cos(t), 2\sin(t), 4t]^T$, find $\vec{r}\,'(0)$ and $\vec{r}\,''(0)$. Then compute $|\vec{r}\,'(0) \times \vec{r}\,''(0)|/|\vec{r}\,'(0)|^3$. We will later call this the curvature.

Problem 7.4: Find the parameterization $\vec{r}(t) = [x(t), y(t), z(t)]^T$ of the curve obtained by intersecting the elliptical cylinder $x^2/16 + y^2/25 = 1$ with the surface $z = x^2 y$. Find the velocity vector $\vec{r}\,'(t)$ at the time $t = \pi/2$.

Problem 7.5: Consider the curve

$$\vec{r}(t) = [x(t), y(t), z(t)]^T = [t^2, 1 + t, 1 + t^3]^T .$$

Check that it passes through the point $(1, 0, 0)$ and find the velocity vector $\vec{r}\,'(t)$, the acceleration vector $\vec{r}\,''(t)$ as well as the jerk vector $\vec{r}\,'''(t)$ at this point.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 8: Arc length and Curvature

Lecture

Definition: If $t \in [a,b] \mapsto \vec{r}(t)$ is a parametrized curve with velocity $\vec{r}\,'(t)$ and speed $|\vec{r}\,'(t)|$, then the number $L = \int_a^b |\vec{r}\,'(t)|\; dt$ is called the **arc length** of the curve.

We justify in class why this formula is reasonable if $\vec{r}$ is differentiable. Written out, the formula is $L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\; dt$.

7.1. Because a parameter change $t = t(s)$ corresponds to a **substitution** in the integration which does not change the integral, we immediately see "path independence of arc length":

The arc length is independent of the parameterization of the curve.

Definition: Define the **unit tangent vector** $\vec{T}(t) = \vec{r}\,'(t)|/|\vec{r}\,'(t)|$ **unit tangent vector**.

Definition: The **curvature** of a curve at the point $\vec{r}(t)$ is defined as $\kappa(t) = \frac{|\vec{T}'(t)|}{|\vec{r}\,'(t)|}$.

7.2. The curvature is the length of the acceleration vector if $\vec{r}(t)$ parametrizes the curve with constant speed 1. A large curvature at a point means that the curve is strongly bent. Unlike the acceleration or the velocity, the curvature does not depend on the parameterization of the curve. You "see" the curvature, while you "feel" the acceleration. We can measure curvature at a point only if $\vec{r}\,'(t)$ is not zero.

The curvature does not depend on the parametrization.

Proof. Let $s(t)$ be an other parametrization, then by the chain rule $d/dtT'(s(t)) = T'(s(t))s'(t)$ and $d/dtr(s(t)) = r'(s(t))s'(t)$. We see that the s' cancels in T'/r'.

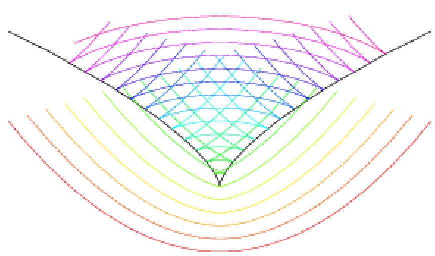
Especially, if the curve is parametrized by arc length, meaning that the velocity vector $r'(t)$ has length 1, then $\kappa(t) = |T'(t)|$. It measures the rate of change of the unit tangent vector.

Definition: If $\vec{r}(t)$ is a curve which has nonzero speed at t, then we can define $\vec{T}(t) = \frac{\vec{r}\,'(t)}{|\vec{r}\,'(t)|}$, the **unit tangent vector**, $\vec{N}(t) = \frac{\vec{T}'(t)}{|\vec{T}'(t)|}$, the **normal vector** and $\vec{B}(t) = \vec{T}(t) \times \vec{N}(t)$ the **bi-normal vector**. The plane spanned by $\vec{N}$ and $\vec{B}$ is called the **normal plane**. It is perpendicular to the curve. The plane spanned by T and N is called the **osculating plane**.

7.3. If we differentiate $\vec{T}(t) \cdot \vec{T}(t) = 1$, we get $\vec{T}'(t) \cdot \vec{T}(t) = 0$ and see that $\vec{N}(t)$ is perpendicular to $\vec{T}(t)$. Because $\vec{B}$ is automatically normal to $\vec{T}$ and $\vec{N}$, we have shown:

The three vectors $(\vec{T}(t), \vec{N}(t), \vec{B}(t))$ are unit vectors orthogonal to each other.

7.4. Here is an application of curvature: if a curve $\vec{r}(t)$ represents a **wave front** and $\vec{n}(t)$ is a **unit vector normal** to the curve at $\vec{r}(t)$, then $\vec{s}(t) = \vec{r}(t) + \vec{n}(t)/\kappa(t)$ defines a new curve called the **caustic** of the curve. Geometers call it the **evolute** of the original curve.



A useful formula for curvature is

$$\kappa(t) = \frac{|\vec{r}\,'(t) \times \vec{r}\,''(t)|}{|\vec{r}\,'(t)|^3}$$

7.5. We prove this in class. Finally, lets mention that curvature is important also in **computer vision**. If the gray level value of a picture is modeled as a function $f(x, y)$ of two variables, places where the level curves of f have maximal curvature corresponds to **corners** in the picture. This is useful when **tracking** or **identifying** objects.

In this picture of John Harvard, software was looking for level curves for each color and started to draw at points where the curvature of the curves is large then follow the level curve.

Examples

Problem: The arc length of the **circle** $\vec{r}(t) = [R\cos(t), R\sin(t)]^T$ is $2\pi R$. The speed $|\vec{r}'(t)|$ is constant and equal to R.

Problem: The **helix** $\vec{r}(t) = [\cos(t), \sin(t), t]^T$ has velocity $\vec{r}\,'(t) = [-\sin(t), \cos(t), 1]^T$ and constant speed $|\vec{r}\,'(t)| = |[-\sin(t), \cos(t), 1]^T| = \sqrt{2}$.

Problem: What is the arc length of the curve

$$\vec{r}(t) = [\sqrt{2}t, \log(t), t^2/2]^T$$

for $1 \leq t \leq 2$? Answer: Because $\vec{r}\,'(t) = [\sqrt{2}, 1/t, t]^T$, we have $\vec{r}'(t) = \sqrt{2 + \frac{1}{t^2} + t^2} = |\frac{1}{t} + t|$ and $L = \int_1^2 \frac{1}{t} + t \; dt = \log(t) + \frac{t^2}{2}|_1^2 = \log(2) + 2 - 1/2$. This curve does not have a name. But because it is constructed in such a way that the arc length can be computed, we an call it "opportunity".

Problem: Find the arc length of the curve $\vec{r}(t) = [3t^2, 6t, t^3]^T$ from $t = 1$ to $t = 3$.

Problem: What is the arc length of the curve $\vec{r}(t) = [\cos^3(t), \sin^3(t)]^T$? Answer: We have $|\vec{r}\ '(t)| = 3\sqrt{\sin^2(t)\cos^4(t) + \cos^2(t)\sin^4(t)} = (3/2)|\sin(2t)|$. Therefore, $\int_0^{2\pi} (3/2)\sin(2t)\, dt = 6$.

Problem: Find the arc length of $\vec{r}(t) = [t^2/2, t^3/3]^T$ for $-1 \leq t \leq 1$. This cubic curve satisfies $y^2 = x^3 8/9$ and is an example of an **elliptic curve**. Because $\int x\sqrt{1+x^2}\, dx = (1+x^2)^{3/2}/3$, the integral can be evaluated as $\int_{-1}^{1} |x|\sqrt{1+x^2}\ dx = 2\int_0^1 x\sqrt{1+x^2}\ dx = 2(1+x^2)^{3/2}/3|_0^1 = 2(2\sqrt{2}-1)/3$.

Problem: The arc length of an **epicycle** $\vec{r}(t) = [t + \sin(t), \cos(t)]^T$ parameterized by $0 \leq t \leq 2\pi$. We have $|\vec{r'}(t)| = \sqrt{2 + 2\cos(t)}$. so that $L = \int_0^{2\pi} \sqrt{2+2\cos(t)}\ dt$. A **substitution** $t = 2u$ gives $L = \int_0^{\pi} \sqrt{2+2\cos(2u)}\ 2du = \int_0^{\pi} \sqrt{2 + 2\cos^2(u) - 2\sin^2(u)}\ 2du = \int_0^{\pi} \sqrt{4\cos^2(u)}\ 2du = 4\int_0^{\pi} |\cos(u)|\ du = 8$.

Problem: Find the arc length of the **catenary** $\vec{r}(t) = [t, \cosh(t)]^T$, where $\cosh(t) = (e^t + e^{-t})/2$ is the **hyperbolic cosine** and $t \in [-1, 1]$. We have

$$\cosh^2(t)^2 - \sinh^2(t) = 1 \ ,$$

where $\sinh(t) = (e^t - e^{-t})/2$ is the **hyperbolic sine**. Solution: We have $|\vec{r}\,'(t)| = \sqrt{1 + \sinh^2(t)} = \cosh(t)$ and $\int_{-1}^{1} \cosh(t)\ dt = 2\sinh(1)$.

Problem: Often, there is no closed formula for the arc length of a curve. For example, the **Lissajous figure** $\vec{r}(t) = [\cos(3t), \sin(5t)]^T$ leads to the arc length integral $\int_0^{2\pi} \sqrt{9\sin^2(3t) + 25\cos^2(5t)}\ dt$ which can only be evaluated numerically.

Problem: The curve $\vec{r}(t) = [t, f(t)]^T$, which is the graph of a function f has the velocity $\vec{r}\,'(t) = (1, f'(t))$ and the unit tangent vector $\vec{T}(t) = (1, f'(t))/\sqrt{1 + f'(t)^2}$. After some simplification we get

$$\kappa(t) = |\vec{T}'(t)|/|\vec{r}\,'(t)| = |f''(t)|/\sqrt{1 + f'(t)^2}^3$$

For example, for $f(t) = \sin(t)$, then $\kappa(t) = |\sin(t)|/|\sqrt{1 + \cos^2(t)}^3$.

Homework

This homework is due on Tuesday, 7/9/2019.

Problem 8.1: a) Find the arc length of the curve

$$\vec{r}(t) = [t^2/2, t^3/3, 1]^T$$

from $t = -2$ to $t = 2$.
b) Find the arc length of

$$\vec{r}(t) = [4t, 4\sin(3t), 4\cos(3t), 2]^T ,$$

with $0 \leq t \leq \pi$.

Problem 8.2: Find the curvature of

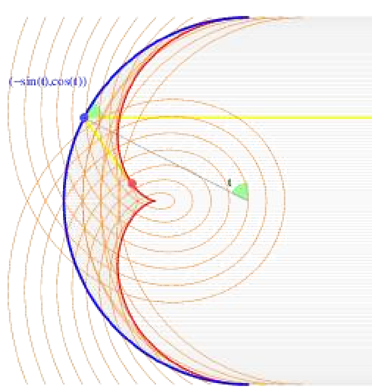
$$\vec{r}(t) = [e^t \cos(t), e^t \sin(t), t]^T$$

at the point $(1, 0, 0)$.

Problem 8.3: Find the vectors $\vec{T}(t), \vec{N}(t)$ and $\vec{B}(t))$ for the curve $\vec{r}(t) = [t^2, t^3, 0]^T$ for $t = 2$. Explore whether the vectors depend continuously on t for all t.

Problem 8.4: Let $\vec{r}(t) = [t, t^2]^T$. Find the equation for the **caustic** $\vec{s}(t) = \vec{r}(t) + \frac{\vec{N}(t)}{\kappa(t)}$. It is known also as the **evolute** of the curve.

Problem 8.5: If $\vec{r}(t) = [-\sin(t), \cos(t)]^T$ is the boundary of a coffee cup and light enters in the direction $[-1, 0]^T$, then light focuses inside the cup on a curve which is called the **coffee cup caustic**. The light ray travels after the reflection for length $\sin(\theta)/(2\kappa)$ until it reaches the caustic. Find a parameterization of the caustic.

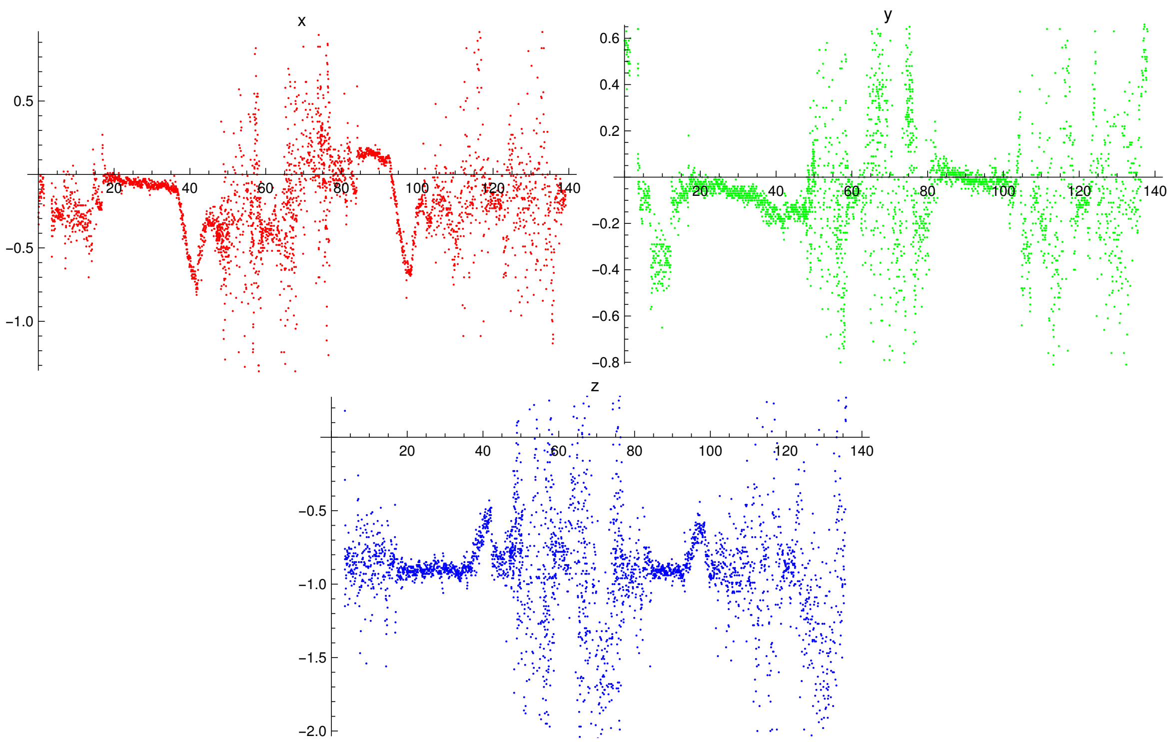


Postscript: Data Visualization

One can use measured acceleration data of a roller coaster to reconstruct the curve $\vec{r}(t) = [x(t), y(t), z(t)]$ which describes the ride. In the blog

$$http://blog.robindeits.com/2013/11/11/roller-coaster-visualizations\ ,$$

Robin Deits has used a smartphone to measure the accelerations during a roller coaster ride on **Cedar point**, the foremost American roller coaster park. Robin made the data available on the website *https://github.com/rdeits/coasters*.
Here are the **x-acceleration**, the **y-acceleration** and the **z-acceleration** data plotted for the Mine ride roller coaster:



One can visualize the actual shape of the roller coaster from these data.

Image source: https://coasterforce.com/mine-train

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 9: Partial derivatives

Lecture

9.1. For functions of several variables we can differentiate to any of them:

Definition: If $f(x, y)$ is a function of the two variables x and y, the **partial derivative** $\frac{\partial}{\partial x} f(x, y)$ is defined as the derivative of the function $g(x) = f(x, y)$ with respect to x, where y is considered a constant. The partial derivative with respect to y is the derivative with respect to y where x is fixed.

9.2. The short hand notation $f_x(x, y) = \frac{\partial}{\partial x} f(x, y)$ is convenient. When iterating derivatives, the notation is similar: we write for example $f_{xy} = \frac{\partial}{\partial x}\frac{\partial}{\partial y} f$. The number $f_x(x_0, y_0)$ gives the slope of the graph sliced at (x_0, y_0) in the x direction. The second derivative f_{xx} is a measure of concavity in that direction. The meaning of f_{xy} is the rate of change of the x-slope if you move the cut along the y-axis.

9.3. The notation $\partial_x f, \partial_y f$ was introduced by Carl Gustav Jacobi. Before that, Josef Lagrange used the term "partial differences". For functions of three or more variables, the partial derivatives are defined in the same way. We write for example $f_x(x, y, z)$ or $f_{xxz}(x, y, z)$.

Theorem: Clairaut's theorem: If f_{xy} and f_{yx} are both continuous, then $f_{xy} = f_{yx}$.

9.4. Proof. Following Euler, we first look at the difference quotients and say that if the "Planck constant" h is positive, then $f_x(x, y) = [f(x+h, y) - f(x, y)]/h$. For $h = 0$, we mean the usual partial derivative f_x. Comparing the two sides of the equation for fixed $h > 0$ shows

$$hf_x(x, y) = f(x + h, y) - f(x, y) \qquad hf_y(x, y) = f(x, y + h) - f(x, y).$$

$$h^2 f_{xy}(x, y) = f(x+h, y+h) - f(x, y+h) - (f(x+h, y) - f(x, y)) \qquad h^2 f_{yx}(x, y) = f(x+h, y+h) - f(x+h, y) - (f(x, y+h) - f(x, y))$$

9.5. Without having taken any limits we established an identity which holds for all $h > 0$: the discrete derivatives f_x, f_y satisfy the relation $f_{xy} = f_{yx}$ for any $h > 0$. We could fancy it as **"quantum Clairaut" formula**. If the classical derivatives f_{xy}, f_{yx} are both continuous, it is possible to take the limit $h \to 0$. The classical Clairaut's theorem can be seen as a "classical limit". The quantum Clairaut holds however for **all** functions $f(x, y)$ of two variables. Not even continuity is needed. [1]

9.6. An equation for an unknown function $f(x, y)$ which involves partial derivatives with respect to at least two different variables is called a **partial differential equation**. We abbreviate PDE. If only the derivative with respect to one variable appears, it is an **ordinary differential equation**, abbreviated ODE.

Examples

9.7. For $f(x, y) = x^4 - 6x^2y^2 + y^4$, we have $f_x(x, y) = 4x^3 - 12xy^2$, $f_{xx} = 12x^2 - 12y^2$, $f_y(x, y) = -12x^2y + 4y^3$, $f_{yy} = -12x^2 + 12y^2$ and see that $\Delta f = f_{xx} + f_{yy} = 0$. A function which satisfies $\Delta f = 0$ is also called **harmonic**. The equation $f_{xx} + f_{yy} = 0$ is a PDE:

Definition: A **partial differential equation** (PDE) is an equation for an unknown function $f(x, y)$ which involves partial derivatives with respect to more than one variables.

9.8.

The **wave equation** $f_{tt}(t, x) = f_{xx}(t, x)$ governs the motion of light or sound. The function $f(t, x) = \sin(x - t) + \sin(x + t)$ satisfies the wave equation.

The **heat equation** $f_t(t, x) = f_{xx}(t, x)$ describes diffusion of heat or spread of an epidemic. The function $f(t, x) = \frac{1}{\sqrt{t}} e^{-x^2/(4t)}$ satisfies the heat equation.

The **Laplace equation** $f_{xx} + f_{yy} = 0$ determines the shape of a membrane. The function $f(x, y) = x^3 - 3xy^2$ is an example satisfying the Laplace equation.

The **advection equation** $f_t = f_x$ is used to model transport in a wire. The function $f(t, x) = e^{-(x+t)^2}$ satisfies the advection equation.

[1] For a full proof of Clairaut's theorem, see *www.math.harvard.edu/knill/teaching/math22a2018/handouts/lecture14.pdf*.

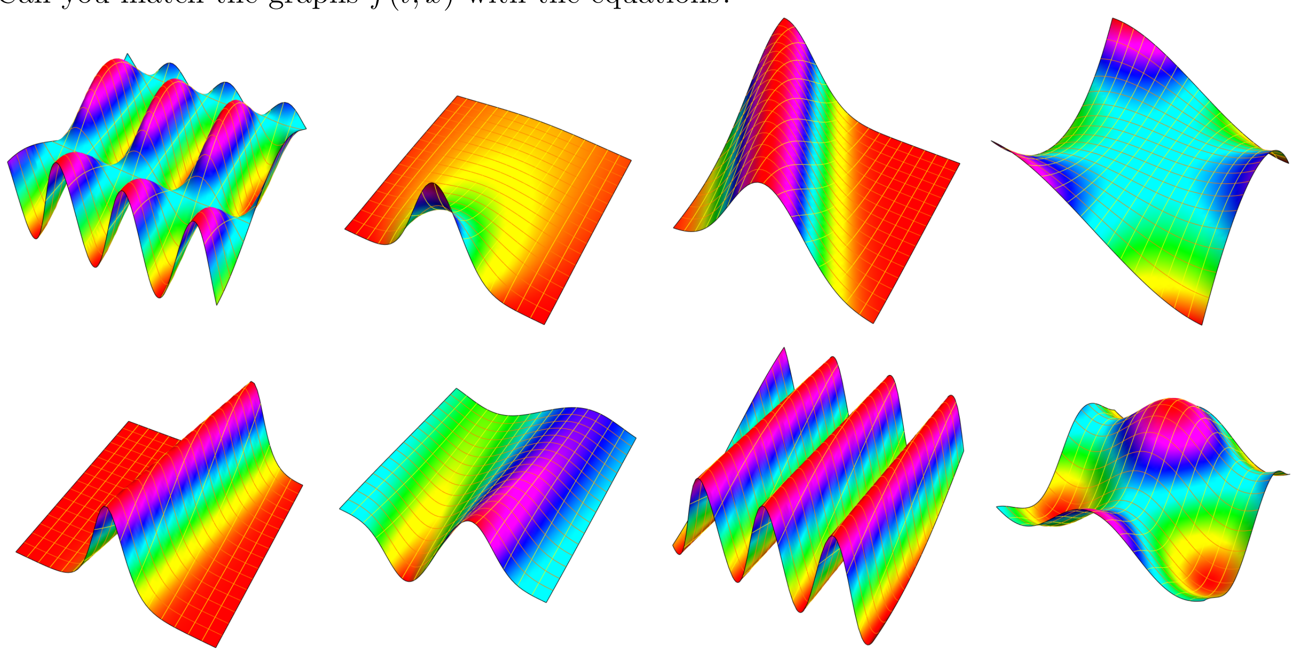
The **eiconal equation** $f_x^2 + f_y^2 = 1$ is used to see the evolution of wave fronts in optics. The function $\boxed{f(x,y) = \cos(x) + \sin(y)}$ satisfies the eiconal equation.

The **Burgers equation** $f_t + f f_x = f_{xx}$ describes waves at the beach which break. The function $\boxed{f(t,x) = \frac{x}{t} \frac{\sqrt{\frac{1}{t}} e^{-x^2/(4t)}}{1+\sqrt{\frac{1}{t}} e^{-x^2/(4t)}}}$ satisfies the Burgers equation.

The **KdV equation** $f_t + 6 f f_x + f_{xxx} = 0$ models **water waves** in a narrow channel. The function $\boxed{f(t,x) = \frac{a^2}{2} \cosh^{-2}(\frac{a}{2}(x - a^2 t))}$ satisfies the KdV equation.

The **Schrödinger equation** $f_t = \frac{i\hbar}{2m} f_{xx}$ is used to describe a **quantum particle** of mass m. The function $\boxed{f(t,x) = e^{i(kx - \frac{\hbar}{2m} k^2 t)}}$ solves the Schrödinger equation. [Here $i^2 = -1$ is the imaginary i and $\hbar$ is the **Planck constant** $\hbar \sim 10^{-34} Js$.]

Can you match the graphs $f(t,x)$ with the equations?



9.9. In all examples, we just see one possible solution to the partial differential equation. There are in general many solutions and additional initial or boundary conditions then determine the solution uniquely. If we know $f(0,x)$ for the Burgers equation, then the solution $f(t,x)$ is determined.

Homework

This homework is due on Tuesday, 7/16/2019.

Problem 9.1: Verify that $f(t,x) = \cos^2(t+x) + e^{e^{\sin(t+x)}}$ is a solution of the transport equation $f_t(t,x) = f_x(t,x)$.

Problem 9.2: a) Verify that $f(x,y) = \sin(x)(\cos(7y)+\sin(7y))$ satisfies the **Klein Gordon equation** $u_{xx} - u_{yy} = 48u$. This PDE is useful in quantum mechanics.
b) Verify that $4\arctan(e^{(x-t/2)2/\sqrt{3}})$ satisfies the **Sine Gordon equation** $u_{tt} - u_{xx} = -\sin(u)$. Use technology. (If you can do it without technology, show it to Oliver).

Problem 9.3: Verify that for any real constant b, the function $f(x,t) = e^{-bt}\cos(x+t)$ satisfies the driven transport equation $f_t(x,t) = f_x(x,t) - bf(x,t)$ This PDE is sometimes called the **advection equation** with damping b.

Problem 9.4: The differential equation

$$f_t = f - xf_x - x^2 f_{xx}$$

is a version of the **infamous Black-Scholes equation**. Here $f(x,t)$ is the prize of a **call option** and x the stock prize ad t is time. Find a function $f(x,t)$ solving it which depends both on x and t. The examples $f(x,y) = x$ or $f(x) = e^t$ do not qualify as they depend only on one variable.

Problem 9.5: The partial differential equation $f_t + ff_x = f_{xx}$ is called **Burgers equation** and describes waves at the beach. In higher dimensions, it leads to the Navier-Stokes equation which are used to describe the weather. Verify that

$$f(t,x)\frac{\left(\frac{1}{t}\right)^{3/2} x e^{-\frac{x^2}{4t}}}{\sqrt{\frac{1}{t}}e^{-\frac{x^2}{4t}} + 1}$$

solves the Burgers equation. You also here might want to get help with technology.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 10: Linearization

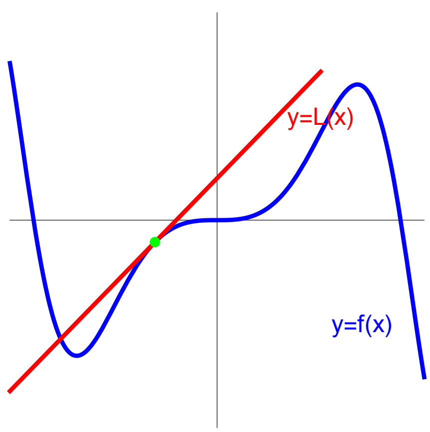
LECTURE

10.1. In single variable calculus, you have seen the following definition for a differentiable function:

Definition: The **linear approximation** of $f(x)$ at a is the affine function

$$L(x) = f(a) + f'(a)(x - a) .$$

10.2. If you have seen **Taylor series**, this is the part of the series $f(x) = \sum_{k=0}^{\infty} f^{(k)}(a)(x-a)^k/k!$ where only the $k = 0$ and $k = 1$ term are considered. We think about the linear approximation L as a function and not as a graph because we will also look at linear approximations for functions of three variables, where we can not draw graphs.



10.3. The graph of the function L is close to the graph of f at a. What about higher dimensions?

Definition: The **linear approximation** of $f(x, y)$ at (a, b) is the affine function

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

The **linear approximation** of a function $f(x, y, z)$ at (a, b, c) is

$$L(x, y, z) = f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c) .$$

10.4. Using the **gradient**

$$\nabla f(x,y) = [f_x, f_y]^T, \qquad \nabla f(x,y,z) = [f_x, f_y, f_z]^T ,$$

the linearization can be written more compactly as

$$L(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) .$$

10.5. How do we justify the linearization? If the second variable $y = b$ is fixed, we have a one-dimensional situation, where the only variable is x. Now $f(x,b) = f(a,b) + f_x(a,b)(x-a)$ is the linear approximation. Similarly, if $x = x_0$ is fixed y is the single variable, then $f(x_0,y) = f(x_0,y_0) + f_y(x_0,y_0)(y-y_0)$. Knowing the linear approximations in both the x and y variables, we can get the general linear approximation by $f(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$.

Examples

10.6. What is the linear approximation of the function $f(x,y) = \sin(\pi xy^2)$ at the point $(1,1)$? Answer: We have $[f_x(x,y), f_y(x,y)]^T = [\pi y^2 \cos(\pi xy^2), 2xy\pi\cos(\pi xy^2)]^T$ which is at the point $(1,1)$ equal to $\nabla f(1,1) = [\pi\cos(\pi), 2\pi\cos(\pi)]^T = [-\pi, -2\pi]^T$. The function is $L(x,y) = 0 + (-\pi)(x-1) - 2\pi(y-1)$.

10.7. Linearization can be used to estimate functions near a point. In the previous example,

$$f(1+0.01, 1+0.01) = -0.0095$$

$$L(1+0.01, 1+0.01) = -\pi 0.01 - 2\pi 0.01 = -3\pi/100 = -0.00942 .$$

10.8. Here is an example in three dimensions: find the linear approximation to $f(x,y,z) = xy+yz+zx$ at the point $(1,1,1)$. Since $f(1,1,1) = 3$, and $\nabla f(x,y,z) = (y+z, x+z, y+x)$, $\nabla f(1,1,1) = (2,2,2)$. we have $L(x,y,z) = f(1,1,1) + (2,2,2)\cdot(x-1, y-1, z-1) = 3 + 2(x-1) + 2(y-1) + 2(z-1) = 2x + 2y + 2z - 3$.

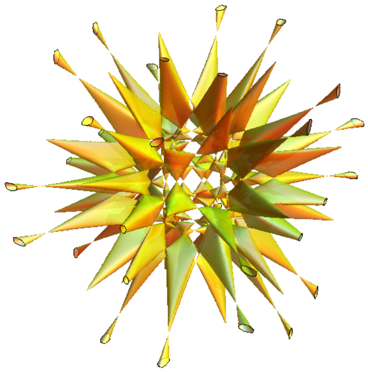
10.9. Estimate $f(0.01, 24.8, 1.02)$ for $f(x,y,z) = e^x\sqrt{y}z$.
Solution: take $(x_0,y_0,z_0) = (0,25,1)$, where $f(x_0,y_0,z_0) = 5$. The gradient is $\nabla f(x,y,z) = (e^x\sqrt{y}z, e^x z/(2\sqrt{y}), e^x\sqrt{y})$. At the point $(x_0,y_0,z_0) = (0,25,1)$ the gradient is the vector $(5, 1/10, 5)$. The linear approximation is $L(x,y,z) = f(x_0,y_0,z_0) + \nabla f(x_0,y_0,z_0)(x-x_0, y-y_0, z-z_0) = 5 + (5,1/10,5)(x-0, y-25, z-1) = 5x + y/10 + 5z - 2.5$. We can approximate $f(0.01, 24.8, 1.02)$ by $5 + (5,1/10,5)\cdot(0.01, -0.2, 0.02) = 5 + 0.05 - 0.02 + 0.10 = 5.13$. The actual value is $f(0.01, 24.8, 1.02) = 5.1306$, very close to the estimate.

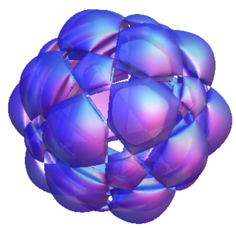
10.10. Find the tangent line to the graph of the function $g(x) = x^2$ at the point $(2,4)$.
Solution: the level curve $f(x,y) = y - x^2 = 0$ is the graph of a function $g(x) = x^2$ and the tangent at a point $(2, g(2)) = (2,4)$ is obtained by computing the gradient $[a,b]^T = \nabla f(2,4) = [-g'(2), 1]^T = [-4, 1]^T$ and forming $-4x + y = d$, where $d = -4\cdot 2 + 1\cdot 4 = -4$. The answer is $\boxed{-4x + y = -4}$ which is the line $y = 4x - 4$ of slope 4.

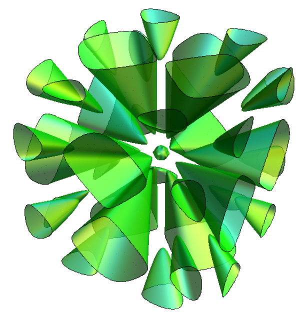
10.11. The **Barth surface** is defined as the level surface $f = 0$ of

$$\begin{aligned} f(x,y,z) &= (3+5t)(-1+x^2+y^2+z^2)^2(-2+t+x^2+y^2+z^2)^2 \\ &+ 8(x^2-t^4y^2)(-(t^4x^2)+z^2)(y^2-t^4z^2)(x^4-2x^2y^2+y^4-2x^2z^2-2y^2z^2+z^4)\,, \end{aligned}$$

where $t = (\sqrt{5}+1)/2$ is a constant called the **golden ratio**. If we replace t with $1/t = (\sqrt{5}-1)/2$ we see the surface to the middle. For $t = 1$, we see to the right the surface $f(x,y,z) = 8$. Find the tangent plane of the later surface at the point $(1,1,0)$.
Answer: We have $\nabla f(1,1,0) = [64,64,0]^T$. The surface is $x+y = d$ for some constant d. By plugging in $(1,1,0)$ we see that $x+y = 2$.



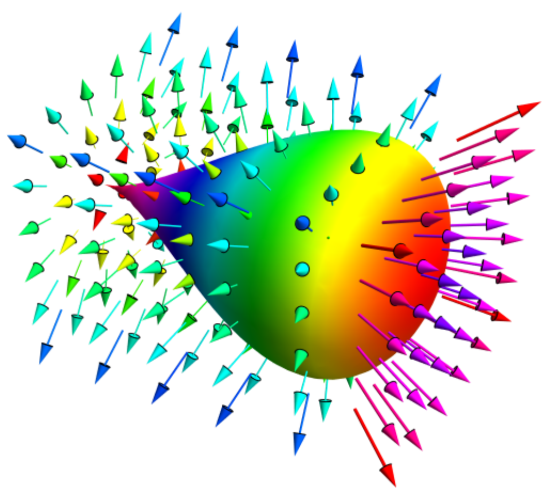




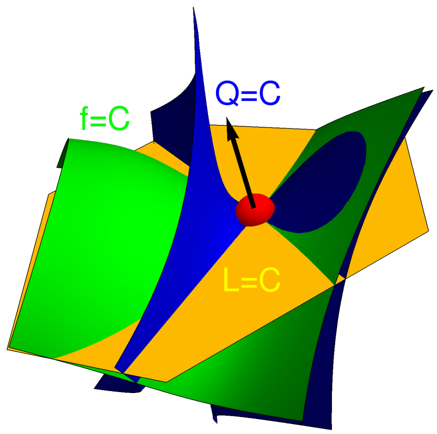
The quartic surface

$$f(x,y,z) = x^4 - x^3 + y^2 + z^2 = 0$$

is called the **piriform**. What is the equation for the tangent plane at the point $P = (2,2,2)$ of this pair shaped surface? We get $[a,b,c]^T = [20,4,4]^T$ and so the equation of the plane $20x+4y+4z = 56$, where we have obtained the constant to the right by plugging in the point $(x,y,z) = (2,2,2)$.



10.12. Finally, we want to point out that linearization is just the first step. One could do quadratic approximations for example. In one dimension, one has $Q(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2!}$. In two dimensions, this becomes $Q(x,y) = L(x,y) + H(a,b)[x-a, y-b]^T \cdot [x-a, y-b]^T/2$, where H is the **Hessian matrix** $H(a,b) = \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}$. We will see this matrix when we maximize or minimze functions.



Homework

This homework is due on Tuesday, 7/16/2019.

Problem 10.1: Estimate $200'000'000'000'000^{1/11}$ using linear approximation of $f(x) = x^{1/11}$ near $x_0 = 20^{11}$.

Problem 10.2: Given $f(x,y) = 3yx/\pi - \cos(x)$. Estimate $f(\pi + 0.01, \pi - 0.03)$ using linearization

Problem 10.3: Estimate $f(0.003, 0.9999)$ for $f(x,y) = \cos(\pi y) + \sin(x + \pi y)$ using linearization.

Problem 10.4: Find the linear approximation $L(x,y)$ of the function

$$f(x,y) = \sqrt{10 - x^2 - 5y^2}$$

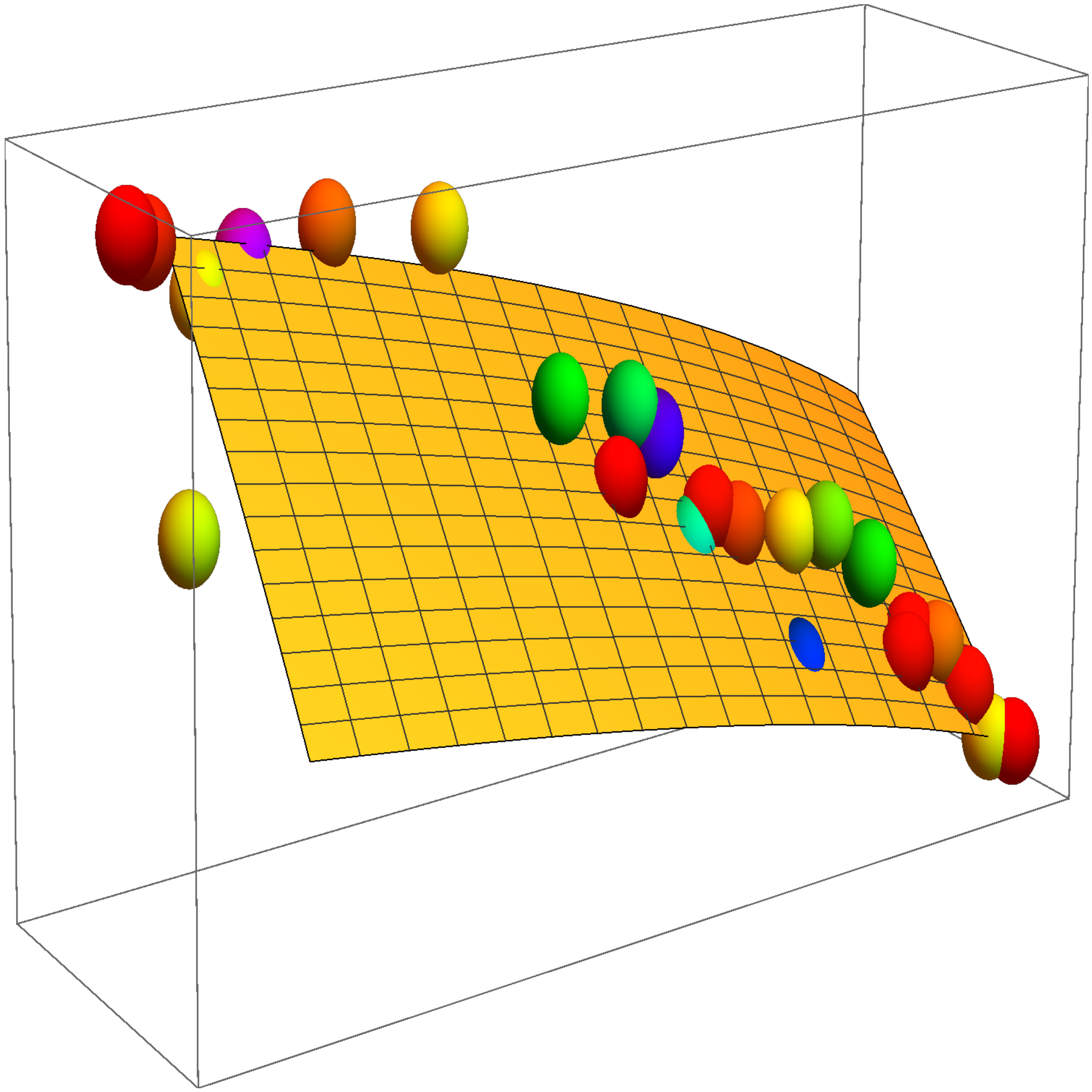
at $(2,1)$ and use it to estimate $f(1.95, 1.04)$.

Problem 10.5: Estimate $(99^3 * 101^2)$ by linearizing the function $f(x,y) = x^3y^2$ at $(100,100)$. What is the difference between $L(100,100)$ and $f(100,100)$?

Data illustration Cobb-Douglas

10.13. The mathematician and economist **Charles W. Cobb** at Amherst college and the economist and politician **Paul H. Douglas** who was also teaching at Amherst found in 1928 empirically a formula $F(K, L) = L^\alpha K^\beta$ which fits the **total production** F of an economic system as a function of the **capital investment** K and the **labor** L. The two authors used logarithms variables and assumed linearity to find α, β. Below are the data normalized so that the date for year 1899 has the value 100. On the website, we give you access to these historical data as well as to the original Nobel prize winning article which already made this assumption.

$Year$	K	L	P
1899	100	100	100
1900	107	105	101
1901	114	110	112
1902	122	118	122
1903	131	123	124
1904	138	116	122
1905	149	125	143
1906	163	133	152
1907	176	138	151
1908	185	121	126
1909	198	140	155
1910	208	144	159
1911	216	145	153
1912	226	152	177
1913	236	154	184
1914	244	149	169
1915	266	154	189
1916	298	182	225
1917	335	196	227
1918	366	200	223
1919	387	193	218
1920	407	193	231
1921	417	147	179
1922	431	161	240



The graph of $F(L,K) = L^{3/4}K^{1/4}$ fits pretty well that data set.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 11: Chain rule

Lecture

11.1. If f and g are functions of a single variable t, the **single variable chain rule** tells us that $d/dt f(g(t)) = f'(g(t))g'(t)$. For example, $d/dt \sin(\log(t)) = \cos(\log(t))/t$. The rule can be proven by linearizing the functions f and g and verifying the chain rule in the linear case. The **chain rule** is also useful:

11.2. To find $\arccos'(x)$ for example, we differentiate $x = \cos(\arccos(x))$ to get $1 = d/dx \cos(\arccos(x)) = -\sin(\arccos(x)) \arccos'(x) = -\sqrt{1-\cos^2(\arccos(x))} \arccos'(x) = -\sqrt{1-x^2} \arccos'(x)$ so that $\arccos'(x) = -1/\sqrt{1-x^2}$.

Definition: Define the **gradient** $\nabla f(x,y) = [f_x(x,y), f_y(x,y)]^T$ or $\nabla f(x,y,z) = [f_x(x,y,z), f_y(x,y,z), f_z(x,y,z)]^T$.

11.3. If $\vec{r}(t)$ is curve and f is a function of several variables we get a function $t \mapsto f(\vec{r}(t))$ of one variable. Similarly, if $\vec{r}(t)$ is a parametrization of a planar curve f is a function of two variables, then $t \mapsto f(\vec{r}(t))$ is a function of one variable.

Theorem: $\frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t)$.

Proof. When written out in two dimensions, it is

$$\frac{d}{dt} f(x(t), y(t)) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \ .$$

The identity

$$\frac{f(x(t+h),y(t+h))-f(x(t),y(t))}{h} = \frac{f(x(t+h),y(t+h))-f(x(t),y(t+h))}{h} + \frac{f(x(t),y(t+h))-f(x(t),y(t))}{h}$$

holds for every $h > 0$. The left hand side converges to $\frac{d}{dt} f(x(t), y(t))$ in the limit $h \to 0$ and the right hand side to $f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$ using the single variable chain rule twice. Here is the proof of the later, when we differentiate f with respect to t and y is treated as a constant:

$$\frac{f(\text{x(t+h)}) - f(x(t))}{h} = \frac{[f(\text{x(t) + (x(t+h)-x(t))}) - f(x(t))]}{\text{[x(t+h)-x(t)]}} \cdot \frac{\text{[x(t+h)-x(t)]}}{h} \ .$$

Write $H(t) =$ x(t+h)-x(t) in the first part on the right hand side.

$$\frac{f(x(t+h)) - f(x(t))}{h} = \frac{[f(x(t)+H) - f(x(t))]}{H} \cdot \frac{x(t+h) - x(t)}{h} .$$

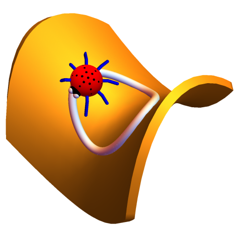
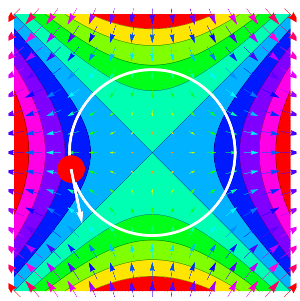
As $h \to 0$, we also have $H \to 0$ and the first part goes to $f'(x(t))$ and the second factor to $x'(t)$.

11.4. The chain rule is powerful because it implies other differentiation rules like the addition, product and quotient rule in one dimensions: $f(x,y) = x+y, x = u(t), y = v(t), d/dt(x+y) = f_x u' + f_y v' = u' + v'$.
$f(x,y) = xy, x = u(t), y = v(t), d/dt(xy) = f_x u' + f_y v' = vu' + uv'$.
$f(x,y) = x/y, x = u(t), y = v(t), d/dt(x/y) = f_x u' + f_y v' = u'/y - v'u/v^2$.

11.5. As in one dimensions, the chain rule follows from linearization. If f is a linear function $f(x,y) = ax+by-c$ and if the curve $\vec{r}(t) = [x_0+tu, y_0+tv]^T$ parametrizes a line. Then $\frac{d}{dt} f(\vec{r}(t)) = \frac{d}{dt}(a(x_0+tu)+b(y_0+tv)) = au+bv$ and this is the dot product of $\nabla f = (a,b)$ with $\vec{r}\,'(t) = (u,v)$. Since the chain rule only refers to the derivatives of the functions which agree at the point, the chain rule is also true for general functions.

Examples

11.6. A ladybug moves on a circle $\vec{r}(t) = [\cos(t), \sin(t)]^T$ on a table with temperature distribution $f(x,y) = x^2 - y^3$. Find the rate of change of the temperature $\nabla f(x,y) = (2x, -3y^2)$, $\vec{r}'(t) = (-\sin(t), \cos(t))$ $d/dt f(\vec{r}(t)) = \nabla T(\vec{r}(t)) \cdot \vec{r}'(t) = (2\cos(t), -3\sin(t)^2) \cdot (-\sin(t), \cos(t)) = -2\cos(t)\sin(t) - 3\sin^2(t)\cos(t)$.

11.7. From $f(x,y) = 0$, one can express y as a function of x, at least near a point where f_y is not zero. From $d/df(x,y(x)) = \nabla f \cdot (1, y'(x)) = f_x + f_y y' = 0$, we obtain $y' = -f_x/f_y$. Even so, we do not know $y(x)$, we can compute its derivative! Implicit differentiation works also in three variables. The equation $f(x,y,z) = c$ defines a surface. Near a point where f_z is not zero, the surface can be described as a graph $z = z(x,y)$. We can compute the derivative z_x without actually knowing the function $z(x,y)$. To do so, we consider y a fixed parameter and compute, using the chain rule

$$f_x(x,y,z(x,y))1 + f_z(x,y) z_x(x,y) = 0$$

so that $z_x(x,y) = -f_x(x,y,z)/f_z(x,y,z)$. This works at points where f_z is not zero.

11.8. The surface $f(x,y,z) = x^2 + y^2/4 + z^2/9 = 6$ is an ellipsoid. Compute $z_x(x,y)$ at the point $(x,y,z) = (2,1,1)$.
Solution: $z_x(x,y) = -f_x(2,1,1)/f_z(2,1,1) = -4/(2/9) = -18$.

This homework is due on Tuesday, 7/16/2019.

Problem 11.1: You know that $d/dt f(\vec{r}(t)) = 25$ at $t = 0$ if $\vec{r}(t) = [t, t]^T$ and $d/dt f(\vec{r}(t)) = 11$ at $t = 0$. $\vec{r}(t) = [t, -t]^T$. Find the gradient of f at $(0, 0)$.

Problem 11.2: The pressure in the space at the position (x, y, z) is $p(x, y, z) = x^2 + y^2 - z^3$ and the trajectory of an observer is the curve $\vec{r}(t) = [t, t, 1/t]^T$. Using the chain rule, compute the rate of change of the pressure the observer measures at time $t = 2$.

Problem 11.3: The chain rule is closely related to linearization. Lets get back to linearlization a bit: A farm costs $f(x, y)$, where x is the number of cows and y is the number of ducks. There are 10 cows and 20 ducks and $f(10, 20) = 1000000$. We know that $f_x(x, y) = 2x$ and $f_y(x, y) = y^2$ for all x, y. Estimate $f(12, 19)$.

Here is a song out of this:
"Old MacDonald had a million dollar farm, E-I-E-I-O,
and on that farm he had $x = 10$ cows, E-I-E-I-O,
and on that farm he had $y = 20$ ducks, E-I-E-I-O,
with $f_x = 2x$ here and $f_y = y^2$ there,
and here two cows more, and there a duck less,
how much does the farm cost now, E-I-E-I-O?"

Problem 11.4: Find, using implicit differentiation the derivative $d/dx \operatorname{arctanh}(x)$, where

$$\tanh(x) = \sinh(x)/\cosh(x) .$$

The **hyperbolic sine** and **hyperbolic cosine** are defined as are $\sinh(x) = (e^x - e^{-x})/2$ and $\cosh(x) = (e^x + e^{-x})/2$. We have $\sinh' = \cosh$ and $\cosh' = \sinh$ and $\cosh^2(x) - \sinh^2(x) = 1$.

Problem 11.5: The equation $f(x, y, z) = e^{xyz} + z = 1 + e$ implicitly defines z as a function $z = g(x, y)$ of x and y. Find formulas (in terms of x,y and z) for $g_x(x, y)$ and $g_y(x, y)$. Estimate $g(1.01, 0.99)$ using linear approximation.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 12: Tangent spaces

Lecture

12.1. The notion of **gradient** is the derivative of a scalar function of many variables. It produces a vector. This vector is useful for example to compute tangent lines or tangent planes.

Definition: The **gradient** of a function $f(x, y)$ is defined as

$$\nabla f(x,y) = [f_x(x,y), f_y(x,y)]^T \ .$$

For functions of three variables, define

$$\nabla f(x,y,z) = [f_x(x,y,z), f_y(x,y,z), f_z(x,y,z)]^T \ .$$

12.2. The symbol ∇ is spelled "Nabla" and named after an Egyptian or Assyrian harp. Early on, the name "Atled" was suggested. But the textbook of 1901 of Gibbs used Nabla was too persuasive. Here is a very important fact, which is true in any dimension. We only formulate it in dimension 2:

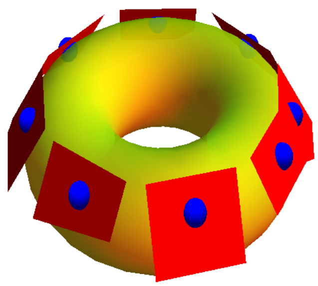
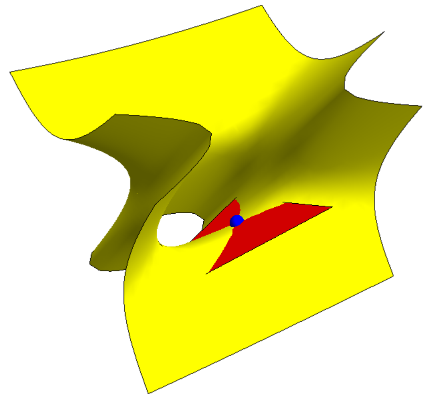
Theorem: Gradient Theorem: $\nabla f(x_0, y_0)$ is perpendicular to the level curve $\{(x,y) \mid f(x,y) = c\}$ containing (x_0, y_0).

Proof. Every curve $\vec{r}(t)$ on the level curve or level surface satisfies $\frac{d}{dt} f(\vec{r}(t)) = 0$. By the chain rule, $\nabla f(\vec{r}(t))$ is perpendicular to the tangent vector $\vec{r}'(t)$. QED.

12.3. Because $\vec{n} = \nabla f(p,q) = [a,b]^T$ is perpendicular to the level curve $f(x,y) = c$ through (p,q), the equation for the tangent line is $ax + by = d$, $a = f_x(p,q)$, $b = f_y(p,q)$, $d = ap + bq$. Compactly written, this is

$$\nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = 0$$

and means that the gradient of f is perpendicular to any vector $(\vec{x} - \vec{x}_0)$ in the plane. It is one of the most important statements in multivariable calculus as it gives a crucial link between calculus and geometry. The just mentioned gradient theorem is also useful. We can immediately compute tangent planes and tangent lines, without linearization!

Definition: If f is a function of several variables and $\vec{v}$ is a unit vector then $D_{\vec{v}}f = \nabla f \cdot \vec{v}$ is called the **directional derivative** of f in the direction $\vec{v}$.

The name "directional derivative" is related to the fact that every unit vector gives a direction. If $\vec{v}$ is a unit vector, then the chain rule tells us $\frac{d}{dt}D_{\vec{v}}f = \frac{d}{dt}f(x + t\vec{v})$.

The directional derivative tells us how the function changes when we move in a given direction. Assume for example that $T(x, y, z)$ is the temperature at position (x, y, z). If we move with velocity $\vec{v}$ through space, then $D_{\vec{v}}T$ tells us at which rate the temperature changes for us. If we move with velocity $\vec{v}$ on a hilly surface of height $h(x, y)$, then $D_{\vec{v}}h(x, y)$ gives us the slope we drive on.

12.4. If $\vec{r}(t)$ is a curve with velocity $\vec{r}\,'(t)$ and the speed is 1, then $D_{\vec{r}'(t)}f = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$ is the temperature change, one measures at $\vec{r}(t)$. The chain rule told us that this is $d/dt f(\vec{r}(t))$.

12.5. For $\vec{v} = (1, 0, 0)$, then $D_{\vec{v}}f = \nabla f \cdot v = f_x$, the directional derivative is a generalization of the partial derivatives. It measures the rate of change of f, if we walk with unit speed into that direction. But as with partial derivatives, it is a **scalar**.

12.6. The directional derivative satisfies $\boxed{|D_{\vec{v}}f| \leq |\nabla f||\vec{v}|}$ because

$$\nabla f \cdot \vec{v} = |\nabla f||\vec{v}||\cos(\phi)| \leq |\nabla f||\vec{v}| \,.$$

Definition: The direction $\vec{v} = \nabla f/|\nabla f|$ is the direction, where f **increases** most. It is the direction of **steepest ascent**.

12.7. If $\vec{v} = \nabla f/|\nabla f|$, then the directional derivative is $\nabla f \cdot \nabla f/|\nabla f| = |\nabla f|$. This means f **increases**, if we move into the direction of the gradient. The slope in that direction is $|\nabla f|$.

Definition: If f is a function of several variables and $\vec{v}$ is a unit vector then $D_{\vec{v}}f = \nabla f \cdot \vec{v}$ is called the **directional derivative** of f in the direction $\vec{v}$.

12.8. The name "directional derivative" is related to the fact that every unit vector gives a direction. If $\vec{v}$ is a unit vector, then the chain rule tells us $\frac{d}{dt} D_{\vec{v}} f = \frac{d}{dt} f(x + t\vec{v})$. The directional derivative tells us how the function changes when we move in a given direction. Assume for example that $T(x, y, z)$ is the temperature at position (x, y, z). If we move with velocity $\vec{v}$ through space, then $D_{\vec{v}} T$ tells us at which rate the temperature changes for us. If we move with velocity $\vec{v}$ on a hilly surface of height $h(x, y)$, then $D_{\vec{v}} h(x, y)$ gives us the slope we drive on.

12.9. If $\vec{r}(t)$ is a curve with velocity $\vec{r}\,'(t)$ and the speed is 1, then $D_{\vec{r}'(t)} f = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$ is the temperature change, one measures at $\vec{r}(t)$. The chain rule told us that this is $d/dt f(\vec{r}(t))$.

12.10. For $\vec{v} = (1, 0, 0)$, then $D_{\vec{v}} f = \nabla f \cdot v = f_x$, the directional derivative is a generalization of the partial derivatives. It measures the rate of change of f, if we walk with unit speed into that direction. But as with partial derivatives, it is a **scalar**.

12.11. The directional derivative satisfies $\boxed{|D_{\vec{v}} f| \leq |\nabla f||\vec{v}|}$ because $\nabla f \cdot \vec{v} = |\nabla f||\vec{v}|| \cos(\phi)| \leq |\nabla f||\vec{v}|$.

Definition: The direction $\vec{v} = \nabla f / |\nabla f|$ is the direction, where f **increases** most. It is the direction of **steepest ascent**.

12.12. If $\vec{v} = \nabla f / |\nabla f|$, then the directional derivative is $\nabla f \cdot \nabla f / |\nabla f| = |\nabla f|$. This means f **increases**, if we move into the direction of the gradient. The slope in that direction is $|\nabla f|$.

12.13. The directional derivative has the same properties than any derivative: $D_v(\lambda f) = \lambda D_v(f)$, $D_v(f + g) = D_v(f) + D_v(g)$ and $D_v(fg) = D_v(f)g + f D_v(g)$.
We will see later that points with $\nabla f = \vec{0}$ are candidates for **local maxima** or **minima** of f. Points (x, y), where $\nabla f(x, y) = (0, 0)$ are called **critical points** and help to understand the function f.

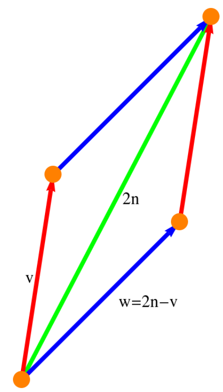
Examples

12.14. Compute the tangent plane to the surface $3x^2 y + z^2 - 4 = 0$ at the point $(1, 1, 1)$. **Solution:** $\nabla f(x, y, z) = [6xy, 3x^2, 2z]^T$. And $\nabla f(1, 1, 1) = [6, 3, 2]^T$. The plane is $6x + 3y + 2z = d$ where d is a constant. We can find the constant d by plugging in a point and get $6x + 3y + 2z = 11$.

12.15. Problem: reflect the ray $\vec{r}(t) = [1 - t, -t, 1]^T$ at the surface

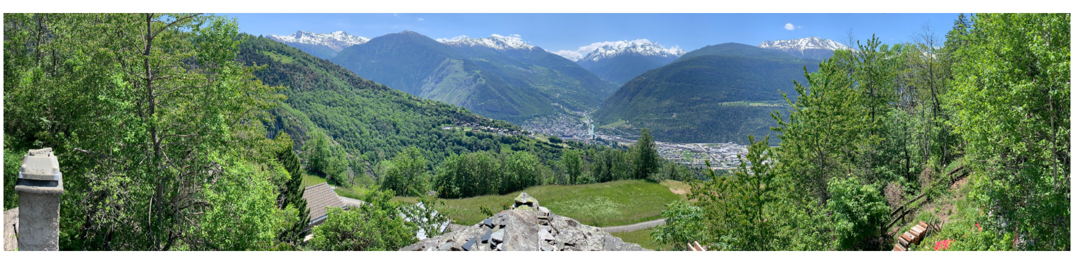
$$x^4 + y^2 + z^6 = 6 \ .$$

Solution: $\vec{r}(t)$ hits the surface at the time $t = 2$ in the point $(-1, -2, 1)$. The velocity vector in that ray is $\vec{v} = [-1, -1, 0]^T$. The normal vector at this point is $\nabla f(-1, -2, 1) = [-4, -4, 6]^T = \vec{n}$. The reflected vector is $R(\vec{v} = 2\text{Proj}_{\vec{n}}(\vec{v}) - \vec{v}$. We have $\text{Proj}_{\vec{n}}(\vec{v}) = 8/68[-4, -4, 6]^T$. Therefore, the reflected ray is $\vec{w} = (4/17)[-4, -4, 6]^T - [-1, -1, 0]^T$.



12.16. You are on a trip in a air-ship over Cambridge at $(1,2)$ and you want to avoid a thunderstorm, a region of low pressure. The pressure is given by a function $p(x,y) = x^2 + 2y^2$. In which direction do you have to fly so that the pressure change is largest? **Solution:** The gradient $\nabla p(x,y) = [2x, 4y]^T$ at the point $(1,2)$ is $[2,8]^T$. Normalize to get the direction $[1,4]^T/\sqrt{17}$.

12.17. The "Dom" is a mountain in Switzerland with an altitude of 4'545 meters. In suitable units on the ground, the height $f(x,y)$ is approximated by the quadratic function $f(x,y) = 4000 - x^2 - y^2$. At height $f(-10,10) = 3800$, at the point $(-10,10,3800)$, you rest. The climbing route continues into the south-east direction $v = [1,-1]^T/\sqrt{2}$. Calculate the rate of change in that direction. We have $\nabla f(x,y) = [-2x,-2y]^T$, so that $[20,-20]^T \cdot [1,-1]^T/\sqrt{2} = 40/\sqrt{2}$. This is a place, with a ladder, where you climb $40/\sqrt{2}$ meters up when advancing 1m forward. The rate of change in all directions is zero if and only if $\nabla f(x,y) = 0$: if $\nabla f \neq \vec{0}$, we can choose $\vec{v} = \nabla f/|\nabla f|$ and get $D_{\nabla f} f = |\nabla f|$.



Dom as seen from the Alp Salmenfee in Switzerland

12.18. Assume we know $D_v f(1,1) = 3/\sqrt{5}$ and $D_w f(1,1) = 5/\sqrt{5}$, where $v = [1,2]^T/\sqrt{5}$ and $w = [2,1]^T/\sqrt{5}$. Find the gradient of f. Note that we do not know anything else about the function f. **Solution:** Let $\nabla f(1,1) = [a,b]^T$. We know $a + 2b = 3$ and $2a + b = 5$. This allows us to get $a = 7/3, b = 1/3$.

Homework

This homework is due on Tuesday, 7/16/2019.

Problem 12.1: Find the directional derivative $D_{\vec{v}} f(2,1) = \nabla f(2,1) \cdot \vec{v}$ into the direction $\vec{v} = [3,-4]^T/5$ for the function $f(x,y) = 7 + x^5 y + y^3 + y$.

Problem 12.2: A surface $x^2 + y^2 - z = 1$ radiates light away. It can be parametrized as $\vec{r}(x,y) = [x, y, x^2 + y^2 - 1]^T$. Find the parametrization of the wave front $\vec{r}(x,y) + \vec{n}(x,y)$, which is distance 1 from the surface. Here $\vec{n}$ is a unit vector normal to the surface.

Problem 12.3: Assume $f(x,y) = 1 - x^2 + y^2$. Compute the directional derivative $D_{\vec{v}}(x,y)$ at $(0,0)$, where $\vec{v} = [\cos(t), \sin(t)]^T$ is a unit vector. Now compute

$$D_v D_v f(x,y)$$

at $(0,0)$, for any unit vector. For which values t is this **second directional derivative** positive?

Problem 12.4: The **Kitchen-Rosenberg formula** gives the curvature of a level curve $f(x,y) = c$ as

$$\kappa = \frac{f_{xx} f_y^2 - 2 f_{xy} f_x f_y + f_{yy} f_x^2}{(f_x^2 + f_y^2)^{3/2}}$$

Use this formula to find the curvature of the ellipse $f(x,y) = x^2 + 2y^2 = 1$ at the point $(1,0)$.

This formula is useful in computer vision. If you want to derive the formula, you can check that the angle

$$g(x,y) = \arctan(f_y/f_x)$$

of the gradient vector has κ as the directional derivative in the direction $\vec{v} = [-f_y, f_x]^T/\sqrt{f_x^2 + f_y^2}$ tangent to the curve.

Problem 12.5: One can find the maximum of a function numerically by moving in the direction of the gradient. This is called the **steepest ascent method**. You start at a point (x_0, y_0) then move in the direction of the gradient for some time c to be at $(x_1, y_1) = (x_0, y_0) + c\nabla f(x_0, y_0)$. Repeat to $(x_2, y_2) = (x_1, y_1) + c\nabla f(x_1, y_1)$ etc. It can be a bit difficult if the function has a flat ridge like in the **Rosenbrock function**

$$f(x,y) = 1 - (1-x)^2 - 100(y - x^2)^2 .$$

Plot the contour map of this function on $-0.6 \leq x \leq 1, -0.1 \leq y \leq 1.1$, then and find the directional derivative at $(1/5, 0)$ in the direction $(1,1)/\sqrt{2}$.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 13: Extrema

Lecture

13.1. In many applications we are led to the task to **maximize** or **minimize** a function f. As in single variable calculus we first search for points where the "derivative" is zero. This is the **Fermat principle** In one dimensions, like for $f(x) = 3x^5 - 5x^3$ we can use the second derivative test to classify extrema, like the local max at -1 and the local min at 1.

> **Definition:** A point (a, b) in the plane is called a **critical point** of a function $f(x, y)$ if $\nabla f(a, b) = [0, 0]^T$.

13.2. The **Fermat principle** predates the discovery of calculus and says:

> If $\nabla f(x, y)$ is not zero, then (x, y) is not a critical point.

13.3. Proof. Take the directional derivative in the direction $v = \nabla f/|\nabla f|$. Then $D_{\vec{v}} f = \nabla f \cdot \vec{v} = |\nabla f| > 0$. You can also see it with linear approximation. If $\nabla f \neq 0$, then the linear approximation L is not constant and f is neither maximal nor minimal. QED.

13.4. Note that in our definition we do **not** include points, where f or its derivative is not defined. Without stating otherwise, we always assume that a function can be differentiated arbitrarily often. Points where the function has no derivatives are not considered part of the domain and need to be studied separately. For the continuous function $f(x, y) = 1/\log(|xy|)$ for example, we would have to look at the points on the coordinate axes as well as the points $xy = 1$ separately.

13.5. In one dimension, we used the condition $f'(x) = 0, f''(x) > 0$ to get a local minimum and $f'(x) = 0, f''(x) < 0$ to assure a local maximum. If $f'(x) = 0, f''(x) = 0$, the nature of the critical point is undetermined and could be a maximum like for $f(x) = -x^4$, or a minimum like for $f(x) = x^4$ or a flat **inflection point** like for $f(x) = x^3$.

Definition: If $f(x,y)$ is a function of two variables with a critical point (a,b), the number $D = f_{xx}f_{yy} - f_{xy}^2$ is called the **discriminant** of the critical point.

13.6. The discriminant can be remembered better if seen as the determinant of the **Hessian matrix** $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$. As of default, we always assume that functions are twice continuously differentiable. Here is the **second derivative test**:

Theorem: Assume (a,b) is a critical point for $f(x,y)$.
If $D > 0$ and $f_{xx}(a,b) > 0$ then (a,b) is a local minimum.
If $D > 0$ and $f_{xx}(a,b) < 0$ then (a,b) is a local maximum.
If $D < 0$ then (a,b) is a saddle point.

13.7. In the case $D = 0$, we need higher derivatives or ad-hoc methods to determine the nature of the critical point.

13.8. To determine the maximum or minimum of $f(x,y)$ on a domain, determine all critical points **in the interior the domain**, and compare their values with maxima or minima **at the boundary**. We will see in the next unit how to get extrema on the boundary.
Sometimes, we want to find the overall maximum and not only the local ones.

Definition: A point (a,b) in the plane is called a **global maximum** of $f(x,y)$ if $f(x,y) \leq f(a,b)$ for all (x,y). For example, the point $(0,0)$ is a global maximum of the function $f(x,y) = 1 - x^2 - y^2$. Similarly, we call (a,b) a **global minimum**, if $f(x,y) \geq f(a,b)$ for all (x,y).

Examples

13.9. Find the critical points of $f(x,y) = x^4 + y^4 - 4xy + 2$. The gradient is $\nabla f(x,y) = [4(x^3 - y), 4(y^3 - x)]^T$ with critical points $(0,0), (1,1), (-1,-1)$.

13.10. $f(x,y) = \sin(x^2 + y) + y$. The gradient is $\nabla f(x,y) = [2x\cos(x^2 + y), \cos(x^2 + y) + 1]^T$. For a critical points, we must have $x = 0$ and $\cos(y) + 1 = 0$ which means $\pi + k2\pi$. The critical points are at $\dots (0,-\pi), (0,\pi), (0,3\pi), \dots$. There are infinitely many.

13.11. The graph of $f(x,y) = (x^2 + y^2)e^{-x^2-y^2}$ looks like a volcano. The gradient $\nabla f = [2x - 2x(x^2 + y^2), 2y - 2y(x^2 + y^2)]^T e^{-x^2-y^2}$ vanishes at $(0,0)$ and on the circle $x^2 + y^2 = 1$. This function has a continuum of critical points.

13.12. The function $f(x,y) = y^2/2 - g\cos(x)$ is the energy of the pendulum. The variable g is a constant. We have $\nabla f = (y, -g\sin(x)) = [(0,0]^T$ for

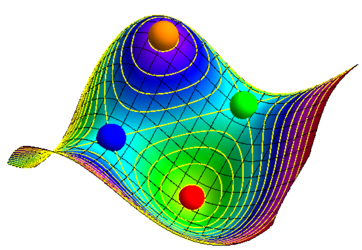
$$(x,y) = \dots, (-\pi,0), (0,0), (\pi,0), (2\pi,0), \dots .$$

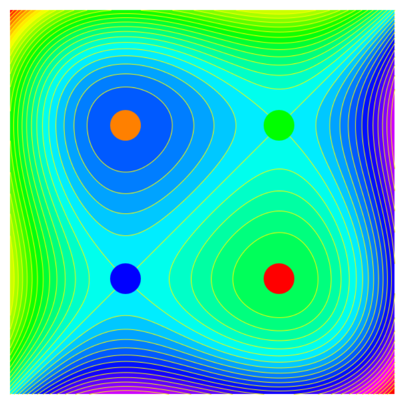
These points are equilibrium points, the angles for which the pendulum is at rest.

13.13. The function $f(x,y) = a\log(y) - by + c\log(x) - dx$ is a function which is invariant by the flow of the **Volterra-Lodka** differential equation $\dot{x} = ax - bxy, \dot{y} = -cy + dxy$. The point $(c/d, a/b)$ is a critical point of f and an equilibrium point of the system.

13.14. The function $f(x,y) = |x| + |y|$ is smooth on the first quadrant $\{x > 0, y > 0\}$. It does not have critical points there. The function has a minimum at $(0,0)$ but it is not in the domain, where f and ∇f are defined. We have to look at the points on the coordinate axis separately. For $y = 0$, we see that $x = 0$ is a minimum. For $x = 0$ we see that $y = 0$ is a minimum. Indeed $(0,0)$ is a minimum of f. This minimum was not detected using derivatives.

13.15. The function $f(x,y) = x^3/3 - x - (y^3/3 - y)$ has a graph which looks like a "napkin". It has the gradient $\nabla f(x,y) = [x^2 - 1, -y^2 + 1]^T$. There are 4 critical points $(1,1)$,$(-1,1)$,$(1,-1)$ and $(-1,-1)$. The Hessian matrix which includes all partial derivatives is $H = \begin{bmatrix} 2x & 0 \\ 0 & -2y \end{bmatrix}$.
For $(1,1)$ we have $D = -4$ and so a saddle point,
For $(-1,1)$ we have $D = 4$, $f_{xx} = -2$ and so a local maximum,
For $(1,-1)$ we have $D = 4$, $f_{xx} = 2$ and so a local minimum.
For $(-1,-1)$ we have $D = -4$ and so a saddle point. The function has a local maximum, a local minimum as well as 2 saddle points.





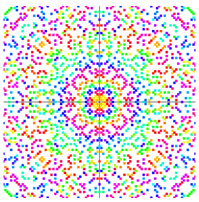
13.16. Find the maximum of $f(x,y) = 2x^2 - x^3 - y^2$ on $y \geq -1$. With $\nabla f(x,y) = 4x - 3x^2, -2y)$, the critical points are $(4/3, 0)$ and $(0,0)$. The Hessian is $H(x,y) = \begin{bmatrix} 4 - 6x & 0 \\ 0 & -2 \end{bmatrix}$. At $(0,0)$, the discriminant is -8 so that this is a saddle point. At $(4/3, 0)$, the discriminant is 8 and $H_{11} = 4/3$, so that $(4/3, 0)$ is a local maximum. We have now also to look at the boundary $y = -1$ where the function is $g(x) = f(x,-1) = 2x^2 - x^3 - 1$. Since $g'(x) = 0$ at $x = 0, 4/3$, where 0 is a local minimum, and $4/3$ is a local maximum on the line $y = -1$. Comparing $f(4/3, 0), f(4/3, -1)$ shows that $(4/3, 0)$ is the global maximum.

13.17. Find the global maxima and minima of $f(x,y) = x^4 + y^4 - 2x^2 - 2y^2$ **Solution:** the function has no global maximum. This can be seen by restricting the function to the x-axis, where $f(x,0) = x^4 - 2x^2$ is a function without maximum. The function has four global minima however. They are located on the 4 points $(\pm 1, \pm 1)$. The best way to see this is to note that $f(x,y) = (x^2 - 1)^2 + (y - 1)^2 - 2$ which is minimal when $x^2 = 1, y^2 = 1$.

Homework

This homework is due on Tuesday, 7/23/2019.

Problem 13.1: a) Find the critical points of $f(x,y) = (3x^4 - 8x^3 - 6x^2 + 24x) + (3y^4 - 8y^3 - 6y^2 + 24y)$.
b) Find all the extrema of the function $f(x,y) = xy + x^2y + xy^2$ and determine whether they are maxima, minima or saddle points.



Remark: b) is a **Gaussian Goldbach function** $f_n = \sum_{k+im \text{prime}, k,m \leq n} x^k y^m$. For $n = 2$ we sum over the Gaussian primes $1+i, 2+i, 1+2i$. (A complex integer $a+ib$ with $a, b \in \mathbb{N}$ is **prime** if a^2+b^2 is a usual prime. A 2D **Goldbach conjecture** claims that all partial derivatives $g^{(p,q)}(0,0)$ with $1 < p, q \leq n$ of $g(x,y) = f_{2n}^2(x,y)$ at $(0,0)$ are non-zero if $p+q$ is even. Equivalently, every Gaussian integer $a+ib$ with $a+b$ even and $a > 1, b > 1$ is a sum of two Gaussian primes in $Q = \{a+ib \mid a > 0, b > 0\}$.

Problem 13.2: Where on the parametrized surface $\vec{r}(u,v) = [1+u^3, v^2, uv]^T$ is the temperature $T(x,y,z) = 2 + x + 12y - 12z$ minimal? To find the minimum, look where the function $f(u,v) = T(\vec{r}(u,v))$ has an extremum. Find all local maxima, local minima or saddle points of f.

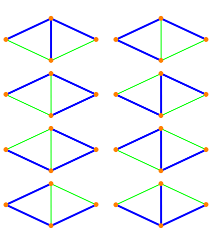
Problem 13.3: Find and classify all the extrema of the function $f(x,y) = e^{-x^2-y^2}(x^2+2y^2)$.

Problem 13.4: Find all extrema of the function $f(x,y) = 70 + x^3 + y^3 - 3x - 12y$ on the plane and characterize them. Do you find a global maximum or global minimum among them?

Problem 13.5: Graph theorists look at the **Tutte polynomial** $f(x,y)$ of a network. We work with the Tutte polynomial

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

of the **Kite network**. Classify using the second derivative test.



Remark. The polynomial is useful: $xf(1-x,0)$ tells in how many ways one can color the nodes of the network with x colors and $f(1,1)$ tells how many spanning trees there are. This picture illustrates that the number of spanning trees of the kite graph is $f(1,1) = 8$ as you see the 8 possible trees.

Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2019

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 14: Lagrange

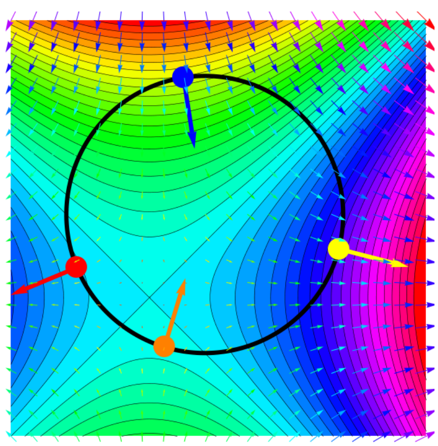
Lecture

14.1. When looking for maxima and minima of a function $f(x,y)$ in the presence of a **constraint** $g(x,y) = 0$, a necessary condition is that the gradients of f and g are parallel because otherwise, we can move along the curve $g = c$ and increase the value of f: the directional derivative of f in the direction tangent to the level curve is zero if and only if the tangent vector to g is perpendicular to the gradient of f. This can also include the case $\nabla g = [0,0]^T$.

Definition: The system of equations $\nabla f(x,y) = \lambda \nabla g(x,y), g(x,y) = 0$ for the three unknowns x, y, λ are the **Lagrange equations**. The variable λ is a **Lagrange multiplier**.

Theorem: A maximum or minimum of $f(x,y)$ on the curve $g(x,y) = c$ is either a solution of the Lagrange equations or is a critical point of g.

Proof. The condition that ∇f is parallel to ∇g either means $\nabla f = \lambda \nabla g$ or $\nabla f = 0$ or $\nabla g = 0$. The case $\nabla f = 0$ can be included in the Lagrange equation case with $\lambda = 0$. The case $\nabla g = 0$ which would lead to $\lambda = \infty$ has to be included separately. QED.



14.2. In higher dimensions the statement is exactly the same: extrema of $f(\vec{x})$ under the constraint $g(\vec{x}) = c$ are either solutions of the Lagrange equations $\nabla f = \lambda \nabla g, g = c$ or points where $\nabla g = \vec{0}$. But we also can have more than one constraint. For example:

Theorem: Extrema of $f(x,y,z)$ under the constraint $g(x,y,z) = c, h(x,y,z) = d$ are either solutions of the Lagrange equations $\nabla f = \lambda \nabla g + \mu \nabla h, g = c, h = d$ or solutions to $\nabla g = 0, \nabla f(x,y,z) = \mu \nabla h, h = d$ or solutions to $\nabla h = 0, \nabla f = \lambda \nabla g, g = c$ or solutions to $\nabla g = \nabla h = 0$.

14.3. Remarks.
1) The conditions in the Lagrange theorem are equivalent to $\nabla f \times \nabla g = 0$ in dimensions 2 or 3.
2) With $g(x,y) = 0$, the Lagrange equations can also be written as $\nabla F(x,y,\lambda) = 0$, where $F(x,y,\lambda) = f(x,y) - \lambda g(x,y)$.
3) The two conditions in the theorem are equivalent to "$\nabla g = \lambda \nabla f$ or f has a critical point".
4) Constrained optimization problems work also in higher dimensions.
5) Can we avoid Lagrange? Sometimes. It is often done in single variable calculus. In order to maximize xy under the constraint $2x + 2y = 4$ for example, we solve for y in the second equation and extremize the single variable problem $f(x, y(x))$. This needs to be done carefully and the boundaries must be considered. To extremize $f(x,y) = y$ on $x^2 + y^2 = 1$ for example we need to maximize $\sqrt{1 - x^2}$. We can differentiate to get the critical points but also have to look at the cases $x = 1$ and $x = -1$, where the actual minima and maxima occur. In general also, we can not do the substitution. To extremize $f(x,y) = x^2 + y^2$ with constraint $g(x,y) = x^4 + 3y^2 - 1 = 0$ for example, we solve $y^2 = (1 - x^4)/3$ and minimize $h(x) = f(x, y(x)) = x^2 + (1 - x^4)/3$. $h'(x) = 0$ gives $x = 0$. The find the maximum $(\pm 1, 0)$, we had to maximize $h(x)$ on $[-1, 1]$, which occurs at ± 1.
To extremize $f(x,y) = x^2 + y^2$ under the constraint $g(x,y) = p(x) + p(y) = 1$, where p is a complicated function in x which satisfies $p(0) = 0, p'(1) = 2$,the Lagrange equations $2x = \lambda p'(x), 2y = \lambda p'(y), p(x) + p(y) = 1$ can be solved with $x = 0, y = 1, \lambda = 1$. We can not solve $g(x,y) = 1$ however for y in an explicit way.
6) How do we determine whether a solution of the Lagrange equations is a maximum or minimum? Instead of using a second derivative test, we make a list of critical points and pick the maximum and minimum. A second derivative test can be designed using second directional derivative in the direction of the tangent.
7) The Lagrange method also works with more constraints. The constraints $g = c, h = d$ define a curve in space. The gradient of f must now be in the plane spanned by the gradients of g and h because otherwise, we could move along the curve and increase f:

Examples

14.4. Minimize $f(x,y) = x^2 + 2y^2$ under the constraint $g(x,y) = x + y^2 = 1$. **Solution:** The Lagrange equations are $2x = \lambda, 4y = \lambda 2y$. If $y = 0$ then $x = 1$. If $y \neq 0$ we can divide the second equation by y and get $2x = \lambda, 4 = \lambda 2$ again showing $x = 1$. The point $x = 1, y = 0$ is the only solution.

14.5. Find the shortest distance from the origin to the curve $x^6 + 3y^2 = 1$. **Solution:** Minimize the function $f(x,y) = x^2 + y^2$ under the constraint $g(x,y) = x^6 + 3y^2 = 1$. The gradients are $\nabla f = [2x, 2y]^T, \nabla g = [6x^5, 6y]^T$. The Lagrange equations $\nabla f = \lambda \nabla g$ lead to the system $2x = \lambda 6x^5, 2y = \lambda 6y, x^6 + 3y^2 - 1 = 0$. We get $\lambda = 1/3, x = x^5$, so that either $x = 0$ or 1 or -1. From the constraint equation $g = 1$, we obtain

$y = \sqrt{(1-x^6)/3}$. So, we have the solutions $(0, \pm\sqrt{1/3})$ and $(1,0), (-1,0)$. To see which is the minimum, just evaluate f on each of the points. $(0, \pm\sqrt{1/3})$ are the minima.

14.6. Which cylindrical soda cans of height h and radius r has minimal surface for fixed volume? **Solution:** The volume is $V(r,h) = h\pi r^2 = 1$. The surface area is $A(r,h) = 2\pi rh + 2\pi r^2$. With $x = h\pi, y = r$, you need to optimize $f(x,y) = 2xy + 2\pi y^2$ under the constrained $g(x,y) = xy^2 = 1$. Calculate $\nabla f(x,y) = (2y, 2x + 4\pi y), \nabla g(x,y) = (y^2, 2xy)$. The task is to solve $2y = \lambda y^2, 2x + 4\pi y = \lambda 2xy, xy^2 = 1$. The first equation gives $y\lambda = 2$. Putting that in the second one gives $2x + 4\pi y = 4x$ or $2\pi y = x$. The third equation finally reveals $2\pi y^3 = 1$ or $y = 1/(2\pi)^{1/3}, x = 2\pi(2\pi)^{1/3}$. This means $h = 0.54.., r = 2h = 1.08$. Remark: Other factors can influence the shape. For example, the can has to withstand a pressure up to 100 psi. A typical can of "Coca-Cola classic" with 3.7 volumes of CO_2 dissolve has at 75F an internal pressure of 55 psi, where PSI stands for pounds per square inch.

14.7. On the curve $g(x,y) = x^3 - y^2$ the function $f(x,y) = x$ obviously has a minimum $(0,0)$. The Lagrange equations $\nabla f = \lambda \nabla g$ have no solutions. This is a case where the minimum is a solution to $\nabla g(x,y) = 0$.

14.8. Find the extrema of $f(x,y,z) = z$ on the sphere $g(x,y,z) = x^2 + y^2 + z^2 = 1$. Solution: compute the gradients $\nabla f(x,y,z) = (0,0,1)$, $\nabla g(x,y,z) = (2x, 2y, 2z)$ and solve $(0,0,1) = \nabla f = \lambda \nabla g = (2\lambda x, 2\lambda y, 2\lambda z), x^2 + y^2 + z^2 = 1$. The case $\lambda = 0$ is excluded by the third equation $1 = 2\lambda z$ so that the first two equations $2\lambda x = 0, 2\lambda y = 0$ give $x = 0, y = 0$. The 4'th equation gives $z = 1$ or $z = -1$. The minimum is the south pole $(0,0,-1)$ the maximum the north pole $(0,0,1)$.

14.9. A dice shows k eyes with probability p_k with k in $\Omega = \{1,2,3,4,5,6\ \}$. A probability distribution is a non-negative function p on Ω which sums up to 1. It can be written as a vector $(p_1, p_2, p_3, p_4, p_5, p_6)$ with $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1$. The **entropy** of the probability vector $\vec{p}$ is defined as $f(\vec{p}) = -\sum_{i=1}^{6} p_i \log(p_i) = -p_1 \log(p_1) - p_2 \log(p_2) - ... - p_6 \log(p_6)$. Find the distribution p which maximizes entropy under the constrained $g(\vec{p}) = p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1$. **Solution:** $\nabla f = (-1 - \log(p_1), \dots, -1 - \log(p_n))$, $\nabla g = (1, \dots, 1)$. The Lagrange equations are $-1 - \log(p_i) = \lambda, p_1 + \cdots + p_6 = 1$, from which we get $p_i = e^{-(\lambda+1)}$. The last equation $1 = \sum_i \exp(-(\lambda+1)) = 6\exp(-(\lambda+1))$ fixes $\lambda = -\log(1/6) - 1$ so that $p_i = 1/6$. The distribution, where each event has the same probability is the distribution of maximal entropy. Maximal entropy means **least information content**. An unfair dice allows a cheating gambler or casino to gain profit. Cheating through asymmetric weight distributions can be avoided by making the dices transparent.

14.10. Assume that the probability that a chemical system is in a state k is p_k and the energy of the state k is E_k. Nature tries to minimize the **free energy** $f(p_1, \dots, p_n) = -\sum_i [p_i \log(p_i) - E_i p_i]$ if the energies E_i are fixed. The probability distribution p_i satisfying $\sum_i p_i = 1$ minimizing the free energy is called a **Gibbs distribution**. Find this distribution in general if E_i are given. **Solution:** $\nabla f = (-1 - \log(p_1) - E_1, \dots, -1 - \log(p_n) - E_n)$, $\nabla g = (1, \dots, 1)$. The Lagrange equation are $\log(p_i) = -1 - \lambda - E_i$, or $p_i = \exp(-E_i)C$, where $C = \exp(-1-\lambda)$. The constraint $p_1 + \cdots + p_n = 1$

gives $C(\sum_i \exp(-E_i)) = 1$ so that $C = 1/(\sum_i e^{-E_i})$. The Gibbs solution is $p_k = \exp(-E_k)/\sum_i \exp(-E_i)$.

15. Homework

This homework is due on Tuesday, 7/23/2019.

Problem 14.1: Find the cylindrical basket which is open on the top has has the largest volume for fixed area π. If x is the radius and y is the height, we have to extremize $f(x, y) = \pi x^2 y$ under the constraint $g(x, y) = 2\pi xy + \pi x^2 = \pi$. Use the method of Lagrange multipliers.

Problem 14.2: Find the extrema of the same function

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

you have seen in the last homework set but now on the entire disc $\{x^2 + y^2 \leq 4\ \}$ of radius 2. Besides the already found extrema **inside** the disk, now find also the extrema on the boundary.

Problem 14.3: Motivated by the Disney movie "Tangled", we want to build a hot air balloon with a cuboid mesh of dimension x, y, z which together with the top and bottom fortifications uses wires of total length $g(x, y, z) = 6x + 6y + 4z = 32$. Find the balloon with maximal volume $f(x, y, z) = xyz$.

Problem 14.4: A **solid bullet** made of a half sphere and a cylinder has the volume $V = 2\pi r^3/3 + \pi r^2 h$ and surface area $A = 2\pi r^2 + 2\pi rh + \pi r^2$. Doctor Manhattan designs a bullet with fixed volume and minimal area. With $g = 3V/\pi = 1$ and $f = A/\pi$ he therefore minimizes $f(h, r) = 3r^2 + 2rh$ under the constraint $g(h, r) = 2r^3 + 3r^2 h = 1$. Use the Lagrange method to find a local minimum of f under the constraint $g = 1$.

Problem 14.5: Which pyramid of height h over a square $[-a, a] \times [-a, a]$ with surface area is $4a\sqrt{h^2 + a^2} + 4a^2 = 4$ has maximal volume $V(h, a) = 4ha^2/3$? By using new variables (x, y) and multiplying V with a constant, we get to the equivalent problem to maximize $f(x, y) = yx^2$ over the constraint $g(x, y) = x\sqrt{y^2 + x^2} + x^2 = 1$. Use the later variables.

MULTIVARIABLE CALCULUS

MATH S-21A

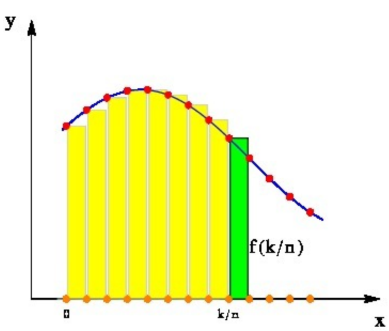
Unit 15: Double Integrals

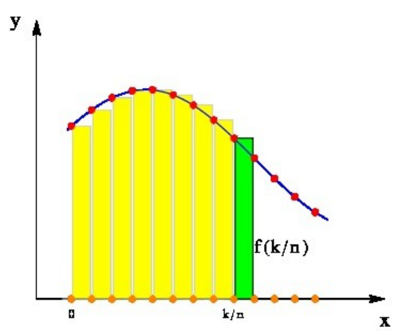
Lecture

15.1. If $f(x)$ is a continuous function, the **Riemann integral** $\int_a^b f(x)\ dx$ is defined as the limit of the **Riemann sums** $S_n f(x) = \frac{1}{n}\sum_{k/n \in [a,b]} f(k/n)$ for $n \to \infty$. The **derivative** of f is the limit of **difference quotients** $D_n f(x) = n[f(x+1/n) - f(x)]$ as $n \to \infty$. The integral $\int_a^b f(x)\ dx$ is the **signed area** under the graph of f and above the x-axes, where "signed" indicates that area below the x-axes has negative sign. The function $F(x) = \int_0^x f(y)\ dy$ is called an **anti-derivative** of f. It is determined up a constant. The **fundamental theorem of calculus** states

$$F'(x) = f(x), \int_0^x f(x) = F(x) - F(0) .$$

It allows to compute integrals by inverting differentiation so that **differentiation rules** become **integration rules**: the product rule leads to integration by parts, the chain rule becomes partial integration.





Definition: If $f(x,y)$ is continuous on a region R, the integral $\iint_R f(x,y)\ dxdy$ is defined as the limit of the Riemann sum

$$\frac{1}{n^2}\sum_{(\frac{i}{n},\frac{j}{n})\in R} f(\frac{i}{n},\frac{j}{n})$$

when $n \to \infty$. We write also $\iint_R f(x,y)\ dA$, where $dA = dxdy$ is a notation standing for "an area element".

15.2. Fubini's theorem allows to switch the order of integration over a rectangle if the function f is continuous:

Theorem: $\int_a^b \int_c^d f(x,y)\ dxdy = \int_c^d \int_a^b f(x,y)\ dydx$.

Proof. For every n, there is the "quantum Fubini identity"

$$\sum_{\frac{i}{n}\in[a,b]} \sum_{\frac{j}{n}\in[c,d]} f(\frac{i}{n},\frac{j}{n}) = \sum_{\frac{j}{n}\in[c,d]} \sum_{\frac{i}{n}\in[a,b]} f(\frac{i}{n},\frac{j}{n})$$

holding for all functions. Now divide both sides by n^2 and take the limit $n \to \infty$. This is possible for continuous functions. Fubini's theorem only holds for rectangles. We extend the class of regions now to so called Type I and Type II regions:

Definition: A **type I region** is of the form

$$R = \{(x,y) \mid a \leq x \leq b,\ c(x) \leq y \leq d(x)\ \} .$$

An integral over a type I region is called a **type I integral**

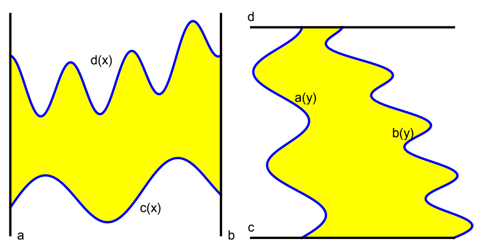
$$\iint_R f\ dA = \int_a^b \int_{c(x)}^{d(x)} f(x,y)\ dydx .$$

A **type II region** is of the form

$$R = \{(x,y) \mid c \leq y \leq d, a(y) \leq x \leq b(y)\ \} .$$

An integral over such a region is called a **type II integral**

$$\iint_R f\ dA = \int_c^d \int_{a(y)}^{b(y)} f(x,y)\ dxdy .$$



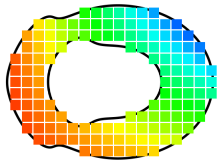
15.3. Similarly as we could see in one dimensions an integral as a signed area, one can interpret $\int\int_R f(x,y)\ dydx$ as the **signed volume** of the solid below the graph of f and above R in the xy plane. As in 1D integration, the volume of the solid below the xy-plane is counted negatively.

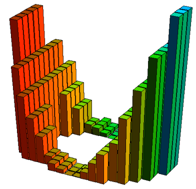
Examples

15.4. If we integrate $f(x,y) = xy$ over the unit square we can sum up the Riemann sum for fixed $y = j/n$ and get $y/2$. Now perform the integral over y to get $1/4$. This example shows how to reduce double integrals to single variable integrals.

15.5. If $f(x,y) = 1$, then the integral is the **area** of the region R. The integral is the limit $L(n)/n^2$, where $L(n)$ is the number of lattice points $(i/n, j/n)$ contained in R.

15.6. The value $\iint_R f(x,y)\ dA / \iint_R 1\ dA$ is the **average** value of f.





15.7. Integrate $f(x,y) = x^2$ over the region bounded above by $\sin(x^3)$ and bounded below by the graph of $-\sin(x^3)$ for $0 \leq x \leq \pi$. The value of this integral has a physical meaning. It is called **moment of inertia**.

$$\int_0^{\pi^{1/3}} \int_{-\sin(x^3)}^{\sin(x^3)} x^2\ dydx = 2\int_0^{\pi^{1/3}} \sin(x^3)x^2\ dx$$

We have now an integral, which we can solve by substitution $-\frac{2}{3}\cos(x^3)|_0^{\pi^{1/3}} = \frac{4}{3}$.

15.8. Integrate $f(x,y) = y^2$ over the region bound by the x-axes, the lines $y = x+1$ and $y = 1-x$. The problem is best solved as a type I integral. As you can see from the picture, we would have to compute 2 different integrals as a type I integral. To do so, we have to write the bounds as a function of y: they are $x = y-1$ and $x = 1-y$

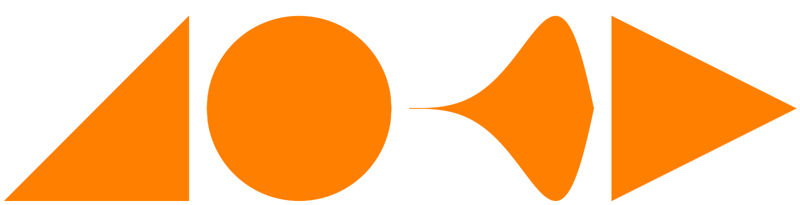
$$\int_0^1 \int_{x-1}^{1-x} y^2\ dy\ dx = 1/6\ .$$

15.9. Let R be the triangle $1 \geq x \geq 0, 0 \leq y \leq x$. What is

$$\int\int_R e^{-x^2}\ dxdy\ ?$$

The type II integral $\int_0^1 [\int_y^1 e^{-x^2}\ dx]dy$ can not be solved because e^{-x^2} has no anti-derivative in terms of elementary functions.
The type I integral $\int_0^1 [\int_0^x e^{-x^2}\ dy]\ dx$ however can be solved:

$$= \int_0^1 xe^{-x^2}\ dx = -\frac{e^{-x^2}}{2}|_0^1 = \frac{(1-e^{-1})}{2} = 0.316...\ .$$



15.10. The area of a disc of radius R is $\int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} 1\ dydx = \int_{-R}^{R} 2\sqrt{R^2-x^2}\ dx$. Substitute $x = R\sin(u), dx = R\cos(u)$, to get $\int_{-\pi/2}^{\pi/2} 2\sqrt{R^2 - R^2\sin^2(u)}R\cos(u)\ du = \int_{-\pi/2}^{\pi/2} 2R^2\cos^2(u)\ du = R^2\pi$.

Homework

This homework is due on Tuesday, 7/23/2019.

Problem 15.1: a) (4 points) Find the iterated integral

$$\int_0^1 \int_0^2 3xy/\sqrt{x^2+(y^2/2)}\, dy\, dx \; .$$

b) (4 points) Now compute

$$\int_0^1 \int_0^2 3xy/\sqrt{x^2+y^2/2}\, dx\, dy \; .$$

c) (2 points) Wouldn't Fubini assure that a) and b) are the same? What change would be needed in b) to make the results agree?

Problem 15.2: Find the area of the region

$$R = \{(x,y) \mid 0 \le x \le 2\pi, \sin(x) - 1 \le y \le \cos(x) + 2\}$$

and use it to compute the average value $\int\int_R f(x,y)\; dxdy/\text{area}(R)$ of $f(x,y) = y$ over that region.

Problem 15.3: Find the volume of the solid lying under the paraboloid $z = x^2 + y^2$ and above the rectangle $R = [-2,2] \times [-3,6] = \{(x,y) \mid -2 \le x \le 2, -3 \le y \le 6\ \}$.

Problem 15.4: Calculate the iterated integral $\int_0^1 \int_x^{2-x} (x^2 - y)\; dydx$. Sketch the corresponding type I region. Write this integral as integral over a type II region and compute the integral again.

Problem 15.5: There is only one way to identify zombies: throw two difficult integrals at them and see whether they can solve them. Prove that you are not a zombie!
a) (6 points) Find the integral

$$\int_0^1 \int_{\sqrt{y}}^{y^2} \frac{3x^7}{\sqrt{x} - x^2}\, dx\, dy \; .$$

b) (4 points) Integrate

$$\int_0^1 \int_0^{\sqrt{1-y^2}} 11(x^2+y^2)^{10}\; dxdy \; .$$

You might want to "time travel" one lecture forward, where polar coordinates are known to solve this problem.

Unit 16: Surface Integration

Lecture

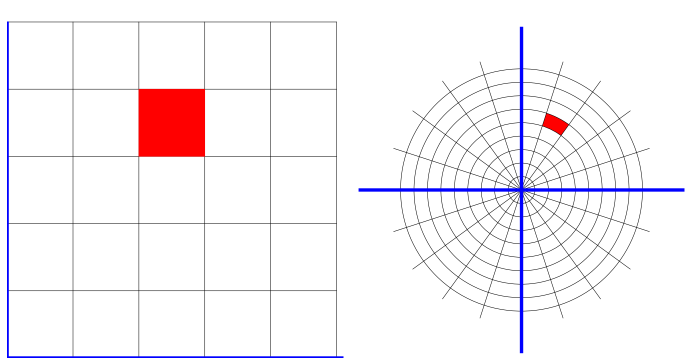
16.1. For certain regions, it is better to use different coordinate system. A reparametrization $(x, y) = \vec{r}(u, v)$ often helps. This works then also in higher dimensions, when suraces are parametrized as $(x, y, z) = \vec{r}(u, v)$. But first to the two dimensional case, where polar coordinates $(x, y) = (r\cos(\theta), r\sin(\theta)$ are an important example

Definition: A **polar region** is a planar region bound by a simple closed curve given in polar coordinates as the curve $(r(t), \theta(t))$. The most common case is $\theta(t) = t$. In Cartesian coordinates the parametrization of the boundary of a polar region is $\vec{r}(t) = [r(t)\cos(\theta(t), r(t)\sin(\theta(t)]^T$, a **polar graph** like the spiral with $\theta(t) = t$.

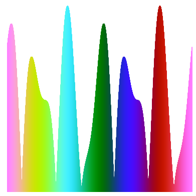
Theorem: To integrate in polar coordinates, we evaluate the integral

$$\iint_R f(x, y)\ dxdy = \iint_R f(r\cos(\theta), r\sin(\theta)r\ drd\theta$$

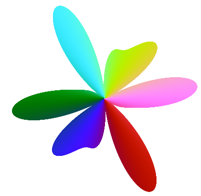
16.2. Why do we have to include the factor r, when we move to polar coordinates? The reason is that a small rectangle R with dimensions $d\theta dr$ in the (r, θ) plane is mapped by $T : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$ to a sector segment S in the (x, y) plane. It has the area $r\ d\theta dr$. If you have seen some linear algebra, note that the Jacobean matrix dT has the determinant r.



16.3. We can now integrate over type I or type II regions in the (θ, r) plane. like **flowers**: $\{(\theta, r) \,|0 \leq r \leq f(\theta)\}$ where $f(\theta)$ is a periodic function of θ.



A polar region shown in polar coordinates. It is a type I region.

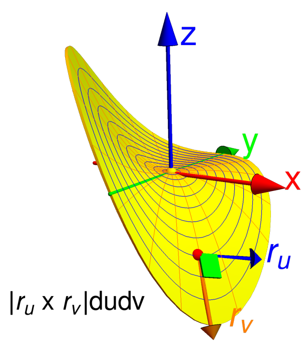


The same region in the xy coordinate system is not type I or II.

Theorem: A surface $\vec{r}(u,v)$ parametrized on a parameter domain R has the **surface area**

$$\int\int_R |\vec{r}_u(u,v) \times \vec{r}_v(u,v)| \; dudv \; .$$

Proof. The vector $\vec{r}_u$ is tangent to the grid curve $u \mapsto \vec{r}(u,v)$ and $\vec{r}_v$ is tangent to $v \mapsto \vec{r}(u,v)$, the two vectors span a parallelogram with area $|\vec{r}_u \times \vec{r}_v|$. A small rectangle $[u, u+du] \times [v, v+dv]$ is mapped by $\vec{r}$ to a parallelogram spanned by $[\vec{r}, \vec{r}+\vec{r}_u]$ and $[\vec{r}, \vec{r}+\vec{r}_v]$ which has the area $|\vec{r}_u(u,v) \times \vec{r}_v(u,v)| \; dudv$.

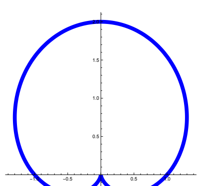


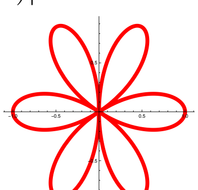
Examples

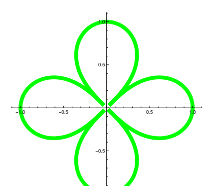
16.4. The polar graph defined by $r(\theta) = |\cos(3\theta)|$ belongs to the class of **roses** $r(t) = |\cos(nt)|$. Regions enclosed by this graph are also called **rhododenea**.

16.5. The polar curve $r(\theta) = 1 + \sin(\theta)$ is called a **cardioid**. It looks like a heart. It belongs to the class of **limacon** curves $r(\theta) = 1 + b\sin(\theta)$.

16.6. The polar curve $r(\theta) = |\sqrt{\cos(2t)}|$ is called a **lemniscate**.







16.7. Integrate

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

over the unit disc. We have $f(x,y) = f(r\cos(\theta), r\sin(\theta)) = r^2 + r^2\cos(\theta)\sin(\theta)$ so that $\iint_R f(x,y)\ dxdy = \int_0^1 \int_0^{2\pi} (r^2 + r^2\cos(\theta)\sin(\theta)) r\ d\theta dr = 2\pi/4$.

16.8. We have earlier computed area of the disc $\{x^2 + y^2 \leq 1\}$ using substitution. It is more elegant to do this integral in polar coordinates:

$$\int_0^{2\pi} \int_0^1 r\ drd\theta = 2\pi r^2/2|_0^1 = \pi \ .$$

16.9. Integrate the function $f(x,y) = 1$ $\{(\theta, r(\theta)) \mid r(\theta) \leq \ |\ \cos(3\theta)|\ \}$.

$$\int\int_R 1\ dxdy = \int_0^{2\pi} \int_0^{\cos(3\theta)} r\ dr\ d\theta = \int_0^{2\pi} \frac{\cos(3\theta)^2}{2}\ d\theta = \pi/2 \ .$$

16.10. Integrate $f(x,y) = y\sqrt{x^2+y^2}$ over the region $R = \{(x,y) \mid 1 < x^2 + y^2 < 4, y > 0\}$.

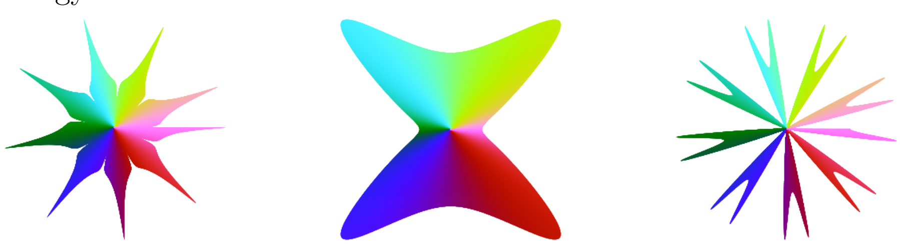
$$\int_1^2 \int_0^{\pi} r\sin(\theta) r\ r\ d\theta dr = \int_1^2 r^3 \int_0^{\pi} \sin(\theta)\ d\theta dr = \frac{(2^4 - 1^4)}{4} \int_0^{\pi} \sin(\theta)\ d\theta = 15/2$$

For integration problems, where the region is part of an annular region, or if you see function with terms $x^2 + y^2$ try to use polar coordinates $x = r\cos(\theta), y = r\sin(\theta)$.

16.11. The Belgian Biologist **Johan Gielis** defined in 1997 with the family of curves given in polar coordinates as

$$r(\phi) = \left(\frac{|\cos(\frac{m\phi}{4})|^{n_1}}{a} + \frac{|\sin(\frac{m\phi}{4})|^{n_2}}{b}\right)^{-1/n_3}$$

This so called **super-curve** can produce a variety of shapes like circles, square, triangle, stars. It can also be used to produce "super-shapes". The super-curve generalizes the **super-ellipse** which had been discussed in 1818 by Lamé and helps to **describe forms** in biology. [1]



16.12. The parametrized surface $\vec{r}(u,v) = [2u, 3v, 0]^T$ is part of the xy-plane. The parameter region G just gets stretched by a factor 2 in the x coordinate and by a factor 3 in the y coordinate. $\vec{r}_u \times \vec{r}_v = [0,0,6]^T$ and we see for example that the area of $\vec{r}(G)$ is 6 times the area of G.

[1]"Gielis, J. A 'generic geometric transformation that unifies a wide range of natural and abstract shapes'. American Journal of Botany, 90, 333 - 338, (2003).

16.13. The map $\vec{r}(u,v) = [L\cos(u)\sin(v), L\sin(u)\sin(v), L\cos(v)]^T$ maps the rectangle $G = [0,2\pi]\times[0,\pi]$ onto the sphere of radius L. We compute $\vec{r}_u \times \vec{r}_v = L\sin(v)\vec{r}(u,v)$. So, $|\vec{r}_u \times \vec{r}_v| = L^2|\sin(v)|$ and $\int\int_R 1 \, dS = \int_0^{2\pi}\int_0^{\pi} L^2\sin(v)\, dvdu = 4\pi L^2$

16.14. For graphs $(u,v) \mapsto [u,v,f(u,v)]^T$, we have $\vec{r}_u = (1,0,f_u(u,v))$ and $\vec{r}_v = (0,1,f_v(u,v))$. The cross product $\vec{r}_u \times \vec{r}_v = (-f_u, -f_v, 1)$ has the length $\sqrt{1+f_u^2+f_v^2}$. The area of the surface above a region G is $\int\int_G \sqrt{1+f_u^2+f_v^2}\, dudv$.

16.15. Lets take a surface of revolution $\vec{r}(u,v) = [v, f(v)\cos(u), f(v)\sin(u)]^T$ on $R = [0,2\pi]\times[a,b]$. We have $\vec{r}_u = (0, -f(v)\sin(u), f(v)\cos(u))$, $\vec{r}_v = (1, f'(v)\cos(u), f'(v)\sin(u))$ and $\vec{r}_u \times \vec{r}_v = (-f(v)f'(v), f(v)\cos(u), f(v)\sin(u)) = f(v)(-f'(v), \cos(u), \sin(u))$. The surface area is $\int\int |\vec{r}_u \times \vec{r}_v|\, dudv = 2\pi \int_a^b |f(v)|\sqrt{1+f'(v)^2}\, dv$.

Homework

This homework is due on Tuesday, 7/23/2019.

Problem 16.1: a) A city near the sea is modeled by a half disk $D = \{(x,y) \mid x^2+y^2 \leq 49, x \geq 0$ with center the origin and radius 7. What is the average distance of a point in D to the origin? in other words, what is the integral $\int\int_D \sqrt{x^2+y^2}\, dxdy / \int\int_D 1\, dxdy$.
b) The distance to the beach is x. Find the average distance $\int\int_D x\, dxdy / \int\int_D 1\, dxdy$ to the beach.

Problem 16.2: Find $\int\int_R (x^2+y^2)^{44}\, dA$, where R is the part of the unit disc $\{x^2+y^2 \leq 1\}$ for which $y > x$.

Problem 16.3: What is the area of the region which is bounded by the following three curves, first by the polar curve $r(\theta) = \theta$ with $\theta \in [0,2\pi]$, second by the polar curve $r(\theta) = 2\theta$ with $\theta \in [0,2\pi]$ and third by the positive x-axis?

Problem 16.4: The average of a function f on a region is defined as

$$\frac{\int_R f\, dxdy}{\int_R 1\, dxdy}.$$

Find the average value of $f(x,y) = 2(x^2+y^2)$ on the annular region $R: 1 \leq |(x,y)| \leq 2$.

Problem 16.5: Find the surface area of the part of the paraboloid $x = y^2+z^2$ which is inside the cylinder $y^2+z^2 \leq 16$.

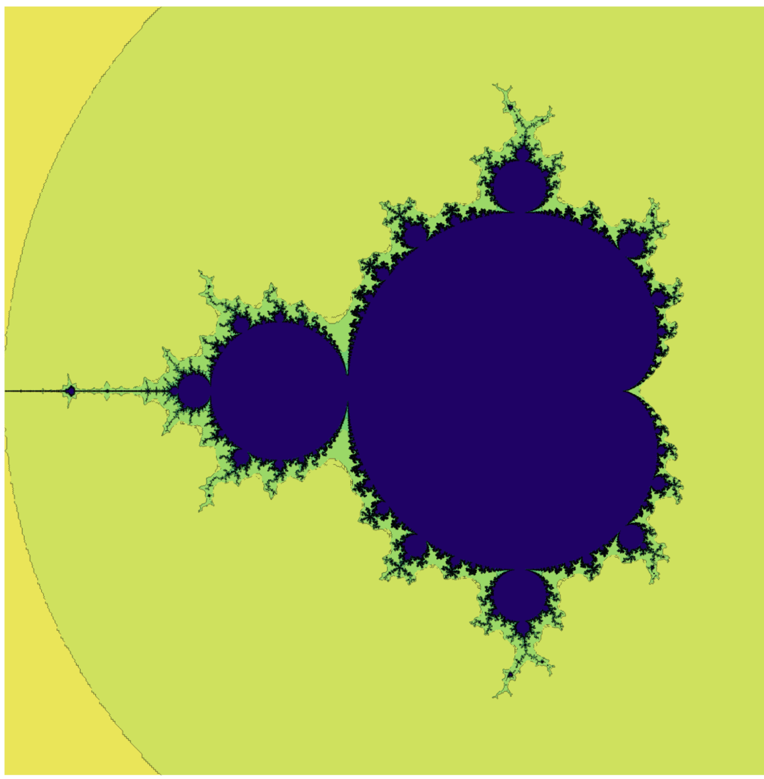
Postscript: Area of the Mandelbrot set

16.16. Often, when we deal with real data, we do not have analytic expressions for the region or function we want to integrate. We want to elaborate here on an example mentioned already in the text. It is the problem to find the area of Mandelbrot set

$$M = \{c = a + ib \in \mathbb{C} \in \mathbb{R}^2 \mid T_c(0)^n \text{ stays bounded } \} ,$$

where $T_c(z) = z^+c$ (as complex numbers, which is written out in real coordinates the map $T_c(x, y) = (x^2 - y^2 + a, 2xy + b)$.

16.17. Here is a picture: it can also be visualized as a function which is 1 on the Mandelbrot set and 0 else.



16.18. What is the area of the Mandelbrot set? We know it is contained in the rectangle $x \in [-2, 1]$ and $y \in [-3/2, 3/2]$. We now just randomly shoot into this rectangle and see whether we are in the Mandelbrot set or not after 1000 iterations. Here is some Mathematica code which allows you to compute things. When we ran it, it gave a value of about 1.515.... More accurate measurements reported hint for a slightly smaller value like 1.506.... Others have given bounds $[1.50311, 1.5613027]$.

```
M=Compile[{x,y},Module[{z=x+I y,k=0},
    While[Abs[z]<2.&&k<1000,z=N[z^2+x+I y];++k];Floor[k/1000]]];
9*Sum[M[-2+3 Random[],-1.5+3 Random[]],{1000000}]/1000000
```

How accurately can you compute the area of the Mandelbrot set? It is a data problem unless somebody comes up with a formula.

Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2019

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 17: Triple integrals

LECTURE

17.1. Integrating over higher dimensional regions is done in the same way than in two dimensions. Three dimensional regions are referred to as **solids**.

Definition: If $f(x, y, z)$ is continuous and E is a **bounded solid** in $\mathbb{R}^3$, then $\iint_E f(x, y, z)\ dxdydz$ is defined as the $n \to \infty$ limit of the Riemann sum

$$\frac{1}{n^3} \sum_{(\frac{i}{n},\frac{j}{n},\frac{k}{n}) \in E} f(\frac{i}{n}, \frac{j}{n}, \frac{k}{n})\ .$$

Triple integrals can be evaluated by iterated single integrals:

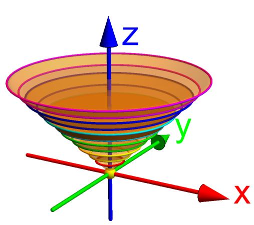
17.2. If E is the box $\{x \in [1, 2], y \in [0, 1], z \in [0, 1]\}$ and $f(x, y, z) = 24x^2y^3z$.

$$\int_0^1 \int_0^1 \int_0^1 24x^2y^3z\ dz\ dy\ dx.$$

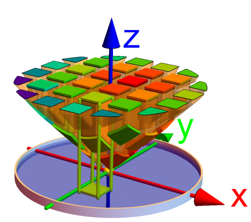
To evaluate the integral, start from the inside $\int_0^1 24x^2y^3z\ dz = 12x^3y^3$, then then integrate the middle layer, $\int_0^1 12x^3y^3\ dy = 3x^2$ and finally and finally handle the most outer layer: $\int_1^2 3x^2 dx = 7$.

For the inner integral, $x = x_0$ and $y = y_0$ are fixed. The middle integral now computes the contribution over a slice $z = z_0$ intersected with R. The outer integral sums up all these slice contributions.

17.3. There are two reductions possible to compute triple integrals:



The **burger method** slices the solid a line and computes $\int_a^b \iint_{R(z)} f(x, y, z)\ dA\ dz$, where $g(z)$ is a double integral giving the values when integrating over cheese, meet or tomato. The **fries method** eats up fries going from $g(x, y)$ to $h(x, y)$ over a region R. We have $\iint_R [\int_{g(x,y)}^{h(x,y)} f(x, y, z)\ dz]\ dA$.



17.4. A special case is the **signed volume**

$$\int\int_R \int_0^{f(x,y)} 1 \; dzdxdy \; .$$

below the graph of a function $f(x,y)$ and above a region R, considered part of the xy-plane. It is the integral $\int\int_R f(x,y) \; dA$. The triple integral which is more natural when considering physical units as volume is measured in cubic meters for example. The triple integral also allows for flexibility: we can replace 1 with a function $f(x,y,z)$. If interpreted as a **charge density**, then the integral is the total charge.

17.5. The problem of computing volumes has been tackled early. **Archimedes (287-212 BC)** already developed an integration method which allowed him to find areas, volumes and surface areas in many cases without calculus. His method of **exhaustion** is close to the numerical method of integration by Riemann sum. In our terminology, Archimedes used the **washer method** to reduce the problem to a single variable problem. The **Archimedes principle** states that any body submerged in a water is acted upon by an upward force which is equal to the weight of the displaced water. This provides a practical way to compute volumes of complicated bodies. A second method, the **displacement method** is a **comparison technique**: the area of a sphere is the area of the cylinder enclosing it. The volume of a sphere is the volume of the complement of a cone in that cylinder. Modern rearrangement techniques use this still today in modern analysis. Heureka! **Cavalieri (1598-1647)** would build on Archimedes ideas and determine area and volume using tricks now called the **Cavalieri principle**. An example already due to Archimedes is the computation of the volume the half sphere of radius R, cut away a cone of height and radius R from a cylinder of height R and radius R. At height z, this body has a cross section with area $R^2\pi - r^2\pi$. If we cut the half sphere at height z, we obtain a disc of area $(R^2 - r^2)\pi$. Because these areas are the same, the volume of the half-sphere is the same as the cylinder minus the cone: $\pi R^3 - \pi R^3/3 = 2\pi R^3/3$ and the volume of the sphere is $4\pi R^3/3$. **Newton (1643-1727) and Leibniz(1646-1716)** developed calculus independently. It provided a new tool which made it possible to compute integrals through "anti-derivation". Suddenly, it became possible to find integrals using analytic tools. We can do this also in higher dimensions.

Examples

17.6. Find the volume of the unit sphere. **Solution:** The sphere is sandwiched between the graphs of two functions obtained by solving for z. Let R be the unit disc in the xy plane. If we use the **sandwich method**, we get

$$V = \int\int_R [\int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} 1 dz] dA \; .$$

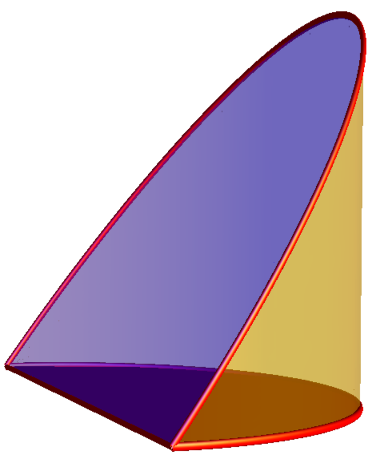
which gives a double integral $\int\int_R 2\sqrt{1-x^2-y^2} \; dA$ which is of course best solved in polar coordinates. We have $\int_0^{2\pi}\int_0^1 \sqrt{1-r^2} r \; drd\theta = 4\pi/3$.
With the **washer method** which is in this case also called **disc method**, we slice along the z axes and get a disc of radius $\sqrt{1-z^2}$ with area $\pi(1-z^2)$. This is a method suitable for single variable calculus because we get directly $\int_{-1}^1 \pi(1-z^2) \; dz = 4\pi/3$.

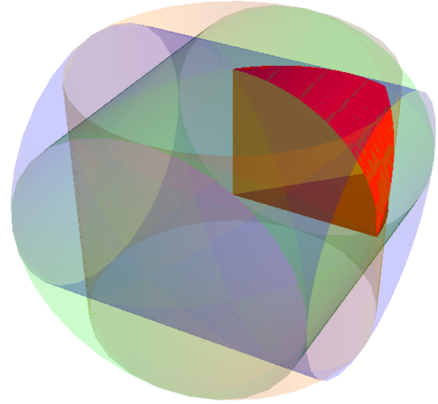
17.7. The mass of a body with mass density $\rho(x,y,z)$ is defined as $\int\int\int_R \rho(x,y,z)\; dV$. For bodies with constant density ρ, the mass is ρV, where V is the volume. Compute the mass of a body which is bounded by the parabolic cylinder $z = 4 - x^2$, and the planes $x = 0, y = 0, y = 6, z = 0$ if the density of the body is z. **Solution:**

$$\int_0^2 \int_0^6 \int_0^{4-x^2} z\; dz\; dy\; dx = \int_0^2 \int_0^6 (4-x^2)^2/2\; dydx$$

$$= \quad 6\int_0^2 (4-x^2)^2/2\; dx = 6(\frac{x^5}{5} - \frac{8x^3}{3} + 16x)|_0^2 = 2 \cdot 512/5$$

17.8. The solid region bound by $x^2 + y^2 = 1$, $x = z$ and $z = 0$ is called the **hoof of Archimedes**. It is historically significant because it is one of the first examples, on which Archimedes probed a Riemann sum integration technique. It appears in every calculus text book. Find the volume of the hoof. **Solution**. Look from the situation from above and picture it in the $x - y$ plane. You see a half disc R. It is the floor of the solid. The roof is the function $z = x$. We have to integrate $\int\int_R x\; dxdy$. We got a double integral problems which is best done in polar coordinates; $\int_{-\pi/2}^{\pi/2} \int_0^1 r^2 \cos(\theta)\; drd\theta = 2/3$.





17.9. Finding the volume of the solid region bound by the three cylinders $x^2+y^2 = 1$, $x^2 + z^2 = 1$ and $y^2 + z^2 = 1$ is one of the most famous volume integration problems going back to Archimedes.
Solution: look at 1/16'th of the body given in cylindrical coordinates $0 \le \theta \le \pi/4, r \le 1, z > 0$. The roof is $z = \sqrt{1-x^2}$ because above the "one eighth disc" R only the cylinder $x^2 + z^2 = 1$ matters. The polar integration problem

$$16\int_0^{\pi/4} \int_0^1 \sqrt{1 - r^2\cos^2(\theta)} r\; drd\theta$$

has an inner r-integral of $(16/3)(1 - \sin(\theta)^3)/\cos^2(\theta)$. Integrating this over θ can be done by integrating $(1 + \sin(x)^3)\sec^2(x)$ by parts using $\tan'(x) = \sec^2(x)$ leading to the anti derivative $\cos(x) + \sec(x) + \tan(x)$. The result is $16 - 8\sqrt{2}$.

Homework

This homework is due on Tuesday, 7/30/2019.

Problem 17.1: Evaluate the triple integral

$$\int_0^3 \int_0^z \int_0^{4y} 7e^{-y^2} z \, dxdydz \ .$$

Problem 17.2: Find the volume of the solid bounded by the paraboloids $z = x^2 + y^2$ and $z = 36 - (x^2 + y^2)$ and satisfying $x \geq 0, y \geq 0$.

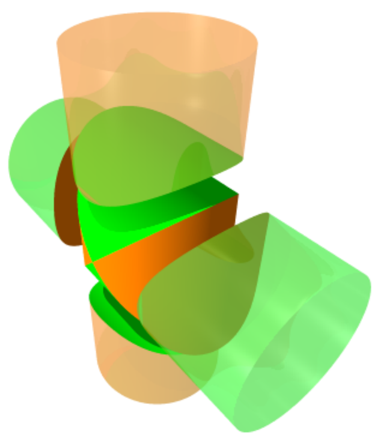
Problem 17.3: Find the **moment of inertia** $\int\int\int_E (x^2 + y^2) \, dV$ of a cone

$$E = \{x^2 + y^2 \leq z^2 \ 0 \leq z \leq 15 \ \} \ ,$$

which has the z-axis as its center of symmetry.

Problem 17.4: Integrate $f(x, y, z) = x^2 + y^2 - z$ over the tetrahedron with vertices
$(0, 0, 0), (4, 4, 0), (0, 4, 0), (0, 0, 12)$.

Problem 17.5: This is a classic problem of Archimedes: What is the volume of the body obtained by intersecting the solid cylinders $x^2+z^2 \leq 9$ and $y^2 + z^2 \leq 9$?



Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2019

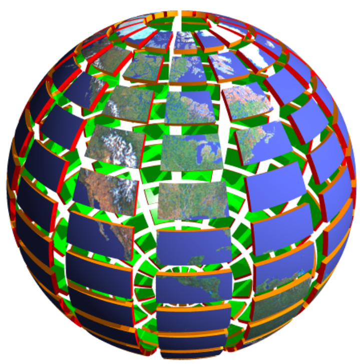
MULTIVARIABLE CALCULUS

MATH S-21A

Unit 18: Spherical integrals

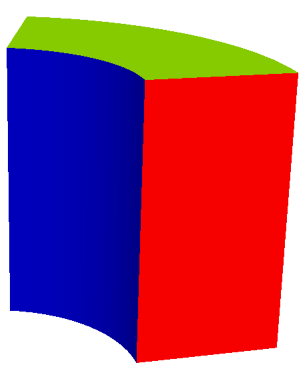
Lecture

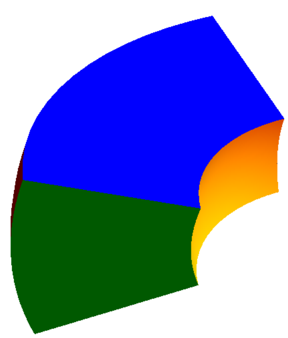
17.1. Cylindrical and spherical coordinate systems help to integrate in many situations.



Definition: **Cylindrical coordinates** are space coordinates where polar coordinates are used in the xy-plane and where the z-coordinate is untouched. The coordinate change transformation $T(r, \theta, z) = (r\cos(\theta), r\sin(\theta), z)$, produces the integration factor $\boxed{r}$. It is the same factor than the factor used in polar coordinates.

$$\iint_{T(R)} f(x, y, z)\; dxdydz = \iint_R g(r, \theta, z)\; \boxed{\mathrm{r}}\; drd\theta dz$$





Definition: Spherical coordinates use ρ, the distance to the origin as well as two **Euler angles**: $0 \leq \theta < 2\pi$ the polar angle and $0 \leq \phi \leq \pi$, the angle between the vector and the z axis. The coordinate change is

$$T : (x,y,z) = (\rho\cos(\theta)\sin(\phi), \rho\sin(\theta)\sin(\phi), \rho\cos(\phi)) .$$

The integration factor measures the volume of a **spherical wedge** which is $d\rho, \rho\sin(\phi)\ d\theta, \rho d\phi = \rho^2\sin(\phi) d\theta d\phi d\rho$.

$$\iint_{T(R)} f(x,y,z)\ dxdydz = \iint_R g(\rho,\theta,z) \boxed{\rho^2\sin(\phi)} \ d\rho d\theta d\phi$$

A sphere of radius R has the volume

$$\int_0^R \int_0^{2\pi} \int_0^{\pi} \rho^2\sin(\phi)\ d\phi d\theta d\rho .$$

The most inner integral $\int_0^\pi \rho^2\sin(\phi) d\phi = -\rho^2\cos(\phi)|_0^\pi = 2\rho^2$. The next layer is, because ϕ does not appear: $\int_0^{2\pi} 2\rho^2\ d\phi = 4\pi\rho^2$. The final integral is $\int_0^R 4\pi\rho^2\ d\rho = 4\pi R^3/3$.

Definition: The moment of inertia of a body G with respect to an axis L is defined as the triple integral $\int\int\int_G r(x,y,z)^2\ dzdydx$, where $r(x,y,z) = \rho\sin(\phi)$ is the distance from the axis L.

Examples

17.2. For a sphere of radius R we obtain with respect to the z-axis:

$$I = \int_0^R \int_0^{2\pi} \int_0^{\pi} \rho^2\sin^2(\phi)\rho^2\sin(\phi)\ d\phi d\theta d\rho$$

$$= (\int_0^\pi \sin^3(\phi)\ d\phi)(\int_0^R \rho^4\ dr)(\int_0^{2\pi}\ d\theta)$$

$$= (\int_0^\pi \sin(\phi)(1-\cos^2(\phi))\ d\phi)(\int_0^R \rho^4\ dr)(\int_0^{2pi}\ d\theta)$$

$$= (-\cos(\phi) + \cos(\phi)^3/3)|_0^\pi (L^5/5)(2\pi) = \frac{4}{3}\cdot\frac{R^5}{5}\cdot 2\pi = \frac{8\pi R^5}{15} .$$

17.3. If the sphere rotates with angular velocity ω, then $I\omega^2/2$ is the **kinetic energy** of that sphere. The moment of inertia of the earth for example is $8\cdot 10^{37} kgm^2$. The angular velocity is $\omega = 2\pi/day = 2\pi/(86400s)$. The rotational energy is $8\cdot 10^{37}kgm^2/(7464960000s^2) \sim 10^{29}J \sim 2.510^{24}kcal$.

17.4. Find the volume and the center of mass of a diamond, the intersection of the unit sphere with the cone given in cylindrical coordinates as $z = \sqrt{3}r$.
Solution: we use spherical coordinates to find the center of mass

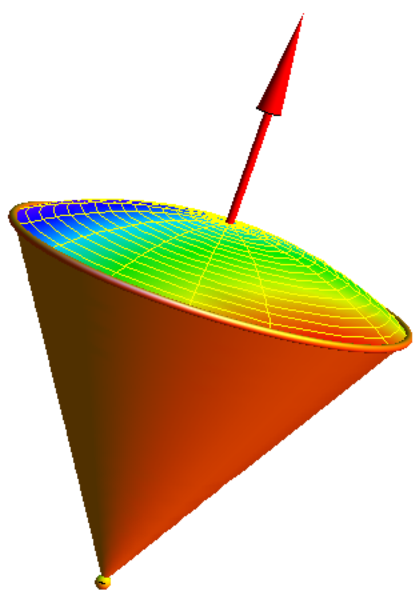
$$
\begin{aligned}
\overline{x} &= \int_0^1 \int_0^{2\pi} \int_0^{\pi/6} \rho^3 \sin^2(\phi)\cos(\theta)\ d\phi d\theta d\rho \frac{1}{V} = 0 \\
\overline{y} &= \int_0^1 \int_0^{2\pi} \int_0^{\pi/6} \rho^3 \sin^2(\phi)\sin(\theta)\ d\phi d\theta d\rho \frac{1}{V} = 0 \\
\overline{z} &= \int_0^1 \int_0^{2\pi} \int_0^{\pi/6} \rho^3 \cos(\phi)\sin(\phi)\ d\phi d\theta d\rho \frac{1}{V} = \frac{2\pi}{32V}
\end{aligned}
$$

17.5. Find $\int\int\int_R z^2\ dV$ for the solid obtained by intersecting $\{1 \leq x^2 + y^2 + z^2 \leq 4\}$ with the double cone $\{z^2 \geq x^2 + y^2\}$.
Solution: since the result for the double cone is twice the result for the single cone, we work with the diamond shaped region R in $\{z > 0\}$ and multiply the result at the end with 2. In spherical coordinates, the solid R is given by $1 \leq \rho \leq 2$ and $0 \leq \phi \leq \pi/4$. With $z = \rho\cos(\phi)$, we have

$$
\int_1^2 \int_0^{2\pi} \int_0^{\pi/4} \rho^4 \cos^2(\phi)\sin(\phi)\ d\phi d\theta d\rho
$$

$$
= (\frac{2^5}{5} - \frac{1^5}{5}) 2\pi (\frac{-\cos^3(\phi))}{3}|_0^{\pi/4} = 2\pi\frac{31}{5}(1 - 2^{-3/2})\ .
$$

The result for the double cone is $\boxed{4\pi(31/5)(1 - 1/\sqrt{2}^3)}$.



Homework

This homework is due on Tuesday, 7/30/2019.

Problem 18.1: Assume the density of a solid $E = x^2 + y^2 - z^2 < 1, -1 < z < 1$ is given by the 10's power of the distance to the z-axes: $\sigma(x,y,z) = r^{10} = (x^2+y^2)^5$. Find its mass

$$M = \int\int\int_E (x^2+y^2)^5 \; dxdydz \; .$$

Problem 18.2: Find the moment of inertia $\int\int\int_E (x^2+y^2) \; dV$ of the body E whose volume is given by the integral

$$\mathrm{Vol}(E) = \int_0^{\pi/4} \int_0^{\pi/2} \int_0^3 \rho^2 \sin(\phi) \; d\rho d\theta d\phi \; .$$

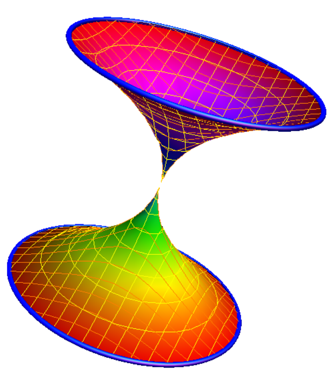
Problem 18.3: A solid is described in spherical coordinates by the inequality $\rho \leq 2\sin(\phi)$. Find its volume.

Problem 18.4: Integrate the function

$$f(x,y,z) = e^{(x^2+y^2+z^2)^{3/2}}$$

over the solid which lies between the spheres $x^2+y^2+z^2 = 1$ and $x^2+y^2+z^2 = 4$, which is in the first octant and which is above the cone $x^2+y^2 = z^2$.

Problem 18.5: Find the volume of the solid $x^2+y^2 \leq z^4, z^2 \leq 4$.



Unit 19: Vectors fields

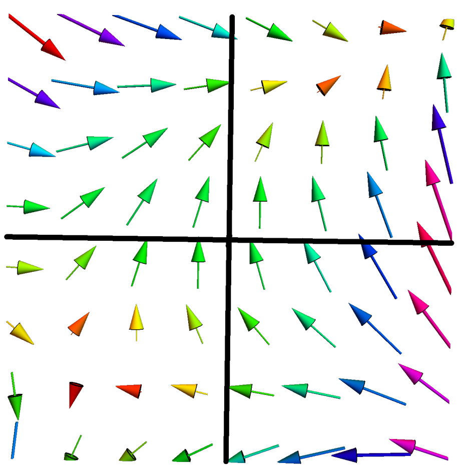
Lecture

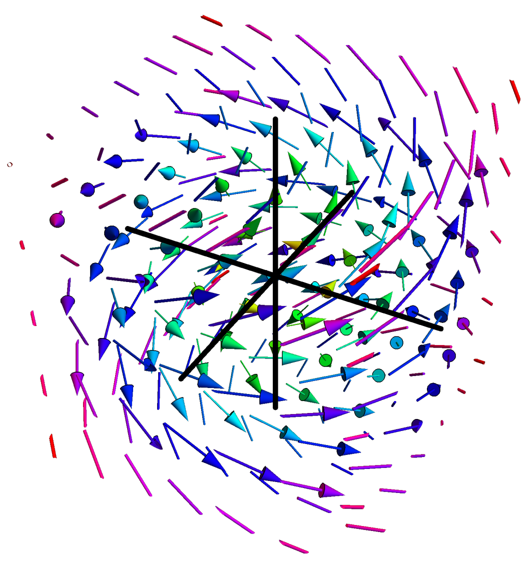
17.1. We have already seen geometries like curves, surfaces or solids. Then we have seen functions, which we consider to be scalar fields fields. If the function becomes vector valued we have **vector field**.

Definition: A **planar vector field** is a map F which assigns to a point $(x, y) \in \mathbb{R}^2$ a vector $\vec{F}(x, y) = [P(x, y), Q(x, y)]^T$. A **vector field in space** is a map, which assigns to each point $(x, y, z) \in \mathbb{R}^3$ a vector $\vec{F}(x, y, z) = [P(x, y, z), Q(x, y, z), R(x, y, z)]^T$.

17.2. Here are examples of vector fields in two and three dimensions

$$\vec{F}(x, y) = \begin{bmatrix} y - \sin(x) \\ x^3 + \cos(2y) \end{bmatrix}, \vec{F}(x, y, z) = \begin{bmatrix} -y \\ x \\ \sin(z) \end{bmatrix}.$$





Definition: If $f(x, y)$ is a function of two variables, then $\vec{F}(x, y) = \nabla f(x, y)$ is called a **gradient field**. Gradient fields in space are of the form $\vec{F}(x, y, z) = \nabla f(x, y, z)$. They are important!

17.3. When is a vector field a gradient field? $\vec{F}(x,y) = [P(x,y), Q(x,y)]^T = \nabla f(x,y)$ implies $Q_x(x,y) = P_y(x,y)$. If this does not hold at some point, $\vec{F}$ is no gradient field.

Clairaut test: If $Q_x(x,y) - P_y(x,y)$ is not zero at some point, then $\vec{F}(x,y) = [P(x,y), Q(x,y)]^T$ is not a gradient field.

17.4. We will see next week that $\text{curl}(\vec{F}) = Q_x - P_y = 0$ is also sufficient for $\vec{F}$ to be a gradient field if $\vec{F}$ is defined everywhere. How do we get f the function with $\vec{F} = \nabla f$? We will look at examples in class.

Examples

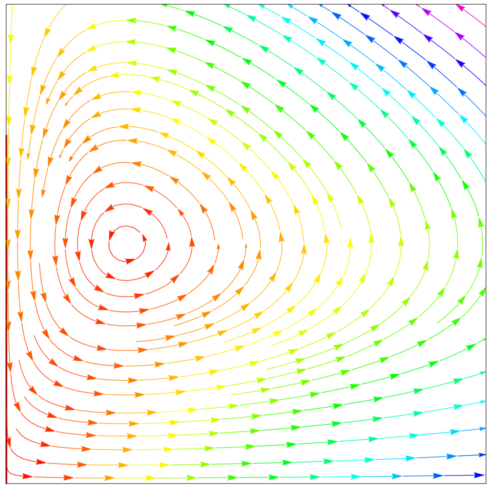
17.5. Is the vector field $\vec{F}(x,y) = [P,Q]^T = [3x^2y + y + 2, x^3 + x - 1]^T$ a gradient field? **Solution**: the Clairaut test shows $Q_x - P_y = 0$. We integrate the equation $f_x = P = 3x^2y + y + 2$ and get $f(x,y) = 2x + xy + x^3y + c(y)$. Now take the derivative of this with respect to y to get $x + x^3 + c'(y)$ and compare with $x^3 + x - 1$. We see $c'(y) = -1$ and so $c(y) = -y + c$. We see the solution $\boxed{x^3y + xy - y + 2x}$.

17.6. Is the vector field $\vec{F}(x,y) = [xy, 2xy^2]^T$ a gradient field? **Solution:** No: $Q_x - P_y = 2y^2 - x$ is not zero.
Vector fields appear naturally when studing differential equations. Here is an example in population dynamics:

17.7. If $x(t)$ is the population of a "prey species" like tuna fish and $y(t)$ is the population size of a "predator" like sharks. We have $x'(t) = ax(t) + bx(t)y(t)$ with positive a, b because both more predators and more prey species will lead to prey consumption. The rate of change of $y(t)$ is $y'(t) = -cy(t) + dxy$, where c, d are positive. This can be written using a vector field $\vec{r}' = \vec{F}(\vec{r}(t))$. We have a negative sign in the first part because predators would die out without food. The second term is explained because both more predators as well as more prey leads to a growth of predators through reproduction. A concrete example is the **Volterra-Lodka system**

$$\begin{aligned} \dot{x} &= 0.4x - 0.4xy \\ \dot{y} &= -0.1y + 0.2xy \ , \end{aligned}$$

where $\vec{F}(x,y) = [0.4x - 0.4xy$
$-0.1y + 0.2xy]^T$. Volterra explained with such systems the oscillation of fish populations in the Mediterranean sea. At any specific point $\vec{r}(x,y) = [x(t), y(t)]^T$, there is a curve $= \vec{r}(t) = [x(t), y(t)]^T$ through that point for which the tangent $\vec{r}\,'(t) = (x'(t), y'(t)$ is the vector field.

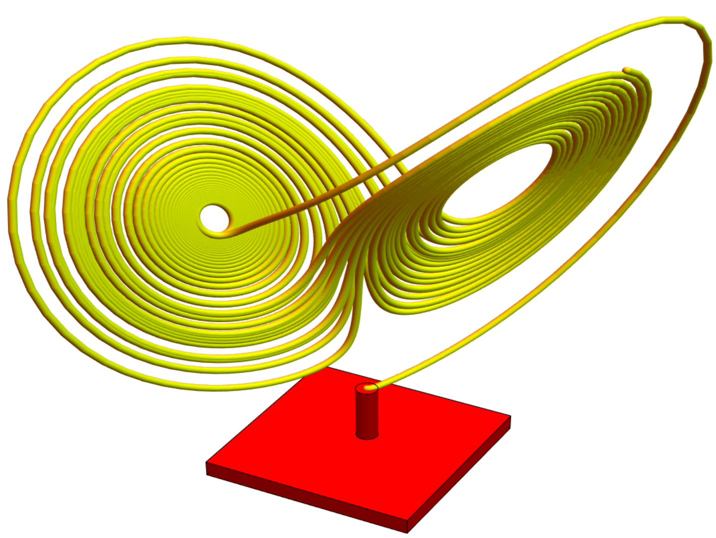


17.8. In mechanics the class of **Hamiltonian fields** plays an important role: if $H(x,y)$ is a function of two variables, then $[H_y(x,y), -H_x(x,y)]^T$ is called a **Hamiltonian vector field**. An example is the harmonic oscillator $H(x,y) = (x^2 + y^2)/2$. Its vector field $(H_y(x,y), -H_x(x,y)) = (y, -x)$. The flow lines of a Hamiltonian vector fields are located on the level curves of H.

17.9. Here is a famous example. It is the **Lorenz vector field**

$$\vec{F}(x,y,z) = \begin{bmatrix} 10y - 10x \\ -xz + 28x - y \\ xy - \frac{8}{3}z \end{bmatrix} .$$

It features a so called **strange attractor**.



Homework

This homework is due on Tuesday, 7/30/2019.

Problem 19.1:
a) Draw the gradient vector field of $f(x,y) = \sin(x^2 - y^2)$.
b) Draw the gradient vector field of $f(x,y) = (x-1)^2 + (y-2)^2$.
In both cases, draw a contour map of f and use gradients to draw the vector field $F(x,y) = \nabla f$.

Problem 19.2: The vector field

$$\vec{F}(x,y) = \begin{bmatrix} \frac{x}{(x^2+y^2)^{(3/2)}} \\ \frac{y}{(x^2+y^2)^{(3/2)}} \end{bmatrix}$$

appears in electrostatics. Find a function $f(x,y)$ such that $\vec{F} = \nabla f$.

Problem 19.3:
a) Is the vector field $\vec{F}(x,y) = \begin{bmatrix} xy \\ x^2 \end{bmatrix}$ a gradient field?
b) Is the vector field $\vec{F}(x,y) = \begin{bmatrix} \sin(x) + y \\ \cos(y) + x \end{bmatrix}$ a gradient field?
In both cases, find $f(x,y)$ satisfying $\nabla f(x,y) = \vec{F}(x,y)$ or give a reason, why it does not exist.

Problem 19.4: Find conditions such that a vector field $\vec{F}(x,y,z)$ is a gradient field. Then check it in the following cases. If there is a gradient field, find f such that $\vec{F} = \nabla f$.
a) $\vec{F}(x,y,z) = [x^{11}, y^9, z]^T$.
b) $\vec{F}(x,y,z) = [y, x, z^3]^T$.
c) $\vec{F}(x,y,z) = [10y + 10x, 10x + 10y, x]^T$.
d) $\vec{F}(x,y) = [y, z, x]^T$.

Problem 19.5: Find the potential f to

$$\vec{F}(x,y,z) = [5x^4y + z^4 + y\cos(xy), x^5 + x\cos(xy), 4xz^3]^T .$$

Unit 20: Line integral theorem

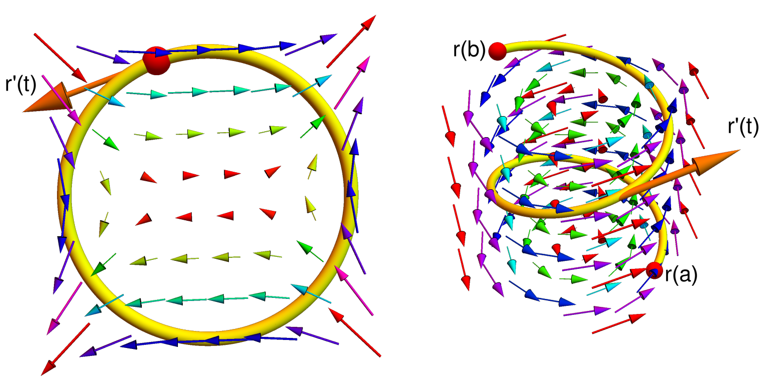
Lecture

17.1. Vector fields can be integrated along curves. If the vector field is a derivative, that is if it is a gradient field, then there is a fundamental theorem of line integrals which generalizes the fundamental theorem of calculus.

Definition: If $\vec{F}$ is a vector field in $\mathbb{R}^2$ or $\mathbb{R}^3$ and $C : t \mapsto \vec{r}(t)$ is a curve, then

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

is called the **line integral** of $\vec{F}$ along the curve C.



17.2. We use also the short-hand notation $\int_C \vec{F} \cdot \vec{dr}$. In physics, if $\vec{F}(x, y, z)$ is a **force field**, then $\vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)$ is called **power** and the line integral $\int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\ dt$ is **work**. In electrodynamics, if $\vec{F}(x, y, z)$ is an electric field, then the line integral $\int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\ dt$ is the **electric potential**.

17.3. Let $C : t \mapsto \vec{r}(t) = [\cos(t), \sin(t)]^T$ be a circle with parameter $t \in [0, 2\pi]$ and let $\vec{F}(x, y) = [-y, x]^T$. Calculate the line integral $I = \int_C \vec{F}(\vec{r}) \cdot \vec{dr}$.
Solution: We have $I = \int_0^{2\pi} \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\ dt = \int_0^{2\pi} (-\sin(t), \cos(t)) \cdot (-\sin(t), \cos(t))\ dt = \int_0^{2\pi} \sin^2(t) + \cos^2(t)\ dt = 2\pi$

17.4. Let $\vec{r}(t)$ be a curve given in polar coordinates as $\vec{r}(t) = \cos(t), \phi(t) = t$ defined on $[0, \pi]$. Let $\vec{F}$ be the vector field $\vec{F}(x, y) = (-xy, 0)$. Calculate the line integral $\int_C \vec{F} \cdot \vec{dr}$.
Solution: In Cartesian coordinates, the curve is $r(t) = (\cos^2(t), \cos(t)\sin(t))$. The velocity vector is then $\vec{r}\,'(t) = [-2\sin(t)\cos(t), -\sin^2(t) + \cos^2(t)) = (x(t), y(t)]^T$. The line integral is

$$\int_0^\pi \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\; dt = \int_0^\pi (\cos^3(t)\sin(t), 0) \cdot (-2\sin(t)\cos(t), -\sin^2(t) + \cos^2(t))\; dt$$

$$= -2\int_0^\pi \sin^2(t)\cos^4(t)\; dt = -2(t/16 + \sin(2t)/64 - \sin(4t)/64 - \sin(6t)/192)|_0^\pi = -\pi/8\,.$$

17.5. The first generalization of the fundamental theorem of calculus to higher dimensions is the **fundamental theorem of line integrals**.

Theorem: Fundamental theorem of line integrals: If $\vec{F} = \nabla f$, then

$$\int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\; dt = f(\vec{r}(b)) - f(\vec{r}(a))\,.$$

17.6. In other words, the line integral is the potential difference between the end points $\vec{r}(b)$ and $\vec{r}(a)$, if $\vec{F}$ is a gradient field.

Examples

17.7. Let $f(x, y, z)$ be the temperature distribution in a room and let $\vec{r}(t)$ the path of a fly in the room, then $f(\vec{r}(t))$ is the temperature, the fly experiences at the point $\vec{r}(t)$ at time t. The change of temperature for the fly is $\frac{d}{dt} f(\vec{r}(t))$. The line-integral of the temperature gradient ∇f along the path of the fly coincides with the temperature difference between the end point and initial point.

17.8. Here are some special cases: If $\vec{r}(t)$ is parallel to the level curve of f, then $d/dt f(\vec{r}(t)) = 0$ and $\vec{r}\,'(t)$ orthogonal to $\nabla f(\vec{r}(t))$. If $\vec{r}(t)$ is orthogonal to the level curve, then $|d/dt f(\vec{r}(t))| = |\nabla f||\vec{r}\,'(t)|$ and $\vec{r}\,'(t)$ is parallel to $\nabla f(\vec{r}(t))$.

17.9. The proof of the fundamental theorem uses the chain rule in the second equality and the fundamental theorem of calculus in the third equality of the following identities:

$$\int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)\; dt = \int_a^b \nabla f(\vec{r}(t)) \cdot \vec{r}'(t)\; dt = \int_a^b \frac{d}{dt} f(\vec{r}(t))\; dt = f(\vec{r}(b)) - f(\vec{r}(a))\,.$$

Theorem: For a gradient field, the line-integral along any closed curve is zero.

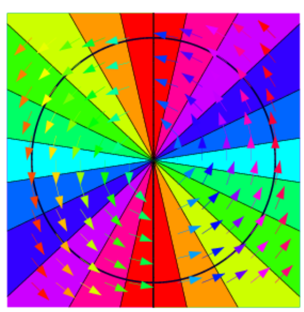
17.10. When is a vector field a gradient field? $\vec{F}(x,y) = \nabla f(x,y)$ implies $P_y(x,y) = Q_x(x,y)$. If this does not hold at some point, $\vec{F} = [P,Q]^T$ is no gradient field. This is called the **component test** or Clairaut test. We will see later that the condition $\text{curl}(\vec{F}) = Q_x - P_y = 0$ implies that the field is conservative, if the region satisfies a certain property.

17.11. Let $\vec{F}(x,y) = [2xy^2 + 3x^2, 2yx^2]^T$. Find a potential f of $\vec{F} = [P,Q]^T$.
Solution: The potential function $f(x,y)$ satisfies $f_x(x,y) = 2xy^2 + 3x^2$ and $f_y(x,y) = 2yx^2$. Integrating the second equation gives $f(x,y) = x^2y^2 + h(x)$. Partial differentiation with respect to x gives $f_x(x,y) = 2xy^2 + h'(x)$ which should be $2xy^2 + 3x^2$ so that we can take $h(x) = x^3$. The potential function is $f(x,y) = x^2y^2 + x^3$. Find g, h from $f(x,y) = \int_0^x P(x,y)\, dx + h(y)$ and $f_y(x,y) = g(x,y)$.

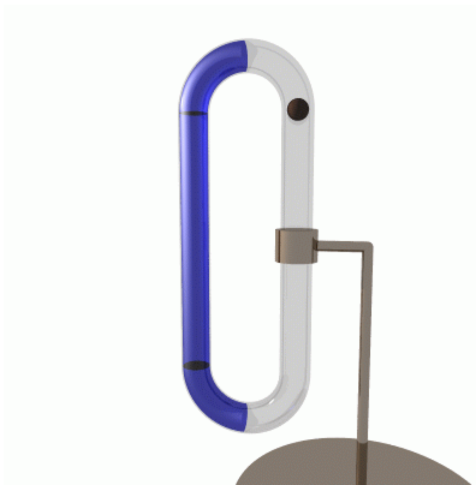
17.12. Let $\vec{F}(x,y) = [P,Q]^T = [\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}]^T$. It is a gradient field because $f(x,y) = \arctan(y/x)$ has the property that $f_x = (-y/x^2)/(1+y^2/x^2) = P, f_y = (1/x)/(1+y^2/x^2) = Q$. However, the line integral $\int_\gamma \vec{F}\, \vec{dr}$, where γ is the unit circle is

$$\int_0^{2\pi} [\frac{-\sin(t)}{\cos^2(t)+\sin^2(t)}, \frac{\cos(t)}{\cos^2(t)+\sin^2}]^T \cdot [-\sin(t), \cos(t)]^T\, dt$$

which is $\int_0^{2\pi} 1\, dt = 2\pi$. What is wrong?
Solution: note that the potential f as well as the vector-field F are not differentiable everywhere. The curl of F is zero except at $(0,0)$, where it is not defined.



17.13. A device which implements a non gradient force field is called a **perpetual motion machine**. It realizes a force field for which the energy gain is positive along some closed loop. The first law of thermodynamics forbids the existence of such a machine. It is informative to contemplate some of the ideas people have come up and to see why they don't work. Here is an example: consider a O-shaped pipe which is filled only on the right side with water. A wooden ball falls on the right hand side in the air and moves up in the water. You find plenty of other futile attempts on youtube.



Homework

This homework is due on Tuesday, 7/30/2019.

Problem 20.1: Let C be the space curve $\vec{r}(t) = [\cos(t), \sin(\sin(t)), t]^T$ for $t \in [0, \pi]$ and let $\vec{F}(x, y, z) = [y, x, 15]^T$. Find the value of the line integral $\int_C \vec{F} \cdot \vec{dr}$. You might want to use a theorem.

Problem 20.2: What is the work done by moving in the force field $\vec{F}(x, y) = [2x^3 + 1, 2y^4]^T$ along the parabola $y = x^2$ from $(-1, 1)$ to $(1, 1)$? a) compute it directly b) use the theorem.

Problem 20.3: Let $\vec{F}$ be the vector field $\vec{F}(x, y) = [-y, x]^T/2$. Compute the line integral of F along an ellipse $\vec{r}(t) = [a\cos(t), b\sin(t)]^T$ with width $2a$ and height $2b$. The result should depend on a and b.

Problem 20.4: It is hot. You unpack your portable swimming pool and place it in Harvard yard. Now, you swim along curve C given by part of the curve $x^{40} + y^{40} = 1$ in the first quadrant, oriented counter clockwise. There is a hose filling in fresh water to the tub so that there is a velocity field $\vec{F}(x, y) = [2x + 5y, 10y^4 + 5x]^T$ inside. Calculate the line integral $\int_C \vec{F} \cdot \vec{dr}$, the energy you gain from the fluid force when dislocating from $(1, 0)$ to $(0, 1)$. Be lazy.

Problem 20.5: Find a closed curve $C : \vec{r}(t)$ for which the vector field

$$\vec{F}(x, y) = [P(x, y), Q(x, y)]^T = [xy, x^2]^T$$

satisfies $\int_C \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt \neq 0$.

MULTIVARIABLE CALCULUS

MATH S-21A

Unit 21: Green's theorem

Lecture

21.1. Green's theorem is the second integral theorem in two dimensions. In this unit, we do multivariable calculus in two dimensions, where we have only **two** derivatives, **two** integral theorems: the **fundamental theorem of line integrals** as well as **Green's theorem**. You might be used to think about the two-dimensional case as a special case of the xy-plane in three space, but we insist on remaining two dimensional. [1]

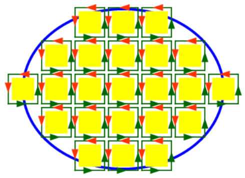
21.2. Remember that the **curl** of a vector field $\vec{F}(x,y) = [P(x,y), Q(x,y)]^T$ is the scalar field $\text{curl}(F)(x,y) = \nabla \times \vec{F} = Q_x(x,y) - P_y(x,y)$. It measures the **vorticity** of the vector field at (x,y). For example, for $\vec{F}(x,y) = [x^3+y^2, y^3+x^2y]^T$, we have $\text{curl}(F)(x,y) = 2xy - 2y$.

Theorem: Green's theorem: If $\vec{F}(x,y) = [P(x,y), Q(x,y)]^T$ is a vector field and G is a region for which the boundary C is a curve parametrized so that G is "to the left", then

$$\int_C \vec{F} \cdot \vec{dr} = \int\int_G \text{curl}(F)\ dxdy\ .$$

21.3. Take a square $G = [x, x+h] \times [y, y+h]$ with small $h > 0$. The line integral of $\vec{F} = [P,Q]^T$ along the boundary is $\int_0^h P(x+t,y)dt + \int_0^h Q(x+h,y+t)\ dt - \int_0^h P(x+t,y+h)\ dt - \int_0^h Q(x,y+t)\ dt$. It measures the "circulation" at the place (x,y). Because $Q(x+h,y) - Q(x,y) \sim Q_x(x,y)h$ and $P(x,y+h) - P(x,y) \sim P_y(x,y)h$, the line integral is $(Q_x - P_y)h^2$ is $\int_0^h \int_0^h \text{curl}(F)\ dxdy$ with an error of the order h^3 or smaller. Now take a region G with area $|G|$ and chop it into small squares of size h. We need about $|G|/h^2$ such squares. Summing up all the line integrals around the boundaries is the sum of the line integral along the boundary of G because of the cancellations in the interior. On the boundary, it is a Riemann sum of the line integral along the boundary. The sum of the curls of the squares is a Riemann sum approximation of the double integral $\int\int_G \text{curl}(F)\ dxdy$. Taking the limit $h \to 0$ gives Greens theorem.

[1] It is better to think about two dimensions as if we were flat-landers unaware about the third dimension. If we speak about "the plane", this is our universe, we are ignorant about 3 space. Edwin Abbot's Flatland is a 1884 romance plays in two dimensions.



21.4. **George Green** lived from 1793 to 1841. Unfortunately, we don't have a single picture of him. He was a physicist, a self-taught mathematician as well as a miller. His work greatly contributed to modern physics.

21.5. Here is a special case: if $\vec{F}$ is a gradient field $\vec{F} = \nabla f$, then both sides of Green's theorem are zero: $\int_C \vec{F} \cdot \vec{dr}$ is zero by the fundamental theorem for line integrals. and $\int\int_G \text{curl}(F) \cdot dA$ is zero because $\text{curl}(F) = \text{curl}(\text{grad}(f)) = 0$.

21.6. If $\vec{F}(x, y) = \nabla f$ is a gradient field then the curl is zero because if $P(x, y) = f_x(x, y), Q(x, y) = f_y(x, y)$ and $\text{curl}(F) = Q_x - P_y = f_{yx} - f_{xy} = 0$ by Clairaut. $\vec{F}(x, y) = [x + y, yx]^T$ for example is not a gradient field because $\text{curl}(F) = y - 1$.

21.7. The already established **Clairaut identity**

$\text{curl}(\text{grad}(f)) = 0$

21.8. This can also be remembered by writing $\text{curl}(\vec{F}) = \nabla \times \vec{F}$ and $\text{curl}(\nabla f) = \nabla \times \nabla f$. Use now that cross product of two identical vectors is 0. Working with ∇ as a vector is called **nabla calculus** which can serve as a mnemonic.

21.9. It had been a consequence of the fundamental theorem of line integrals that:

If $\vec{F}$ is a gradient field then $\text{curl}(F) = 0$ everywhere.

21.10. Is the converse true? Here is the answer:

Definition: A region R is called **simply connected** if every closed loop in R can be pulled together continuously within R to a point inside R.

21.11. $R = \{x^2 + y^2 \leq 1\}$ is simply connected, $O = \{3 \leq x^2 + y^2 \leq 4\}$ is not.

If $\text{curl}(\vec{F}) = 0$ in a simply connected region G, then $\vec{F}$ is a gradient field.

Proof. Given a closed curve C in G enclosing a region R. Green's theorem assures that $\int_C \vec{F}\ \vec{dr} = 0$. So $\vec{F}$ has the closed loop property in G. This is equivalent to the fact that line integrals are path independent. In that case $\vec{F}$ is therefore a gradient field: one can get $f(x, y)$ by taking the line integral from an arbitrary point O to (x, y). In the homework, you look at an example of a not simply connected region where the $\text{curl}(\vec{F}) = 0$ does not imply that $\vec{F}$ is a gradient field.

Examples

21.12. Problem: Find the line integral of $\vec{F}(x, y) = [x^2 - y^2, 2xy]^T = [P, Q]^T$ along the boundary of the rectangle $[0, 2]\times[0, 1]$. **Solution:** $\text{curl}(\vec{F}) = Q_x - P_y = 2y + 2y = 4y$ so that $\int_C \vec{F}\ \vec{dr} = \int_0^2 \int_0^1 4y\ dydx = 2y^2|_0^1 x|_0^2 = 4$.

21.13. Problem: Find the area of the region enclosed by

$$\vec{r}(t) = [\frac{\sin(\pi t)^2}{t}, t^2 - 1]^T$$

for $-1 \leq t \leq 1$. To do so, use Greens theorem with the vector field $\vec{F} = [0, x]^T$.

21.14. Green's theorem allows to express the coordinates of the **centroid** = center of mass

$$(\int\int_G x\ dA/A, \int\int_G y\ dA/A)$$

using line integrals. With the vector field $\vec{F} = [0, x^2]^T$ we have

$$\int\int_G x\ dA = \int_C \vec{F}\ \vec{dr}\ .$$

21.15. An important application of Green is **area computation**: Take a vector field like $\vec{F}(x, y) = [P, Q]^T = [0, x]^T$ which has constant vorticity $\text{curl}(\vec{F})(x, y) = 1$. For $\vec{F}(x, y) = [0, x]^T$, the right hand side in Green's theorem is the **area** $\text{Area}(G) = \int_C \vec{F}(\vec{r}(t))\ \vec{r}'(t)\ dt$.

21.16. Let G be the region below the graph of a function $f(x)$ on $[a, b]$. The line integral around the boundary of G is 0 from $(a, 0)$ to $(b, 0)$ because $\vec{F}(x, y) = [0, 0]^T$ there. The line integral is also zero from $(b, 0)$ to $(b, f(b))$ and $(a, f(a))$ to $(a, 0)$ because $N = 0$. The line integral along the curve $(t, f(t))$ is $-\int_a^b [-y(t), 0]^T \cdot [1, f'(t)]^T\ dt = \int_a^b f(t)\ dt$. Green's theorem confirms that this is the area of the region below the graph.

21.17. An engineering application is the **planimeter**, a mechanical device for measuring areas. We demonstrate it in class. Historically it had been used in medicine to measure the size of the cross-sections of tumors, in biology to measure the area of leaves or wing sizes of insects, in agriculture to measure the area of forests, in engineering to measure the size of profiles.

21.18. There is a vector field $\vec{F}$ associated to the device which is obtained by placing a unit vector perpendicular to the arm). One can prove that $\vec{F}$ has vorticity 1. The planimeter calculates the line integral of $\vec{F}$ along a given curve. Green's theorem assures this is the area.

Homework

This homework is due on Tuesday, 8/6/2019.

Problem 21.1: Given $f(x,y) = x^5 + xy^4$, compute the line integral of $\vec{F}(x,y) = [25y + 6y^2, 12xy + 10y^{44}]^T + \nabla f$ along the boundary of the **Monster region** given in the picture. There are four boundary curves, oriented as shown in the picture: a large ellipse of area 16, two circles of area 1 and 2 as well as a small ellipse (the mouth) of area 3.

Problem 21.2: Given $f(x,y) = x^5 + xy^4$, compute the line integral of $\vec{F}(x,y) = [15y+6y^2, 12xy+y^4]^T + \nabla f$ along the boundary of the **Monster region** given in the picture. There are four boundary curves, oriented as shown in the picture: a large ellipse of area 16, two circles of area 1 and 2 as well as a small ellipse (the mouth) of area 3.

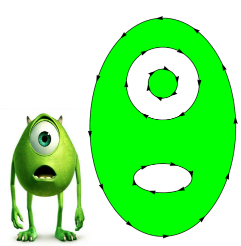
Problem 21.3: Find the area of the region bounded by the **hypocycloid** $\vec{r}(t) = [4\cos^3(t), 4\sin^3(t)]^T, 0 \leq t \leq 2\pi$.

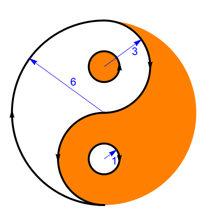
Problem 21.4: Let G be the region $x^6 + y^6 \leq 1$. Mathematica allows us to get the area as Area[ImplicitRegion[$\mathrm{x}^6 + \mathrm{y}^6 <= 1, \{\mathrm{x}, \mathrm{y}\}$]] and tells, it is $A = 3.8552$. Compute the line integral of $\vec{F}(x,y) = [x^{800} + \sin(x) - 55y, y^{12} + \cos(y) + 3x]^T$ along the boundary in terms of A (leave A in the answer).

Problem 21.5: Let C be the boundary curve of the white Yang part of the Ying-Yang symbol in the disc of radius 6. You can see in the image that the curve C has three parts, and that the orientation of each part is given. Find the line integral of the vector field

$$\vec{F}(x,y) = [-y + \sin(e^x), x]^T$$

around C.





Unit 22: Curl and Flux

Lecture

22.1. The **curl** in two dimensions was the scalar field curl$(F) = Q_x - P_y$. By Green's theorem, the curl evaluated at (x, y) is $\lim_{r\to 0} \int_{C_r} \vec{F}\, dr/(\pi r^2)$, where C_r is a small circle of radius r oriented counter clockwise an centered at (x, y). Green's theorem explains so what the curl is: it measures how the field "curls"... As rotations in two dimensions are determined by a single angle, in three dimensions, three parameters are needed. It is a vector whose direction tells the axes of rotation and the length tells the amount of rotation. The curl now becomes a vector:

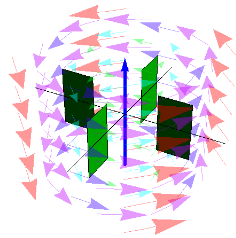
Definition: The **curl** of $\vec{F} = [P, Q, R]^T$ is the vector field

$$\text{curl}(P, Q, R) = [R_y - Q_z, P_z - R_x, Q_x - P_y]^T \ .$$

22.2. In **Nabla calculus**, this is written as curl$(\vec{F}) = \nabla \times \vec{F}$. Note that the third component $Q_x - P_y$ of the curl is for fixed z just the curl of the two-dimensional vector field $\vec{F} = [P, Q]^T$. While the curl in two dimensions is a scalar field, it is a vector field in 3 dimensions. In n dimensions, it would have $n(n-1)/2$ components, as this is the number of 2-dimensional coordinate planes. The curl measures the "vorticity" of the field and each component measures this in one of the two dimensional coordinate planes.

Definition: If a field has zero curl everywhere, the field is called **irrotational**.

22.3. The curl is frequently visualized using a "paddle wheel". If the rotation axes points into direction v, the rotation speed is $\vec{F} \cdot \vec{v}$. The direction in which the wheel turns fastest, is the direction of curl$(\vec{F})$. The angular velocity of the wheel is the length of the curl.



22.4. In two dimensions, we had two derivatives, the gradient and curl. In three dimensions, there are three fundamental derivatives: the **gradient**, the **curl** and the **divergence**.

Definition: The **divergence** of $\vec{F} = [P, Q, R]^T$ is the scalar field $\text{div}([P, Q, R]^T) = \nabla \cdot \vec{F} = P_x + Q_y + R_z$.

22.5. The divergence can also be defined in two dimensions, but it is there not as fundamental as it is not an "exterior derivatives". We want in d dimensions to have d fundamental derivatives and d fundamental integrals and d fundamental theorems. Distinguishing dimensions helps to organize the integral theorems. While Green looks like Stokes, we urge you to look at it as a different theorem taking place in "flatland". It is a small matter but it is much clearer to have in every dimension d a separate calculus. This prevents mixing up the theorems and makes things easier.

Definition: In two dimensions, the **divergence** of $\vec{F} = [P, Q]^T$ is defined as $\text{div}(P, Q) = \nabla \cdot \vec{F} = P_x + Q_y$.

22.6. In two dimensions, the divergence can be written as the curl of a -90 degrees rotated field $\vec{G} = [Q, -P]^T$ because $\text{div}(\vec{G}) = Q_x - P_y = \text{curl}(\vec{F})$. The divergence measures the "expansion" of a field. If a field has zero divergence everywhere, the field is called **incompressible**.

22.7. With the "vector" $\nabla = [\partial_x, \partial_y, \partial_z]^T$, we can write $\text{curl}(\vec{F}) = \nabla \times \vec{F}$ and $\text{div}(\vec{F}) = \nabla \cdot \vec{F}$. Formulating formulas using the "Nabla vector" and using rules from geometry is called **Nabla calculus**. This works both in 2 and 3 dimensions even so the ∇ vector is not an actual vector but an operator. The following combination of divergence and gradient often appears in physics:

Definition:

$$\Delta f = \text{div}(\text{grad}(f)) = f_{xx} + f_{yy} + f_{zz} \; .$$

is called the **Laplacian** of f. One can write $\Delta f = \nabla^2 f$.

22.8. Mathematicians know Δ it as a 'form Laplacian". Here are some identities:

$\text{div}(\text{curl}(\vec{F})) = 0.$
$\text{curlgrad}(\vec{F}) = \vec{0}$
$\text{curl}(\text{curl}(\vec{F})) = \text{grad}(\text{div}(\vec{F}) - \Delta(\vec{F})).$

Examples

22.9. Question: Is there a vector field $\vec{G}$ such that $\vec{F} = [x + y, z, y^2]^T = \text{curl}(\vec{G})$? **Answer:** No, because $\text{div}(\vec{F}) = 1$ is incompatible with $\text{div}(\text{curl}(\vec{G})) = 0$.

22.10. Show that in simply connected region, every irrotational and incompressible field can be written as a vector field $\vec{F} = \text{grad}(f)$ with $\Delta f = 0$. Proof. Since $\vec{F}$ is irrotational, there exists a function f satisfying $F = \text{grad}(f)$. As the region is simply connected, we can deform any path between two points without changing the result. Now, $\text{div}(F) = 0$ implies $\text{divgrad}(f) = \Delta f = 0$.

22.11. If we rotate the vector field $\vec{F} = [P, Q]^T$ by 90 degrees $= \pi/2$, we get a new vector field $\vec{G} = [-Q, P]^T$. The integral $\int_C F \ \cdot ds$ becomes a **flux** $\int_\gamma G \cdot dn$ of G through the boundary of R, where dn is a normal vector with length $|r'|dt$. With $\text{div}(\vec{F}) = (P_x + Q_y)$, we see that

$$\text{curl}(\vec{F}) = \text{div}(\vec{G}) \ .$$

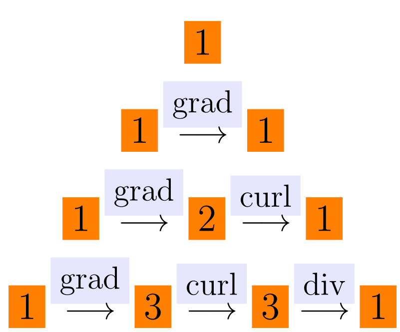
Green's theorem now becomes

$$\int \int_R \text{div}(\vec{G}) \ dxdy = \int_C \vec{G} \cdot \vec{dn} \ ,$$

where $dn(x, y)$ is a normal vector at (x, y) orthogonal to the velocity vector $\vec{r}\,'(x, y)$ at (x, y). This new theorem has a generalization to three dimensions, where it is called Gauss theorem or divergence theorem. Don't treat this however as a different theorem in two dimensions. It is just Green's theorem in disguise.

In two dimensions, the divergence at a point (x, y) is the average flux of the field through a small circle of radius r around the point in the limit when the radius of the circle goes to zero.

We have now all the derivatives we need. In dimension d, there are d fundamental derivatives.



Homework

This homework is due on Tuesday, 8/6/2019.

Problem 22.1: Construct your own nonzero vector field $\vec{F}(x, y) = [P(x, y), Q(x, y)]^T$ in each of the following cases:
a) $\vec{F}$ is irrotational but not incompressible.
b) $\vec{F}$ is incompressible but not irrotational.
c) $\vec{F}$ is irrotational and incompressible.
d) $\vec{F}$ is not irrotational and not incompressible.

Problem 22.2: The vector field $\vec{F}(x, y, z) = [x, y, -2z]^T$ satisfies $\text{div}(\vec{F}) = 0$. Can you find a vector field $\vec{G}(x, y, z)$ such that $\text{curl}(\vec{G}) = \vec{F}$? Such a field $\vec{G}$ is called a **vector potential**.
Hint. Write $\vec{F}$ as a sum $[x, 0, -z]^T + [0, y, -z]^T$ and find vector potentials for each of the parts using a vector field you have seen on the blackboard in class.

Problem 22.3: Evaluate the flux integral $\int\int_S [0, 0, yz]^T \cdot \vec{dS}$, where S is the surface with parametric equation $x = uv, y = u + v, z = u - v$ on $R : u^2 + v^2 \leq 4$ and $u > 0$.

Problem 22.4: Evaluate the flux integral $\int\int_S \text{curl}(F) \cdot \vec{dS}$ for

$$\vec{F}(x, y, z) = [3xy, 3yz, 3zx]^T .$$

where S is the part of the paraboloid $z = 4 - x^2 - y^2$ that lies above the square $[0, 2] \times [0, 2]$ and has an upward orientation.

Problem 22.5: a) What is the relation between the flux of the vector field $\vec{F} = \nabla g/|\nabla g|$ through the surface $S : \{g = 1\}$ with $g(x, y, z) = x^6 + y^4 + 2z^8$ and the surface area of S?
b) Find the flux of the vector field $\vec{G} = \nabla g \times [0, 0, 2]^T$ through the surface S.

[1]

Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2019

[1] Both a and b) do not need any computation. You can answer each question with one sentence. In part a) compare $\vec{F} \cdot \vec{dS}$ with dS in that case.

Unit 23: Stokes Theorem

LECTURE

22.1. We work with a surface S parametrized as $\vec{r}(u,v) = [x(u,v), y(u,v), z(u,v)]^T$ over a domain R in the uv-plane. Remember that the **flux integral** of $\vec{F}$ through S is defined as the double integral

$$\iint_R \vec{F}(\vec{r}(u,v)) \cdot (\vec{r}_u \times \vec{r}_v) \, dudv \, .$$

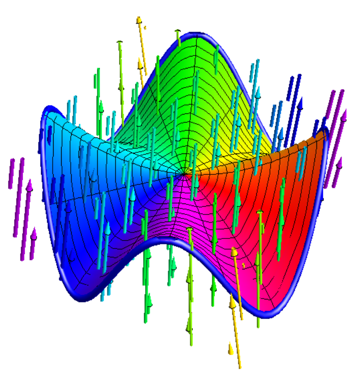
The following theorem is the second fundamental theorem of calculus in three dimensions:

Definition: The **boundary** of a surface S consists of all points P where even arbitrary small circle $S_r(P) \cap S$ around the point is not closed.

22.2. The boundary is a collection of curves oriented so that the surface is to the "left" if the normal vector to the surface is pointing "up". In other words, the velocity vector v, a vector w pointing towards the surface and the normal vector n to the surface form a right handed coordinate system.

Theorem: **Stokes theorem**: Let S be a surface bounded by a curve C and $\vec{F}$ be a vector field. Then

$$\iint_S \text{curl}(\vec{F}) \cdot d\vec{S} = \int_C \vec{F} \cdot d\vec{r} \, .$$



Proof. Stokes theorem is proven in the same way than Green's theorem. Chop up S into a union of small triangles. As before, the sum of the fluxes through all these triangles adds up to the flux through the surface and the sum of the line integrals along the boundaries adds up to the line integral of the boundary of S. Stokes theorem for a small triangle can be reduced to Green's theorem because with a coordinate system such that the triangle is in the xy plane, the flux of the field is the double integral of $\text{curl}\vec{F}\ d\vec{S} = \text{curl}F(\vec{r}) \cdot \vec{n} dudv = (Q_x - P_y)\cos(\theta) dudv$, where θ is the angle between the normal vector and $\vec{F} = [P, Q, R]^T$. On the other hand, since the power $\vec{F}(\vec{r}) \cdot \vec{r}'(t)dt = (P(\vec{r})\cos(\theta)x'(t) + Q(\vec{r})\cos(\theta)y'(t))dt$ also has everything multiplied by $\cos(\theta)$, the result for each space triangle follows from Green. Stokes theorem now follows by making the triangulation finer and finer. On both sides we have a Riemann sum approximation to the integrals. □

Examples

22.3. Let $\vec{F}(x, y, z) = [-y, x, 0]^T$ and let S be the upper semi hemisphere, then $\text{curl}(\vec{F})(x, y, z) = [0, 0, 2]^T$. The surface is parameterized by

$$\vec{r}(u, v) = [\cos(u)\sin(v), \sin(u)\sin(v), \cos(v)]^T$$

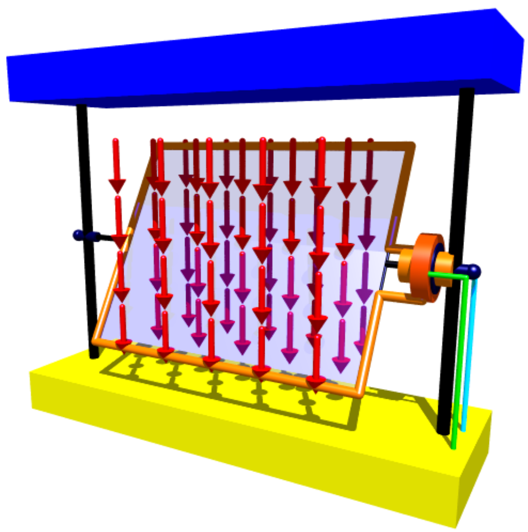
on $R = [0, 2\pi] \times [0, \pi/2]$ and $\vec{r}_u \times \vec{r}_v = \sin(v)\vec{r}(u, v)$ so that $\text{curl}(\vec{F})(x, y, z) \cdot \vec{r}_u \times \vec{r}_v = \cos(v)\sin(v)2$. The integral $\int_0^{2\pi} \int_0^{\pi/2} \sin(2v)\ dvdu = 2\pi$.
The boundary C of S is parameterized by $\vec{r}(t) = [\cos(t), \sin(t), 0]^T$ so that $\vec{dr} = \vec{r}\ '(t)dt = [-\sin(t), \cos(t), 0]^T\ dt$ and $\vec{F}(\vec{r}(t))\ \vec{r}\ '(t)dt = \sin(t)^2 + \cos^2(t) = 1$. The line integral $\int_C \vec{F} \cdot \vec{dr}$ along the boundary is 2π.

22.4. If S is a surface in the xy-plane and $\vec{F} = [P, Q, 0]^T$ has zero z component, then $\text{curl}(\vec{F}) = [0, 0, Q_x - P_y]^T$ and $\text{curl}(\vec{F}) \cdot d\vec{S} = Q_x - P_y\ dxdy$. We see that for a surface which is flat, Stokes theorem is a consequence of Green's theorem. If we put the coordinate axis so that the surface is in the xy-plane, then the vector field F induces a vector field on the surface such that its $2D$ curl is the normal component of $\text{curl}(F)$. The reason is that the third component $Q_x - P_y$ of $\text{curl}(\vec{F})[R_y - Q_z, P_z - R_x, Q_x - P_y]^T$ is the two dimensional curl: $\vec{F}(\vec{r}(u, v)) \cdot [0, 0, 1]^T = Q_x - P_y$. If C is the boundary of the surface, then $\int\int_S \vec{F}(\vec{r}(u, v)) \cdot [0, 0, 1]^T\ dudv = \int_C \vec{F}(\vec{r}(t))\ \vec{r}\ '(t)dt$.

22.5. Calculate the flux of the curl of $\vec{F}(x, y, z) = [-y, x, 0]^T$ through the surface parameterized by $\vec{r}(u, v) = [\cos(u)\cos(v), \sin(u)\cos(v), \cos^2(v) + \cos(v)\sin^2(u + \pi/2)]^T$. Because the surface has the same boundary as the upper half sphere, the integral is again 2π as in the above example.

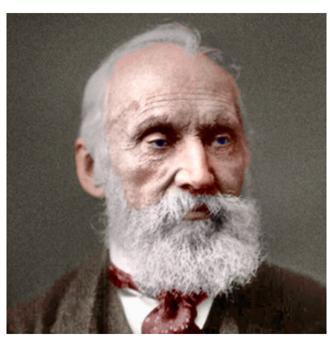
22.6. For every surface bounded by a curve C, the flux of $\text{curl}(\vec{F})$ through the surface is the same. Proof. The flux of the curl of a vector field through a surface S depends only on the boundary of S. Compare this with the earlier statement that for every curve between two points A, B the line integral of $\text{grad}(f)$ along C is the same. The line integral of the gradient of a function of a curve C depends only on the end points of C.

22.7. Electric and magnetic fields are linked by the **Maxwell equation** $\text{curl}(\vec{E}) = -\frac{1}{c}\dot{B}$. These are examples of partial differential equations. If a closed wire C bounds a surface S then $\iint_S B \cdot dS$ is the flux of the magnetic field through S. Its change can be related with a voltage using Stokes theorem: $d/dt \iint_S B \cdot dS = \iint_S \dot{B} \cdot dS = \iint_S -c\ \text{curl}(\vec{E}) \cdot d\vec{S} = -c \int_C \vec{E}\ d\vec{r} = U$, where U is the voltage. If we change the flux of the magnetic field through the wire, then this induces a voltage. The flux can be changed by changing the amount of the magnetic field but also by changing the direction. If we turn around a magnet around the wire or the wire inside the magnet, we get an electric voltage. This happens in a power-generator, like the alternator in a car. Stokes theorem explains why we can generate electricity from motion.



22.8. The history of Stokes theorem is a bit hazy. [1]. A version of Stokes theorem appeared to be known by **André Ampère** in 1825. **William Thomson** (Lord Kelvin) mentioned the theorem to Stokes in 1850). **George Gabriel Stokes** (1819-1903) who found parts of the identity earlier 1840 formulated it in a prize exam from 1854 (the proof is one of the exam problems). The first pushed proof is by Hermann Hankel in 1861.

George Gabriel Stokes



William Thomson (Kelvin)

André Marie Ampere

[1]See V. Katz, the History of Stokes theorem, Mathematics Magazine 52, 1979, p 146-156

Homework

This homework is due on Tuesday, 8/6/2019.

Problem 23.1: Assume S is the surface $x^8 + y^4 + z^6 = 100$ and $\vec{F} = [e^{e^{xyz}}, x^2yz, x - y - \sin(zx)]^T$. Explain why $\int\int_S \mathrm{curl}(\vec{F}) \cdot d\vec{S} = 0$.

Problem 23.2: Find $\int_C \vec{F} \cdot d\vec{r}$, where $\vec{F}(x,y,z) = [12x^2y, 4x^3, 12xy]^T$ and C is the curve of intersection of the hyperbolic paraboloid $z = y^2 - x^2$ and the cylinder $x^2 + y^2 = 1$, oriented counterclockwise as viewed from above.

Problem 23.3: Evaluate the flux integral $\int\int_S \mathrm{curl}(\vec{F}) \cdot d\vec{S}$, where

$$\vec{F}(x,y,z) = [xe^{y^2}z^3 + 2xyze^{x^2+z}, x + z^2e^{x^2+z}, ye^{x^2+z} + ze^x]^T$$

and where S is the part of the ellipsoid $x^2 + y^2/4 + (z+1)^2 = 2$, $z > 0$ oriented so that the normal vector points upwards.

Problem 23.4: Find the line integral $\int_C \vec{F}\ d\vec{r}$, where C is the circle of radius 5 in the xz-plane oriented counter clockwise when looking from the point $(0,1,0)$ onto the plane and where $\vec{F}$ is the vector field

$$\vec{F}(x,y,z) = [9x^2z + x^5, \cos(e^y), -9xz^2 + \sin(\sin(z))]^T .$$

Use a convenient surface S which has C as a boundary.

Problem 23.5: Find the flux integral $\int\int_S \mathrm{curl}(\vec{F}) \cdot d\vec{S}$, where $\vec{F}(x,y,z) =$

$$[2\cos(\pi y)e^{2x} + z^2, x^2\cos(z\pi/2) - \pi\sin(\pi y)e^{2x}, 2xz]^T$$

and S is the surface parametrized by

$$\vec{r}(s,t) = [(1 - s^{1/3})\cos(t) - 4s^2, (1 - s^{1/3})\sin(t), 5s]^T$$

with $0 \leq t \leq 2\pi, 0 \leq s \leq 1$ and oriented so that the normal vectors point to the outside of the thorn.

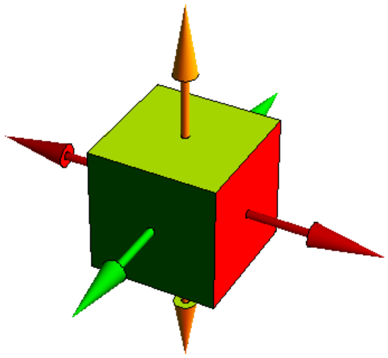
Unit 24: Divergence Theorem

Lecture

22.1. There are three integral theorems in three dimensions. We have already seen the fundamental theorem of line integrals and Stokes theorem. The **divergence theorem** completes the list of integral theorems in three dimensions:

> **Theorem: Divergence Theorem**. If E be a solid with boundary surface S oriented so that the normal vector points outside and if $\vec{F}$ be a vector field, then
> $$\iiint_E \operatorname{div}(\vec{F})\, dV = \iint_S \vec{F} \cdot dS .$$

22.2. To prove this, take a small box $[x, x + dx] \times [y, y + dy] \times [z, z + dz]$. The flux of $\vec{F} = [P, Q, R]^T$ through the faces perpendicular to the x-axes is $[\vec{F}(x + dx, y, z) \cdot [1, 0, 0]^T + \vec{F}(x, y, z) \cdot [-1, 0, 0]^T]dydz = P(x + dx, y, z) - P(x, y, z) \sim P_x\ dxdydz$. Similarly, the flux through the y-boundaries is $P_y\ dydxdz$ and the flux through the two z-boundaries is $P_z\ dzdxdy$. The total flux through the faces of the cube is $(P_x + P_y + P_z)\ dxdydz = \operatorname{div}(\vec{F})\ dxdydz$. A general solid can be approximated as a union of small cubes. The sum of the fluxes through all the cubes consists now of the flux through all faces without neighboring faces. Important is that fluxes through adjacent sides cancel. The sum of all the fluxes of the cubes is the flux through the boundary of the union. The sum of all the $\operatorname{div}(\vec{F})\ dxdydz$ is a Riemann sum approximation for the integral $\iiint_G \operatorname{div}(\vec{F})\ dxdydz$. In the limit, when dx, dy, dz all go to zero, we obtain the divergence theorem.



22.3. The theorem explains what divergence means. If we integrate the divergence over a small cube, it is equal the flux of the field through the boundary of the cube. If this is positive, then more field exits the cube than entering the cube. There is field "generated" inside. The divergence measures the "expansion" of the field.

Examples

22.4. Let $\vec{F}(x,y,z) = [x,y,z]^T$ and let S be the unit sphere. The divergence of $\vec{F}$ is the constant function $\text{div}(\vec{F}) = 3$ and $\iiint_G \text{div}(\vec{F})\ dV = 3 \cdot 4\pi/3 = 4\pi$. The flux through the boundary is $\iint_S \vec{r} \cdot \vec{r}_u \times \vec{r}_v\ dudv = \iint_S |\vec{r}(u,v)|^2 \sin(v)\ dudv = \int_0^{\pi} \int_0^{2\pi} \sin(v)\ dudv = 4\pi$ also. We see that the divergence theorem allows us to compute the area of the sphere from the volume of the enclosed ball or compute the volume from the surface area.

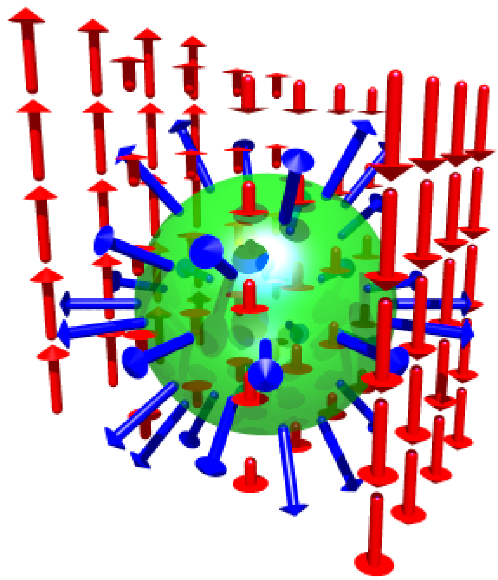
22.5. What is the flux of the vector field $\vec{F}(x,y,z) = [2x, 3z^2+y, \sin(x)]^T$ through the solid $G = [0,3]\times[0,3]\times[0,3]\setminus([0,3]\times[1,2]\times[1,2]\cup[1,2]\times[0,3]\times[1,2]\cup[0,3]\times[0,3]\times[1,2])$ which is a cube where three perpendicular cubic holes have been removed? **Solution:** Use the divergence theorem: $\text{div}(\vec{F}) = 2$ and so $\iiint_G \text{div}(\vec{F})\ dV = 2\iiint_G\ dV = 2\text{Vol}(G) = 2(27-7) = 40$. Note that the flux integral here would be over a complicated surface over dozens of rectangular planar regions.

22.6. Find the flux of $\text{curl}(F)$ through a torus if $\vec{F} = [yz^2, z+\sin(x)+y, \cos(x)]^T$ and the torus has the parametrization

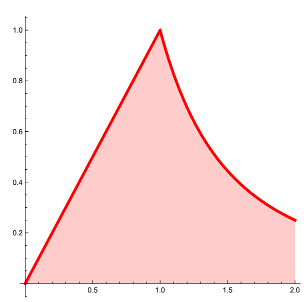
$$\vec{r}(\theta,\phi) = [(2+\cos(\phi))\cos(\theta), (2+\cos(\phi))\sin(\theta), \sin(\phi)]^T .$$

Solution: The answer is 0 because the divergence of $\text{curl}(F)$ is zero. By the divergence theorem, the flux is zero.

22.7. Similarly as Green's theorem allowed to calculate the area of a region by passing along the boundary, the volume of a region can be computed as a flux integral: Take for example the vector field $\vec{F}(x, y, z) = [x, 0, 0]^T$ which has divergence 1. The flux of this vector field through the boundary of a solid region is equal to the volume of the solid: $\iint_{\delta G} [x, 0, 0]^T \cdot d\vec{S} = \text{Vol}(G)$.



22.8. How heavy are we, at distance r from the center of the earth?
Solution: The law of gravity can be formulated as $\text{div}(\vec{F}) = 4\pi\rho$, where ρ is the mass density. We assume that the earth is a ball of radius R. By rotational symmetry, the gravitational force is normal to the surface: $\vec{F}(\vec{x}) = \vec{F}(r)\vec{x}/||\vec{x}||$. The flux of $\vec{F}$ through a ball of radius r is $\iint_{S_r} \vec{F}(x) \cdot d\vec{S} = 4\pi r^2 \vec{F}(r)$. By the **divergence theorem**, this is $4\pi M_r = 4\pi \iiint_{B_r} \rho(x)\ dV$, where M_r is the mass of the material inside S_r. We have $(4\pi)^2 \rho r^3/3 = 4\pi r^2 \vec{F}(r)$ for $r < R$ and $(4\pi)^2 \rho R^3/3 = 4\pi r^2 \vec{F}(r)$ for $r \geq R$. Inside the earth, the gravitational force $\vec{F}(r) = 4\pi\rho r/3$. Outside the earth, it satisfies $\vec{F}(r) = M/r^2$ with $M = 4\pi R^3 \rho/3$.



22.9. To the end we make an overview over the integral theorems and give an other typical example in each case.

The fundamental theorem for line integrals, Green's theorem, Stokes theorem and divergence theorem are all part of **one single theorem** $\int_A dF = \int_{\delta A} F$, where dF is a **exterior derivative** of F and where δA is the **boundary** of A. It generalizes the fundamental theorem of calculus.

Fundamental theorem of line integrals: If C is a curve with boundary $\{A, B\}$ and f is a function, then

$$\int_C \nabla f \cdot \vec{dr} = f(B) - f(A)$$

Remarks.

1) For closed curves, the line integral $\int_C \nabla f \cdot \vec{dr}$ is zero.
2) Gradient fields are **path independent**: if $\vec{F} = \nabla f$, then the line integral between two points P and Q does not depend on the path connecting the two points.
3) The theorem holds in any dimension. In one dimension, it reduces to the **fundamental theorem of calculus** $\int_a^b f'(x)\, dx = f(b) - f(a)$
4) The theorem justifies the name **conservative** for gradient vector fields.
5) The term "potential" was coined by George Green who lived from 1783-1841.

22.10. Example. Let $f(x,y,z) = x^2 + y^4 + z$. Find the line integral of the vector field $\vec{F}(x,y,z) = \nabla f(x,y,z)$ along the path $\vec{r}(t) = [\cos(5t), \sin(2t), t^2]^T$ from $t = 0$ to $t = 2\pi$.

Solution. $\vec{r}(0) = [1,0,0]^T$ and $\vec{r}(2\pi) = [1,0,4\pi^2]^T$ and $f(\vec{r}(0)) = 1$ and $f(\vec{r}(2\pi)) = 1 + 4\pi^2$. The fundamental theorem of line integral gives $\int_C \nabla f\ \vec{dr} = f(r(2\pi)) - f(r(0)) = 4\pi^2$.

Green's theorem. If R is a region with boundary C and $\vec{F}$ is a vector field, then

$$\iint_R \text{curl}(\vec{F})\ dxdy = \int_C \vec{F} \cdot \vec{dr}\ .$$

22.11. Remarks.
1) Greens theorem allows to switch from double integrals to one dimensional integrals.
2) The curve is oriented in such a way that the region is to the left.
3) The boundary of the curve can consist of piecewise smooth pieces.
4) If $C : t \mapsto \vec{r}(t) = [x(t), y(t)]^T$, the line integral is $\int_a^b [P(x(t),y(t)), Q(x(t),y(t))]^T \cdot [x'(t), y'(t)]^T\ dt$.
5) Green's theorem was found by George Green (1793-1841) in 1827 and by Mikhail Ostrogradski (1801-1862).
6) If $\text{curl}(\vec{F}) = 0$ in a simply connected region, then the line integral along a closed curve is zero. If two curves connect two points then the line integral along those curves agrees.

7) Taking $\vec{F}(x,y) = [-y,0]^T$ or $\vec{F}(x,y) = [0,x]^T$ gives **area formulas**.

22.12. Example. Find the line integral of the vector field $\vec{F}(x,y) = [x^4 + \sin(x) + y, x + y^3]^T$ along the path $\vec{r}(t) = [\cos(t), 5\sin(t) + \log(1+\sin(t))]^T$, where t runs from $t=0$ to $t=\pi$.

Solution. $\mathrm{curl}(\vec{F}) = 0$ implies that the line integral depends only on the end points $(0,1),(0,-1)$ of the path. Take the simpler path $\vec{r}(t) = [-t,0]^T, -1 \leq t \leq 1$, which has velocity $\vec{r}\,'(t) = [-1,0]^T$. The line integral is $\int_{-1}^{1} [t^4 - \sin(t), -t]^T \cdot [-1,0]^T \; dt = -t^5/5|_{-1}^{1} = -2/5$.

Remark We could also find a potential $f(x,y) = x^5/5 - \cos(x) + xy + y^5/4$. It has the property that $\mathrm{grad}(f) = F$. Again, we get $f(0,-1) - f(0,1) = -1/5 - 1/5 = -2/5$.

Stokes theorem. If S is a surface with boundary C and $\vec{F}$ is a vector field, then

$$\int\int_S \mathrm{curl}(\vec{F}) \cdot dS = \int_C \vec{F} \cdot \vec{dr} \; .$$

22.13. Remarks.
1) Stokes theorem allows to derive Greens theorem: if $\vec{F}$ is z-independent and the surface S is contained in the xy-plane, one obtains the result of Green.
2) The orientation of C is such that if you walk along C and have your head in the direction of the normal vector $\vec{r}_u \times \vec{r}_v$, then the surface to your left.
3) Stokes theorem was found by André Ampère (1775-1836) in 1825 and rediscovered by George Stokes (1819-1903).
4) The flux of the curl of a vector field does not depend on the surface S, only on the boundary of S.
5) The flux of the curl through a closed surface like the sphere is zero: the boundary of such a surface is empty.

22.14. Example. Compute the line integral of $\vec{F}(x,y,z) = [x^3 + xy, y, z]^T$ along the polygonal path C connecting the points $(0,0,0),(2,0,0),(2,1,0),(0,1,0)$ in that order.
Solution. The path C bounds a surface $S : \vec{r}(u,v) = [u,v,0]^T$ parameterized by $R = [0,2] \times [0,1]$. By Stokes theorem, the line integral is equal to the flux of $\mathrm{curl}(\vec{F})(x,y,z) = [0,0,-x]^T$ through S. The normal vector of S is $\vec{r}_u \times \vec{r}_v = [1,0,0]^T \times [0,1,0]^T = [0,0,1]^T$ so that $\int\int_S \mathrm{curl}(\vec{F}) \; \vec{dS} = \int_0^2 \int_0^1 [0,0,-u]^T \cdot [0,0,1]^T \; dudv = \int_0^2 \int_0^1 -u \; dudv = -2$.

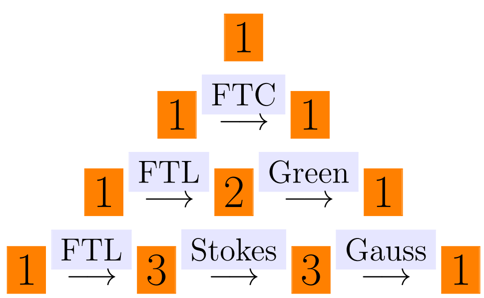
Divergence theorem: If S is the boundary of a region E in space and $\vec{F}$ is a vector field, then

$$\iiint_B \operatorname{div}(\vec{F})\, dV = \iint_S \vec{F} \cdot d\vec{S} \;.$$

22.15. Remarks.
1) The divergence theorem is also called **Gauss theorem**.
2) It is useful to determine the flux of vector fields through surfaces.
3) It can be used to compute volume.
3) It was discovered in 1764 by Joseph Louis Lagrange (1736-1813), later it was rediscovered by Carl Friedrich Gauss (1777-1855) and by George Green.
4) For divergence free vector fields $\vec{F}$, the flux through a closed surface is zero. Such fields $\vec{F}$ are also called **incompressible** or **source free**.

22.16. Example. Compute the flux of the vector field $\vec{F}(x, y, z) = [-x, y, z^2]^T$ through the boundary S of the rectangular box $[0, 3] \times [-1, 2] \times [1, 2]$.
Solution. By Gauss theorem, the flux is equal to the triple integral of $\operatorname{div}(F) = 2z$ over the box: $\int_0^3 \int_{-1}^2 \int_1^2 2z \; dxdydz = (3 - 0)(2 - (-1))(4 - 1) = 27$.

22.17. How do these theorems fit together? In n-dimensions, there are n theorems. We have here seen the situation in dimension $n = 2$ and $n = 3$, but one could continue. The fundamental theorem of line integrals generalizes directly to higher dimensions. Also the divergence theorem generalizes directly since an n-dimensional integral in n dimensions. The generalization of curl and flux would need more explanation as for $n = 4$ already, the curl of a vector field is a 6-dimensional object. It is a $n(n - 1)/2$ dimensional object in general.



22.18. In one dimension, there is one derivative $f(x) \to f'(x)$ from scalar to scalar functions. It corresponds to the entry $1 - 1$ in the Pascal triangle. The next entry $1 - 2 - 1$ corresponds to differentiation in two dimensions, where we have the gradient $f \to \nabla f$ mapping a scalar function to a vector field with 2 components as well as the curl, $F \to \operatorname{curl}(F)$ which corresponds to the transition $2 - 1$. The situation in three dimensions is captured by the entry $1 - 3 - 3 - 1$ in the Pascal triangle. The first derivative $1 - 3$ is the gradient. The second derivative $3 - 3$ is the curl and the third derivative $3 - 1$ is the divergence. In $n = 4$ dimensions, we would have to look at $1 - 4 - 6 - 4 - 1$. The first derivative $1 - 4$ is still the gradient. Then we have a first curl, which maps a vector field with 4 components into an object with 6 components. Then there is a second curl, which maps an object with 6 components back to a vector field, we would have to look at $1-4-6-4-1$.

22.19. When setting up calculus in dimension n, one talks about **differential forms** instead of scalar fields or vector fields. Functions are 0 forms or n-forms. Vector fields can be described by 1 or $n-1$ forms. The general formalism defines a derivative d called **exterior derivative** on differential forms. It maps k forms to $k+1$ forms. There is also an integration of k-forms on k-dimensional objects. The **boundary operation** δ which maps a k-dimensional object into a $k-1$ dimensional object. This boundary operation is dual to differentiation. They both satisfy the same relation $dd(F)=0$ and $\delta\delta G=0$. Differentiation and integration are linked by the general Stokes theorem:

$$\int_{\delta G} F = \int_G dF$$

22.20. One can see this as a single theorem, the **fundamental theorem of multivariable calculus**. The theorem is simpler in quantum calculus, where geometric objects and fields are on the same footing. There are various ways how one can generalize this. One way is to write it as $<\delta G, F>=<G, dF>$ which in linear algebra would be written as $[A^T v, w]^T = [v, Aw]^T$, where A^T is the transpose of a matrix A $[v,w]^T$ is the dot product. Since traditional calculus we deal with "smooth" functions and fields, we have to pay a prize and consider in turn "singular" objects like points or curves and surfaces. These are idealized objects which have zero diameter, radius or thickness.

22.21. So, it is all about **geometries** and **fields**. Geometries are curves, or surfaces or solids. Fields are scalar functions or vector fields. Geometries G can be "differentiated" by taking the boundary δG. Fields F can be differentiated by applying differential operators dF like grad, curl or div. And then there is integration which pairs up geometries G and fields F. The fundamental theorem $\int_{\delta G} F = \int_G dF$ tells that taking the boundary on the object corresponds to taking the derivative of the field.

22.22. Nature likes simplicity and elegance [1] and therefore found a quantum mathematics to be more fundamental. But the symmetry in which **geometry and fields become indistinguishable** manifests only in the very small.

[1]Leibniz: 1646-1716

Homework

This homework is due on Tuesday, 8/6/2019.

Problem 24.1: Compute using the divergence theorem the flux of the vector field $\vec{F}(x,y,z) = [9y, 2xy, 4yz + 187xy]^T$ through the unit cube $[0,1] \times [0,1] \times [0,1]$.

Problem 24.2: Find the flux of the vector field $\vec{F}(x,y,z) = [xy, yz, zx]^T$ through the cylinder $x^2 + y^2 \leq 1$, $0 < z \leq 2$ without the bottom disk $z = 0$. (Still use the divergence theorem by closing it and computing the flux through the bottom.)

Problem 24.3: Use the divergence theorem to calculate the flux of $\vec{F}(x,y,z) = [x^3, y^3, z^3]^T$ through the sphere $S : x^2 + y^2 + z^2 = 1$, where the sphere is oriented so that the normal vector points outwards.

Problem 24.4: Assume the vector field

$$\vec{F}(x,y,z) = [5x^3 + 12xy^2, y^3 + e^y \sin(z), 5z^3 + e^y \cos(z)]^T$$

is the magnetic field of the **sun** whose surface is a sphere of radius 3 oriented with the outward orientation. Compute the magnetic flux $\iint_S \vec{F} \cdot d\vec{S}$.

Problem 24.5: Find $\int \int_S \vec{F} \cdot d\vec{S}$, where $\vec{F}(x,y,z) = [-37x, 22y, 25z]^T$ and S is the boundary of the solid built with the 18 cubes shown in the picture.