Multivariable Calculus Math S21A, 2025

Oliver Knill **Harvard University**

Welcome to the 2025 multi-variable course Math S 21a of the Summer school! We cover 6 chapters in 6 weeks. Each chapter has 4 sections. While you may well consult with a text book or online resources, it is important to focus on what we do in lectures, homework and practice exams.

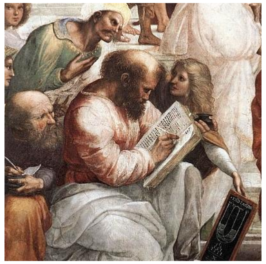
Chapter 1. Geometry and Space

Section 1.1: Space and Distance

René Descartes

The arena of multi-variable calculus uses two or higher dimensions. We mostly work in the two-dimensional plane or in three-dimensional **space**. **Points** P in space are described by triples of **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 $d(P, Q) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$. While motivated by the **Pythagoras theorem**, we will derive and prove this theorem. We then explore some geometric objects like **lines**, **cylinders**, **planes** or **spheres**.

Section 1.2: Vectors and Dot product



Pythagoras

Two points P, Q define a **vector** $\vec{PQ} = -\vec{QP}$. Vectors can describe many things, like **forces**, **velocity**, **acceleration**, **color** or **data**. The **components** of $\vec{PQ}$ connecting $P = (a, b, c)$ with $Q = (x, y, z)$ are the entries of the **vector** $[x-a, y-b, z-c]$. The **zero vector** is $\vec{0} = [0, 0, 0]$. The **standard basis vectors** are $\vec{i} = [1, 0, 0], \vec{j} = [0, 1, 0], \vec{k} = [0, 0, 1]$. **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**, **Pythagoras** $|\vec{v} + \vec{w}|^2 = |\vec{v}|^2 + |\vec{w}|^2$ follows.

Section 1.3: Cross and Triple Product

Rowan Hamilton

The **cross product** $\vec{v} \times \vec{w}$ of $\vec{v} = [a, b, c]$ and $\vec{w} = [p, q, r]$ is defined as $[br - cq, cp - ar, aq - bp]$. It is a vector perpendicular to $\vec{v}$ and $\vec{w}$. In two dimensions, the cross product is defined as the scalar $[a, b] \times [p, q] = aq - bp$. 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 $|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}| \sin(\alpha)$, then define it to be 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 $\vec{u}, \vec{v}, \vec{w}$ are in a common plane.

Section 1.4: Lines and Planes

Arthur Cayley

Because $[a, b, c] = \vec{n} = \vec{u} \times \vec{v}$ 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 the equation $ax + by + cz = d$. We often know the normal vector $\vec{n} = [a, b, c]$ 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 a surface. We introduce **lines** by the parameterization $\vec{r}(t) = \vec{OP} + t\vec{v}$, where P is a point on the line and $\vec{v} = [a, b, c]$ 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**.

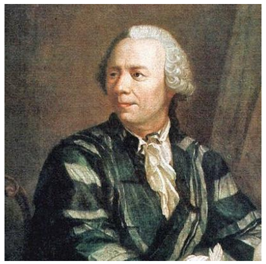
Chapter 2. Curves and Surfaces

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 of 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**.

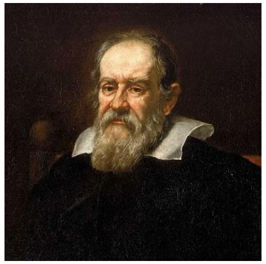
Section 2.2: Parametric Surfaces



Leonhard Euler

Surfaces are described implicitly or parametric. Examples of implicit descriptions $g(x, y, z) = 0$ are $x^2 + y^2 + z^2 - 1 = 0$. Examples of **parametrizations** $\vec{r}(u, v) = [x(u, v), y(u, v), z(u, v)]$ are the **sphere** $\vec{r}(\theta, \phi) = [\rho \cos(\theta) \sin(\phi), \rho \sin(\theta) \sin(\phi), \rho \cos(\phi)]$, where ρ is fixed and ϕ, θ are the **Euler angles**. Computers can draw also complicated surfaces. Parametrization of surfaces is important in **geodesy**. We use longitude and latitude in maps for example. In **computer generated imaging**, a parameterization $\vec{r}(u, v)$ is called the **"uv-map"**. Parametrizations can use **cylindrical coordinates** (r, θ, z), where $r \geq 0$ is the distance to the z-axes and $0 \leq \theta < 2\pi$ is the **polar angle**. **spherical coordinates** (ρ, θ, ϕ) use ρ, the distance to $(0, 0, 0)$ and θ, ϕ, the **Euler angles**.

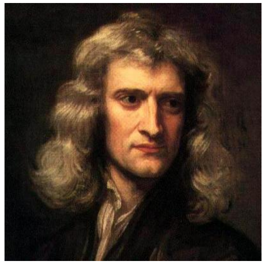
Section 2.3: Parametric Curves



Johannes Kepler

The parametrization $\vec{r}(t) = [x(t), y(t), z(t)]$, 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**. Curves also appear as **grid curves** on parametrized surfaces. 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 gives the **acceleration** vector $\vec{r}\,''(t)$. The **speed** $|r'(t)|$ is a scalar. Integration allows to get us from $\vec{r}''(t)$ and $\vec{r}\,'(0)$ and $\vec{r}(0)$ the location $\vec{r}(t)$ at a later time. 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 the limiting length of polygons and is the 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 at $\vec{r}(t)$. Acceleration and curvature involve second derivatives. Curvature is a quantity which does not depend on parameterizations. One "feels" acceleration but "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 in multi-variable calculus is even 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 still not 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 functions that have continuous second partial derivatives. We look then at some **partial differential equations** (PDE's). 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

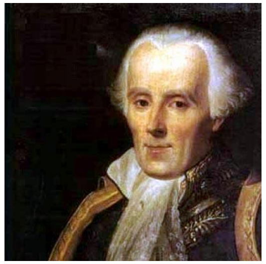
Linearization is an important concept because many physical laws are linearization of more complicated laws. Linearization is also useful for estimating quantities. After a review of linearization of functions of one variables, we introduce the **linearization** of a function $f(x, y)$ of two variables at a point (p, q) 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 $a = f_x, b = f_y$. In three dimensions, the definition is analog. It allows to compute the **tangent plane** $ax + by + cz = d$. The key is the **gradient** $f' = \nabla f = [f_x, f_y, f_z]$. We will mention that 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 possible second derivatives of f at the point $\vec{p}$.

Section 3.3: Implicit Differentiation

Gottfried Wilhelm Leibniz

The **chain rule** $d/dt f(g(t)) = f'(g(t))g'(t)$ in one dimension generalizes to higher dimensions. It becomes $d/dt f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\ \ '(t)$, where $\nabla f = [f_x, f_y, f_z]$ 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. A nice application of the chain rule is **implicit differentiation**: if $f(x, y, z) = 0$ defines a surface which near a point looks like $z = g(x, y)$ for an unknown g, then $f_x + f_z z' = 0$ gives us $g_x = -f_x/f_z$ 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 $\nabla g = [a, b, c]$ is constant. 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**. This is important in machine learning and AI.

Chapter 4. Extrema and Double integrals

Section 4.1: Maxima and Minima


Pierre de Fermat

To **maximize** $f(x, y)$, first locate **critical points**, points where the gradient vanishes: $\nabla f(x, y) = [0, 0]$. 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 $f_{xx} < 0$, we have a local maximum, if $D > 0$ and $f_{xx} > 0$, 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 $f(x_0, y_0) \geq f(x, y)$ 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 **bottom-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 $\iint_G f dA = \iint_G f(x, y)\ dxdy$ is the **signed volume** under the graph of the function of two variables. Double integrals define **area** if $f(x, y) = 1$. By **changing the order of integration** we sometimes can integrate integrals which appear to be 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** to the positive x-axes. **Polar region** like $0 \leq r \leq g(\theta)$ can trace flower-like shapes in the plane. An other application of double integrals is **surface area**. We derive the formula $\iint_R |r_u \times r_v|\ dudv$ and compute it for examples like graphs, surfaces of revolution and the sphere. Similar as for arc length, the surface area can not always be computed in closed form. 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

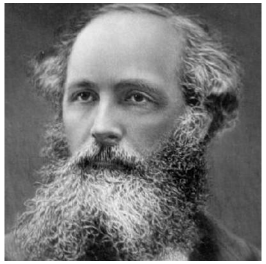
Triple integrals can measure volume, moments of inertia, or the center of mass of a solid. It is first introduced for **cuboids**, then to more general regions like solids that are sandwiched between the graphs of two functions $g(x, y)$ and $h(x, y)$. Applications are computations of **mass** $\iiint_E \delta(x, y, z)\ dxdydz$, **moment of inertia** $\iiint_E (x^2 + y^2 + z^2)\ dxdydz$, **center of mass**, $\iiint [x, y, z]\ dV$ the **expectation** $\mathrm{E}[X] = \iiint X(x, y, z)\ dV / \iiint\ dV$ 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). These 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** $\int_C \vec{F}(\vec{r}(t))\ \vec{r}\,'(t)\ dt$ 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 connectedness**: any closed curve can be contracted 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 $[2\cos(t), 2\sin(t)]$ can not be pulled together to a point. In two dimensions, the curl of a field $\text{curl}([P,Q]) = Q_x - P_y$ 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 C of a region G with a double integral of the curl in G: $\iint_G \text{curl}(\vec{F})(x,y)\ dxdy = \int_C \vec{F}(\vec{r}(t))\ \vec{r}\,'(t)\ dt$. The theorem is useful to **compute areas**: take a field $\vec{F} = [0,x]$ which has constant curl 1. It also allows to compute complicated line integrals. Greens theorem implies that if $\text{curl}(F) = Q_x - P_y = 0$ everywhere in the plane, then the field has the **closed loop property** and is therefore conservative. The **curl** of a field $\vec{F} = [P,Q,R]$ in three dimensions is a new vector field which can be computed as $\nabla \times \vec{F}$. The 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 $\iint_S \vec{F} \cdot (\vec{r}_u \times \vec{r}_v)\ dudv$. The angle between $\vec{F}$ and the normal vector $\vec{n} = \vec{r}_u \times \vec{r}_v$ determines the sign of $d\vec{S} = \vec{F} \cdot \vec{n}\ dudv$. Many concepts enter this definition: the parametrization of surfaces, vector fields, the dot and the 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: $\int_C \vec{F}\ dr = \iint_S \text{curl}(F) d\vec{S}$. 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. This creative project helps to illustrate connections of mathematics with art.

Section 6.4 Divergence Theorem

Carl Friedrich Gauss

The integral of the **divergence** of a vector field $\vec{F} = [P,Q,R]$ over a **solid** E is the flux of $\vec{F}$ through the boundary S. This **divergence theorem** equates the "local expansion rate" integrated over the solid $\iiint_E \text{div}(\vec{F})\ dV$ of a vector field $\vec{F}$ with the flux $\iint_S \vec{F} \cdot d\vec{S}$ 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.