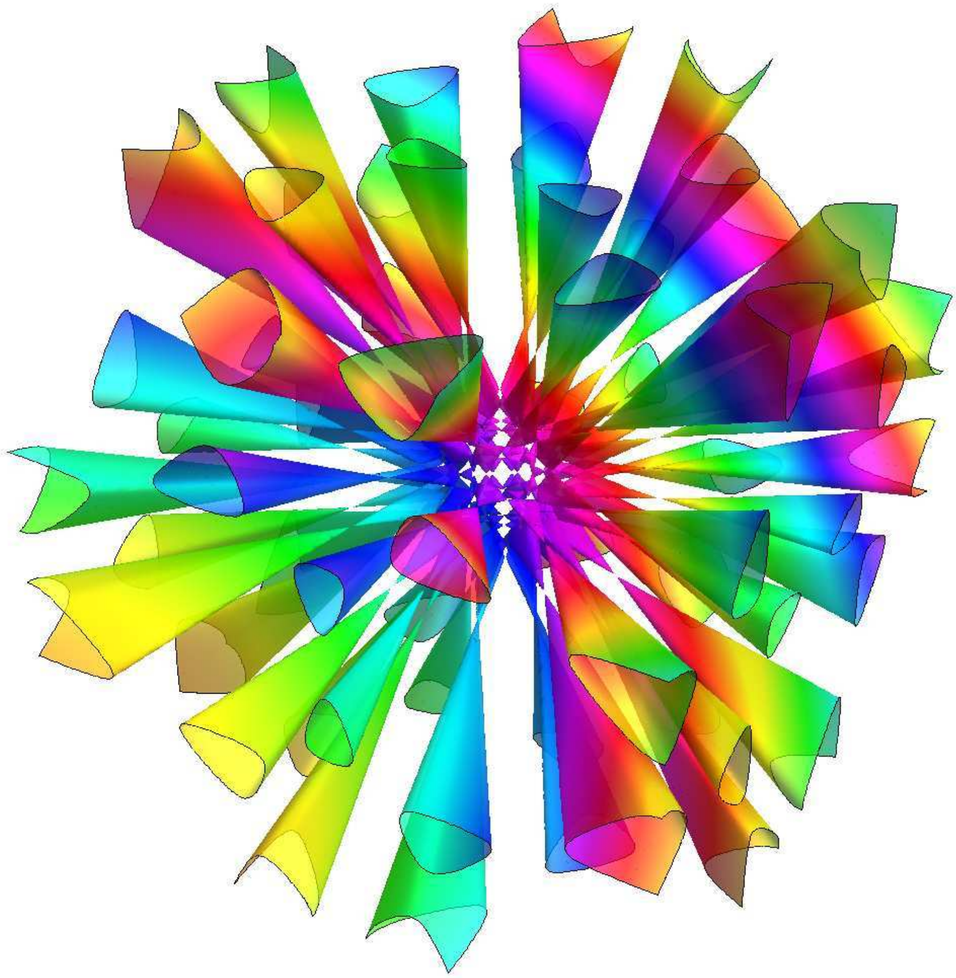
Multivariable Calculus

Oliver Knill

Math 21a, Fall 2014



These notes contain condensed "two pages per lecture" notes with essential information.

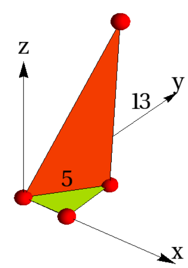
1: Geometry and Distance

A point in the **plane** has two **coordinates** $P = (x, y)$ like $P = (2, -3)$. A point in space is determined by three coordinates $P = (x, y, z)$ like $P = (1, 2, 4)$. The plane is divided into 4 **quadrants**, and space is divided into 8 **octants**. The point $P = (1, 2, 4)$ is in the **first octant**. These regions by intersect at the **origin** $O = (0, 0)$ or $O = (0, 0, 0)$ and are separated by **coordinate axes** $\{y = 0\}$ and $\{x = 0\}$ or **coordinate planes** $\{x = 0\}, \{y = 0\}, \{z = 0\}$.

1 Describe the location of the points $P = (1, 2, 3), Q = (0, 0, -5), R = (1, 2, -3)$ in words. **Possible Answer:** $P = (1, 2, 3)$ is in the positive octant of space, where all coordinates are positive. The point $R = (0, 0, -5)$ is on the negative z axis. The point $S = (1, 2, -3)$ is below the xy-plane. When projected onto the xy-plane it is in the first quadrant.

2 **Problem.** Find the midpoint M of $P = (1, 2, 5)$ and $Q = (-3, 4, 9)$. **Answer.** The midpoint is obtained by taking the average of each coordinate $M = (P + Q)/2 = (-1, 3, 7)$.

The **Euclidean distance** between two points $P = (x, y, z)$ and $Q = (a, b, c)$ in space is defined as $d(P, Q) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$.



The Euclidean distance is **motivated** by the **Pythagorean theorem**. [1]

3 Find the distance $d(P, Q)$ between the points $P = (1, 2, 5)$ 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 $d(P, M) + d(M, Q) = d(P, Q)$.

A **circle** of radius r centered at $P = (a, b)$ is the collection of points in the plane which have distance r from P.

A **sphere** of radius ρ centered at $P = (a, b, c)$ is the collection of points in space which have distance ρ from P. The equation of a sphere is $(x-a)^2+(y-b)^2+(z-c)^2 = \rho^2$.

4 Is the point $(3, 4, 5)$ outside or inside the sphere $(x-2)^2 + (y-6)^2 + (z-2)^2 = 16$? **Answer**: The distance of the point to the center of the sphere is $\sqrt{1 + 4 + 9}$ which is smaller than 4 the radius of the sphere. The point is inside.

[1] It appears in an appendix to "Geometry" of "Discours de la méthode" from 1637, **René Descartes** (1596-1650). More about Descartes in Aczel's book "Descartes Secret Notebook".

The **completion of the square** of an equation $x^2 + bx + c = 0$ is the idea to add $(b/2)^2 - c$ on both sides to get $(x + b/2)^2 = (b/2)^2 - c$. Solving for x gives the solution $x = -b/2 \pm \sqrt{(b/2)^2 - c}$.

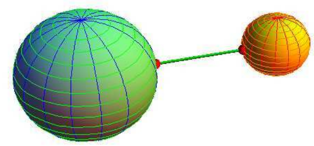
[2]

5 Solve $2x^2 - 10x + 12 = 0$. **Answer.** The equation is equivalent to $x^2 + 5x = -6$. Adding $(5/2)^2$ on both sides gives $(x + 5/2)^2 = 1/4$ so that $x = 2$ or $x = 3$.

6 Find the center of the sphere $x^2 + 5x + y^2 - 2y + z^2 = -1$. **Answer:** Complete the square to get $(x+5/2)^2 - 25/4 + (y-1)^2 - 1 + z^2 = -1$ or $(x-5/2)^2 + (y-1)^2 + z^2 = (5/2)^2$. We see a sphere **center** $(5/2, 1, 0)$ and **radius** $5/2$.

Al-Khwarizai

Rene Descartes



Distance between spheres

7 Find the set of points $P = (x, y, z)$ in space which satisfy $x^2 + y^2 = 9$. **Answer:** This is a cylinder of radius 3 around the z-axes parallel to the y axis.

8 What is $x^2 + y^2 = z^2$. **Answer**: this is the set of points for which the distance to the z axes is equal to the distance to the xy-plane. It must be a cone.

9 Find the distances of $P = (12, 5, 3)$ to the xy-plane. **Answer:** 3. Find the distance of $P = (12, 5, 0)$ to z axes. **Answer:** 13.

10 Describe $x^2 + 2x + y^2 - 16y + z^2 + 10z + 54 = 0$. **Answer:** Complete the square to get a sphere $(x + 2)^2 + (y - 8)^2 + (z + 5)^2 = 36$ with center $(-2, 8, -5)$ and radius 6.

11 Describe the set $xz = x$. **Answer:** We either must have $x = 0$ or $z = 1$. The set is a union of two coordinate planes.

12 Find an equation for the set of points which have the same distance to $(1, 1, 1)$ and $(0, 0, 0)$. **Answer:** $(x - 1)^2 + (y - 1)^2 + (z - 1)^2 = x^2 + y^2 + z^2$ gives $-2x + 1 - 2y + 1 - 2z + 1 = 0$ or $2x + 2y + 2z = 3$. This is the equation of a plane.

13 Find the distance between the spheres $x^2 + (y-12)^2 + z^2 = 1$ and $(x-3)^2 + y^2 + (z-4)^2 = 9$. **Answer:**The distance between the centers is $\sqrt{3^2 + 4^2 + 12^2} = 13$. The distance between the spheres is $13 - 3 - 1 = 9$.

[2]Due to **Al-Khwarizmi** (780-850) in *"Compendium on Calculation by Completion and Reduction"* The book "The mathematics of Egypt, Mesopotamia,China, India and Islam, a Source book, Ed Victor Katz, contains translations of some of this work.

2: Vectors and Dot product

Two points $P = (a, b, c)$ and $Q = (x, y, z)$ in space define a **vector** $\vec{PQ} = \vec{v} = \langle x-a, y-b-z-c \rangle$ pointing from P to Q. The real numbers v_1, v_2, v_3 in $\vec{v} = \langle v_1, v_2, v_3 \rangle$ are called the **components** of $\vec{v}$.

Similar definitions hold in two dimensions, where vectors have two components. Vectors can be drawn **everywhere** in space but two vectors with the same components are considered **equal**. [1]

The **addition** of two vectors is $\vec{u} + \vec{v} = \langle u_1, u_2, u_3 \rangle + \langle v_1, v_2, v_3 \rangle = \langle u_1+v_1, u_2+v_2, u_3+v_3 \rangle$. The **scalar multiple** $\lambda \vec{u} = \lambda \langle u_1, u_2, u_3 \rangle = \langle \lambda u_1, \lambda u_2, \lambda u_3 \rangle$. The difference $\vec{u} - \vec{v}$ can best be seen as the addition of $\vec{u}$ and $(-1) \cdot \vec{v}$.

The addition and scalar multiplication of vectors satisfy the laws you know from **arithmetic**. **commutativity** $\vec{u}+\vec{v} = \vec{v}+\vec{u}$, **associativity** $\vec{u}+(\vec{v}+\vec{w}) = (\vec{u}+\vec{v})+\vec{w}$ and $r*(s*\vec{v}) = (r*s)*\vec{v}$ as well as **distributivity** $(r+s)\vec{v} = \vec{v}(r+s)$ and $r(\vec{v}+\vec{w}) = r\vec{v}+r\vec{w}$, where $*$ is scalar multiplication.

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**.

1 $|\langle 3, 4 \rangle| = 5$ and $|\langle 3, 4, 12 \rangle| = 13$. Examples of unit vectors are $|\vec{i}| = |\vec{j}| = \vec{k}| = 1$ and $\langle 3/5, 4/5 \rangle$ and $\langle 3/13, 4/13, 12/13 \rangle$. The only vector of length 0 is the zero vector $|\vec{0}| = 0$.

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

The dot product determines distance and distance determines the dot product.

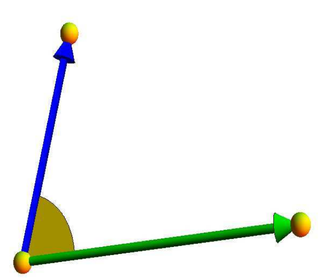
Proof: Using the dot product one can express the length of $\vec{v}$ as $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$. On the other hand, $(\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w} + 2(\vec{v} \cdot \vec{w})$ allows to solve for $\vec{v} \cdot \vec{w}$:

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

The **Cauchy-Schwarz inequality** tells $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$.

Proof. We only need to show the case $|\vec{w}| = 1$. Plug in $a = \vec{v} \cdot \vec{w}$ into the equation $0 \leq (\vec{v} - a\vec{w}) \cdot (\vec{v} - a\vec{w})$ to get $0 \leq (\vec{v} - (\vec{v} \cdot \vec{w})\vec{w}) \cdot (\vec{v} - (\vec{v} \cdot \vec{w})\vec{w}) = |\vec{v}|^2 + (\vec{v} \cdot \vec{w})^2 - 2(\vec{v} \cdot \vec{w})^2 = |\vec{v}|^2 - (\vec{v} \cdot \vec{w})^2$ which means $(\vec{v} \cdot \vec{w})^2 \leq |\vec{v}|^2$.

The Cauchy-Schwarz inequality allows us to define what an "angle" is.



The **angle** between two nonzero vectors is defined as the unique $\alpha \in [0, \pi]$ which satisfies $\vec{v} \cdot \vec{w} = |\vec{v}| \cdot |\vec{w}| \cos(\alpha)$.

[1] We cover 2300 years of math from Pythagoras (500 BC), Al Kashi (1400), Cauchy (1800) to Hamilton (1850).

Al Kashi's theorem: A triangle ABC with side lengths a, b, c and angle α opposite to c satisfies $a^2 + b^2 = c^2 + 2ab\cos(\alpha)$.

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 $c^2 = |\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}| \cdot |\vec{w}| \cos(\alpha) = a^2 + b^2 - 2ab\cos(\alpha)$.

The **triangle inequality** tells $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$

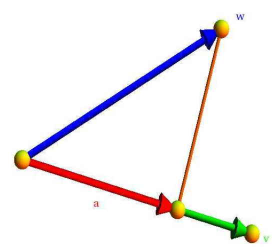
Proof: $|\vec{u}+\vec{v}|^2 = (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v}) = \vec{u}^2+\vec{v}^2+2\vec{u}\cdot\vec{v} \leq \vec{u}^2+\vec{v}^2+2|\vec{u}\cdot\vec{v}| \leq \vec{u}^2+\vec{v}^2+2|\vec{u}|\cdot|\vec{v}| = (|\vec{u}|+|\vec{v}|)^2$.

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, $\vec{v} = \langle 2, 3 \rangle$ is orthogonal to $\vec{w} = \langle -3, 2 \rangle$.

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

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]

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 plus or minus 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 $\vec{w}$.



2 Find the projection of $\vec{v} = \langle 0, -1, 1 \rangle$ onto $\vec{w} = \langle 1, -1, 0 \rangle$. **Answer:** $\mathrm{P}(\vec{v}) = \langle 1/2, -1/2, 0 \rangle$.

3 A wind force $\vec{F} = \langle 2, 3, 1 \rangle$ is applied to a car which drives in the direction of the vector $\vec{w} = \langle 1, 1, 0 \rangle$. Find the projection of $\vec{F}$ onto $\vec{w}$, the force which accelerates or slows down the car. **Answer**: $\vec{w}(\vec{F} \cdot \vec{w}/|\vec{w}|^2) = \langle 5/2, 5/2, 0 \rangle$.

4 How can we visualize the dot product? **Answer**: Difficult task but lets try. The absolute value of the dot product is the length of the projection. The dot product is positive if $\vec{v}$ and $\vec{w}$ form an acute angle, negative if that angle is obtuse.

5 Given $\vec{v} = \langle 2, 1, 2 \rangle$ and $\vec{w} = \langle 3, 4, 0 \rangle$. Find a vector which is in the plane defined by $\vec{v}$ and $\vec{w}$ and which bisects the angle between these two vectors. **Answer**. Normalize the two vectors to make them unit vectors then add them to get $\langle 13, 17, 10 \rangle/15$.

6 Given two vectors $\vec{v}, \vec{w}$ which are perpendicular. Under which condition is $\vec{v}+\vec{w}$ perpendicular to $\vec{v} - \vec{w}$? **Answer**: Find the dot product of $\vec{v} + \vec{w}$ with $\vec{v} - \vec{w}$ and set it zero.

7 Is the angle between $\langle 1, 2, 3 \rangle$ and $\langle -15, 2, 4 \rangle$ acute or obtuse? **Answer**: the dot product is 1. Cute!

[2] We have just proven Pythagoras and Al Kashi. Distance and angle were defined, not deduced.

3: Cross product

The **cross product** of two vectors $\vec{v} = \langle v_1, v_2, v_3 \rangle$ and $\vec{w} = \langle w_1, w_2, w_3 \rangle$ in space is defined as the vector

$$\vec{v} \times \vec{w} = \langle v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1 \rangle \,.$$

[1]

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

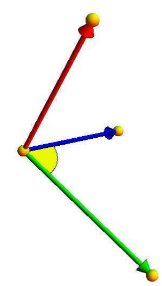
$$\begin{bmatrix} i & j & k \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} i & & \\ & v_2 & v_3 \\ & w_2 & w_3 \end{bmatrix} - \begin{bmatrix} & j & \\ v_1 & & v_3 \\ w_1 & & w_3 \end{bmatrix} + \begin{bmatrix} & & k \\ v_1 & v_2 & \\ w_1 & w_2 & \end{bmatrix}$$

which is $\vec{i}(v_2w_3 - v_3w_2) - \vec{j}(v_1w_3 - v_3w_1) + \vec{k}(v_1w_2 - v_2w_1)$.

1 The cross product of $\langle 1, 2, 3 \rangle$ and $\langle 4, 5, 1 \rangle$ is the vector $\langle -13, 11, -3 \rangle$.

The cross product $\vec{v} \times \vec{w}$ is anti-commutative. The resulting vector is orthogonal to both $\vec{v}$ and $\vec{w}$.

Proof. We verify for example that $\vec{v} \cdot (\vec{v} \times \vec{w}) = 0$ and look at the definition.



The **sin** formula: $|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}| \sin(\alpha)$.

Proof: We verify the **Lagrange's identity** $|\vec{v} \times \vec{w}|^2 = |\vec{v}|^2|\vec{w}|^2 - (\vec{v} \cdot \vec{w})^2$ by direct computation. Now, $|\vec{v} \cdot \vec{w}| = |\vec{v}||\vec{w}| \cos(\alpha)$.

The absolute value respectively length $|\vec{v} \times \vec{w}|$ defines the **area of the parallelogram** spanned by $\vec{v}$ and $\vec{w}$.

$\vec{v} \times \vec{w}$ is zero exactly if $\vec{v}$ and $\vec{w}$ are **parallel**, that is if $\vec{v} = \lambda\vec{w}$ for some real λ.

Proof. This follows immediately from the sin formula and the fact that $\sin(\alpha) = 0$ if $\alpha = 0$ or $\alpha = \pi$.

The cross product can therefore be used to check whether two vectors are parallel or not. Note that v and $-v$ are also considered parallel even so sometimes one calls this anti-parallel.

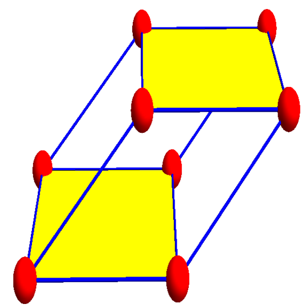
[1] It was Hamilton who found in 1843 a multiplication $*$ of 4 vectors. It contains intrinsically both dot and cross product because $(0, v_1, v_2, v_3) * (0, w_1, w_2, w_3) = (-vw, v \times w)$.

The **trigonometric sin-formula**: if a, b, c are the side lengths of a triangle and α, β, γ are the angles opposite to a, b, c then $a/\sin(\alpha) = b/\sin(\beta) = c/\sin(\gamma$.

Proof. Twice the area of the triangle is $ab\sin(\gamma) = bc\sin(\alpha) = ac\sin(\beta)$ 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.

2 If $\vec{v} = \langle a, 0, 0\rangle$ and $\vec{w} = \langle b\cos(\alpha), b\sin(\alpha), 0\rangle$, then $\vec{v} \times \vec{w} = \langle 0, 0, ab\sin(\alpha)\rangle$ which has length $|ab\sin(\alpha)|$.

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 number $|[\vec{u}, \vec{v}, \vec{w}]|$ defines the **volume of the parallelepiped** spanned by $\vec{u}, \vec{v}, \vec{w}$ and the **orientation** of three vectors is the sign of $[\vec{u}, \vec{v}, \vec{w}]$.



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 indeed 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.

3 **Problem:** Find the volume of a **cuboid** of width a length b and height c. **Answer**. The cuboid is a parallelepiped spanned by $\langle a, 0, 0\rangle$ $\langle 0, b, 0\rangle$ and $\langle 0, 0, c\rangle$. The triple scalar product is abc.

4 **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 it is spanned by the vectors $\vec{u} = \langle 1, 2, 1\rangle$, $\vec{v} = \langle 3, 2, 1\rangle$, and $\vec{w} = \langle 0, 3, 1\rangle$. We get $\vec{v} \times \vec{w} = \langle -1, -3, 9\rangle$ and $\vec{u} \cdot (\vec{v} \times \vec{w}) = 2$. The volume is 2.

5 **Problem:** 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 $R = (1, 1, 2)$. To do so find a vector $\vec{n} = \langle a, b, c\rangle$ normal to the and noting $(\vec{x} - \vec{OP}) \cdot \vec{n} = 0$. **Answer:** A normal vector $\vec{n} = \langle 1, -2, 2\rangle = \langle a, b, c\rangle$ of the plane $ax + by + cz = d$ is obtained as the cross product of $\vec{PQ}$ and $\vec{RQ}$ With $d = \vec{n} \cdot \vec{OP} = \langle 1, -2, 2\rangle \cdot \langle 1, 2, 3\rangle = 3$, we get the equation $x - 2y + 2z = 3$.

The cross product appears in physics, like for the angular momentum, the Lorentz force or the Coriolis force. We will however mainly use the cross product for constructions like to get the equation of a plane through 3 points A, B, C.

4: Lines and Planes

A point $P = (p, q, r)$ and a vector $\vec{v} = \langle a, b, c \rangle$ define the **line**

$$L = \{ \langle p, q, r \rangle + t \langle a, b, c \rangle, t \in \mathbf{R} \} .$$

The line is obtained by adding a multiple of the vector $\vec{v}$ to the vector $\vec{OP} = \langle p, q, r \rangle$. Every vector contained in the line is necessarily parallel to $\vec{v}$. We think about the parameter t as "time". For $t = 0$, we are at P and for $t = 1$ we are at $\vec{OP} + \vec{v}$.

If t is restricted to the **parameter interval** $[s, u]$, then $L = \{ \langle p, q, r \rangle + t \langle a, b, c \rangle, s \leq t \leq u \}$ is a **line segment** connecting $\vec{r}(s)$ with $\vec{r}(u)$.

1 **Problem.** Get the line through $P = (1, 1, 2)$ and $Q = (2, 4, 6)$. **Solution.** with $\vec{v} = \vec{PQ} = \langle 1, 3, 4 \rangle$ we get get $L = \{ \langle x, y, z \rangle = \langle 1, 1, 2 \rangle + t \langle 1, 3, 4 \rangle ; \}$ which is $\vec{r}(t) = \langle 1 + t, 1 + 3t, 2 + 4t \rangle$. Since $\langle x, y, z \rangle = \langle 1, 1, 2 \rangle + t \langle 1, 3, 4 \rangle$ consists of three equations $x = 1 + 2t, y = 1 + 3t, z = 2 + 4t$ we can solve each for t to get $t = (x - 1)/2 = (y - 1)/3 = (z - 2)/4$.

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

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

Proof. Each of these expressions is equal to t. These symmetric equations have to be modified a bit if 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 $z = r$.

2 Find the symmetric equations for the line through the two points $P = (0, 1, 1)$ and $Q = (2, 3, 4)$ **Solution.** first first form the parametric equations $\langle x, y, z \rangle = \langle 0, 1, 1 \rangle + t \langle 2, 2, 3 \rangle$ or $x = 2t, y = 1 + 2t, z = 1 + 3t$ and solve for t to get $x/2 = (y - 1)/2 = (z - 1)/3$.

3 **Problem:** Find the symmetric equation for the z axes. **Answer:** This is a situation where $a = b = 0$ and $c = 1$. The symmetric equations are simply $x = 0, y = 0$. If two of the numbers a, b, c are zero, we have a coordinate plane. If one of the numbers are zero, then the line is contained in a coordinate plane.

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 $\}$.

4 An example is $\Sigma = \{ \langle x, y, z \rangle = \langle 1, 1, 2 \rangle + t \langle 2, 4, 6 \rangle + s \langle 1, 0, -1 \rangle \}$. This is called the **parametric description** of a plane.

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 $\vec{n} = \langle a, b, c \rangle$, this means $ax + by + cz = ax_0by_0 + cz_0 = d$. The plane is therefore described by a single equation $ax + by + cz = d$. We have just shown

> The equation of the plane $\vec{x} = \vec{x}_0 + t\vec{v} + s\vec{w}$
>
> $$ax + by + cz = d \ ,$$
>
> where $\langle a, b, c \rangle = \vec{v} \times \vec{w}$ and d is obtained by plugging in $\vec{x}_0$.

5 **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 $\vec{v} = \langle 1, 2, 0 \rangle$ and $\vec{w} = \langle 2, 2, 2 \rangle$. We have $\vec{n} = \langle 4, -2, -2 \rangle$ and the equation is $4x - 2y - 2z = d$. The constant d is obtained by plugging in the coordinates of a point to the left. In our case, it is $4x - 2y - 2z = -4$.

6 **Problem:** Find the angle between the planes $x + y = -1$ and $x + y + z = 2$. **Answer:** find the angle between $\vec{n} = \langle 1, 1, 0 \rangle$ and $\vec{m} = \langle 1, 1, 1 \rangle$. It is $\arccos(2/\sqrt{6})$.

Finally, lets look at some distance functions.

> The distance between P and $\Sigma : \vec{n} \cdot \vec{x} = d$ containing Q is $d(P, \Sigma) = \frac{|\vec{PQ} \cdot \vec{n}|}{|\vec{n}|}$.

Proof. Project PQ onto $\vec{n}$.

> The distance between P and the line L is $d(P, L) = \frac{|(\vec{PQ}) \times \vec{u}|}{|\vec{u}|}$.

Proof: the area of the parallelogram spanned by PQ and $\vec{u}$ divided by the base length $|\vec{u}|$.

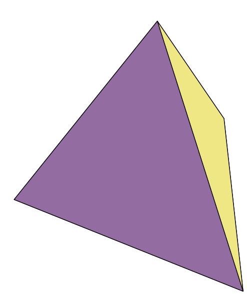
> The lines $L : \vec{r}(t) = Q + t\vec{u}, M : \vec{s}(t) = P + t\vec{v}$ have distance $d(L, M) = \frac{|(\vec{PQ}) \cdot (\vec{u} \times \vec{v})|}{|\vec{u} \times \vec{v}|}$.

Proof. Project PQ onto $\vec{n} = \vec{u} \times \vec{v}$.

> The distance between two planes $\vec{n} \cdot \vec{x} = d$ and $\vec{n} \cdot \vec{x} = e$ is $d(\Sigma, \Pi) = \frac{|e-d|}{|\vec{n}|}$.

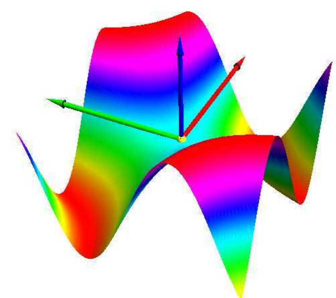
Proof: If P is on the first and Q on the second plane, the distance is the scalar projection of $\vec{PQ}$ onto $\vec{n}$. Note that $\vec{PQ} \cdot \vec{n} = d - e$.

7 A tetrahedron has vertices at the points $O = (0, 0, 0)$ and $A = (0, 0, 2)$,$B = (0, 2, 0)$, $C = (0, 0, 2)$. Find the distance between two edges which do not intersect. Solution: You can either use the distance formula or just see that the distance is $\boxed{\sqrt{2}}$ by noticing that the line AC is perpendicular to the x-axes. Now look at the triangle OAC.



Lecture 5: Functions

A **function of two variables** $f(x,y)$ is a rule which assigns to two numbers x, y a third number $f(x,y)$. For example, the function $f(x,y) = x^2y + 2x$ assigns to $(3,2)$ the number $3^2 2 + 6 = 24$. The **domain** D of a function is set of points where f is defined, the range is $\{f(x,y) \mid (x,y) \in D\}$. The **graph** of $f(x,y)$ is the surface $\{(x,y,f(x,y)) \mid (x,y) \in D\}$ in space Graphs allow to visualize functions.

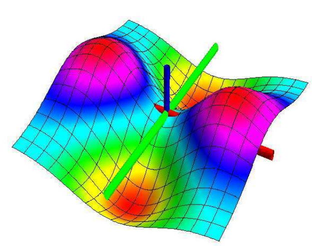


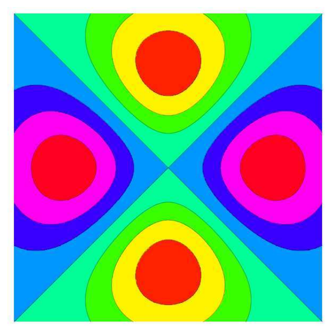
1 The graph of $f(x,y) = \sqrt{1-(x^2+y^2)}$ on the domain $D = \{x^2 + y^2 < 1\}$ is a half sphere. The range is the interval $[0,1]$.

The set $f(x,y) = c = \text{const}$ is called a **contour curve** or **level curve** of f. For example, for $f(x,y) = 4x^2 + 3y^2$, the level curves $f = c$ are ellipses if $c > 0$. The collection of all contour curves $\{f(x,y) = c\}$ is called the **contour map** of f.

2 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 curve $x^2 - y^2 = 1$ is made of two hyperbola with with their "noses" at the point $(-1,0)$ and $(1,0)$. The curve $x^2 - y^2 = -1$ consists of two hyperbola with their noses at $(0,1)$ and $(0,-1)$.

3 For $f(x,y) = (x^2 - y^2)e^{-x^2-y^2}$, we can not find explicit expressions for the contour curves $(x^2 - y^2)e^{-x^2-y^2} = c$. but we can draw the curves with the computer:



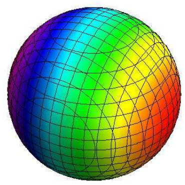


A function of three variables $g(x,y,z)$ assigns to three variables x, y, z a real number $g(x,y,z)$. We can visualize it by **contour surfaces** $g(x,y,z) = c$, where c is constant. It is helpful to look at the **traces**, the intersections of the surfaces with the coordinate planes $x = 0, y = 0$ or $z = 0$.

4 For $g(x,y,z) = z - f(x,y)$, the level surface $g = 0$ which is the graph $z = f(x,y)$ of a function of two variables. For example, for $g(x,y,z) = z - x^2 - y^2 = 0$, we have the graph $z = x^2 + y^2$ of the function $f(x,y) = x^2 + y^2$ which is a paraboloid. Most surfaces $g(x,y,z) = c$ are not graphs.

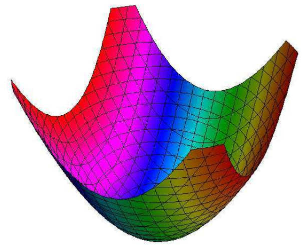
5 If $f(x,y,z)$ is a polynomial and $f(x,x,x)$ is quadratic in x, then $\{f = c\}$ is a **quadric**.

Sphere



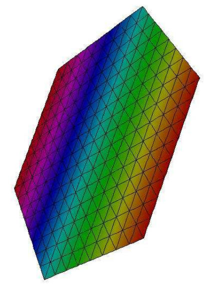
$x^2 + y^2 + z^2 = 1$

Paraboloid



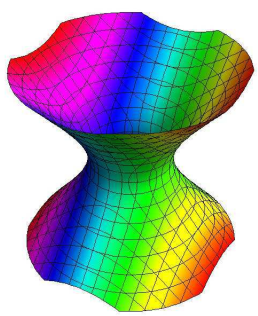
$x^2 + y^2 - c = z$

Plane



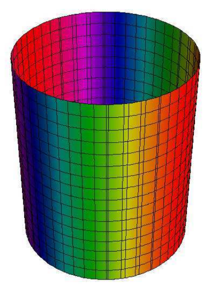
$ax + by + cz = d$

One sheeted Hyperboloid



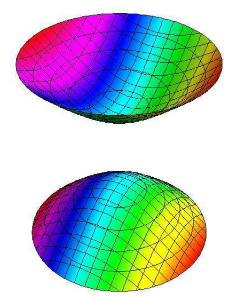
$x^2 + y^2 - z^2 = 1$

Cylinder



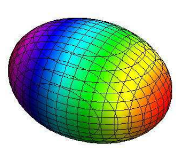
$x^2 + y^2 = r^2$

Two sheeted Hyperboloid



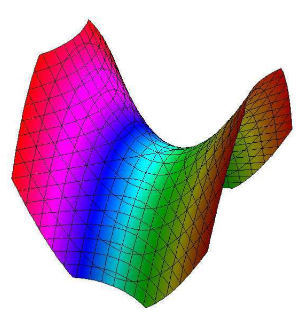
$x^2 + y^2 - z^2 = -1$

Ellipsoid



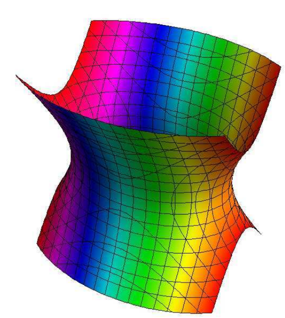
$x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$

Hyperbolic paraboloid



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

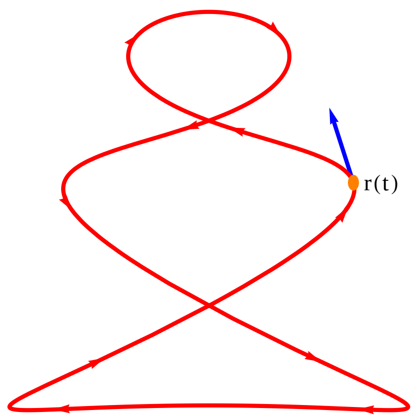
Elliptic hyperboloid

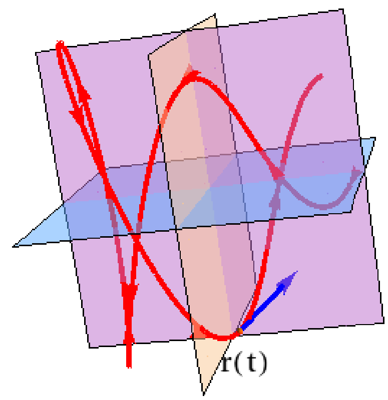


$x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$

Lecture 6: Curves

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





We always think of the **parameter** t as **time**. For a fixed t, we have a vector $\langle x(t), y(t), z(t) \rangle$ in space. As t varies, the end point of this vector moves along a curve. The parametrization contains more information about the curve then the curve. It tells also how fast and in which direction we trace the curve.

1 The parametrization $\vec{r}(t) = \langle 1 + 3\cos(t), 3\sin(t) \rangle$ is a **circle** of radius 3 centered at $(1, 0)$

2 $\vec{r}(t) = \langle \cos(3t), \sin(5t) \rangle$ defines a **Lissajous curve** example.

3 If $x(t) = t$, $y(t) = f(t)$, the curve $\vec{r}(t) = \langle t, f(t) \rangle$ traces the **graph** of the function $f(x)$. For example, for $f(x) = x^2 + 1$, the graph is a parabola.

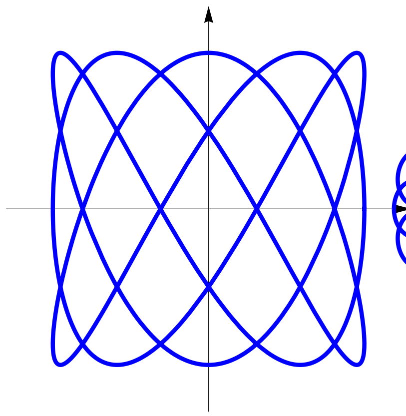
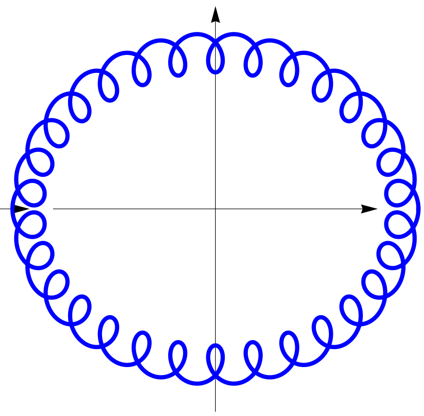
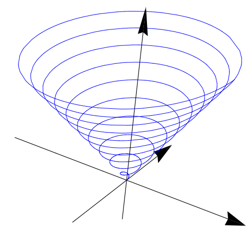
4 With $x(t) = 2\cos(t), y(t) = \sin(t)$, then $\vec{r}(t)$ follows an **ellipse** $x(t)^2/4 + y(t)^2 = 1$.

5 The space curve $\vec{r}(t) = \langle t\cos(t), t\sin(t), t \rangle$ traces a **helix** with increasing radius.

6 If $x(t) = \cos(2t), y(t) = \sin(2t), z(t) = 2t$ is the same curve as before but the **parameterization** has changed.

7 With $x(t) = \cos(-t), y(t) = \sin(-t), z(t) = -t$ it is traced in the **opposite direction**.

8 With $\vec{r}(t) = \langle \cos(t), \sin(t) \rangle + 0.1\langle \cos(17t, \sin(17t) \rangle$ we have an example of an **epicycle**, where a circle turns on a circle. It was used in the **Ptolemaic geocentric system** which predated the Copernican system still using circular orbits and then the modern Keplerian system, where planets move on ellipses and which can be derived from Newton's laws.

If $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ is a curve, then $\vec{r}\,'(t) = \langle x'(t), y'(t), z'(t)\rangle = \langle \dot{x}, \dot{y}, \dot{z}\rangle$ is called the **velocity** at time t. Its length $|\vec{r}'(t)|$ is called **speed** and $\vec{v}/|\vec{v}|$ is called **direction of motion**. The vector $\vec{r}\,''(t)$ is called the **acceleration**. The third derivative $\vec{r}\,'''$ is called the **jerk**.

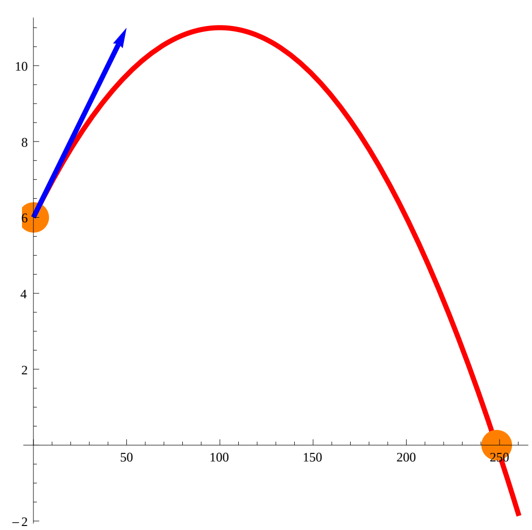
Any vector parallel to $\vec{r}\,'(t)$ is called **tangent** to the curve at $\vec{r}(t)$.

The **addition rule** in one dimension $(f+g)' = f'+g'$, the **scalar multiplication rule** $(cf)' = cf'$ and the **Leibniz rule** $(fg)' = f'g + fg'$ and the **chain rule** $(f(g))' = f'(g)g'$ generalize to vector-valued functions because in each component, we have the single variable rule.
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. In particular, if $\vec{r}''(t) = \langle 0, 0, -10\rangle$, $\vec{r}'(0) = \langle 0, 1000, 2\rangle, \vec{r}(0) = \langle 0, 0, h\rangle$, then $\vec{r}(t) = \langle 0, 1000t, h + 2t - 10t^2/2\rangle$.

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



7: Arc length and curvature

If $t \in [a,b] \mapsto \vec{r}(t)$ is a curve with velocity $\vec{r}\,'(t)$ and speed $|\vec{r}\,'(t)|$, then $L = \int_a^b |\vec{r}\,'(t)|\,dt$ is called the **arc length of the curve**. Written out in three dimensions, this is $L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\,dt$.

1 The arc length of the **circle** of radius R given by $\vec{r}(t) = \langle R\cos(t), R\sin(t)\rangle$ parameterized by $0 \leq t \leq 2\pi$ is 2π because the speed $|\vec{r}'(t)|$ is constant R. The answer is $2\pi R$.

2 The **helix** $\vec{r}(t) = \langle \cos(t), \sin(t), t\rangle$ has velocity $\vec{r}\,'(t) = \langle -\sin(t), \cos(t), 1\rangle$ and constant speed $|\vec{r}\,'(t)| = \langle -\sin(t), \cos(t), 1\rangle = \sqrt{2}$.

3 What is the arc length of the curve

$$\vec{r}(t) = \langle t, \log(t), t^2/2\rangle$$

for $1 \leq t \leq 2$? **Answer:** Because $\vec{r}'(t) = \langle 1, 1/t, t\rangle$, we have $\vec{r}'(t) = \sqrt{1 + \frac{1}{t^2} + t^2} = |\frac{1}{t} + t|$ and $L = \int_1^2 \frac{1}{t} + t\,dt = \log(t) + \frac{t^2}{2}|_1^2 = \log(2) + 2 - 1/2$.

4 Find the arc length of the curve $\vec{r}(t) = \langle 3t^2, 6t, t^3\rangle$ from $t = 1$ to $t = 3$.

5 What is the arc length of the curve $\vec{r}(t) = \langle \cos^3(t), \sin^3(t)\rangle, 0 \leq t \leq 2\pi$? **Answer:** We have $|\vec{r}'(t)| = 3\sqrt{\sin^2(t)\cos^4(t) + \cos^2(t)\sin^4(t)} = (3/2)|\sin(2t)|$. The absolute value forces us to split the integral into 4 intervals. Since $\int_0^{\pi/2} \sin(2t)\,dt = 1$, we have $\int_0^{2\pi} (3/2)|\sin(2t)|\,dt = (3/2)4 = 6$.

6 Find the arc length of $\vec{r}(t) = \langle t^2/2, t^3/3\rangle$ for $-1 \leq t \leq 1$. This cubic curve satisfies $y^2 = x^3 8/9$ and is an example of an **elliptic curve**. The speed is $|\vec{r}(t)| = \sqrt{t^2 + t^4}$. Because $\int x\sqrt{1+x^2}\,dx = (1+x^2)^{3/2}/3$, the arc length integral can be evaluated using substitution by as $\int_{-1}^1 |t|\sqrt{1+t^2}\,dx = 2\int_0^1 t\sqrt{1+t^2}\,dt = 2(1+t^2)^{3/2}/3|_0^1 = 2(2\sqrt{2}-1)/3$.

7 The arc length of an **epicycle** $\vec{r}(t) = \langle t + \sin(t), \cos(t)\rangle$ 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 $L = \int_0^{\pi} \sqrt{2 + 2\cos(2u)}\,2du = \int_0^{\pi} \sqrt{2 + 2\cos^2(u) - 2\sin^2(u)}\,2du = \int_0^{\pi} \sqrt{4\cos^2(u)}\,2du = 4\int_0^{\pi} |\cos(u)|\,du = 8$.

8 Compute the arc length of the **catenary** $\vec{r}(t) = \langle t, e^t + e^{-t}\rangle$ on an interval $[a,b]$ can be computed as $e^b - e^a - e^{-b} + e^{-a}$. By the way, $(e^t + e^{-t})/2$ is called the hyperbolic cosine and denoted by $\cosh(t)$.

Because a parameter change $t = t(s)$ corresponds to a **substitution** in the integration which does not change the integral, we immediately have

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

9 The circle parameterized by $\vec{r}(t) = \langle \cos(t^2), \sin(t^2) \rangle$ on $t = [0, \sqrt{2\pi}]$ has the velocity $\vec{r}\,'(t) = 2t \langle -\sin(t), \cos(t) \rangle$ and speed $2t$. The arc length is still $\int_0^{\sqrt{2\pi}} 2t\ dt = t^2|_0^{\sqrt{2\pi}} = 2\pi$.

10 We do not always have a closed formula for the arc length of a curve. The length of the **Lissajous figure** $\vec{r}(t) = \langle \cos(3t), \sin(5t) \rangle$ leads to $\int_0^{2\pi} \sqrt{9\sin^2(3t) + 25\cos^2(5t)}\ dt$ which needs to be evaluated numerically.

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

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

The curvature is the magnitude of the acceleration vector if $\vec{r}(t)$ traces the curve with constant speed 1. A large curvature at a point means that the curve turns sharply. 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.

The curvature does not depend on the parametrization.

Proof. If $s(t)$ 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.

11 The curve $\vec{r}(t) = \langle t, f(t) \rangle$, 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 $\kappa(t) = |\sin(t)|/|\sqrt{1 + \cos^2(t)}^3$.

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**.

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 B is automatically normal to T and N, we have shown:

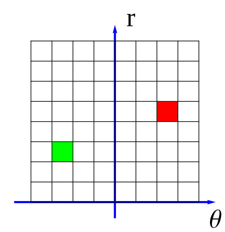
The three vectors $(\vec{T}(t), \vec{N}(t), \vec{B}(t))$ are unit vectors orthogonal to each other.

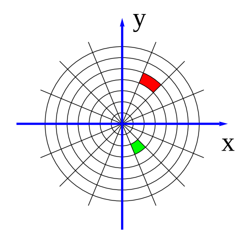
A useful formula for curvature is

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

We prove this in class.

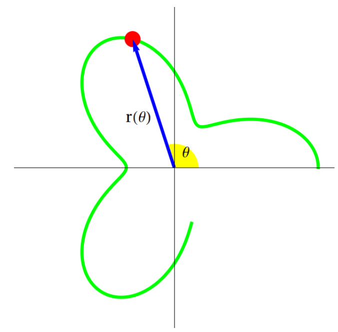
8: Polar and spherical coordinates





A point (x, y) in the plane has the **polar coordinates** $r = \sqrt{x^2 + y^2}, \theta = \arctan(y/x)$ leading to the relation $(x, y) = (r\cos(\theta), r\sin(\theta))$.

The formula $\theta = \text{arctg}(y/x)$ defines the angle θ only up to an addition of π and assigns to (x, y) and $(-x, -y)$ the same θ value. To get the correct θ, choose $\arctan(y/x)$ in $(-\pi/2, \pi/2]$ for $x > 0$ define it to be $\pi/2$ for the positive y axes and $\arctan(y/x) + \pi$ for $x < 0$ equal to $-\pi/2$ on the negative y-axes. For $(x, y) = (0, 0)$, the polar angle θ is not defined.



A curve given in polar coordinates as $r(\theta) = f(\theta)$ is called a **polar curve**. It can in Cartesian coordinates be described as $\vec{r}(t) = \langle f(t)\cos(t), f(t)\sin(t)\rangle$.

1 Describe the curve $r = \theta$ in Cartesian coordinates. **Solution** A formal substitution gives $\sqrt{x^2 + y^2} = \arctan(y/x)$ but we can do better. Remember that for the curve $r(t) = \langle r\cos(t), r\sin(t)$ we have exactly the relation $r = t$. The curve is a spiral.

2 What is the curve $r = |2\sin(\theta)|$? **Solution** Lets ignore the absolute value for a moment and look at $r^2 = 2r\sin(\theta)$. This can be written as $x^2 + y^2 = 2y$ which is $x^2 + y^2 - 2y + 1 = 1$. The curve is a circle of radius 1 centered at $(0, 1)$. Since we have the absolute value, the radius at θ and $\theta + \pi$ is the same and add the circle of radius 1 centered at $(0, -1)$.

If we represent points in space as

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

we speak of **cylindrical coordinates**.

Here are some surfaces described in cylindrical coordinates:

3 $r = 1$ is a **cylinder**,

4 $r = |z|$ is a **double cone**

5 $\theta = 0$ is a **half plane**

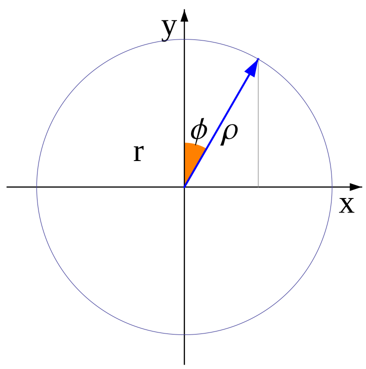
6 $r = \theta$ is a **rolled sheet of paper**

7 $r = 2 + \sin(z)$ is an example of a **surface of revolution**.

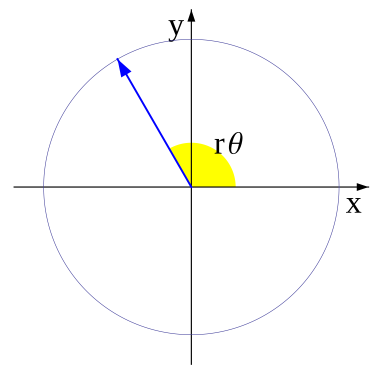
Spherical coordinates use the distance ρ to the origin as well as two angles θ and ϕ. The first angle θ is the polar angle in polar coordinates of the xy coordinates and ϕ is the angle between the vector $\vec{OP}$ and the z-axis. The relation is

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

There are two important figures to see the connection. The distance to the z axes $r = \rho \sin(\phi)$ and the height $z = \rho \cos(\phi)$ can be read off by the left picture the rz-plane, the coordinates $x = r\cos(\theta), y = r\sin(\theta)$ can be seen in the right picture the xy-plane.



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



Here are some level surfaces described in spherical coordinates:

8 $\rho = 1$ is a **sphere**,

9 The surface $\phi = \pi/2$ is a **single cone**

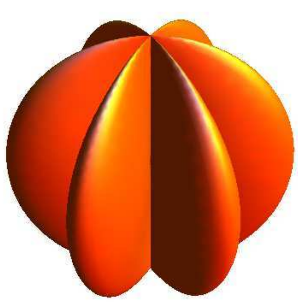
10 The surface $\sin(\theta) = \cos(\phi)$ is a **plane**.

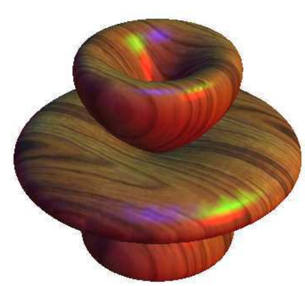
11 $\rho = \phi$ is an **apple shaped surface**

12 $\rho = 2 + \cos(3\theta)\sin(\phi)$ is an example of a **bumpy sphere**.

13 Write $x^2 + y^2 - 5x = z^2$ in cylindrical coordinates! Answer: $r^2 - 5r\cos(\theta) = z^2$.

14 Match the surfaces with $\rho = |\sin(3\phi)|, \rho = |\sin(3\theta)|$ in spherical coordinates (ρ, θ, ϕ). It helps to see this in the rz plane or the xy plane.





9: Parametrized surfaces

We have seen that lanes can be described by implicit equations $x + y + z = 1$ as well as by by parametrization $\vec{r}(t,s) = \langle 1 + t + s, -t, -s \rangle$. Today, we look at parametrizations of more general surfaces.

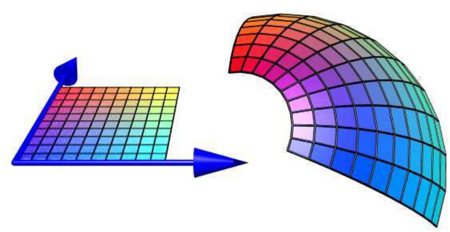
A **parametrization** of a surface is a vector-valued function

$$\vec{r}(u,v) = \langle x(u,v), y(u,v), z(u,v) \rangle \ ,$$

where $x(u,v), y(u,v), z(u,v)$ are three functions of two variables.

Because two parameters u and v are involved, the map $\vec{r}$ from the uv-plane to space is also called uv**-map**.

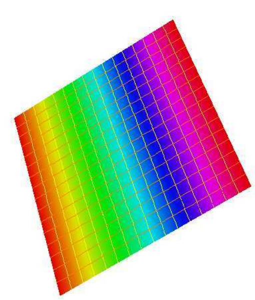
A **parametrized surface** is the image of the uv-map. The domain of the uv-map is called the **parameter domain**.



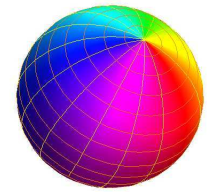
If the first parameter u is kept constant, then $v \mapsto \vec{r}(u,v)$ is a curve on the surface. Similarly, for constant v, the map $u \mapsto \vec{r}(u,v)$ traces a curve on the surface. These curves are called **grid curves**.

A computer draws surfaces using grid curves. The world of parametric surfaces is intriguing. It can be explored with the help of a computer. Keep in mind the following 4 important examples. They cover a wide range of cases.

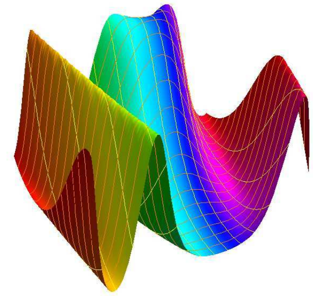
I **Planes**. Parametric: $\vec{r}(s,t) = \vec{OP} + s\vec{v} + t\vec{w}$
Implicit: $ax + by + cz = d$. Parametric to Implicit: find the normal vector $\vec{n} = \vec{v} \times \vec{w}$.
Implicit to Parametric: find two vectors $\vec{v}, \vec{w}$ normal to the vector $\vec{n}$. For example, find three points P, Q, R on the surface and forming $\vec{u} = \vec{PQ}, \vec{v} = \vec{PR}$.



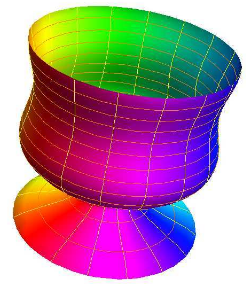
II **Spheres**: Parametric: $\vec{r}(u,v) = \langle a,b,c\rangle + \langle \rho\cos(u)\sin(v), \rho\sin(u)\sin(v), \rho\cos(v)\rangle$.
Implicit: $(x-a)^2 + (y-b)^2 + (z-c)^2 = \rho^2$.
Parametric to implicit: read off the radius and the center
Implicit to parametric: find the center (a,b,c) and the radius r possibly by completing the square.



III **Graphs**:
Parametric: $\vec{r}(u,v) = \langle u, v, f(u,v)\rangle$
Implicit: $z - f(x,y) = 0$. Parametric to implicit: think about $z = f(x,y)$
Implicit to parametric: use x and y as the parameterizations.



IV **Surfaces of revolution**:
Parametric: $\vec{r}(u,v) = \langle g(v)\cos(u), g(v)\sin(u), v\rangle$
Implicit: $\sqrt{x^2+y^2} = r = g(z)$ can be written as $x^2 + y^2 = g(z)^2$.
Parametric to implicit: read off the function $g(z)$ the distance to the z-axis.
Implicit to parametric: again, the function g is the key link.



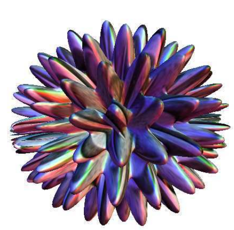
1 Describe the surface $\vec{r}(u,v) = \langle v^5\cos(u), v^5\sin(u), v\rangle$, where $u \in [0,2\pi]$ and $v \in \mathbf{R}$. **Solution.** It is a surface of revolution. We have $r = v^5 = z^5$. Draw this in the rz-plane. You see that r is small the origin making the surface pointy like a needle at the tip.

2 Find a parametrization for the plane which contains the three points $P = (3,7,1)$,$Q = (6,2,1)$ and $R = (0,3,4)$. **Solution.** Take $\vec{r}(s,t) = \vec{OP} + s\vec{QP} + t\vec{RP}$. $\vec{r}(s,t) = (3-3s-3t, 7-5s-4t, 1+3t)$.

3 Parametrize the lower half of the ellipsoid $x^2/4 + y^2/9 + z^2/25 = 1, z < 0$. **Solution**. One solution is to solve for z and write it as a graph $\vec{r}(u,v) = \langle u, v, -\sqrt{25 - 25u^2/4 - 25v^2/9}\rangle$. We can also deform a sphere $\vec{r}(\theta,\phi) = \langle 2\sin(\phi)\cos(\theta), 3\sin(\phi)\sin(\theta), 5\cos(\phi)\rangle$.

4 Parametrize the upper half of the hyperboloid $x^2 + y^2/4 - z^2 = -1$. **Solution.** The round hyperboloid, where $r^2 = z^2 - 1$ is given by $\vec{r}(\theta,z) = \langle\sqrt{z^2-1}\cos(\theta), \sqrt{z^2-1}\sin(\theta), z\rangle$. Deform this now to get $\vec{r}(\theta,z) = \langle\sqrt{z^2-1}\cos(\theta), 2\sqrt{z^2-1}\sin(\theta), z\rangle$.

5 Describe the surface $\vec{r}(u,v) = \langle 2 + \sin(13u) + \sin(17v)\rangle\langle\cos(u)\sin(v), \sin(u)\sin(v), \cos(v)\rangle$. **Solution.** It looks a bit like a hedgehog. It is an example of a "bumpy sphere". The radius ρ is a function of the angles. In spherical coordinates, we have $\rho = (2 + \sin(13\theta) + \sin(17\phi))$.



11: Continuity

A function $f(x, y)$ with domain R is is **continuous at a point** $(a, b) \in R$ if $f(x, y) \to f(a, b)$ whenever $(x, y) \to (a, b)$. The function f is **continuous on** R, if f is continuous for every point (a, b) on R.

1 a) $f(x, y) = x^2 + y^4 + xy + \sin(y + \sin\sin\sin\sin(x)^2$ is continuous on the entire plane. It is built up from functions which are continuous using addition, multiplication or composition of functions which are all continuous everywhere.

2 $f(x, y) = 1/(x^2+y^2)$ is continuous everywhere except at the origin, where it is not defined.

3 $f(x, y) = y + \sin(x)/|x|$ is continuous away from $x = 0$. At every point $(0, y)$ it is discontinuous. $f(1/n, y) \to y + 1$ and $f(-1/n, y) \to y - 1$ for $n \to \infty$.

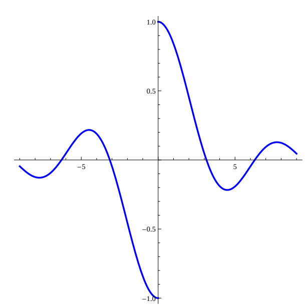
4 $f(x, y) = \sin(1/(x + y))$ is continuous except on the line $x + y = 0$.

5 $f(x, y) = (x^4 - y^4)/(x^2 + y^2)$ is continuous at $(0, 0)$. You can divide out $x^2 + y^2$ to see that the function is equivalent to $x^2 - y^2$ away from $(0, 0)$. After defining $f(0, 0) = 0$ we see that the function is continuous.

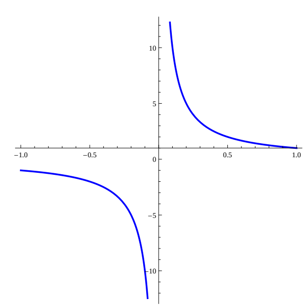
6 There are three reasons, why a function can be discontinuous: it can **jump**, it can **diverge to infinity**, or it can **oscillate**. An example of a jump appears with $f(x) = \sin(x)/|x|$, a pole $g(x) = 1/x$ leads to a vertical asymptote and the function going to infinity. An example of a function discontinuous due to oscillations is $h(x) = \sin(1/x)$. Its graph is the **devil's comb**.

The prototypes in one dimensions are

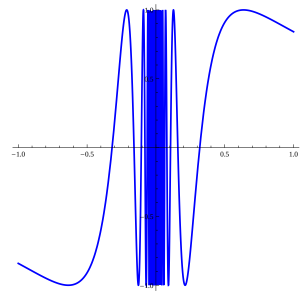
Jump $f(x) = \sin(x)/|x|$



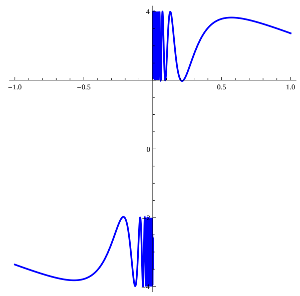
Diverge $g(x) = 1/x$



Oscillate $h(x) = \sin(1/x)$



One can have mixtures of these phenomena like the function $3\sin(x)/|x| + \sin(1/x)$, which jumps and also has an oscillatory problem at $x = 0$.



There are two handy tools to discover a discontinuities:
1) Use polar coordinates with coordinate center at the point to analyze the function.
2) Restrict the function to one dimensional curves and check continuity on that curve, where one has a function of one variables.

7 Determine whether the function $f(x,y) = \frac{\sin(x^2+y^2)}{x^2+y^2}$ is continuous at $(0,0)$. **Solution** Use polar coordinates to write this as $\sin(r^2)/r^2$ which is continuous at 0 (apply l'Hopital twice if you want to verify this).

8 Is the function $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$ continuous at $(0,0)$? **Solution** Use polar coordinates to see that this $\cos(2\theta)$. We see that the value depends on the angle only. Arbitrarily close to $(0,0)$, the function takes any value from -1 to 1.

9 Is the function $f(x,y) = \frac{x^2y}{x^4+y^2}$ continuous? **Solution**. Look on the line $x^2 = y$ to get the function $x^4/(2x^4) = 1/(2x^2)$. It is not continuous at 0. This example is a **real shocker** because it is continuous through each line through the origin: if $y = ax$, then $f(x,ax) = ax^3/(x^4 + a^2x^2) = ax/(x^2 + a^2)$ which is goes to zero for $x \to 0$ as long as $a \neq 0$. If $a = 0$ however, we have $y = 0$ and $f = 0/x^4$ which can be continuously extended to $x = 0$ too.

10 What about the function $f(x,y) = \frac{xy^2+y^3}{x^2+y^2}$? **Solution**. Use polar coordinates and write $r^3\sin^2(\theta)(\cos(\theta) + \sin(\theta)/r^2 = r\sin^2(\theta)(\cos(\theta) + \sin(\theta)$ which shows that the function converges to 0 as $r \to 0$.

11 Is the function $f(x,y) = \frac{\sin(x^2+y^2)}{\sqrt{x^2+y^2}}$ continuous everywhere? **Solution.** Use polar coordinates to see that this is $\sin(r^2)/r^2$. This function is continuous at 0 by Hôpital's theorem.

12: Partial derivatives

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

We also write $f_x(x,y) = \frac{\partial}{\partial x} f(x,y)$. and $f_{yx} = \frac{\partial}{\partial x}\frac{\partial}{\partial y} f$. [1]

1 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 $f_{xx} + f_{yy} = 0$. A function which satisfies this equation is also called **harmonic**. The equation $f_{xx} + f_{yy} = 0$ is an example of a **partial differential equation** for the unknown function $f(x,y)$ involving partial derivatives. The vector $\langle f_x, f_y \rangle$ is called the gradient.

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

Proof: we look at the equations without taking limits first. We extend the definition and say that a background Planck constant h is positive, then $f_x(x,y) = [f(x+h,y) - f(x,y)]/h$. For $h = 0$ we define f_x as before. Compare the two sides for fixed $h > 0$:

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

$hf_y(x,y) = f(x,y+h) - f(x,y)$.
$h^2 f_{yx}(x,y) = f(x+h,y+h) - f(x+h,y) - (f(x,y+h) - f(x,y))$

No limits were taken. We established an identity which holds for all $h > 0$, the discrete derivatives f_x, f_y satisfy $f_{xy} = f_{yx}$. It is a "quantum Clairaut" theorem. If the classical derivatives f_{xy}, f_{yx} are both continuous, the limit $h \to 0$ leads to the classical Clairaut's theorem. The quantum Clairaut theorem holds for **any** functions $f(x,y)$ of two variables. Not even continuity is needed.

2 Find $f_{xxxxxyxxxxx}$ for $f(x) = \sin(x) + x^6y^{10}\cos(y)$. Answer: Do not compute, but think.

3 The continuity assumption for f_{xy} is necessary. The example $f(x,y) = \frac{x^3y - xy^3}{x^2+y^2}$ contradicts Clairaut's theorem:

$f_x(x,y) = (3x^2y - y^3)/(x^2+y^2) - 2x(x^3y - xy^3)/(x^2+y^2)^2$, $f_x(0,y) = -y$, $f_{xy}(0,0) = -1$,

$f_y(x,y) = (x^3 - 3xy^2)/(x^2+y^2) - 2y(x^3y - xy^3)/(x^2+y^2)^2$, $f_y(x,0) = x$, $f_{y,x}(0,0) = 1$.

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**. If only the derivative with respect to one variable appears, it is called an **ordinary differential equation**.

[1] $\partial_x f, \partial_y f$ were introduced by Carl Gustav Jacobi. Josef Lagrange had used the term "partial differences".

Here are some examples of partial differential equations. You should know the first 4 well.

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

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

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

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

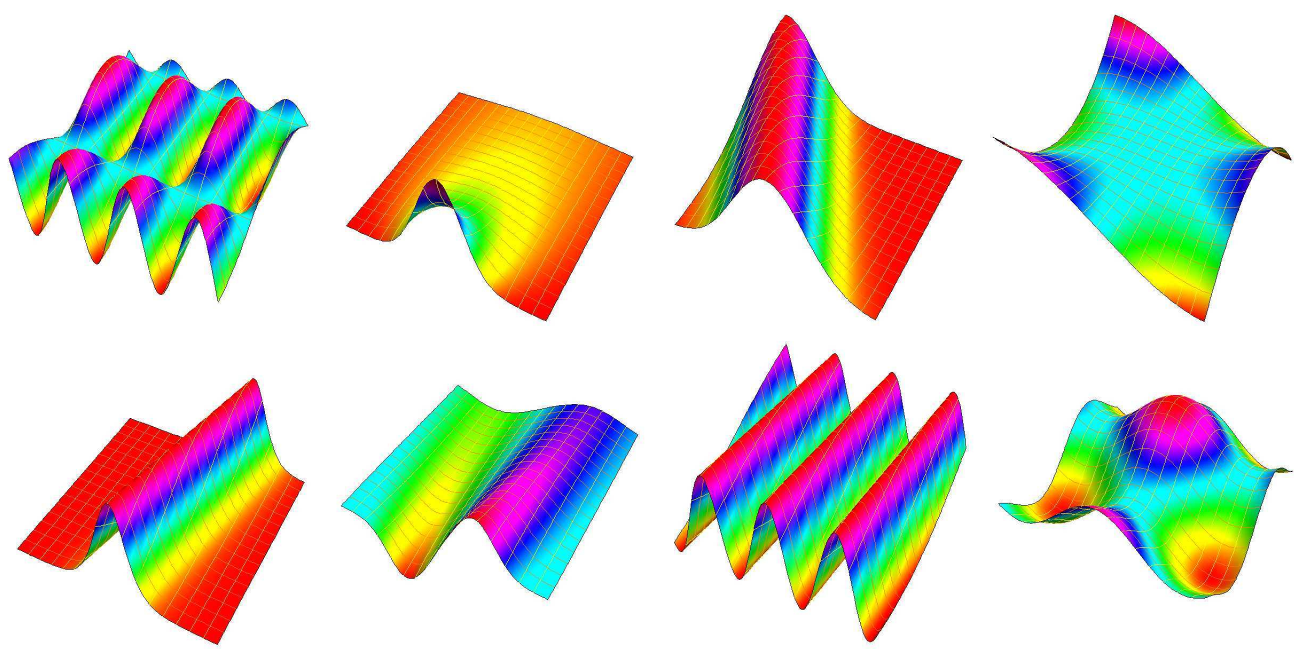
8 The **eiconal equation** $f_x^2 + f_y^2 = 1$ is used to see the evolution of wave fronts in optics. The function $\boxed{f(x,y) = \sqrt{x^2+y^2}}$ satisfies the eiconal equation.

9 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.

10 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.

11 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** $\hbar \sim 10^{-34} Js$.]

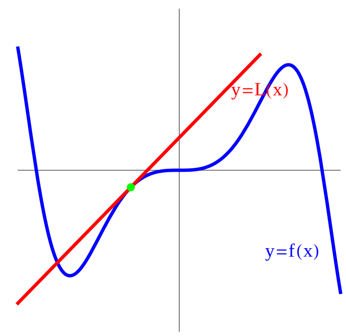
Here are the graphs of the solutions of the equations. Can you match them with the PDE's?



14: Linearization

The **linear approximation** of a function $f(x)$ at a point a is the linear function

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



The graph of the function L is close to the graph of f near a. We generalize this to higher dimensions:

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

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

Using the **gradient** $\nabla f(x, y) = \langle f_x, f_y \rangle$ rsp. $\nabla f(x, y, z) = \langle f_x, f_y, f_z \rangle$, the linearization can be written as $L(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})$. By keeping the second variable $y = b$ is fixed, we get to 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 $f(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$. Please avoid the notion of **differentials**. It is a relict from old times.

1 What is the linear approximation of the function $f(x, y) = \sin(\pi x y^2)$ at the point $(1, 1)$? We have $(f_x(x, y), y_f(x, y) = (\pi y^2 \cos(\pi x y^2), 2y\pi \cos(\pi x y^2))$ which is at the point $(1, 1)$ equal to $\nabla f(1, 1) = \langle \pi \cos(\pi), 2\pi \cos(\pi) \rangle = \langle -\pi, 2\pi \rangle$.

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

$$-0.00943 = f(1+0.01, 1+0.01) \sim L(1+0.01, 1+0.01) = -\pi 0.01 - 2\pi 0.01 + 3\pi = -0.00942 .$$

3 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 $\nabla f(x, y, z) = \langle y + z, x + z, y + x \rangle, \nabla f(1, 1, 1) = \langle 2, 2, 2 \rangle$. we have $L(x, y, z) = f(1, 1, 1) + \langle 2, 2, 2 \rangle \cdot \langle x - 1, y - 1, z - 1 \rangle = 3 + 2(x - 1) + 2(y - 1) + 2(z - 1) = 2x + 2y + 2z - 3$.

4 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 $f(x_0, y_0, z_0) = 5$. **Solution.**The gradient

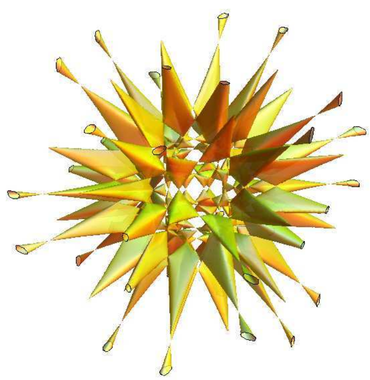
is $\nabla f(x,y,z) = (e^x\sqrt{y}z, e^x z/(2\sqrt{y}), e^x\sqrt{y})$. At the point $(x_0,y_0,z_0) = (0,25,1)$ the gradient is the vector $(5,1/10,5)$. The linear approximation is $L(x,y,z) = f(x_0,y_0,z_0) + \nabla f(x_0,y_0,z_0)(x-x_0,y-y_0,z-z_0) = 5+(5,1/10,5)(x-0,y-25,z-1) = 5x+y/10+5z-2.5$. We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

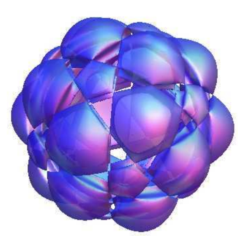
5 Find the tangent line to the graph of the function $g(x) = x^2$ at the point $(2,4)$. **Solution:** the tangent line is the level curve of the linearlization of $L(x,y)$ of $f(x,y) = y - x^2 = 0$ which passes through the point. We compute the gradient $\langle a,b\rangle = \nabla f(2,4) = \langle -g'(2),1\rangle = \langle -4,1\rangle$ and forming $ax+by = -4x+y = d$, where $d = -4\cdot 2+1\cdot 4 = -4$. The answer is $\boxed{-4x+y=-4}$.

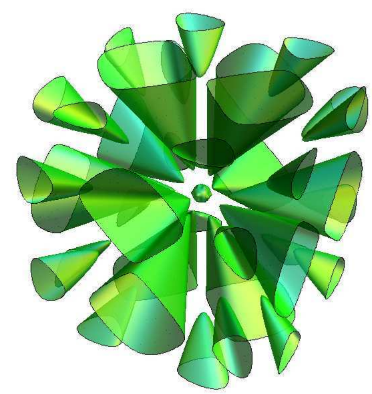
6 The **Barth surface** is defined as the level surface $f = 0$ of

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

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 $(1,1,0)$. **Solution:** We find the level curve of the linearization by computing the gradient $\nabla f(1,1,0) = \langle 64,64,0\rangle$. The surface is $x+y = d$ for some constant d. By plugging in the point $(1,1,0)$ we see that $x+y = 2$.



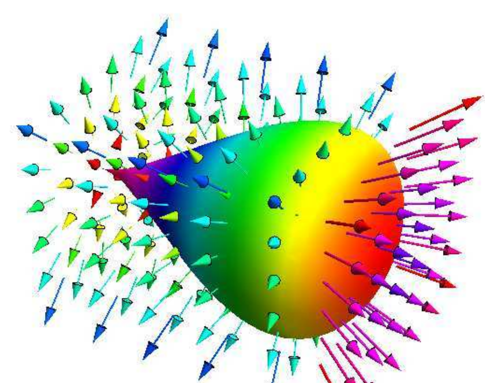




The quartic surface

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

7 is called the **piriform**. What is the equation for the tangent plane at the point $P = (2,2,2)$ of this pair shaped surface? **Solution.** We get $\langle a,b,c\rangle = \langle 20,4,4\rangle$ 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 $(x,y,z) = (2,2,2)$.



15: Chain rule

If f and g are functions of t, then the **single variable chain rule** tells

$$\frac{d}{dt} f(g(t)) = f'(g(t))g'(t) .$$

For example, $d/dt \sin(\log(t)) = \cos(\log(t))/t$. This **chain rule** can be proven by linearizing the functions f and g and verifying the chain rule in the linear case. The rule is useful for finding derivatives like $\arccos'(x)$: write $1 = d/dx \cos(\arccos(x)) = -\sin(\arccos(x)) \arccos'(x) = -\sqrt{1 - \sin^2(\arccos(x))} \arccos'(x) = \sqrt{1 - x^2} \arccos'(x)$ so that $\arccos'(x) = -1/\sqrt{1 - x^2}$.

1 Find the derivative $d/dx \arctan(x)$. **Solution.** We have $\sin' = \cos$ and $\cos' = -\sin$ and from $\cos^2(x) + \sin^2(x) = 1$. follows $1 + \tan^2(x) = 1/\cos^2(x)$. Therefore $d/dx \tan(\arctan(x)) = 1/\cos^2(\arctan(x)) \tan'(x) = x$ Now $1/\cos^2(x) = 1/(1 + \tan^2(x)$ so that $\tan'(x) = 1/(1 + x^2)$.

Define the **gradient** $\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle$ or $\nabla f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle$.

If $\vec{r}(t)$ is curve and f is a function of several variables we can build a function $t \mapsto f(\vec{r}(t))$ of one variable. Similarly, If $\vec{r}(t)$ is a parametrization of a curve in the plane and f is a function of two variables, then $t \mapsto f(\vec{r}(t))$ is a function of one variable.

The **multivariable chain rule** is

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

Proof. When written out in two dimensions, it is

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

Now, 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:

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

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 $x'(t)$.

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

From $f(x,y) = 0$, we can express y as a function of x. From $d/df(x, y(x)) = \nabla f \cdot (1, y'(x)) = f_x + f_y y' = 0$, we get

Implicit differentation: $y' = -f_x/f_y$.

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 $f_x(x,y,z(x,y))1 + f_z(x,y)z_x(x,y) = 0$. This leads to the following

Implicit differentiation:
$z_x(x,y) = -f_x(x,y,z)/f_z(x,y,z)$ $\qquad$ $z_y(x,y) = -f_y(x,y,z)/f_z(x,y,z)$

3 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:** $z_x(x,y) = -f_x(2,1,1)/f_z(2,1,1) = -4/(2/9) = -18$.

4 How does the chain rule relate to other differentiation rules? **Answer.** The chain rule is universal: it implies single variable differentiation rules like the addition, product and quotient rule in one dimensions:
$f(x,y) = x + y, x = u(t), y = v(t), d/dt(x+y) = f_x u' + f_y v' = u' + v'$.
$f(x,y) = xy, x = u(t), y = v(t), d/dt(xy) = f_x u' + f_y v' = vu' + uv'$.
$f(x,y) = x/y, x = u(t), y = v(t), d/dt(x/y) = f_x u' + f_y v' = u'/y - v'u/v^2$.

5 Can one prove the chain rule from linearization and just verifying it for linear functions? **Solution**. Yes, 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 $\vec{r}(t) = \langle x_0 + tu, y_0 + tv \rangle$ 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.

6 Mechanical systems are determined by the energy function $H(x,y)$, which is a function of two variables. The first variable x is the position and the second variable y is the momentum. The equations of motion for the curve $\vec{r}(t) = \langle x(t), y(t) \rangle$ are called **Hamilton equations**:

$$\begin{aligned} x'(t) &= H_y(x,y) \\ y'(t) &= -H_x(x,y) \end{aligned}$$

In a homework you verify that the energy of a Hamiltonian system is preserved: for every path $\vec{r}(t) = \langle x(t), y(t) \rangle$ solving the system, we have $H(x(t), y(t)) = const$.

16: Gradient and Tangent

The **gradient** $\nabla f(x,y) = \langle f_x(x,y), f_y(x,y)\rangle$ or $\nabla f(x,y,z) = \langle f_x(x,y,z), f_y(x,y,z), f_z(x,y,z)\rangle$ in three dimensions is an important object. The symbol ∇ is spelled "Nabla" and named after an Egyptian harp. The following theorem is important because it provides a crucial link between calculus and geometry.

Gradient theorem: Gradients are orthogonal to level curves and level surfaces.

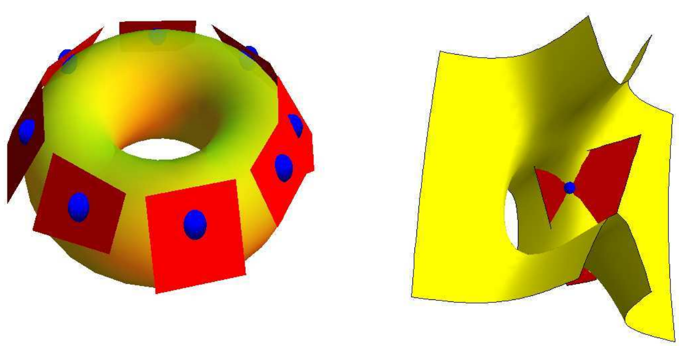
Proof. Every curve $\vec{r}(t)$ on the level curve or level surface satisfies $\frac{d}{dt} f(\vec{r}(t)) = 0$. By the chain rule, $\nabla f(\vec{r}(t))$ is perpendicular to the tangent vector $\vec{r}'(t)$. Because this is true for every curve, the gradient is perpendicular to the surface.

The gradient theorem is useful for example because it allows get tangent planes and tangent lines faster:

The tangent plane through (x_0, y_0, z_0) to a level surface of $f(x,y,z)$ is $ax+by+cz = d$, where $\nabla f(x_0,y_0,z_0) = \langle a,b,c\rangle$ and d is obtained by plugging in the point.

The statement in two dimensions is completely analog.

1 Find the tangent plane to the surface $3x^2y + z^2 - 4 = 0$ at the point $(1,1,1)$. **Solution:** $\nabla f(x,y,z) = \langle 6xy, 3x^2, 2z\rangle$. And $\nabla f(1,1,1) = \langle 6,3,2\rangle$. 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 $6x+3y+2z = 11$.



2 **Problem:** reflect the ray $\vec{r}(t) = \langle 1-t, -t, 1\rangle$ 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 $\vec{v} = \langle -1,-1,0\rangle$ The normal vector at this point is $\nabla f(-1,-2,1) = \langle -4,4,6\rangle = \vec{n}$. The reflected vector is

$$R(\vec{v} = \vec{v} - 2\mathrm{Proj}_{\vec{n}}(\vec{v})$$

We have $\mathrm{Proj}_{\vec{n}}(\vec{v}) = 8/68\langle -4,-4,6\rangle$. Therefore, the reflected ray is $\vec{w} = \langle -1,-1,0\rangle - (4/17)\langle -4,-4,6\rangle - \langle -1,-1,0\rangle$.

Lecture 16: Tangent spaces

1 Lets compute the tangent line at $(\pi, 0)$ to the curve $y = \sin(x)$ directly by determining the slope and making sure the line goes through the point.

2 Look at $f(x, y) = y - \sin(x) = 0$. Find the gradient $\nabla f(\pi, 0) = \langle a, b \rangle$ of f at $(\pi, 0)$. Now find the tangent line again.

3 Find the tangent plane to the surface $x^2 - y^2 + z^2 = -1$ at the point $(2, 3, 2)$.

4 Find a line perpendicular to the surface $x^2 - y^2 + z^2 = -1$ at the point $(2, 3, 2)$

17: Directional Derivative

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 unit vectors are directions. Because of the chain rule $\frac{d}{dt} D_{\vec{v}} f = \frac{d}{dt} f(x + t\vec{v})$, the directional derivative tells us how the function changes when we move in a given direction.
Assume for example that $f(x, y, z)$ is the temperature at position (x, y, z). If we move with velocity $\vec{v}$ through space, then $D_{\vec{v}} f$ tells us at which rate the temperature changes for us. If we move with velocity $\vec{v}$ on a hilly surface of height $f(x, y)$, then $D_{\vec{v}} f(x, y)$ gives us the slope in the direction $\vec{v}$.

1 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 $\frac{d}{dt} f(\vec{r}(t))$.

2 For $\vec{v} = \langle 1, 0, 0 \rangle$, then $D_{\vec{v}} f = \nabla f \cdot v = f_x$. The directional derivative generalizes 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**.

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

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

The gradient points in the direction where f increases most.

At a point where the gradient is not zero, the direction $\vec{v} = \nabla f/|\nabla f|$ is the direction, where f **increases** most. It is the direction of **steepest ascent**.

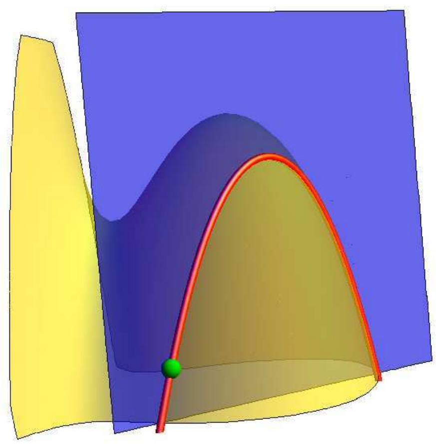
If $\vec{v} = \nabla f/|\nabla f|$, then the directional derivative is $\nabla f \cdot \nabla f/|\nabla f| = |\nabla f|$. This means f **increases**, if we move into the direction of the gradient. The slope in that direction is $|\nabla f|$.

3 You are in an airship at $(1, 2)$ and want to avoid a thunderstorm, a region of low pressure, where pressure is $p(x, y) = x^2 + 2y^2$. In which direction do you have to fly so that the pressure decreases fastest? **Solution:** the pressure gradient is $\nabla p(x, y) = \langle 2x, 4y \rangle$. At the point $(1, 2)$ this is $\langle 2, 8 \rangle$. Normalize to get the direction $\vec{v} = \langle 1, 4 \rangle/\sqrt{17}$. If you want to head into the direction where pressure is lower, go towards $-\vec{v}$.

Directional derivatives satisfy the same properties then any derivative: $D_v(\lambda f) = \lambda D_v(f)$, $D_v(f+g) = D_v(f) + D_v(g)$ and $D_v(fg) = D_v(f)g + fD_v(g)$.

We will see later that points with $\nabla f = \vec{0}$ are candidates for **local maxima** or **minima** of f. Points (x, y), where $\nabla f(x, y) = \langle 0, 0\rangle$ are called **critical points** and help to understand the function f.

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



If you should be interested in higher derivatives. We have seen that we can compute f_{xx}. This can be seen as the second directional derivative in the direction $(1, 0)$.

5 The Matterhorn is a famous mountain in the Swiss alps. Its height is 4'478 meters (14'869 feet). Assume in suitable units on the ground, the height $f(x, y)$ of the Matterhorn is approximated by $f(x, y) = 4000 - x^2 - y^2$. At height $f(-10, 10) = 3800$, at the point $(-10, 10, 3800)$, you rest. The climbing route continues into the south-east direction $\vec{v} = (1, -1)/\sqrt{2}$. Calculate the rate of change in that direction.
We have $\nabla f(x, y) = \langle -2x, -2y\rangle$, so that $(20, -20) \cdot (1, -1)/\sqrt{2} = 40/\sqrt{2}$. This is a place, where you climb $40/\sqrt{2}$ meters up when advancing 1 meter forward.

We can also look at higher derivatives in a direction. It can be used to measure the concavity of the function in the $\vec{v}$ direction.

The second directional derivative in the direction v is $D_{\vec{v}} D_{\vec{v}} f(x, y)$.

6 For the function $f(x, y) = x^2 + y^2$ the first directional derivative at a point in the direction $\langle 1, 2\rangle/\sqrt{5}$ is $\langle 2x, 2y\rangle \cdot \langle 1, 2\rangle = (2x + 4y)/\sqrt{5}$. The second directional derivative in the same direction is $\langle 2, 4\rangle \cdot \langle 1, 2\rangle/5 = 6/5$. This reflects the fact that the graph of f is concave up in the direction $\langle 1, 2\rangle/5$.

18: Extrema

An important problem in multi-variable calculus is to **extremize** a function $f(x,y)$ of two variables. As in one dimensions, in order to look for maxima or minima, we consider points, where the "derivative" is zero.

A point (a,b) is called a **critical point** of $f(x,y)$ if $\nabla f(a,b) = \langle 0,0 \rangle$.

Critical points are candidates for extrema because at critical points, all directional derivatives $D_{\vec{v}}f = \nabla f \cdot \vec{v}$ are zero. We can not increase the value of f by moving into any direction.

[1]

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

2 $f(x,y) = \sin(x^2 + y) + y$. The gradient is $\nabla f(x,y) = \langle 2x\cos(x^2+y), \cos(x^2+y)+1 \rangle$. For a critical points, we must have $x = 0$ and $\cos(y) + 1 = 0$ which means $\pi + k2\pi$. The critical points are at ...$(0,-\pi), (0,\pi), (0,3\pi)$, ...

3 The graph of $f(x,y) = (x^2+y^2)e^{-x^2-y^2}$ looks like a volcano. The gradient $\nabla f = \langle 2x - 2x(x^2+y^2), 2y - 2y(x^2+y^2) \rangle e^{-x^2-y^2}$ vanishes at $(0,0)$ and on the circle $x^2 + y^2 = 1$. This function has infinitely many critical points.

4 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 $\nabla f = (y, -g\sin(x)) = \langle (0,0 \rangle$ for $(x,y) = \dots, (-\pi,0), (0,0), (\pi,0), (2\pi,0), \dots$. These points are angles for which the pendulum is at rest.

5 The function $f(x,y) = |x| + |y|$ is differentiable on the first quadrant. It does not have critical points there. The function has a minimum at $(0,0)$ but it is not in the domain, where the gradient ∇f is defined.

In one dimension, we needed $f'(x) = 0, f''(x) > 0$ to have a local minimum, $f'(x) = 0, f''(x) < 0$ for a local maximum. If $f'(x) = 0, f''(x) = 0$, then the critical point was 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 $f(x) = x^3$.

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

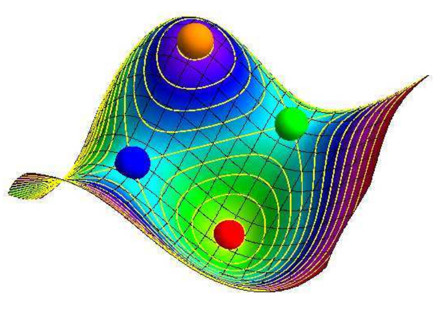
[1] This definition does not include points, where f or its derivative is not defined. We usually assume that functions are nice.

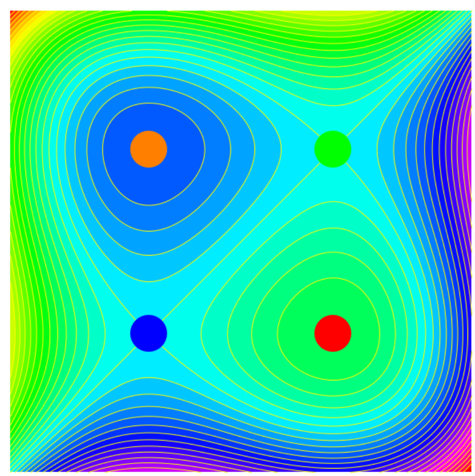
Remark: it can be remembered better if knowing that it is the determinant of the **Hessian matrix** $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$.

Second derivative test. 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.

In the case $D = 0$, we need higher derivatives to determine the nature of the critical point.

6 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 $\nabla f(x, y) = \langle x^2 - 1, -y^2 + 1\rangle$. 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.





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 next time how to look for extrema on the boundary.

7 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.

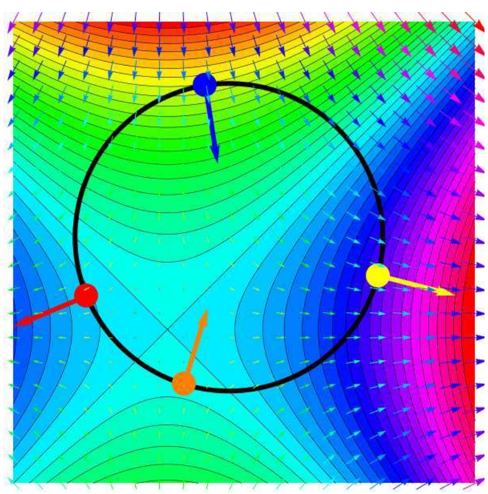
19: Lagrange multipliers

We aim to find maxima and minima of a function $f(x,y)$ in the presence of a **constraint** $g(x,y) = c$. A necessary condition for a critical point is that the gradients of f and g are parallel. The reason is that otherwise moving on the level curve $g = c$ will increase or decrease f: the directional derivative of f in the direction tangent to the level curve $g = c$ is zero if and only if the tangent vector to g is perpendicular to the gradient of f or if there is no tangent vector.

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

Lagrange theorem: Extrema of $f(x,y)$ on the curve $g(x,y) = c$ are either solutions of the Lagrange equations or critical points of g.

Proof. The condition that ∇f is parallel to ∇g either means $\nabla f = \lambda \nabla g$ or $\nabla g = 0$.



1 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.

2 Find the shortest distance from the origin $(0,0)$ to the curve $x^6 + 3y^2 = 1$. **Solution:** Minimize $f(x,y) = x^2 + y^2$ under the constraint $g(x,y) = x^6 + 3y^2 = 1$. The gradients are $\nabla f = \langle 2x, 2y \rangle, \nabla g = \langle 6x^5, 6y \rangle$. 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. We see that $(0, \pm\sqrt{1/3})$ are the minima.

3 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 $h = 0.54.., r = 2h = 1.08$.

4 On the curve $g(x,y) = x^2 - y^3$ 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 $\nabla g(x,y) = 0$.

Remarks.
1) Either of the two properties equated 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) Either of the two properties equated in the Lagrange theorem are equivalent to "$\nabla g = \lambda \nabla f$ or f has a critical point".
4) Constrained optimization problems work also in higher dimensions. The proof is 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 $\nabla g = \vec{0}$.

5 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 $(0,0,1)$.

6 A dice shows k eyes with probability p_k. Introduce the 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 $\vec{p}$ 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 $-1 - \log(p_i) = \lambda, p_1 + ... + p_6 = 1$, 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 fair dice has maximal entropy. Maximal entropy means **least information content**. An unfair dice provides additional information and allows a cheating gambler or casino to gain profit.

7 Assume that the probability that a physical or chemical system is in a state k is p_k and that the energy of the state k is E_k. Nature minimizes 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 $p_k = \exp(-E_k)/\sum_i \exp(-E_i)$. [1]

[1] This example appears in a book of Rufus Bowen, Lecture Notes in Math, 470, 1978

20: Global Extrema

To determine the maximum or minimum of $f(x,y)$ on a a region, we determine all critical points **in the interior the domain**, and compare their values with maxima or minima **at the boundary**. We have to solve both extrema problems with constraints and without constraints.

A point (a,b) 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).

1 Does the function $f(x,y) = x^4 + y^4 - 2x^2 - 2y^2$ have a global maximum or a global minimum? If yes, find them. **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 $x^2 = 1, y^2 = 1$.

2 Find the maximum of $f(x,y) = 2x^2 - x^3 - y^2$ on $y \geq -1$. **Solution.** 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.

3 Find all extrema of the function $f(x,y) = x^3 + y^3 - 3x - 12y + 20$ on the plane and characterize them. Do you find a global maximum or global minimum among them? **Solution.** The critical points satisfy $\nabla f(x,y) = \langle 0,0 \rangle$ or $\langle 3x^2 - 3, 3y^2 - 12 \rangle = \langle 0,0 \rangle$. There are 4 critical points $(x,y) = (\pm 1, \pm 2)$. The discriminant is $D = f_{xx}f_{yy} - f_{xy}^2 = 36xy$ and $f_{xx} = 6x$.

point	D	f_{xx}	classification	value
(-1,-2)	72	-6	maximum	38
(-1, 2)	-72	-6	saddle	6
(1, -2)	-72	6	saddle	34
(1, 2)	72	6	minimum	2

There are no global maxima nor global minima because the function takes arbitrarily large and small values. For $y = 0$ the function is $g(x) = f(x,0) = x^3 - 3x + 20$ which satisfies $\lim_{x \to \pm\infty} g(x) = \pm\infty$.

You can ignore the following questions and answers.
1. **Do global extrema always exist?** Yes, if the region Y is **compact** meaning that for every sequence x_n, y_n we can pick a subsequence which converges in Y. This is equivalent that the domain is **closed and bounded**.

Bolzano's extremal value theorem. Every continuous function on a compact domain has a global maximum and a global minimum.

2. **Why are critical points important?** Critical points are relevant in physics because they represent configurations with lowest energy. Many physical laws describe extrema. The Newton equations $mr''(t)/2 - \nabla V(r(t)) = 0$ describing a particle of mass m moving in a field V along a path $\gamma : t \mapsto \vec{r}(t)$ are equivalent to the property that the path extremizes the are length $S(\gamma) = \int_a^b mr'(t)^2/2 - V(r(t))\, dt$ among all paths.

3. **Why is the second derivative test true?** Assume $f(x, y)$ has the critical point $(0, 0)$ and is a quadratic function satisfying $f(0, 0) = 0$. Then $ax^2 + 2bxy + cy^2 = a(x + \frac{b}{a}y)^2 + (c - \frac{b^2}{a})y^2 = a(A^2 + DB^2)$ with $A = (x + \frac{b}{a}y), B = b^2/a^2$ and discriminant D. You see that if $a = f_{xx} > 0$ and $D > 0$ then $c - b^2/a > 0$ and the function has positive values for all $(x, y) \neq (0, 0)$. The point $(0, 0)$ is a minimum. If $a = f_{xx} < 0$ and $D > 0$, then $c - b^2/a < 0$ and the function has negative values for all $(x, y) \neq (0, 0)$ and the point (x, y) is a local maximum. If $D < 0$, then f takes both negative and positive values near $(0, 0)$. For a general function approximate by a quadratic one.

4. **Is there something cool about critical points?** Yes, assume $f(x, y)$ be the height of an island. Assume there are only finitely many critical points and all of them have nonzero determinant. Label each critical point with a $+1$ if it is a maximum or minimum, and with -1 if it is a saddle point. The sum of all these numbers is 1, independent of the island. [1]

5) **Can we avoid Lagrange by substitution?** To extremize $f(x, y)$ under the constraint $g(x, y) = 0$ we find $y = y(x)$ from the second equation and extremize the single variable problem $f(x, y(x))$. To extremize $f(x, y) = y$ on $x^2 + y^2 = 1$ for example we need to extremize $\sqrt{1 - x^2}$. 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.

6) **Is there a second derivative test for Lagrange?** A second derivative test can be designed using second directional derivative in the direction of the tangent. Instead, we just make a list of critical points and pick the maximum and minimum.

7) **Does Lagrange also work with more constraints?** With two constraints the constraint $g = c, h = d$ defines a curve. 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. Here is a formulation in three dimensions. 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 $\nabla g = \nabla h = 0$.

8) **Why do D and f_{xx} appear in the second derivative test**. They are natural. The discriminant D is a determinant $\det(H)$ of the matrix $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$. If $D > 0$ then the sign of f_{xx} is the same as the sign of the trace $f_{xx} + f_{yy}$ which is coordinate independent too. The determinant is the product $\lambda_1 \lambda_2$ of the eigenvalues of H and the trace is the sum of the eigenvalues.

9) **What does D mean?** The discriminant D is defined also at points where we have no critical point. The number $K = D/(1 + |\nabla f|^2)^2$ is called the **Gaussian curvature** of the surface. At critical points $K = D$. Curvature is remarkable quantity since it only depends on the intrinsic geometry of the surface and not on the way how the surface is embedded in space. [2]

10) **Is there a 2. derivative test in higher dimensions?** Yes. one can form the second derivative matrix H and look at all the eigenvalues of H. If all eigenvalues are negative, we have a local maximum, if all eigenvalues are positive, we have a local minimum. In general eigenvalues have different signs and we have a saddle point type.

---

[1] This follows from the **Poincare-Hopf** theorem.

[2] This is the **Theorema Egregia of Gauss**.

21: Double integrals

The integral $\int\int_R f(x,y)\ dxdy$ is defined as the limit of the Riemann sum

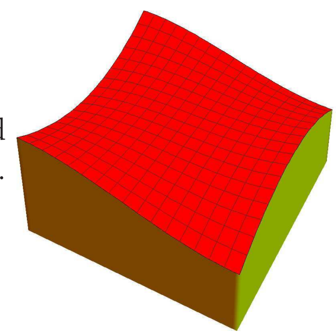
$$\frac{1}{n^2}\sum_{(\frac{i}{n},\frac{j}{n})\in R} f(\frac{i}{n},\frac{j}{n})$$

when $n \to \infty$.

1 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 we can reduce double integrals to single variable integrals.

2 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 $(i/n, j/n)$ inside R.

3 The integral $\int\int_R f(x,y)\ dxdy$ as the **signed volume** of the solid below the graph of f and above the region R in the $x-y$ plane. The volume below the xy-plane is counted negatively.



Fubini's theorem allows to switch the order of integration over a rectangle, if the function f is continuous: $\int_a^b \int_c^d f(x,y)\ dxdy = \int_c^d \int_a^b f(x,y)\ dydx$.

Proof. For every n the "quantum Fubini identity"

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

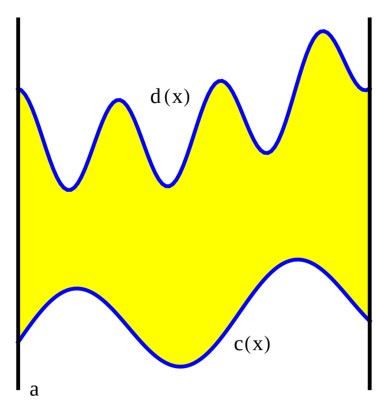
holds for all functions. Now divide both sides by n^2 and take the limit $n \to \infty$.

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

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

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

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

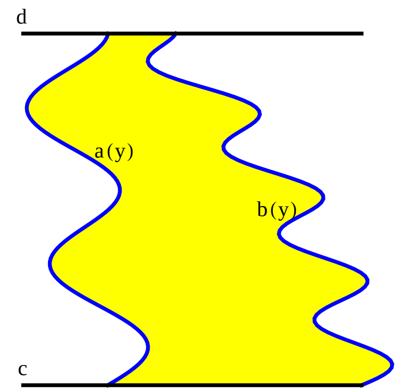


A **type II region** is of the form

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

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

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



4 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 $0 \le x \le \pi$. The value of this integral has a physical meaning. It is called **moment of inertia**.

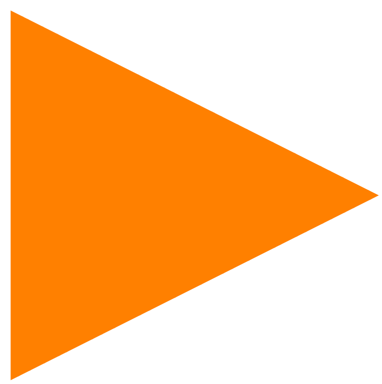
$$\int_0^{\pi^{1/3}} \int_{-\sin(x^3)}^{\sin(x^3)} x^2\ dydx = 2\int_0^{\pi^{1/3}} \sin(x^3)x^2\ dx$$

This can be solved by substitution

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

5 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. because we would have to compute 2 different integrals as a type I integral. The y bounds are $x = y-1$ and $x = 1-y$

$$\int_0^1 \int_{y-1}^{1-y} y^3\ dx\ dy = 2\int_0^1 y^3(1-y)\ dy = 2(\frac{1}{4} - \frac{1}{3}) = \frac{1}{10} \ .$$



6 Let R be the triangle $1 \ge x \ge 0, 0 \le y \le x$. What is

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

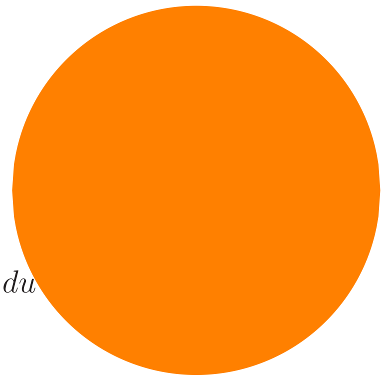
The type II integral $\int_0^1 [\int_y^1 e^{-x^2}\ dx] dy$ can not be solved because e^{-x^2} has no anti-derivative in terms of elementary functions. The type I integral $\int_0^1 [\int_0^x e^{-x^2}\ dy]\ dx$ however can be solved:

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

Lecture 22: Polar integration

1 The area of a disc of radius R is

$$\int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} 1 \; dydx = \int_{-R}^{R} 2\sqrt{R^2-x^2} \; dx \; .$$



This integral can be solved with the substitution $x = R\sin(u), dx = R\cos(u)$

$$\int_{-\pi/2}^{\pi/2} 2\sqrt{R^2 - R^2\sin^2(u)} R\cos(u) \; du = \int_{-\pi/2}^{\pi/2} 2R^2\cos^2(u) \; du$$

Using a double angle formula we get $R^2 \int_{-\pi/2}^{\pi/2} 2\frac{(1+\cos(2u)}{2} \; du = R^2\pi$. We will now see how to do that better in polar coordinates.

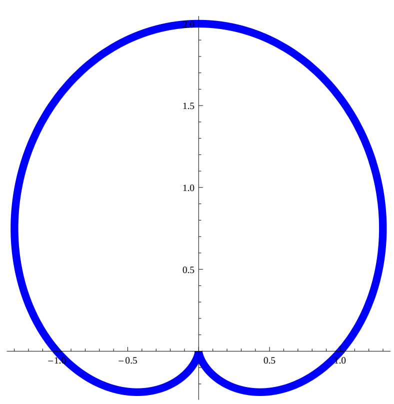
A **polar region** is a region bound by a simple closed curve given in polar coordinates as the curve $(r(t), \theta(t))$.

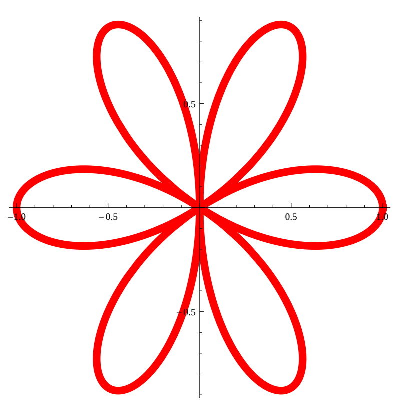
In Cartesian coordinates the parametrization of the boundary curve is $\vec{r}(t) = \langle r(t)\cos(\theta(t), r(t)\sin(\theta(t)) \rangle$. We are especially interested in regions which are bound by **polar graphs**, where $\theta(t) = t$.

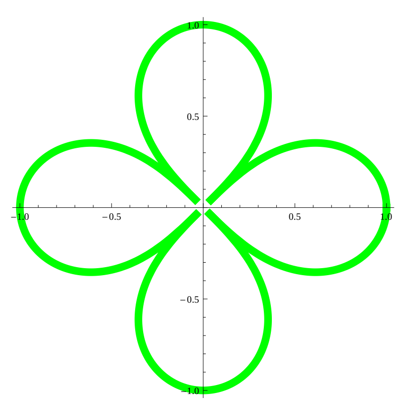
2 The **polar region** defined by $r \leq |\cos(3\theta)|$ belongs to the class of **roses** $r(t) = |\cos(nt)|$ they are also called **rhododenea**. These names reflect that polar regions model flowers well.

3 The polar curve $r(\theta) = 1 + \sin(\theta)$ is called a **cardioid**. It looks like a heart. It is a special case of a **limacon** a polar curve of the form $r(\theta) = 1 + b\sin(\theta)$.

4 The polar curve $r(\theta) = |\sqrt{\cos(2t)}|$ is called a **lemniscate**. It looks like an infinity sign. It encloses a flower with two petals.







To integrate in polar coordinates, we evaluate the integral

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

5 Integrate

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

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 $\iint_R f(x,y)\, dxdy = \int_0^1 \int_0^{2\pi} (r^2 + r^2\cos(\theta)\sin(\theta)) r\, d\theta dr = 2\pi/4$.

6 We have earlier computed area of the disc $\{x^2 + y^2 \leq R^2\ \}$ using substitution. It is more elegant to do this integral in polar coordinates: $\int_0^{2\pi} \int_0^R r\, drd\theta = 2\pi r^2/2|_0^R = \pi R^2$.

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 $r\, d\theta dr$.

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

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

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

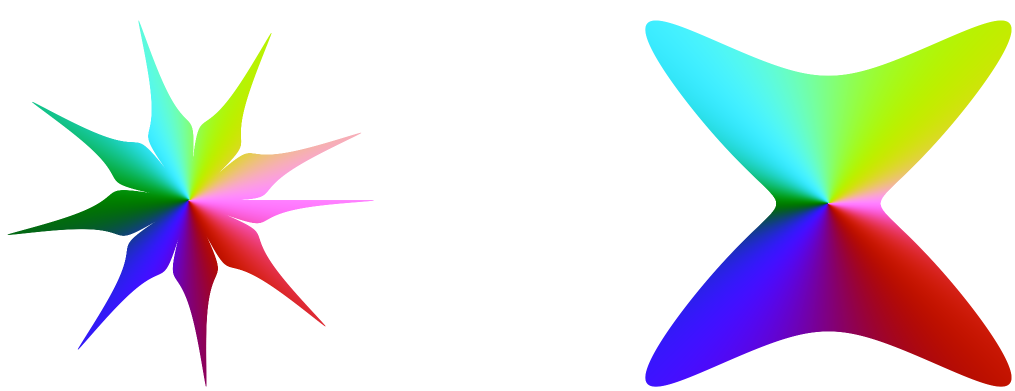
$$\int_1^2 \int_0^\pi r\sin(\theta) r\, r\, d\theta dr = \int_1^2 r^3 \int_0^\pi \sin(\theta)\, d\theta dr = \frac{(2^4 - 1^4)}{4} \int_0^\pi \sin(\theta)\, d\theta = 15/2$$

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

9 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 **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]



[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).

Lecture 22: Surface area

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

1 The parametrized surface $\vec{r}(u, v) = \langle 2u, 3v, 0\rangle$ 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. $\vec{r}_u \times \vec{r}_v = \langle 0, 0, 6\rangle$ and we see for example that the area of $\vec{r}(G)$ is 6 times the area of G.

For a planar region $\vec{r}(s, t) = P + sv + tw$ where $(s, t) \in G$, the surface area is the area of G times $|v \times w|$.

2 The map $\vec{r}(u, v) = \langle L\cos(u)\sin(v), L\sin(u)\sin(v), L\cos(v)\rangle$ 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 $\int\int_R 1 \, dS = \int_0^{2\pi}\int_0^{\pi} L^2 \sin(v) \, dvdu = 4\pi L^2$

For a sphere of radius L, we have $|\vec{r}_u \times \vec{r}_v| = L^2 \sin(v)$ The surface area is $4\pi L^2$.

3 For graphs $(u, v) \mapsto \langle u, v, f(u, v)\rangle$, 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 $\int\int_G \sqrt{1 + f_u^2 + f_v^2} \, dudv$.

For a graph $z = f(x, y)$ parametrized over G, the surface area is

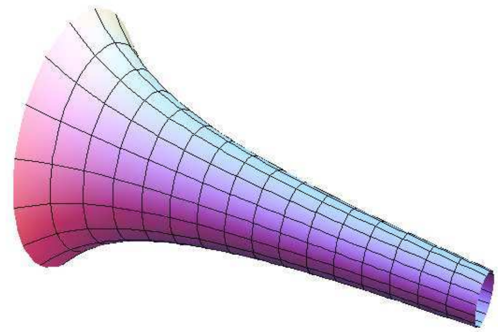
$$\int\int_G \sqrt{1 + f_x^2 + f_y^2} \, dxdy \, .$$

4 Lets take a surface of revolution $\vec{r}(u, v) = \langle v, f(v)\cos(u), f(v)\sin(u)\rangle$ 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 $\int\int |\vec{r}_u \times \vec{r}_v| \, dudv = 2\pi \int_a^b |f(v)|\sqrt{1 + f'(v)^2} \, dv$.

For a surface of revolution $r = f(z)$ with $a \leq z \leq b$, the surface area is

$$2\pi \int_a^b |f(z)|\sqrt{1 + f'(z)^2}\, dz \; .$$

5 Gabriel's trumpet is the surface of revolution where $g(z) = 1/z$, where $1 \leq z < \infty$. Its volume is $\int_1^\infty \pi g(z)^2 \; dz = \pi$. We will compute in class the surface area.



6 Find the surface area of the part of the paraboloid $x = y^2 + z^2$ which is inside the cylinder $y^2 + z^2 \leq 9$. **Solution.** We use polar coordinates in the yz-plane. The paraboloid is parametrized by $(u, v) \mapsto (v^2, v\cos(u), v\sin(u))$ and the surface integral $\int_0^3 \int_0^{2\pi} |\vec{r}_u \times \vec{r}_v| \; dudv$ is equal to $\int_0^3 \int_0^{2\pi} v\sqrt{1 + 4v^2} \; dudv = 2\pi \int_0^3 v\sqrt{1 + 4v^2} \; dv = \pi(37^{3/2} - 1)/6$.

7 In this example we derive the distortion factor r in polar coordinates. To do so, we parametrize a region in the xy plane with $\vec{r}(u, v) = \langle u\cos(v), u\sin(v), 0\rangle$. Given a region G in the uv plane like the rectangle $[0, \pi] \times [1, 2]$, we obtain a region S in the xy plane as the image. The factor $|\vec{r}_u \times \vec{r}_v|$ is equal to the radius u. In our example, the surface area is $\int_0^\pi \int_1^2 u \; dudv = \pi(4 - 1) = 3\pi$. This is the area of the half annulus S. We could have used polar coordinates directly in the xy plane and compute $\int_0^\pi \int_1^2 r \; dr \; d\theta = 3\pi$. But the only thing which has changed are the names of the variables.

8 The surface parametrized by

$\vec{r}(u, v) = \langle (2{+}v\cos(u/2))\cos(u), (2{+}v\cos(u/2))\sin(u), v$

on $G = [0, 2\pi] \times [-1, 1]$ is called a **Möbius strip**. What is its surface area? **Solution.** The calculation of $|\vec{r}_u \times \vec{r}_v|^2 = 4 + 3v^2/4 + 4v\cos(u/2) + v^2\cos(u)/2$ is straightforward but a bit tedious. The integral over $[0, 2\pi] \times [-1, 1]$ can only be evaluated numerically, the result is 25.413....