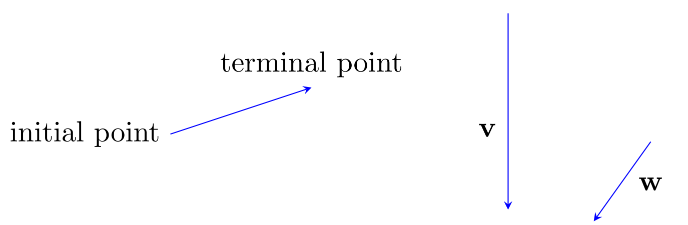
Intro to Calculus of Several Variables Notes

1 Vectors and operations on them

1.1 Vectors in the plane

Definition 1.1. A **vector in the plane** is a finite line segment in the plane whose endpoints are ordered. One endpoint is called the **initial point**, while the other is called the **terminal point**. That is, vectors are line segments with a preferred direction. When drawn, the terminal point is distinguished by an arrowhead.

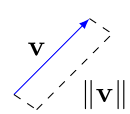
Example 1.2. Vectors are denoted with boldface lowercase letters.



□

Definition 1.3. The **magnitude** or **length** of a vector, $\mathbf{v}$, is the length of the line segment corresponding to $\mathbf{v}$, denoted by $\|\mathbf{v}\|$.

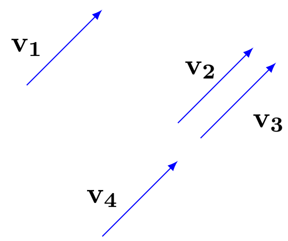
Example 1.4.



□

Definition 1.5. Two vectors are **equivalent** if they have the same magnitude and direction. If $\mathbf{v}$ and $\mathbf{w}$ are equivalent, we write $\mathbf{v} = \mathbf{w}$.

Example 1.6. Below, $\mathbf{v_1} = \mathbf{v_2} = \mathbf{v_3} = \mathbf{v_4}$.



□

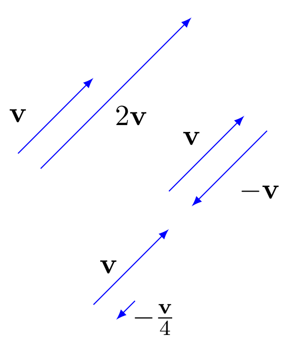
Definition 1.7. The **zero vector**, $\mathbf{0}$, is the vector whose initial and terminal points are the same.

Definition 1.8. If $k \in \mathbb{R}$ ($\in$ denotes membership), that is, if k is a real number, otherwise known as a **scalar**, and $\mathbf{v}$ is a vector in the plane, then we may define a new vector in the plane, $k\mathbf{v}$, as follows:

- If $k > 0$, $k\mathbf{v}$ is the unique vector (up to equivalence) with the same initial point as $\mathbf{v}$, but with magnitude $k\|\mathbf{v}\|$.
- If $k = -1$, denote $(-1)\mathbf{v}$ by $-\mathbf{v}$ and define $-\mathbf{v}$ as the unique vector whose initial point is the terminal point of $\mathbf{v}$, and vice versa.
- If $k < 0$, define $k\mathbf{v} := -k(-\mathbf{v})$.
- If $k = 0$, $k\mathbf{v} := \mathbf{0}$.

The operation sending k and $\mathbf{v}$ to $k\mathbf{v}$ is called **scalar multiplication**. The notation $A := B$ means we're defining object A to be object B. That is, we know what B is, and we're defining A to be that object.

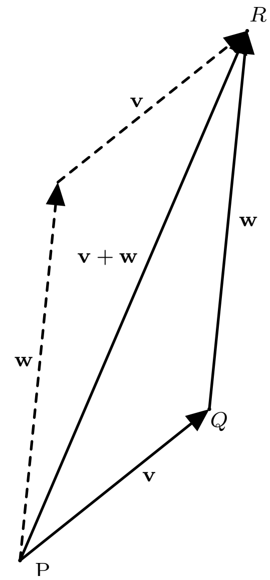
Example 1.9.



□

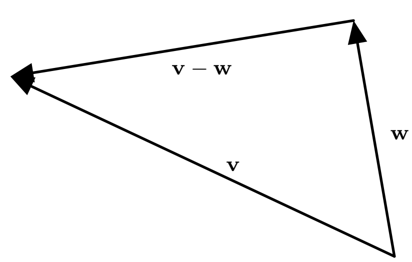
Remark 1.10. If P is the initial point of $\mathbf{v}$, Q its terminal point, we commonly write $\mathbf{v} = \mathbf{PQ}$.

Definition 1.11. Vector Addition. If $\mathbf{v}$ and $\mathbf{w}$ are vectors in the plane, then we may define the sum, $\mathbf{v} + \mathbf{w}$ as follows. If $\mathbf{v} = \mathbf{PQ}$ for some points P and Q, then we can write $\mathbf{w} = \mathbf{QR}$ for some point R. Then we define $\mathbf{v} + \mathbf{w} := \mathbf{PR}$. In words, if the initial point of $\mathbf{w}$ is placed at the terminal point of $\mathbf{v}$, then $\mathbf{v} + \mathbf{w}$ is the vector with $\mathbf{v}$'s initial point and $\mathbf{w}$'s terminal point.



This image also shows $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.

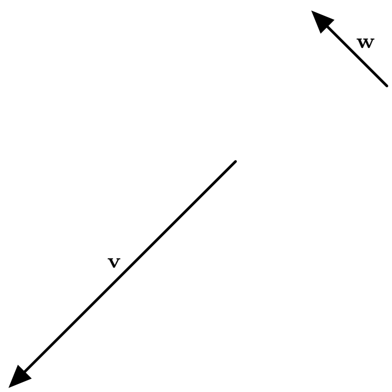
Definition 1.12. The **vector difference** of $\mathbf{v}$ and $\mathbf{w}$ is denoted as $\mathbf{v} - \mathbf{w}$ and defined to be the unique vector $\mathbf{z}$ such that $\mathbf{w} + \mathbf{z} = \mathbf{v}$:



That is, if the initial points of $\mathbf{v}$ and $\mathbf{w}$ are the same, then $\mathbf{v} - \mathbf{w}$ is the vector with initial point the terminal point of $\mathbf{w}$, and terminal point the terminal point of $\mathbf{v}$. Rather, if $\mathbf{v} = \mathbf{PQ}$ and $\mathbf{w} = \mathbf{PR}$, then $\mathbf{v} - \mathbf{w} := \mathbf{RQ}$.

Remark 1.13. We see that $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$.

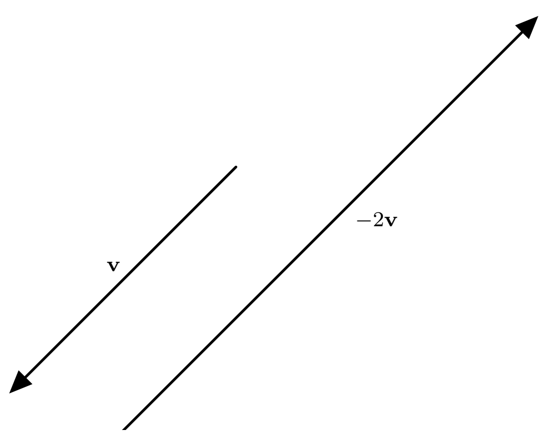
Example 1.14. Given



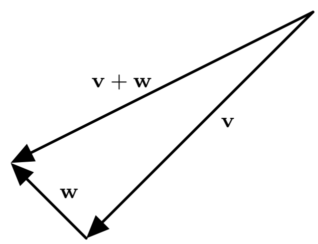
find

1. $-2\mathbf{v}$.
2. $\mathbf{v} + \mathbf{w}$.
3. $\mathbf{v} - 2\mathbf{w}$.

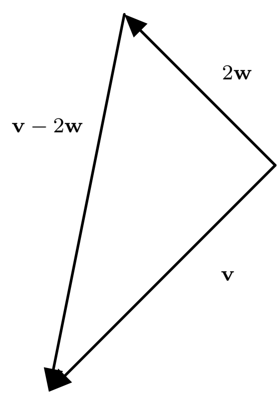
1.



2.



3.

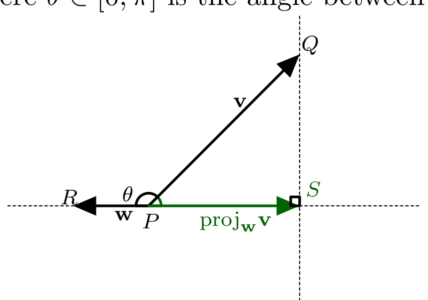


□

Suppose we have two vectors, $\mathbf{v} = \mathbf{PQ}$ and $\mathbf{w} = \mathbf{PR}$, whose initial points are the same. Then drop a perpendicular from Q to the line which contains the line segment corresponding to $\mathbf{w}$, intersecting at say S. Then define $\operatorname{proj}_{\mathbf{w}} \mathbf{v} := \mathbf{PS}$ ($\operatorname{proj}_{\mathbf{w}} \mathbf{v}$ is read: the projection of $\mathbf{v}$ onto $\mathbf{w}$). Notice then that

$$\|\operatorname{proj}_{\mathbf{w}} \mathbf{v}\| = \|\mathbf{v}\| |\cos\theta|$$

where $\theta \in [0, \pi]$ is the angle between $\mathbf{v}$ and $\mathbf{w}$. See the figure below.



Definition 1.15. The **dot product** of two vectors in the plane, $\mathbf{v}$ and $\mathbf{w}$, is given by,

$$\mathbf{v} \cdot \mathbf{w} := \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta,$$

where $\theta \in [0, \pi]$ is the angle between $\mathbf{v}$ and $\mathbf{w}$ if their initial points are the same.

Proposition 1.16. *Properties of the dot product.*

$$(\mathbf{v} + \mathbf{z}) \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} + \mathbf{z} \cdot \mathbf{w}. \tag{1.1}$$

Also from a picture, we see that

$$(k\mathbf{v}) \cdot \mathbf{w} = k(\mathbf{v} \cdot \mathbf{w}). \tag{1.2}$$

And it's evident that

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}, \tag{1.3}$$

along with

$$\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2. \tag{1.4}$$

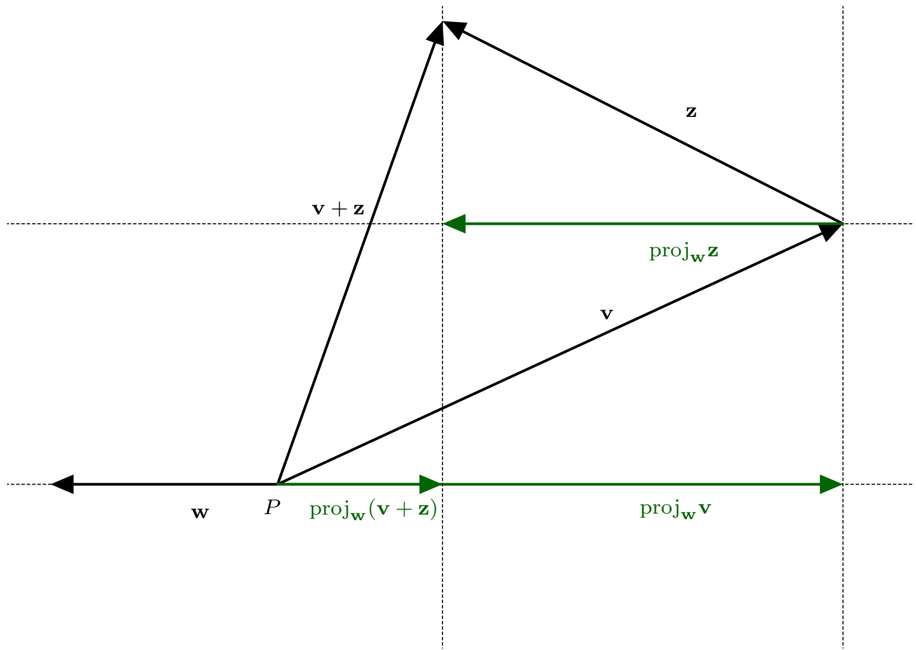
Also,

$$\mathbf{v} \cdot \mathbf{w} = 0 \tag{1.5}$$

if and only if $\mathbf{v}$ *and* $\mathbf{w}$ *are orthogonal (perpendicular).*

Proof. The below image shows an example of the fact that $\operatorname{proj}_{\mathbf{w}}(\mathbf{v} + \mathbf{z}) = \operatorname{proj}_{\mathbf{w}} \mathbf{v} + \operatorname{proj}_{\mathbf{w}} \mathbf{z}$ (notice the initial point of $\operatorname{proj}_{\mathbf{w}} \mathbf{v}$ is P). In this case, we have

$$\|\operatorname{proj}_{\mathbf{w}}(\mathbf{v} + \mathbf{z})\| = \|\operatorname{proj}_{\mathbf{w}} \mathbf{v}\| - \|\operatorname{proj}_{\mathbf{w}} \mathbf{z}\|. \tag{1.6}$$



We have

$$|\mathbf{v}\cdot\mathbf{w}| = \|\operatorname{proj}_{\mathbf{w}}\mathbf{v}\|\|\mathbf{w}\|$$

in general, with the sign depending on the angle between $\mathbf{v}$ and $\mathbf{w}$. In the above image, we have

$$\mathbf{v}\cdot\mathbf{w} = -\|\operatorname{proj}_{\mathbf{w}}\mathbf{v}\|\|\mathbf{w}\|,$$

$$\mathbf{z}\cdot\mathbf{w} = \|\operatorname{proj}_{\mathbf{w}}\mathbf{z}\|\|\mathbf{w}\|,$$

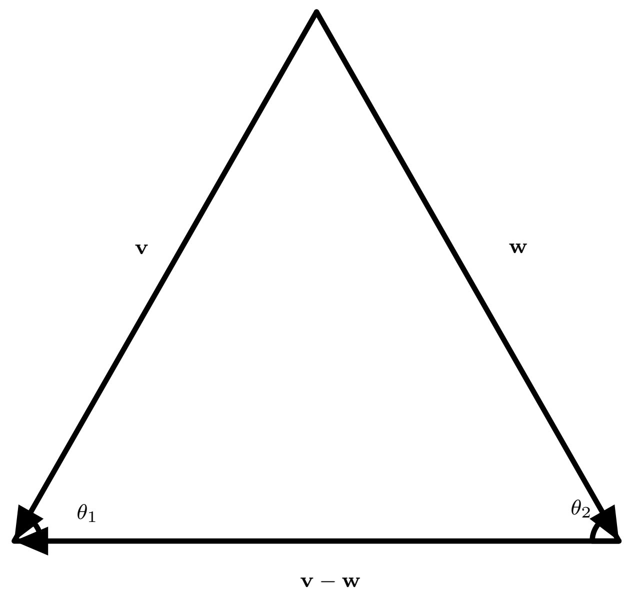
and

$$(\mathbf{v}+\mathbf{z})\cdot\mathbf{w} = -\|\operatorname{proj}_{\mathbf{w}}(\mathbf{v}+\mathbf{z})\|\|\mathbf{w}\|,$$

which, combined with (1.6), gives (1.1). □

Example 1.17. Use vectors to show the angles opposite the sides of same length in an isosceles triangle are the same.

Proof. Suppose $\mathbf{v}$ and $\mathbf{w}$ have the same length. That is, suppose $\|\mathbf{v}\| = \|\mathbf{w}\|$. Then $\mathbf{v}, \mathbf{w}$ and $\mathbf{v}-\mathbf{w}$ form an isosceles triangle. From the image:



we have

$$\mathbf{v} \cdot (\mathbf{v} - \mathbf{w}) = \|\mathbf{v}\| \|\mathbf{v} - \mathbf{w}\| \cos \theta_1,$$

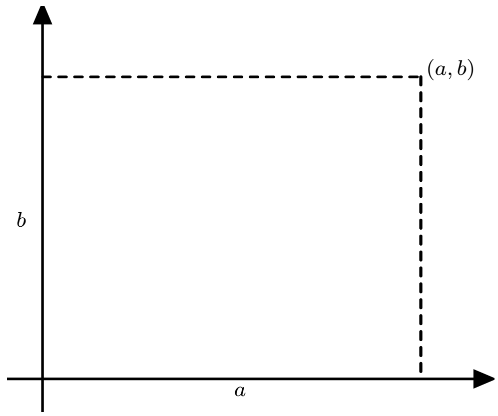
and

$$-\mathbf{w} \cdot (\mathbf{v} - \mathbf{w}) = \| - \mathbf{w}\| \|\mathbf{v} - \mathbf{w}\| \cos \theta_2.$$

And recall that cos is one-to-one in $[0, \pi]$.

□

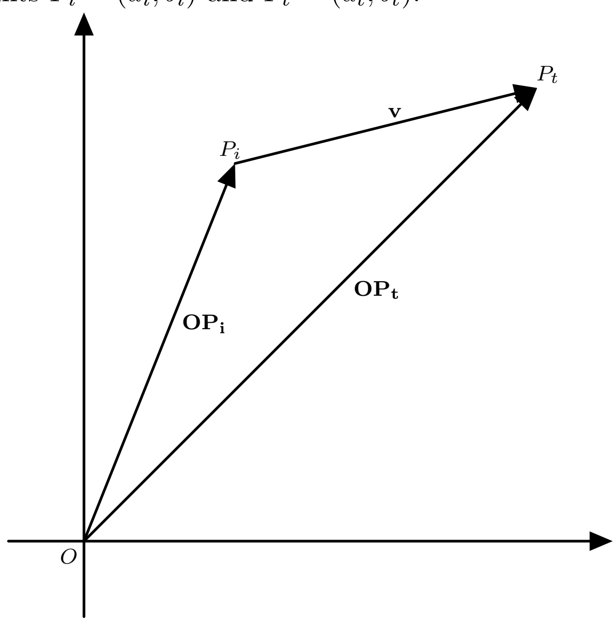
Definition 1.18. Now fix a pair of Cartesian (perpendicular, orthogonal) coordinate axes in the plane so we can label points relative to them.



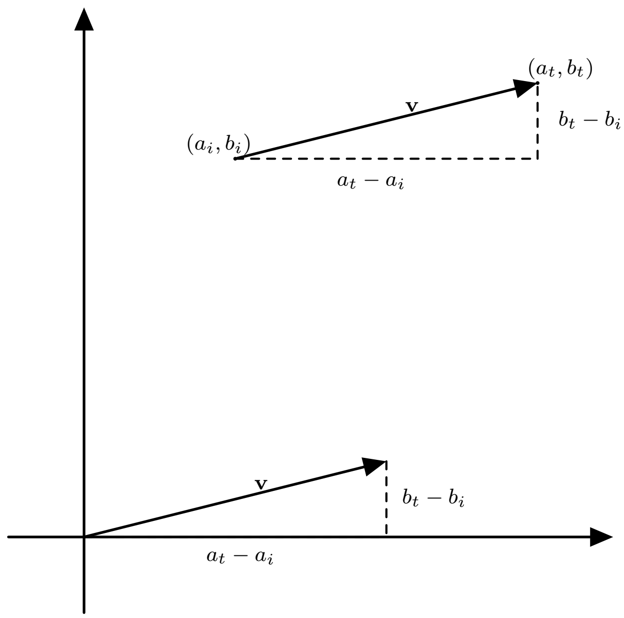
Denote by O the origin. Then, if $P = (a, b)$, denote

$$\langle a, b \rangle := \mathbf{OP}.$$

We say the scalars a, b are the **components** of $\mathbf{OP}$ and $\mathbf{OP}$ is the **position vector** of P. If $\mathbf{v} = \mathbf{OP}$ for some point P, we say $\mathbf{v}$ is in **standard position** relative to this fixed coordinate system. Now suppose $\mathbf{v} = \mathbf{P_i P_t}$ for some points $P_i = (a_i, b_i)$ and $P_t = (a_t, b_t)$.



We see that $\mathbf{v} = \mathbf{OP_t} - \mathbf{OP_i}$. But also $\mathbf{v} = \langle a_t - a_i, b_t - b_i \rangle$, since



That is,

$$\langle a_t, b_t \rangle - \langle a_i, b_i \rangle = \langle a_t - a_i, b_t - b_i \rangle . \tag{1.7}$$

The length of vector $\langle a, b \rangle$ is, by the Pythagorean Theorem,

$$\| \langle a, b \rangle \| = \sqrt{a^2 + b^2} \tag{1.8}$$

From (1.8) it follows that

$$k \langle a, b \rangle = \langle ka, kb \rangle \tag{1.9}$$

for any scalar k.

As a consequence of (1.7) and (1.9), for any two vectors in the plane in standard position, $\langle v_1, v_2 \rangle, \langle w_1, w_2 \rangle$, we have

$$\langle v_1, v_2 \rangle + \langle w_1, w_2 \rangle = \langle v_1 + w_1, v_2 + w_2 \rangle .$$

We gather these results in a proposition.

Proposition 1.19. *The coordinate description of vectors in the plane. For any scalars,* $a, b, k, v_1, v_2, w_1, w_2$,

$$\| \langle a, b \rangle \| = \sqrt{a^2 + b^2}$$

$$k \langle a, b \rangle = \langle ka, kb \rangle$$

$$\langle v_1, v_2 \rangle + \langle w_1, w_2 \rangle = \langle v_1 + w_1, v_2 + w_2 \rangle . \tag{1.10}$$

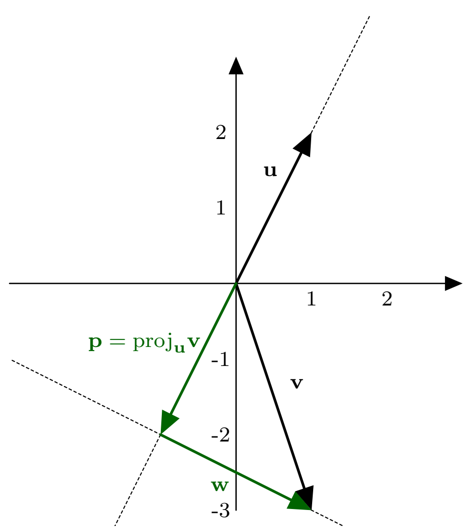
Definition 1.20. In physics, one might come across such notation as $\langle a, b \rangle = a\mathbf{i} + b\mathbf{j}$, where $\mathbf{i} := \langle 1, 0 \rangle$ and $\mathbf{j} := \langle 0, 1 \rangle$. $\mathbf{i}$ and $\mathbf{j}$ are called the **standard unit vectors**. A **unit vector**, $\mathbf{u}$, is any vector with magnitude 1: $\|\mathbf{u}\| = 1$.

Proposition 1.21. *For any scalars v_1, v_2, w_1, w_2,*

$$\langle v_1, v_2 \rangle \cdot \langle w_1, w_2 \rangle = v_1 w_1 + v_2 w_2. \tag{1.11}$$

Proof. This follows from (1.1) and (1.2) and the fact that $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = 1$ and $\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{i} = 0$, since we write, for example, $\langle v_1, v_2 \rangle = v_1 \mathbf{i} + v_2 \mathbf{j}$. □

Example 1.22. Given $\mathbf{u} = \langle 1, 2 \rangle$, $\mathbf{v} = \langle 1, -3 \rangle$, write $\mathbf{v} = \mathbf{p} + \mathbf{w}$, where $\mathbf{p} = \text{proj}_{\mathbf{u}} \mathbf{v}$ and $\mathbf{w} = \mathbf{v} - \mathbf{p}$.



Notice

$$\mathbf{p} = \operatorname{proj}_{\mathbf{u}} \mathbf{v} = -\|\operatorname{proj}_{\mathbf{u}} \mathbf{v}\| \frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{u}\|^2} \mathbf{u}$$

from the fact that

$$\mathbf{v} \cdot \mathbf{u} = -\|\operatorname{proj}_{\mathbf{u}} \mathbf{v}\| \|\mathbf{u}\|$$

in this case.

After some computation, using Proposition 1.21, we obtain $\mathbf{p} = -\mathbf{u}$ and so $\mathbf{w} = \langle 2, -1 \rangle$. □

Remark 1.23. We can also write $\mathbf{v} = \langle \|\mathbf{v}\| \cos\theta, \|\mathbf{v}\| \sin\theta \rangle = \langle \|\mathbf{v}\| \sin(\pi/2 - \theta), \|\mathbf{v}\| \cos(\pi/2 - \theta) \rangle$, where θ is the angle between $\mathbf{v}$ and the x-axis.

Example 1.24 (**OS** Exercise 2.54)**.** A boat is traveling in the water at 30 mph in a direction of N20E (20 degrees East of North). A current is moving at 15 mph in a direction of N45E. What are the new speed and direction of the boat?

Velocity of the boat's motor:

$$\mathbf{b} = 30 \sin(20^\circ)\mathbf{i} + 30 \cos(20^\circ)\mathbf{j}$$

Velocity of the current:

$$\mathbf{c} = 15 \sin(45^\circ)\mathbf{i} + 15 \cos(45^\circ)\mathbf{j}$$

New velocity of the boat:

$$\mathbf{n} = (30 \sin(20^\circ) + 15 \sin(45^\circ))\mathbf{i} + (30 \cos(20^\circ) + 15 \cos(45^\circ))\mathbf{j}$$

Which implies the new speed is $\|\mathbf{n}\|$ and new direction is $180 \arccos(\mathbf{n} \cdot \mathbf{j}/\|\mathbf{n}\|)/\pi$ degrees East of North. □

1.2 Vectors in three dimensions

The definitions and examples for vectors in the plane essentially carry over verbatim to vectors in 3-dimensional space, where we informally describe 3-dimensional space as a plane with an added direction. Formally, we make use of coordinates, where in 3-dimensional space we have 3 axes, all mutually perpendicular.

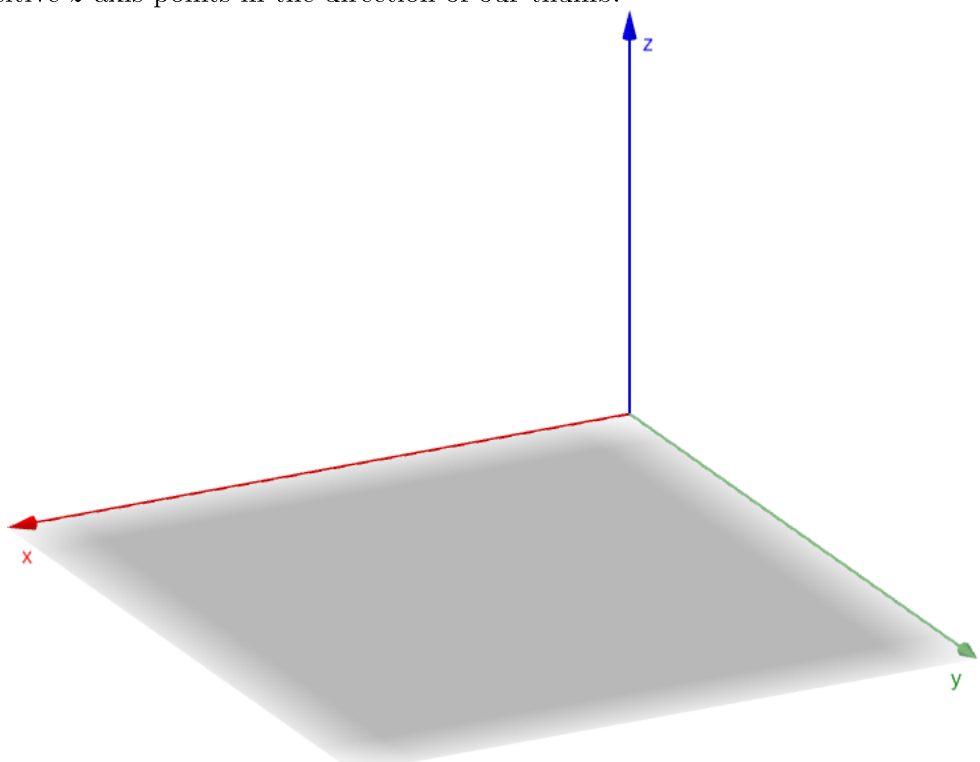
Definition 1.25.

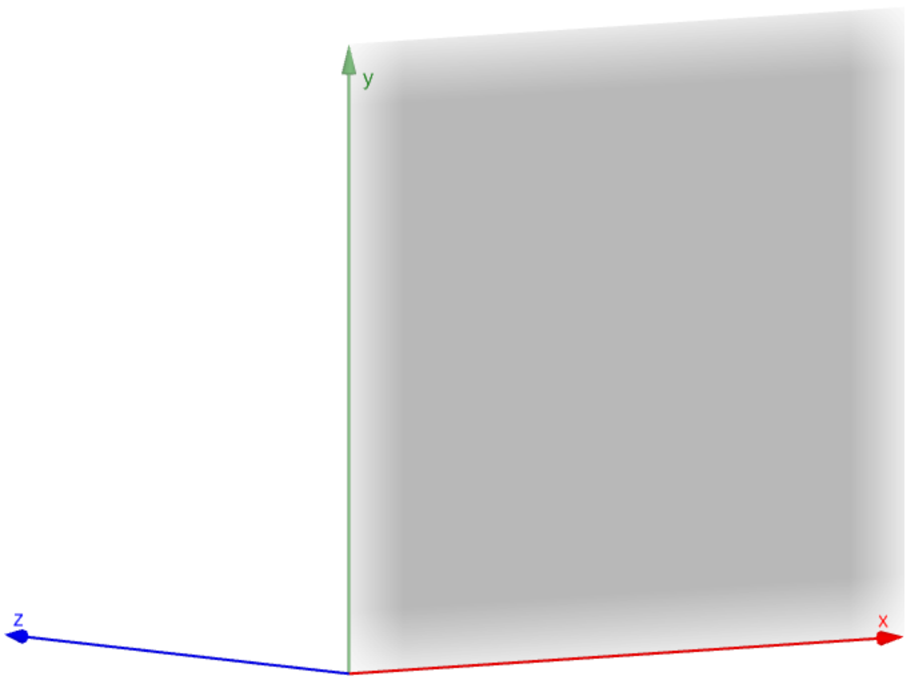
$$\mathbb{R}^2 := \{\langle a, b\rangle \mid a, b \in \mathbb{R}\} = \{\text{set of vectors in the plane}\}$$

$$\mathbb{R}^3 := \{\langle a, b, c\rangle \mid a, b, c \in \mathbb{R}\} = \{\text{set of vectors in 3-dimensional space}\}$$

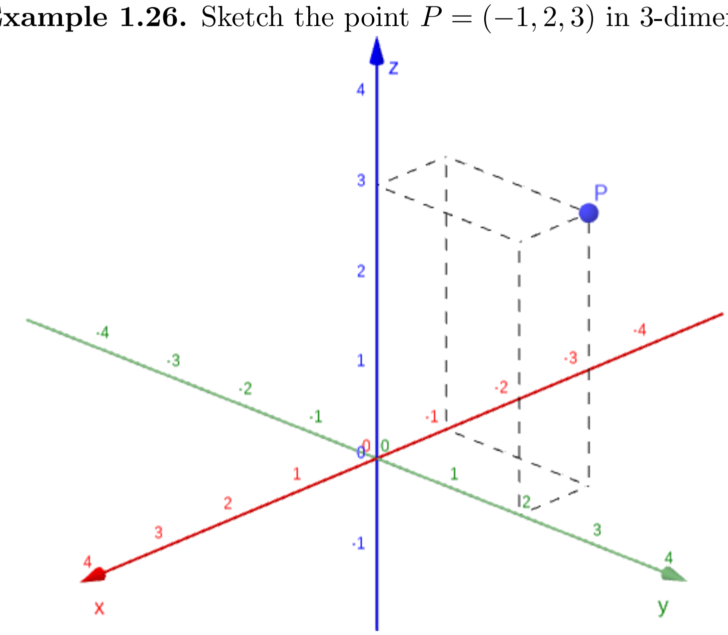
Note: sometimes we will describe $\mathbb{R}^2$ or $\mathbb{R}^3$ as a set of points, not vectors, but this is o.k., due to the one-to-one correspondence, e.g., $P = (a, b) \leftrightarrow \mathbf{OP} = \langle a, b\rangle$ (once we fix an origin). Also, $\mathbb{R}^2$ is usually identified as a subset of $\mathbb{R}^3$ as $\{\langle a, b, 0\rangle \mid a, b \in \mathbb{R}\}$, called the xy-plane. More on this later.

To draw a set of coordinate axes for $\mathbb{R}^3$, we follow the **right-hand rule**: If we take our right hand and point our fingers in the direction of the positive x-axis, and curl our fingers in the direction of the positive y-axis, then the positive z-axis points in the direction of our thumb:





Example 1.26. Sketch the point $P = (-1, 2, 3)$ in 3-dimensional space.

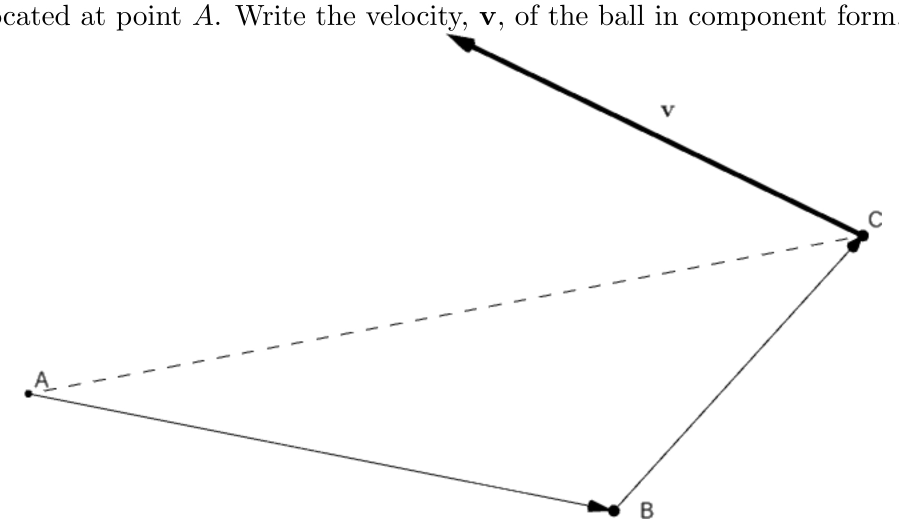


□

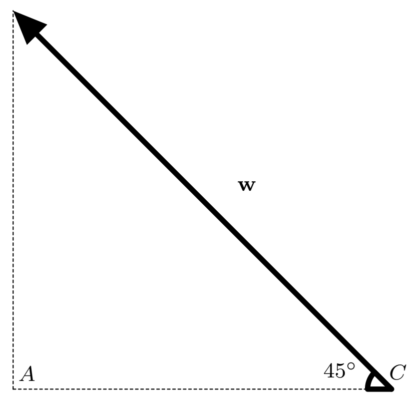
Remark 1.27. (a) The magnitude of $\langle a, b, c \rangle \in \mathbb{R}^3$ is given by $\| \langle a, b, c \rangle \| = \sqrt{a^2 + b^2 + c^2}$, seen from the distance formula in $\mathbb{R}^3$. See the problem set or [**OS** Theorem 2.4].

(b) Sometimes we write $\langle a, b, c \rangle = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, where $\mathbf{i} = \langle 1, 0, 0 \rangle, \mathbf{j} = \langle 0, 1, 0 \rangle, \mathbf{k} = 001$ and are called the **standard unit vectors**.

Example 1.28 (**OS** Exercise 2.120)**.** Two soccer players are practicing for a match. One of them runs 10 m from point A to point B. She then turns left at 90° and runs 10 m until she reaches point C. Then she kicks the ball with a speed of 10 m/s at an upward angle of 45° to her teammate, who is located at point A. Write the velocity, $\mathbf{v}$, of the ball in component form.



Consider the vector $\mathbf{w}$ in the same direction and with the same initial point as $\mathbf{v}$, but terminal point directly above A:



If $A = (0, 0, 0)$, $B = (10, 0, 0)$, $C = (10, 10, 0)$, then $\|\mathbf{AC}\| = 10\sqrt{2}$, and so, from the above image,

$$\|\mathbf{w}\| = \|\mathbf{AC}\| \sec 45^\circ = 20.$$

Also from the above image,

$$\mathbf{AC} + \mathbf{w} = \|\mathbf{w}\| \sin 45^\circ \mathbf{k} = 10\sqrt{2}\mathbf{k},$$

which implies

$$\mathbf{w} = -10\mathbf{i} - 10\mathbf{j} + 10\sqrt{2}\mathbf{k}.$$

Since $\mathbf{v}$ and $\mathbf{w}$ have the same direction, $\mathbf{v} = k\mathbf{w}$ for some $k > 0$. Hence,

$$10 = \|\mathbf{v}\| = k\|\mathbf{w}\| = 20k,$$

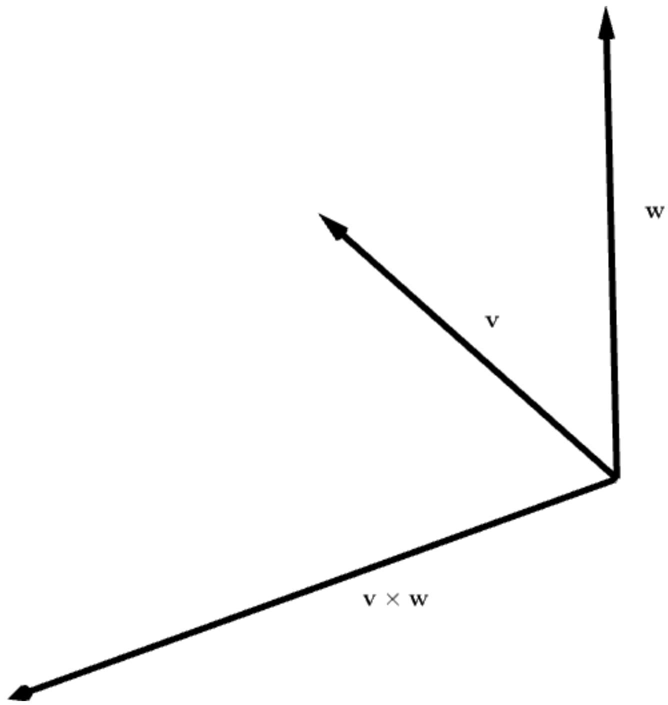
so that

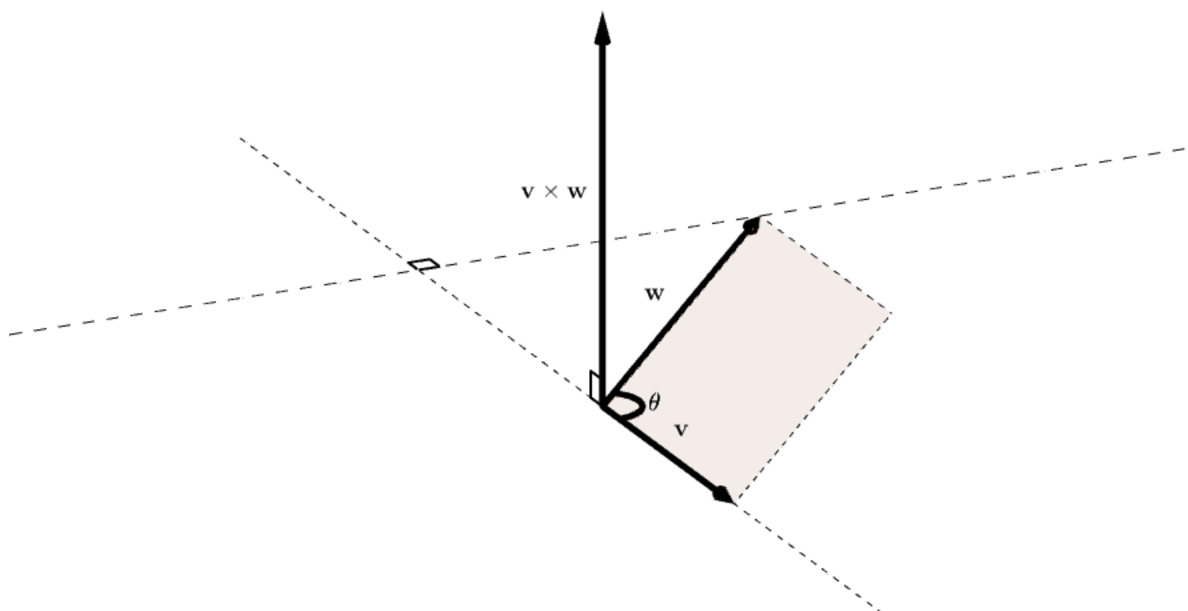
$$\mathbf{v} = -5\mathbf{i} - 5\mathbf{j} + 5\sqrt{2}\mathbf{k}$$

is the velocity of the ball right after the player kicks it. □

Definition 1.29. The **cross product** of two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ is defined to be the unique vector, $\mathbf{v} \times \mathbf{w}$, such that it's orthogonal to both $\mathbf{v}$ and $\mathbf{w}$, its direction is found using the right-hand rule, and its magnitude is equal in value to the area of the parallelogram spanned by $\mathbf{v}$ and $\mathbf{w}$. In the second image below we have,

$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\|\|\mathbf{w}\| \sin \theta.$$


w
v
v × w


v × w
w
θ
v

Proposition 1.30. *If* $\mathbf{v} = \langle v_1, v_2, v_3\rangle$, $\mathbf{w} = \langle w_1, w_2, w_3\rangle$, *then*

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

Proof. See the first page of [**OS** section 2.4] and [**OS** Theorem 2.7]. □

Proposition 1.31. *Properties of the cross product. For any* $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^3$, *for any* $a \in \mathbb{R}$, *then*

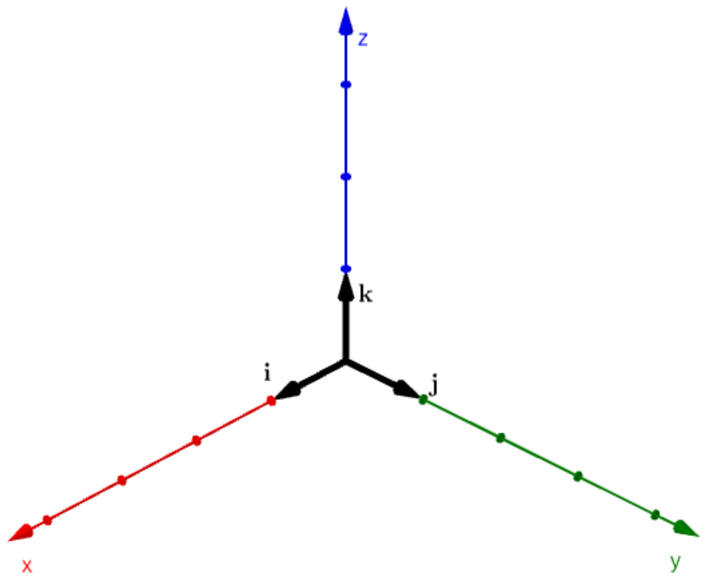
- $(\mathbf{u} + \mathbf{w}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$.
- $a(\mathbf{v} \times \mathbf{w}) = (a\mathbf{v}) \times \mathbf{w} = \mathbf{v} \times (a\mathbf{w})$.
- $\mathbf{v} \times \mathbf{0} = \mathbf{0} \times \mathbf{v} = \mathbf{v} \times \mathbf{v} = \mathbf{0}$.
- $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$.

Proof. The first two properties are referred to as linearity of the cross product and can be proven with the use of Proposition 1.30. The second two can be seen straight from the definition. The last property is referred to as the anticommutativity of the cross product and can seen by the use of the right-hand rule. See [**OS** section 2.4]. □

Example 1.32.

$$\mathbf{i} \times \mathbf{j} = \mathbf{k},$$
$$\mathbf{j} \times \mathbf{k} = \mathbf{i},$$
$$\mathbf{k} \times \mathbf{i} = \mathbf{j},$$

from the right-hand rule and the fact that the standard unit vectors pairwise span a unit square.


z
k
i
j
x
y

□

2 Vector valued functions and surface parametrizations

2.1 Vector valued functions (curve parametrizations)

Definition 2.1. Given real valued f, g, h defined on an open interval $(a, b) = \{t \in \mathbb{R} \mid a < t < b\}$, we call $\mathbf{r} = \langle f, g, h \rangle$ a **vector valued function**. For any $t \in (a, b)$, $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$ is the position vector for the point $(f(t), g(t), h(t))$, so $\mathbf{r}$ is referred to as a **parametrization** of the curve $\mathcal{C} = \{(f(t), g(t), h(t)) \mid t \in (a, b)\}$ in 3-dimensional space, and the equations

$$x = f(t),$$

$$y = g(t),$$

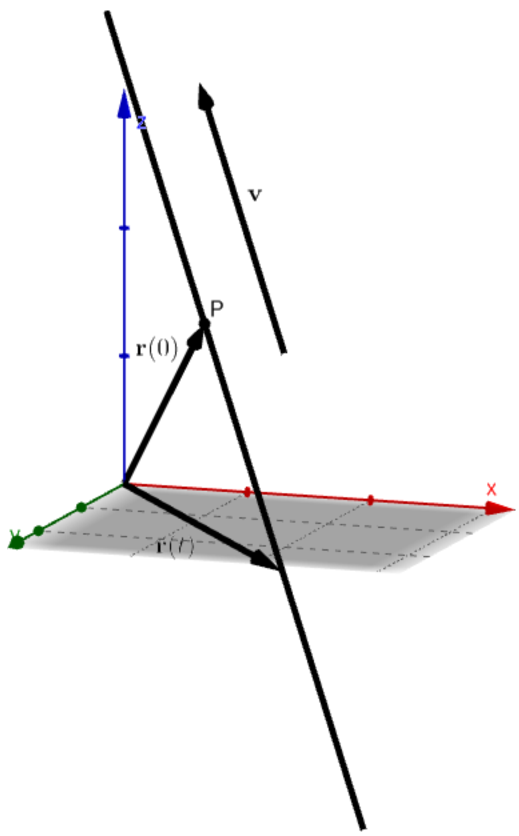
$$z = h(t),$$

are called **parametric equations** for $\mathcal{C}$.

Example 2.2. (a) For any $\mathbf{v} = \langle v_1, v_2, v_3 \rangle \in \mathbb{R}^3$ and any point $P = (P_1, P_2, P_3)$ in space, the vector valued function

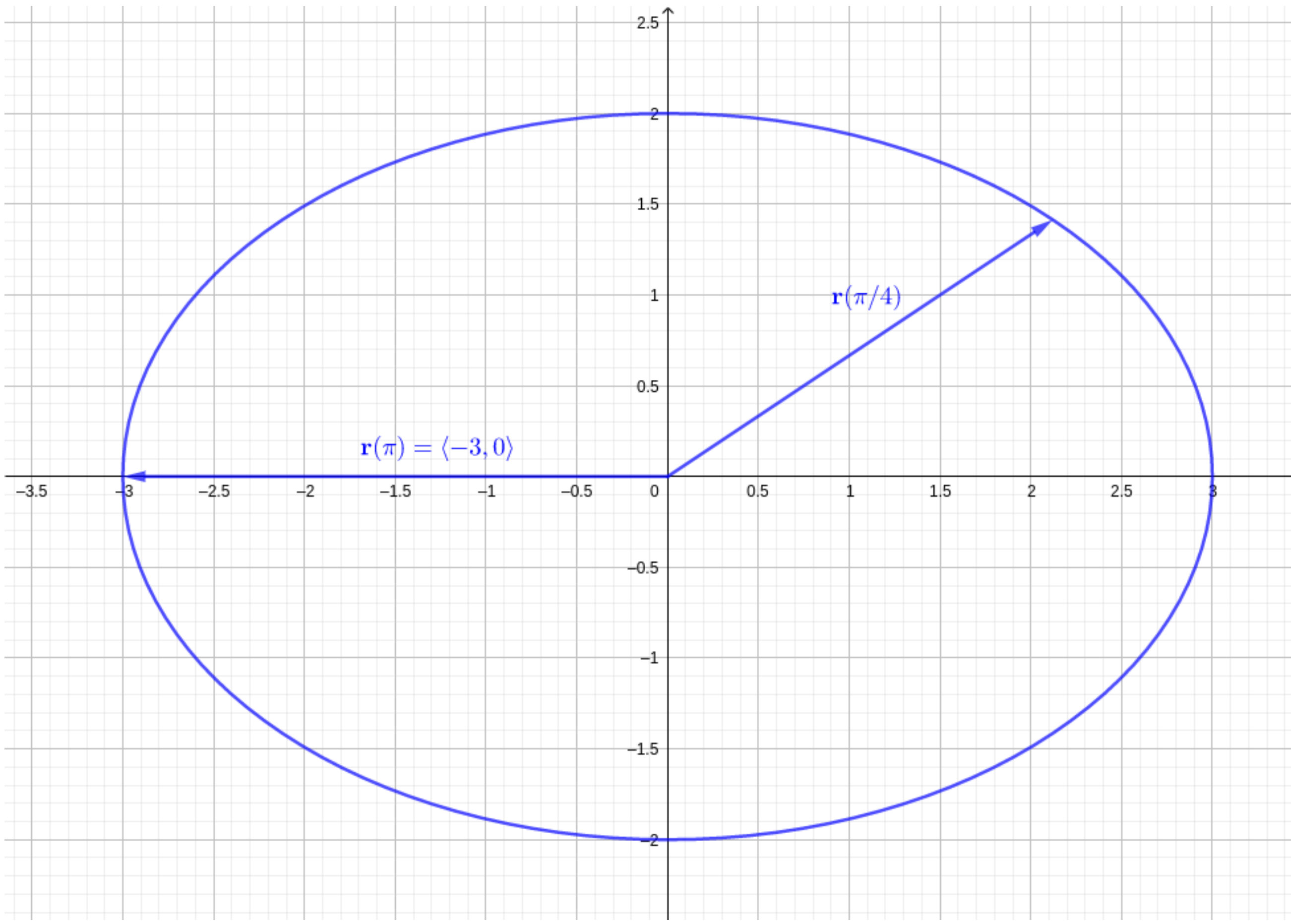
$$\mathbf{r}(t) = \langle tv_1 + P_1, tv_2 + P_2, tv_3 + P_3 \rangle = t\mathbf{v} + \mathbf{OP}$$

for $t \in \mathbb{R}$ parametrizes a line with direction $\mathbf{v}$ passing through P.



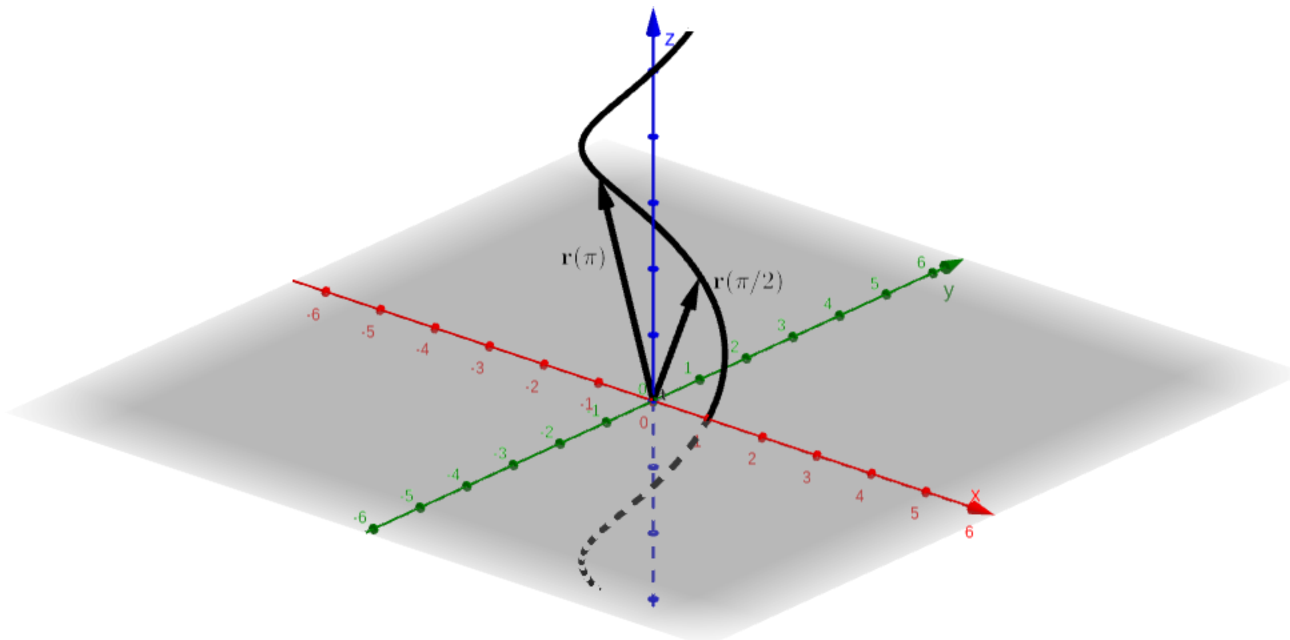
(b) $\mathbf{r}(t) = \langle 3\cos(t), 2\sin t\rangle$ for $t \in [0, 2\pi]$ parametrizes an ellipse. Notice the planar curve described by the equation $\frac{x^2}{9} + \frac{y^2}{4} = 1$ is the ellipse parametrized by $\mathbf{r}$, since if $x(t) = 3\cos t$ and $y(t) = 2\sin t$, then

$$\frac{x^2(t)}{9} + \frac{y^2(t)}{4} = 1.$$



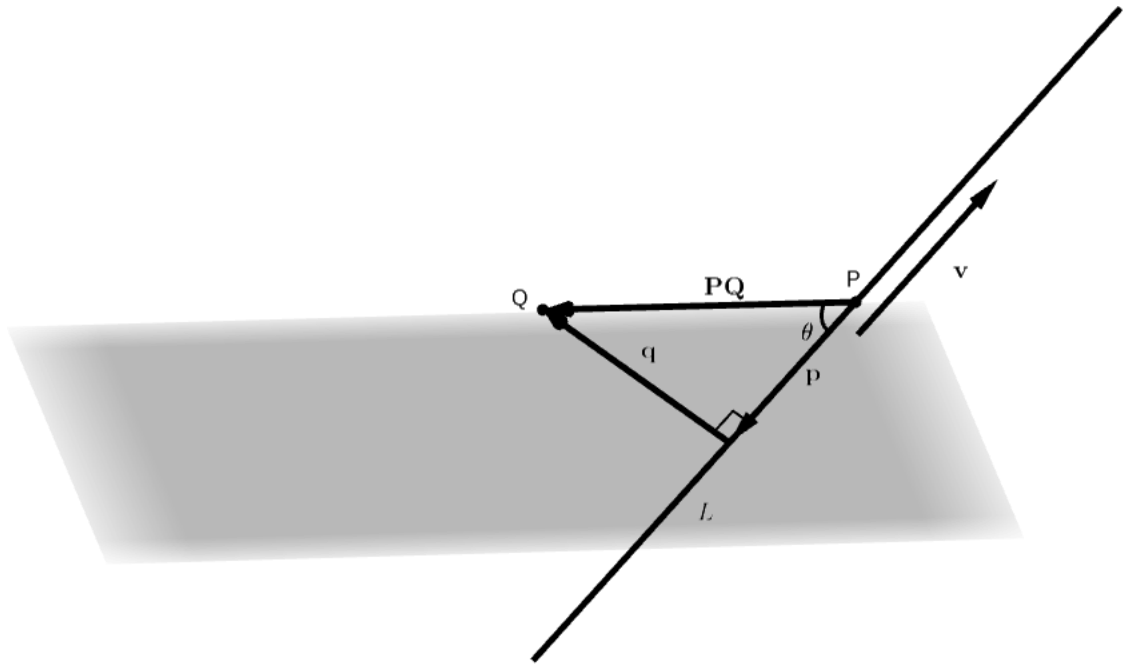
(c) To find the domain of a vector valued function, it's enough to find the domain of each component function. For example, the domain of $\mathbf{r}(t) = \left\langle -\frac{1}{t-1}, \tan(\pi/2 - t), \log t \right\rangle$ is all $t \in \mathbb{R}$ such that $t > 0, t \neq 1$ and $t \neq n\pi$, for positive integers n.

(d) $\mathbf{r}(t) = \langle \cos t, \sin t, t \rangle$, $t \in \mathbb{R}$ parametrizes a **helix**.



$\square$

Example 2.3. We can find the distance between a line, L, and a point, Q, in space as follows. Let $\mathbf{v}$ be a direction vector for L, and P a point on L.



Set $\mathbf{p} = \text{proj}_{\mathbf{v}}\mathbf{PQ} = \frac{\mathbf{PQ}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v}$. Then if $\mathbf{q} = \mathbf{PQ} - \mathbf{p}$, then $\|\mathbf{q}\|$ is the distance from L to Q, by the Pythagorean Theorem.

Now, by Proposition 1.16 and the definition of the dot product,

$$\begin{aligned}
\|\mathbf{q}\|^2 = \|\mathbf{PQ} - \mathbf{p}\|^2 &= (\mathbf{PQ} - \mathbf{p}) \cdot (\mathbf{PQ} - \mathbf{p}) \\
&= \|\mathbf{PQ}\|^2 + \|\mathbf{p}\|^2 - 2\mathbf{PQ} \cdot \mathbf{p} \\
&= \|\mathbf{PQ}\|^2 + \frac{(\mathbf{PQ} \cdot \mathbf{v})^2}{\|\mathbf{v}\|^2} - 2\frac{(\mathbf{PQ} \cdot \mathbf{v})^2}{\|\mathbf{v}\|^2} \\
&= \frac{\|\mathbf{PQ}\|^2\|\mathbf{v}\|^2(1 - \cos^2\theta)}{\|\mathbf{v}\|^2} \\
&= \frac{\|\mathbf{PQ} \times \mathbf{v}\|^2}{\|\mathbf{v}\|^2},
\end{aligned}$$

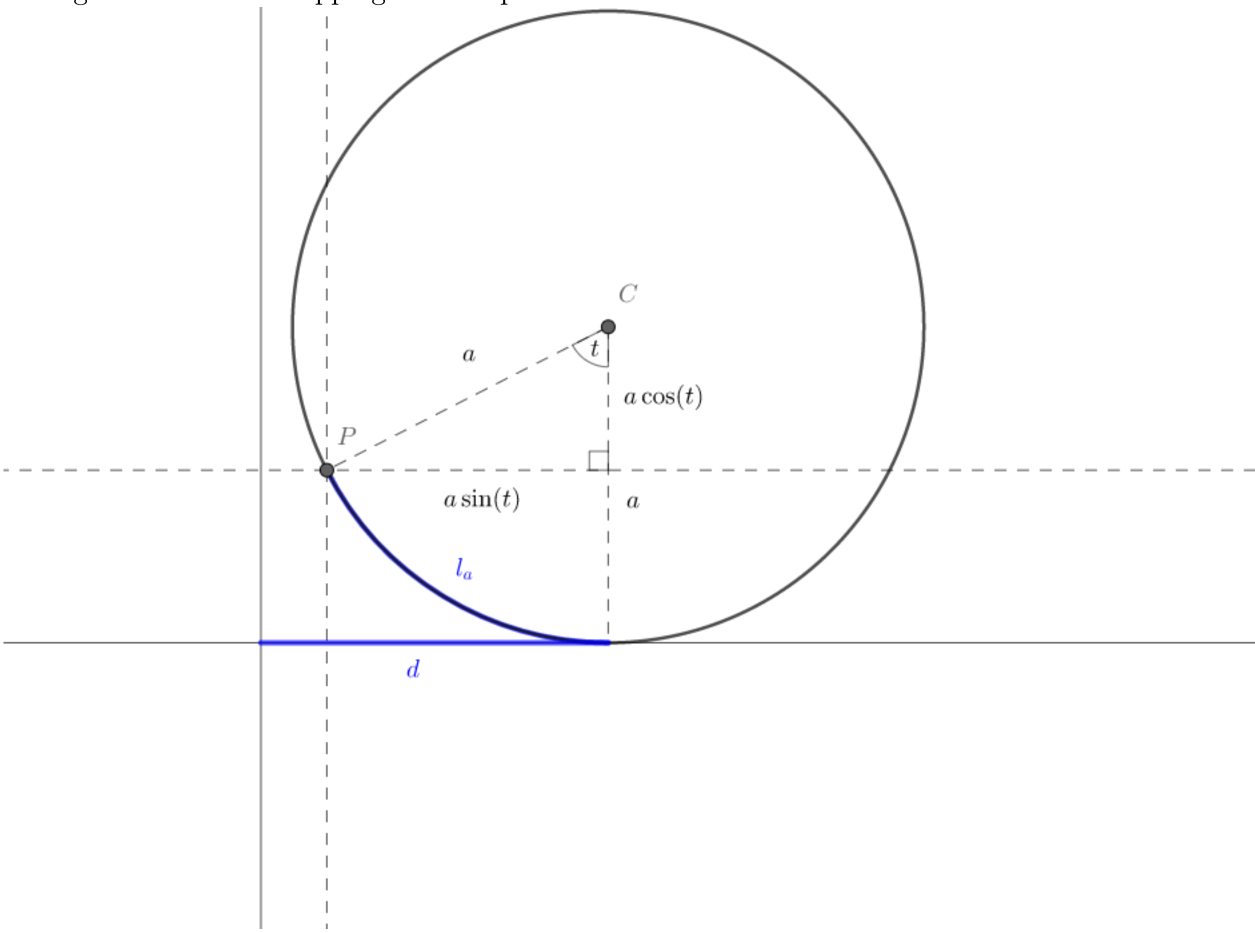
where the last line follows from the definition of the cross product.

That is, the distance between a line L and a point Q is

$$\text{dist}(L, Q) = \|\mathbf{q}\| = \frac{\|\mathbf{PQ} \times \mathbf{v}\|}{\|\mathbf{v}\|},$$

where $\mathbf{q} = \mathbf{PQ} - \text{proj}_{\mathbf{v}}\mathbf{PQ}$, P is an arbitrary point on L, and $\mathbf{v}$ is a direction vector for L. (Opinion: seeing $\text{dist}(L, Q)$ as $\|\mathbf{q}\|$ is more geometric, while seeing $\text{dist}(L, Q)$ as $\frac{\|\mathbf{PQ}\times\mathbf{v}\|}{\|\mathbf{v}\|}$ is more computationally friendly, at least by hand.) $\square$

Example 2.4. A **cycloid** is a path traced out by a point on a circle moving in a straight line without slipping. Find a parametrization for this curve.



From the above image, $d = l_a$ by the no-slip condition. And $l_a = at$, which is the arc-length of the sector corresponding to the angle t the circle with radius a has rotated through. Then we see that $P = (at - a\sin t, a - a\cos t)$. So, a parametrization for a cycloid is

$$\mathbf{r}(t) = \langle at - a\sin t, a - a\cos t\rangle\,.$$

$\square$

Definition 2.5. We say a vector valued function $\mathbf{r}$ approaches a vector $\mathbf{L}$ as t approaches a, written $\lim_{t\to a}\mathbf{r}(t) = \mathbf{L}$, provided $\lim_{t\to a}\|\mathbf{r}(t) - \mathbf{L}\| = 0$.

Theorem 2.6. *If* $\mathbf{r} = \langle f, g, h\rangle$, $\mathbf{L} = \langle L_1, L_2, L_3\rangle$, *then* $\lim_{t\to a}\mathbf{r}(t) = \mathbf{L}$ *if and only if* $\lim_{t\to a} f(t) = L_1$, $\lim_{t\to a} g(t) = L_2$, *and* $\lim_{t\to a} h(t) = L_3$.

Proof. Both directions can be seen by a proof by contradiction. □

Definition 2.7. We say a vector valued function $\mathbf{r}$ is **continuous** at $t = a$ provided

1. $\mathbf{r}(a)$ exists.
2. $\lim_{t\to a}\mathbf{r}(t) = \mathbf{r}(a)$.

We say $\mathbf{r}$ is continuous on an open interval (a, b) if $\mathbf{r}$ is continuous at every point $t \in (a, b)$. Similarly with closed and half-open intervals, with the appropriate conditions on the boundary.

Example 2.8. $\mathbf{r}(t) = \left\langle \sin t\cos t, e^{-t}, t^2 + 1\right\rangle$ is continuous on $\mathbb{R}$ by Theorem 2.6 since its components are continuous on $\mathbb{R}$. □

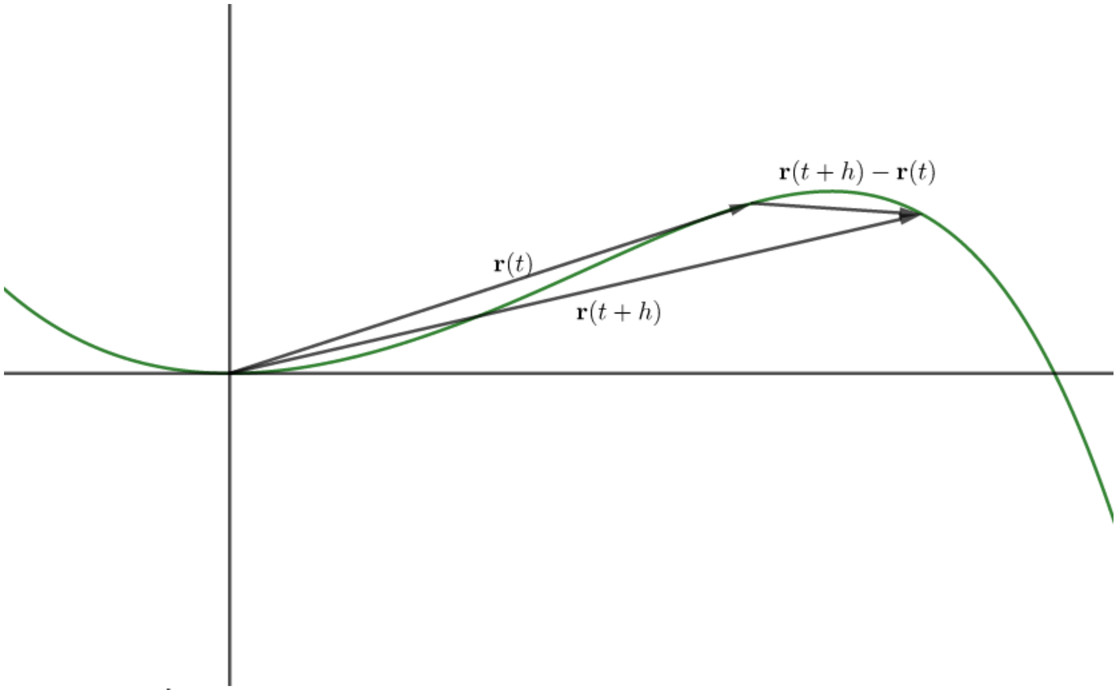
Definition 2.9. We say a vector valued function $\mathbf{r}$ is **differentiable** at t if $\lim_{h\to 0}\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}$ exists. If this is the case, we denote

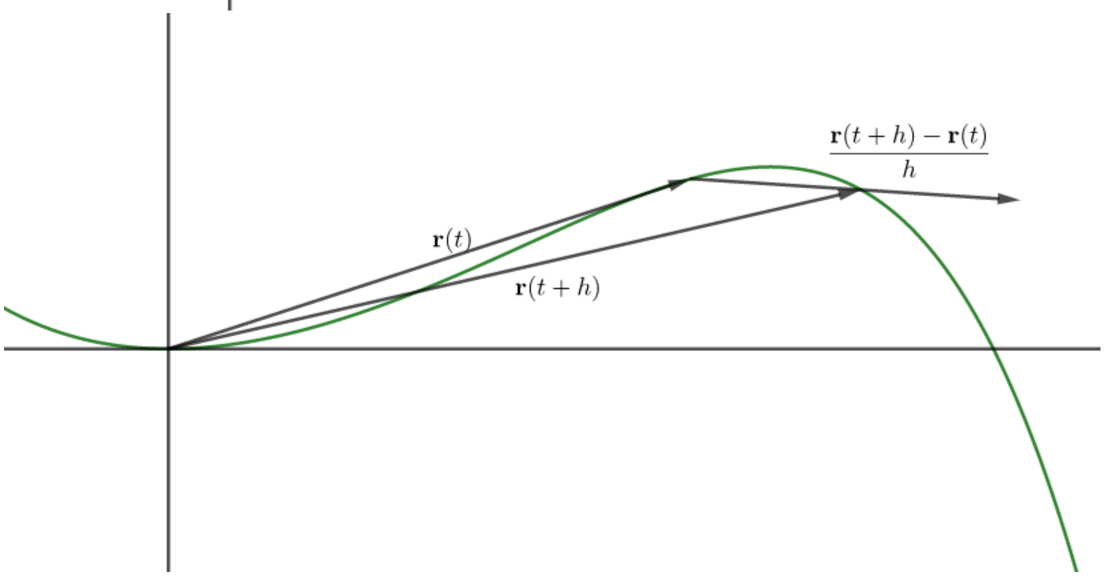
$$\mathbf{r}'(t) := \frac{d}{dt}\mathbf{r}(t) := \lim_{h\to 0}\frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}$$

called the **tangent vector** to $\mathbf{r}$ at t.

We say $\mathbf{r}$ is **differentiable** on (a, b) if $\mathbf{r}$ is differentiable for every $t \in (a, b)$. Similarly for closed and half-open intervals, with the appropriate conditions on the boundary. See [**OS** Section 3.2 page 268].

The **tangent line** to a curve $\mathcal{C}$ at a point P parametrized by $\mathbf{r}$ is the line parametrized by $\mathbf{l}(t) = (t - a)\mathbf{r}'(a) + \mathbf{r}(a)$, where $\mathbf{r}(a) = \mathbf{OP}$.





Theorem 2.10. *A vector valued function* $\mathbf{r} = \langle f, g, h \rangle$ *is differentiable at* t *if and only if* f, g, h *are at* t. *In this case,* $\mathbf{r}'(t) = \langle f'(t), g'(t), h'(t) \rangle$.

Proof. This follows from Theorem 2.6 and Proposition 1.19 (for space vectors). □

Example 2.11. If $\mathbf{r}(t) = \left\langle \log(t-1), e^{2t}+t, \tan(t) \right\rangle$, then $\mathbf{r}'(t) = \left\langle \frac{1}{t-1}, 2e^{2t}+1, \sec^2(t) \right\rangle$. □

Theorem 2.12. *Properties of the derivative. For any real-valued function, f, vector valued functions $\mathbf{c}, \mathbf{r}$, we have*

i. $\frac{d}{dt}(\mathbf{c} \pm \mathbf{r})(t) = \mathbf{c}'(t) \pm \mathbf{r}'(t)$
ii. $\frac{d}{dt}(\mathbf{c} \cdot \mathbf{r})(t) = \mathbf{c}'(t) \cdot \mathbf{r}(t) + \mathbf{c}(t) \cdot \mathbf{r}'(t)$
iii. $\frac{d}{dt}(\mathbf{c} \times \mathbf{r})(t) = \mathbf{c}'(t) \times \mathbf{r}(t) + \mathbf{c}(t) \times \mathbf{r}'(t)$
iv. $\frac{d}{dt}(f\mathbf{r})(t) = f'(t)\mathbf{r}(t) + f(t)\mathbf{r}'(t)$
v. $\frac{d}{dt}\mathbf{r}(f(t)) = f'(t)\mathbf{r}'(f(t))$

Proof. The proofs of these properties are standard following from Theorem 2.10. The first property is one of linearity. The middle three are product rules, while the last is the chain rule for vector valued functions. □

Remark 2.13. The (in)definite integral(s) of a vector valued function are described in the same way (component-wise). See [**OS** pages 274-276].

2.2 Surface parametrizations

Definition 2.14. (informal) A **surface parametrization** is a map sending the point (s, t) to the vector

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

where s and t are both scalars, so f, g, h are functions of two variables. Just as a curve is a set of points in three dimensional space parametrized by a vector valued function (c.f. Definition 2.1), a **surface** is a set of points in three dimensional space parametrized by a surface parametrization.

Example 2.15. For any two vectors $\mathbf{u} = \langle u_1, u_2, u_3 \rangle$, $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$, not scalar multiples of each other, and for any other vector $\mathbf{w} = \langle w_1, w_2, w_3 \rangle$, the map

$$\mathbf{r}(s,t) = \langle su_1 + tv_1 + w_1, su_2 + tv_2 + w_2, su_3 + tv_3 + w_3 \rangle = s\mathbf{u} + t\mathbf{v} + \mathbf{w}$$

parametrizes a **plane**, say, $\mathcal{P}$, with directions $\mathbf{u}$, $\mathbf{v}$, passing through (w_1, w_2, w_3).

This can be seen by example 2.2 (a), since if we fix s or t in the above expression and vary the other, then $\mathbf{r}(s,t)$ parametrizes a line. For example, for any $s_0 \in \mathbb{R}$, the line parametrized by $\mathbf{l}(t) = t\mathbf{v} + s_0\mathbf{u} + \mathbf{w}$ is contained in $\mathcal{P}$.

Let $\mathbf{n} = \mathbf{u} \times \mathbf{v}$. Then notice $\mathbf{n} \neq \mathbf{0}$ since $\mathbf{u}$ and $\mathbf{v}$ are not scalar multiples of each other. If P is any point on $\mathcal{P}$, then $\mathbf{c}(t) = t\mathbf{n} + \mathbf{OP}$ parametrizes the unique line, L, orthogonal to $\mathcal{P}$ passing through P. $\mathbf{n}$ is called a **normal vector** to the plane $\mathcal{P}$. Now, for any $(x,y,z) \in \mathcal{P}$, there exists $s,t \in \mathbb{R}$ such that $\langle x,y,z\rangle = \mathbf{r}(s,t)$, by definition. But this implies we can express $\mathcal{P}$ by the equation

$$\mathbf{u} \times \mathbf{v} \cdot (\langle x,y,z\rangle - \mathbf{w}) = 0.$$

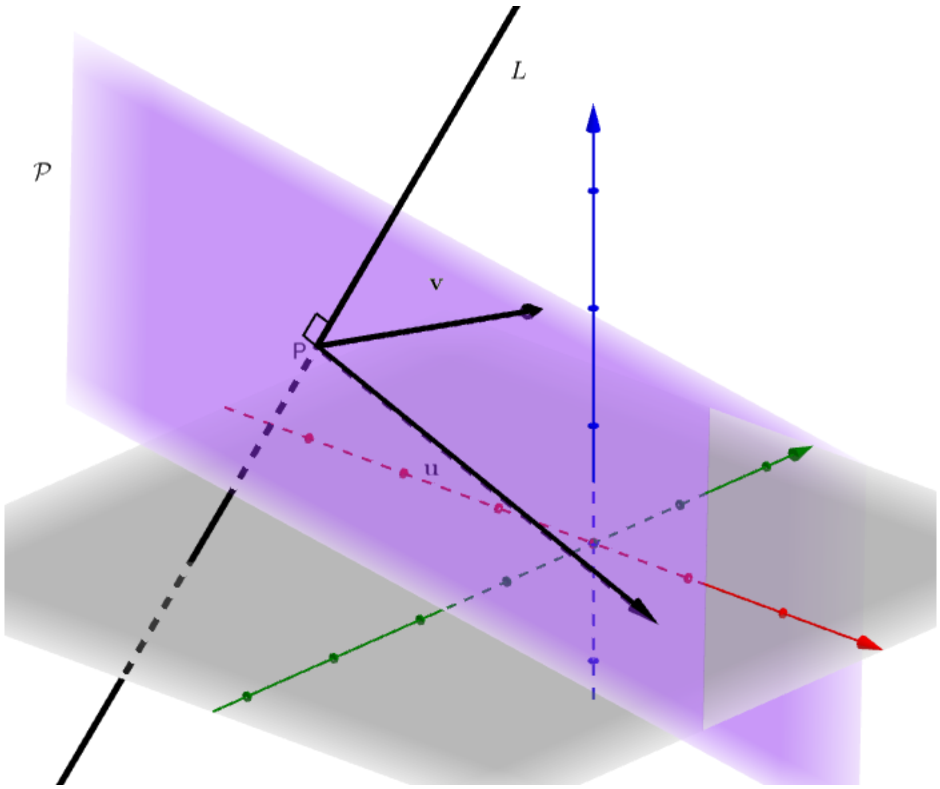
That is,

$$\mathcal{P} = \{(x,y,z) \mid \mathbf{n} \cdot (\langle x,y,z\rangle - \mathbf{OP}) = 0\}.$$

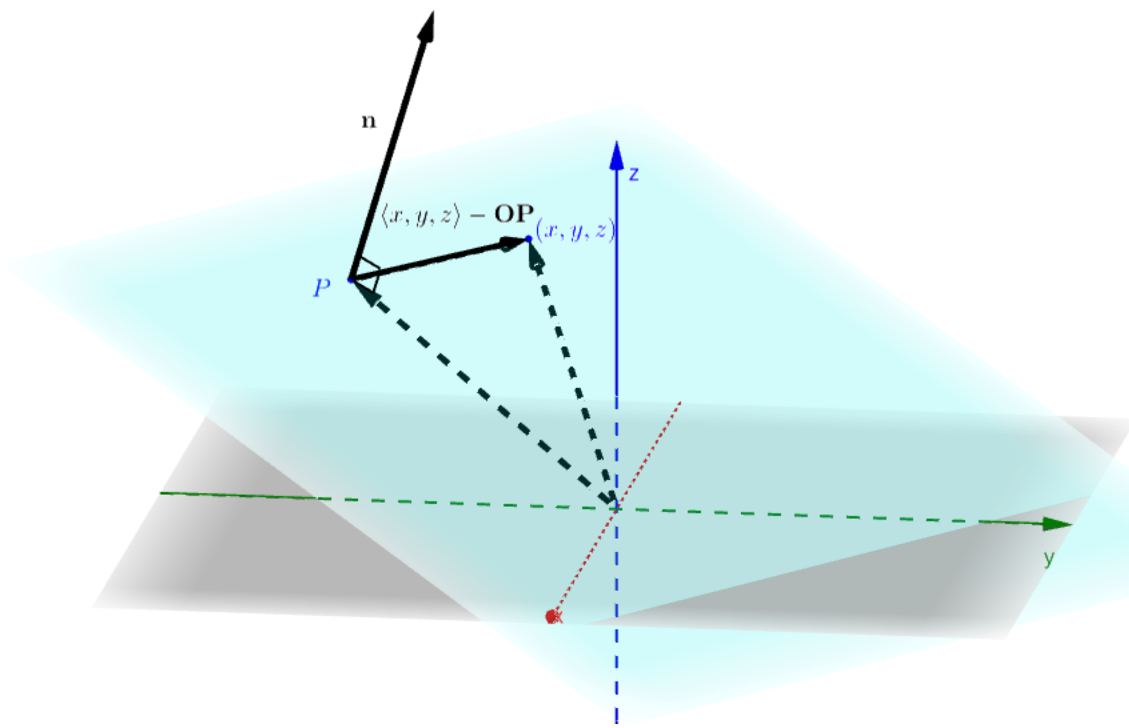
In fact, any plane in 3-dimensional space can be described by a single linear equation with three unknowns:

$$ax + by + cz = d,$$

for scalars a, b, c, d. See [**OS** page 193].



L
𝒫
v
P
u



We say a vector $\mathbf{x} \in \mathbb{R}^3$ **lies on** or **is a direction vector of** $\mathcal{P}$ if there is some $s, t \in \mathbb{R}$ such that $\mathbf{x} = \mathbf{r}(s,t) - \mathbf{OP} = s\mathbf{u} + t\mathbf{v}$. Then $\mathbf{u}$ and $\mathbf{v}$ are certainly direction vectors of $\mathcal{P}$. By the two figures above, convince yourself that this terminology makes sense geometrically. Notice this is different than saying $\mathbf{x} \in \mathcal{P}$. $\mathcal{P}$ is a set of points, not of vectors. We also have, importantly, vector $\mathbf{x}$ is in the direction of $\mathcal{P}$ if and only if it's orthogonal to $\mathbf{n}$. □

Example 2.16. The image below shows a plane parametrized by

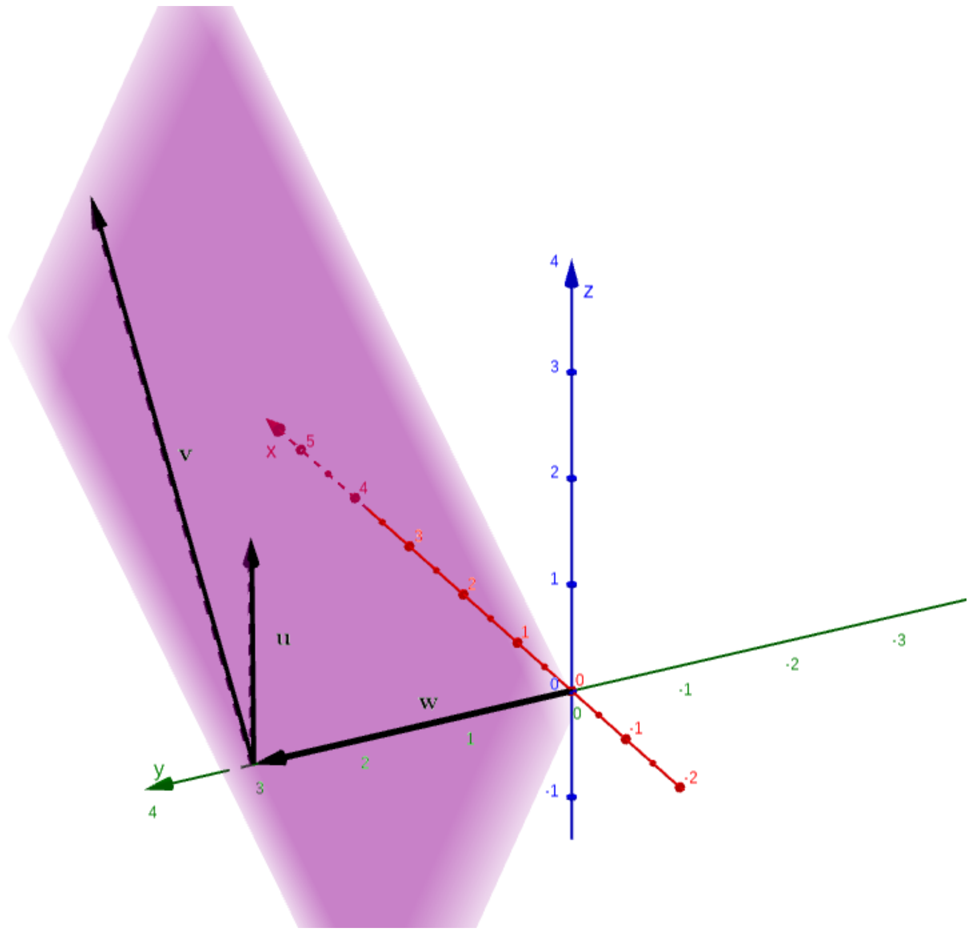
$$\mathbf{r}(s,t) = \langle 2s + 3t, -s + 3, s + 4t \rangle = s\mathbf{u} + t\mathbf{v} + \mathbf{w},$$

where $\mathbf{u} = \langle 2, -1, 1 \rangle$, $\mathbf{v} = \langle 3, 0, 4 \rangle$, $\mathbf{w} = \langle 0, 3, 0 \rangle$. In this case, $\mathbf{n} = \mathbf{u} \times \mathbf{v} = \langle -4, -5, 3 \rangle$, so this plane can also be described by as the set of all points (x, y, z) such that

$$\mathbf{n} \cdot (\langle x, y, z \rangle - \mathbf{w}) = 0.$$

That is, this plane is the set of all points (x, y, z) such that

$$4x + 5y - 3z = 15.$$



□

Example 2.17. The **coordinate planes** are:

The xy-plane: Parametrization: $\mathbf{r}(s,t) = s\mathbf{i} + t\mathbf{j}$.

The origin lies in this plane, and a normal vector is $\mathbf{k}$. So, an equation: $z = 0$.

The xz-plane: Parametrization: $\mathbf{r}(s,t) = s\mathbf{i} + t\mathbf{k}$. The origin lies in this plane, and a normal vector is $\mathbf{j}$. So, an equation: $y = 0$.

The yz-plane: Parametrization: $\mathbf{r}(s,t) = s\mathbf{j} + t\mathbf{k}$. The origin lies in this plane, and a normal vector is $\mathbf{i}$. So, an equation: $x = 0$. See [**OS** page 123].

□

Example 2.18. Find a parametrization and an equation for the plane $\mathcal{P}$ containing the line L with parametric equations

$$x = tv_1 + a_1$$
$$y = tv_2 + a_2$$
$$z = tv_3 + a_3$$

and the point $P = (P_1, P_2, P_3)$ not on L.

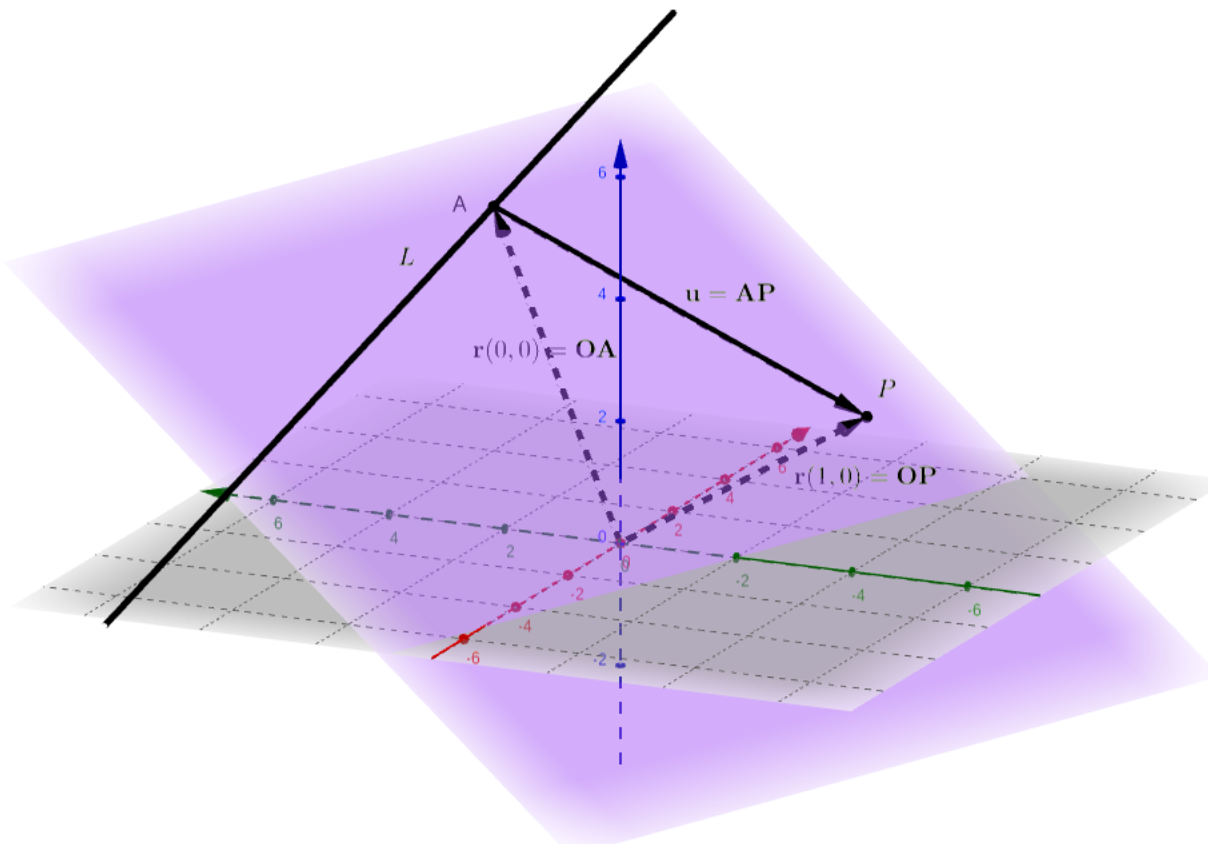
Let $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$, $A = (a_1, a_2, a_3)$. Then L has parametrization $\mathbf{l}(s) = s\mathbf{v} + \mathbf{OA}$.

If $\mathbf{u} = \mathbf{AP}$, then the line parametrized by $t\mathbf{u} + \mathbf{OA}$ is contained in $\mathcal{P}$.

Hence, a parametrization of $\mathcal{P}$ is given by

$$\mathbf{r}(s, t) = s\mathbf{u} + t\mathbf{v} + \mathbf{OA},$$

since then $\mathbf{r}(0, s) = \mathbf{l}(s)$ and $\mathbf{r}(1, 0) = \mathbf{OP}$.



So, an equation for $\mathcal{P}$ is then given by

$$\mathbf{u} \times \mathbf{v} \cdot (\langle x, y, z \rangle - \mathbf{OA}) = 0.$$

That is,

$$ax + by + cz = d$$

if $\mathbf{u} \times \mathbf{v} = \langle a, b, c \rangle$ and $d = \mathbf{u} \times \mathbf{v} \cdot \mathbf{OA}$.

In the above image: $A = (4, 4, 4), \mathbf{v} = \langle 5, 0, 1 \rangle, P = (5, -2, 1)$ and so $\mathcal{P} = \{(x, y, z) \mid 3x + 8y - 15z = -16\}$. $\square$

Example 2.19. Quadric surfaces. See [**OS** pages 213-221]. For $\theta \in [0, 2\pi]$ and scalars a, b, c,

(a) $\mathbf{r}(\varphi, \theta) = \langle a \cos\theta \sin\varphi, b \sin\theta \sin\varphi, c \cos\varphi \rangle$, for $\varphi \in [0, \pi]$, parametrizes an **ellipsoid**: the set of all points (x, y, z) such that

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

It is so called because its **traces**, that is, its intersections with planes parallel to coordinate planes: $z = t, y = s, x = r$, are all ellipses. For example, the intersection of $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ with the plane $y = s$ is the ellipse $\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1 - \frac{s^2}{b^2}$, provided $|s| < |b|$. Otherwise the trace is empty.

(b) $\mathbf{r}(s, \theta) = \langle a \cos\theta \cosh s, b \sin\theta \cosh s, c \sinh s \rangle$ (here, cosh, sinh are the hyperbolic sin and cos functions), for $s \in \mathbb{R}$, parametrizes a **hyperboloid of one sheet**: the set of all points (x, y, z) such that

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

It is so called because its traces are hyperbolas when $x = r$ and $y = s$, and an ellipse when $z = t$.

(c) $\mathbf{r}(s, \theta) = \langle a \cos\theta \sinh s, b \sin\theta \sinh s, c \cosh s \rangle$, for $s \in \mathbb{R}$, parametrizes the upper sheet of a **hyperboloid of two sheets**: the set of all points (x, y, z) such that

$$\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

It is so called because its traces are hyperbolas when $x = r$ and $y = s$, and an ellipse when $z = t$.

(d) $\mathbf{r}(s,\theta) = \langle as\cos\theta, bs\sin\theta, cs\rangle$ parametrizes a **double cone**: the set of all points (x, y, z) such that

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

A double cone can be thought of as a degenerate hyperboloid, since its trace when $x = 0$ and when $y = 0$ is a pair of orthogonal lines, and it intersects the plane $z = 0$ at a point (the origin), but when, e.g., $x = r \neq 0$, we recover a hyperbola, and when $z = t \neq 0$, we recover an ellipse.

For (b)-(d), see this link. Vary ρ, d, and s separately and see how the trace changes. Notice, in particular, that when $s = d = 0$, the trace is the intersection of the surface $x^2 + y^2 - z^2 = \rho$ with the plane $y = 0$.

In this case, when $\rho < 0$, the surface is a hyperboloid of two sheets and the trace is a hyperbola opening in the direction of the z-axis.

When $\rho > 0$, the surface is a hyperboloid of one sheet and the trace is a hyperbola opening in the direction of the x-axis.

When $\rho = 0$, the surface is a cone and the trace is a pair of lines. The hyperbola maps to its asymptotes. □

Example 2.20. Tangent planes. For $\theta \in [0, 2\pi]$, and $\varphi \in [0, \pi]$, $\mathbf{r}(\varphi, \theta) = \langle\cos\theta\sin\varphi, \sin\theta\sin\varphi, \cos\varphi\rangle$, for $\varphi \in [0, \pi]$, parametrizes the **unit sphere**, denoted $\mathbb{S}^2$: the set of all points (x, y, z) such that

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

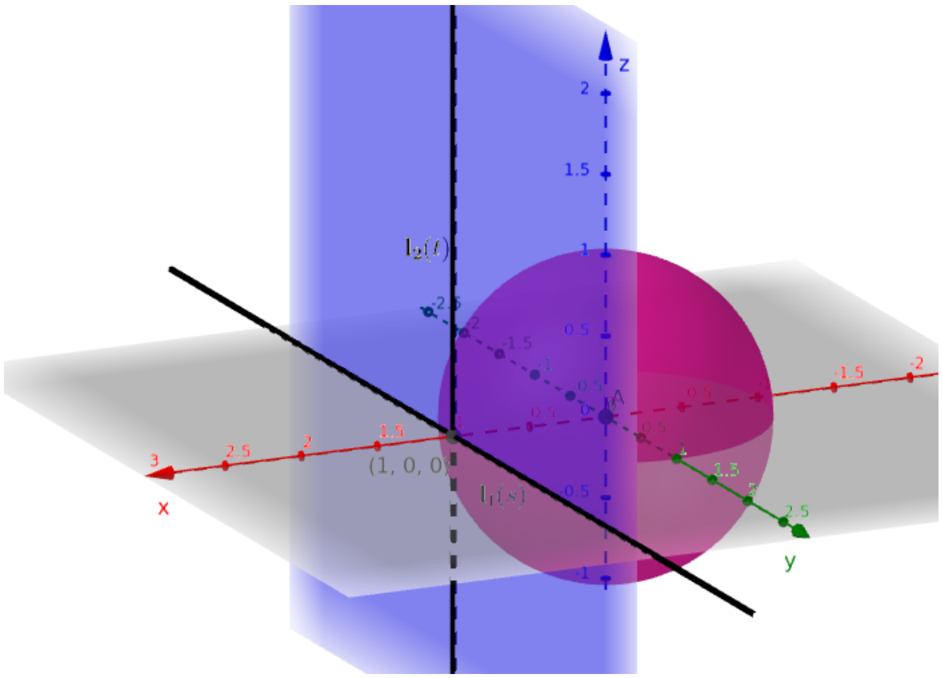
This set is also described as the set of all points unit distance from the origin.

Notice $\mathbf{r}(\pi/2, 0) = \langle 1, 0, 0\rangle$.

The vector valued function $\mathbf{c_1}(\theta) = \mathbf{r}(\pi/2, \theta) = \langle\cos\theta, \sin\theta, 0\rangle$ parametrizes the unit circle in the xy-plane. The unit circle has tangent line L_1, with parametrization $\mathbf{l_1}(s) = s\frac{d}{d\theta}\mathbf{c_1}(0) + \mathbf{c_1}(0) = \langle 1, s, 0\rangle$ at $(1, 0, 0)$.

The vector valued function $\mathbf{c_2}(\varphi) = \mathbf{r}(\varphi, 0) = \langle\sin\varphi, 0, \cos\varphi\rangle$ parametrizes half of the unit circle in the xz-plane. This unit circle has tangent line L_2, with parametrization $\mathbf{l_2}(t) = t\frac{d}{d\theta}\mathbf{c_2}(\pi/2) + \mathbf{c_1}(\pi/2) = \langle 1, 0, t\rangle$ at $(1, 0, 0)$.

In this case, the unique plane containing the tangent lines L_1 and L_2 is referred to as the *tangent plane* to $\mathbb{S}^2$ at $(1, 0, 0)$. This plane is parametrized by $\mathbf{p}(s, t) = s\mathbf{j} + t\mathbf{k} + \mathbf{i} = \langle 1, s, t\rangle$, which has corresponding equation $x = 1$. See Definition 3.6.


z
x
y
(1, 0, 0)
A
$l_2(t)$
$l_1(s)$

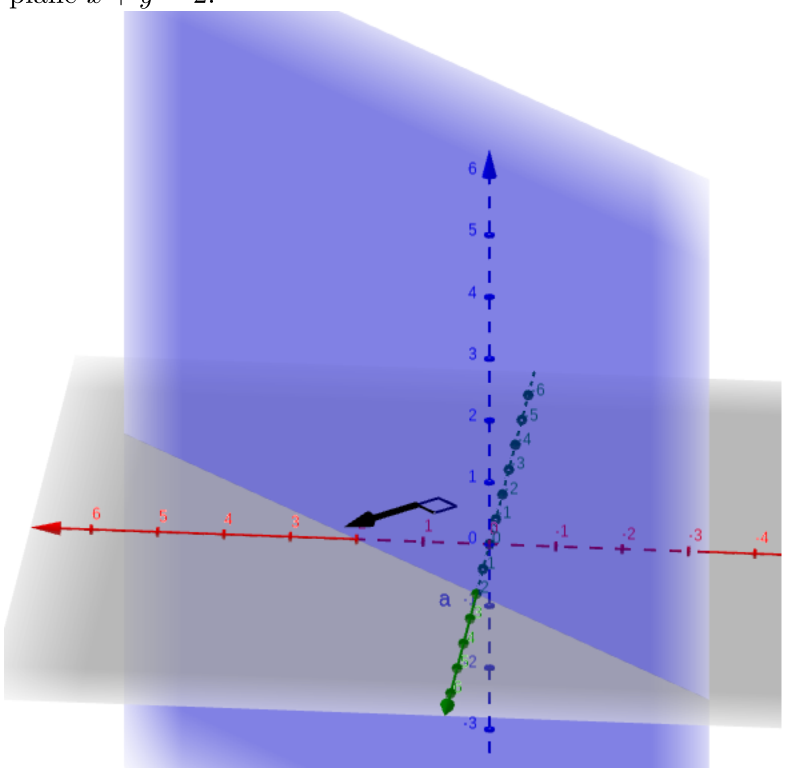
□

3 Functions of several variables

Recall (Example 2.15), we determined a plane by the equation

$$ax + by + cz = d$$

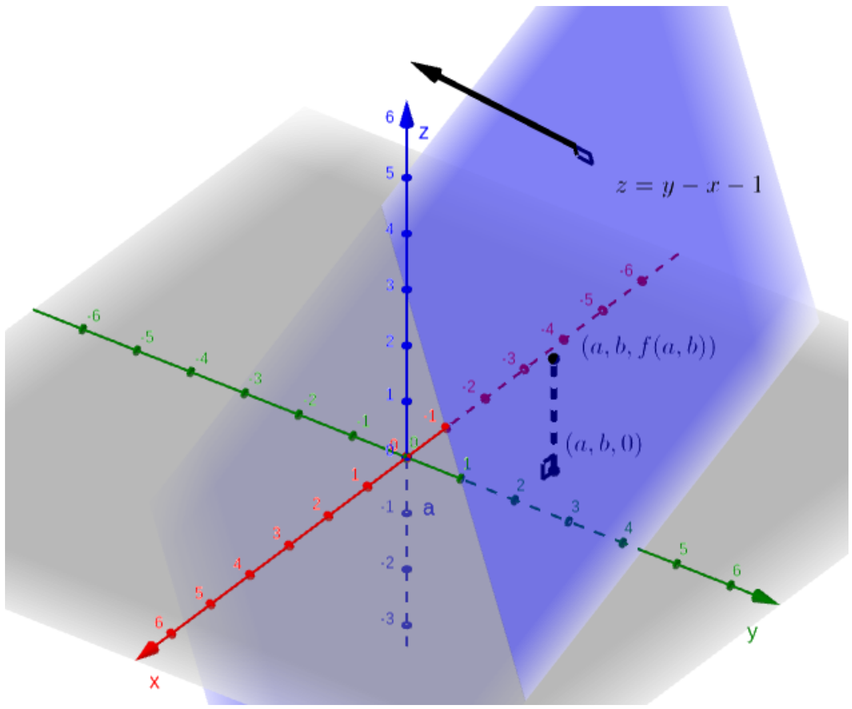
where a, b, c, d are fixed real numbers. If $c = 0$, the the plane $ax + by = d$ is perpendicular to the xy-plane, in that their normal vectors are orthogonal: $\langle a, b, 0\rangle \cdot \langle 0, 0, 1\rangle = 0$. See the figure in the above page for an example. Below: the plane $x + y = 2$.



If $c \neq 0$, then we may write

$$z = c^{-1}(d - ax - by).$$

That is, we can write $z = f(x, y)$ if $f(x, y) = c^{-1}(d - ax - by)$.



3.1 Functions of two variables

Definition 3.1. A **function of two variables** is a real-valued function with domain a subset of the xy-plane.

The domain of a function of two variables is usually taken to be the largest subset of points in the plane which generates a real number. That is, we exclude division by zero and points which lead to complex numbers (paraphrase of Bob Hingtgen). The **range** of a function f with domain D is the set $R = \{z \in \mathbb{R} \mid z = f(x, y) \text{ for some } (x, y) \in D\}$.

Example 3.2. (a) If $f(x, y) = 3x + y^3$, then $f(3, -2) = 3(3) + (-2)^3 = 1$.

(b) (i) If $f(x, y) = \sqrt{x^2 + y^2}$, then the domain of f is the entire plane, since we only need to ensure that $x^2 + y^2 \geq 0$, and this is true for any pair (x, y). Its range is $[0, \infty)$, since, for example, if $x = 0$, then $f(x, y) = |y|$, which takes on all non-negative real values. And $f(x, y) \geq |y|$ for all ordered pairs (x, y).

(ii) If $g(x, y) = \frac{1}{xy}$, then the domain of g is wherever $xy \neq 0$. And $xy = 0$

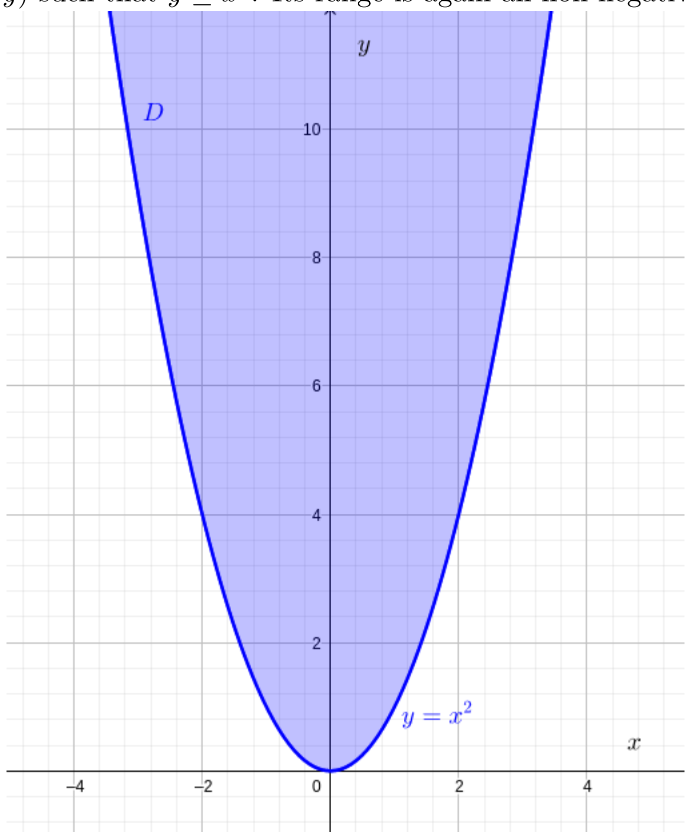
if and only if $x = 0$ or $y = 0$. That is, the domain of g is

$$D = \{(x, y) \mid x \neq 0 \text{ and } y \neq 0\};$$

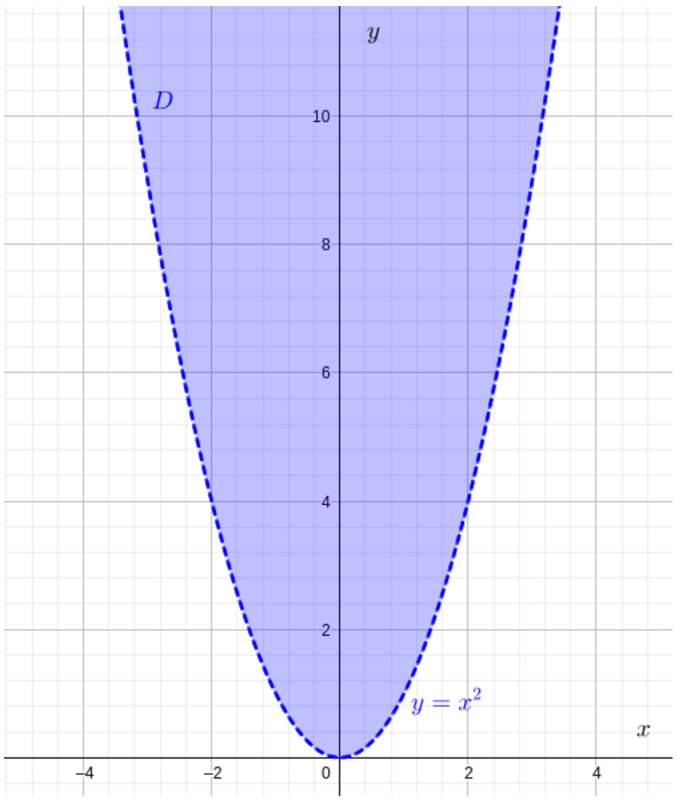
everything but the coordinate axes in the xy-plane.

The range of g is $R = (-\infty, 0) \cup (0, \infty)$, since, for example, if $y = y_0$ is fixed, then $g(x, y) = \frac{1}{y_0 x}$. And the range of this function of one variable is R, for any $y_0 \neq 0$.

(iii) If $h(x, y) = \sqrt{y - x^2}$, then the domain, D, of h consists of all pairs (x, y) such that $y \geq x^2$. Its range is again all non-negative real numbers.



(iv) If $k(x, y) = \frac{1}{\sqrt{y - x^2}}$, then the domain, D, of k, is similar to that of h in (iii), except we must exclude its boundary curve $y = x^2$: $D = \{(x, y) \mid y > x^2\}$. We draw a dashed line along $y = x^2$ instead of a solid one above to denote this:



□

Definition 3.3. The **graph** of a function of two variables f with domain D is the following subset of three dimensional space:

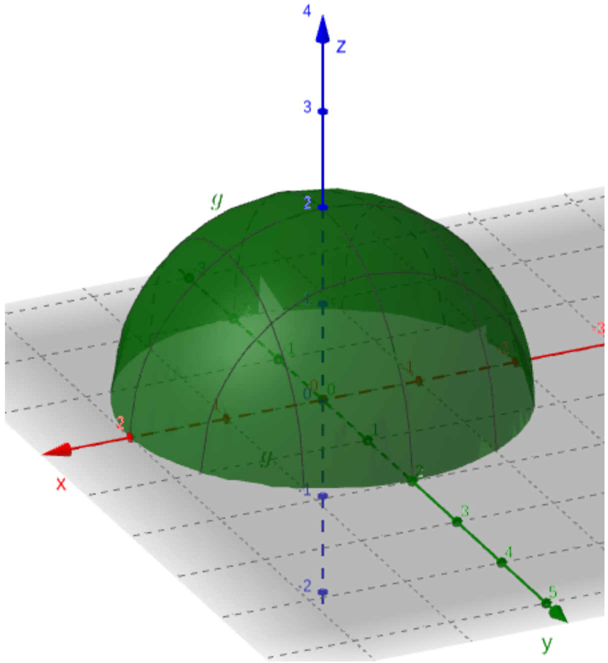
$$\operatorname{graph} f := \{(x, y, f(x, y)) \mid (x, y) \in D\}$$

Notice the graph of the function $z = y - x - 1$ is the plane described by this equation, while the plane $x + y = 2$ is not the graph of any function of two variables with domain in the xy-plane. (See the discussion at the beginning of this section.)

Remark 3.4. A graph of a function of two variables is a surface, since it has surface parametrization $\mathbf{r}(x, y) = \langle x, y, f(x, y)\rangle$.

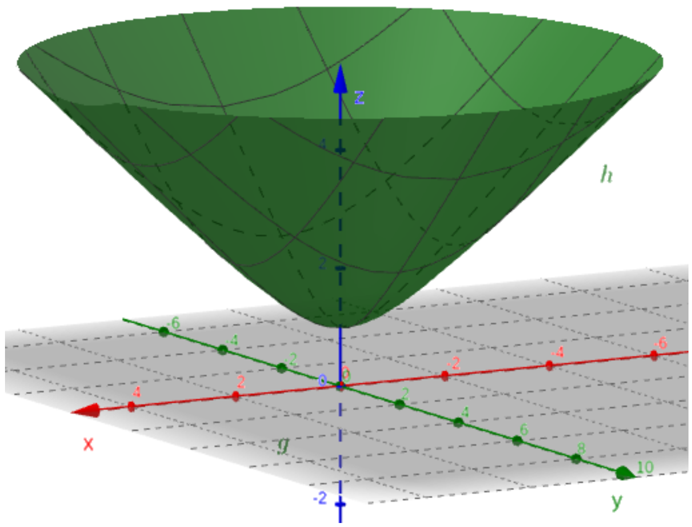
Example 3.5. (a) To describe the graph of $g(x, y) = \sqrt{4 - x^2 - y^2}$, we can rewrite this expression to make it look more familiar: that is, if $z = g(x, y)$, then $g(x, y) = \sqrt{4 - x^2 - y^2}$ if and only if $x^2 + y^2 + z^2 = 4$ and $z \geq 0$.

We recognize this as the upper hemisphere of the sphere of radius 2, center origin:



Notice the domain of g is the **closed disk** described by $x^2 + y^2 \leq 4$. Sketch this domain in the xy-plane. What is the range of g?

(b) To describe the graph of $h(x, y) = \sqrt{1 + x^2 + y^2}$, we can again rewrite this expression to make it look more familiar: that is, if $z = h(x, y)$, then $h(x, y) = \sqrt{1 + x^2 + y^2}$ if and only if $z^2 - y^2 - y^2 = 1$ and $z \geq 0$. We recognize this as the upper sheet of a hyperboloid of two sheets:



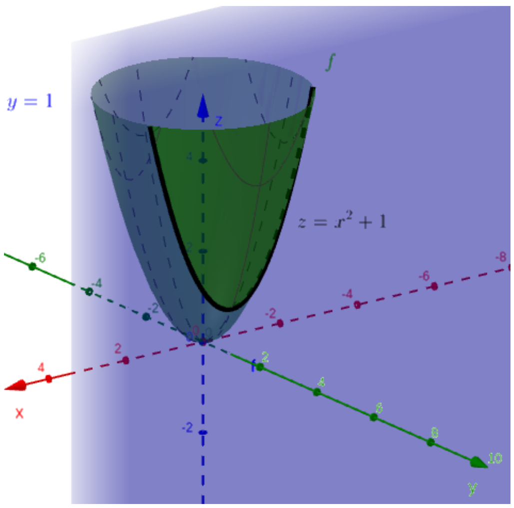
What is the domain and range of h? □

The examples in 3.5 are special in the sense that we recognized their graphs as quadric surfaces. But recall (Example 2.19) how we described quadric surfaces. (With their traces.) We play a similar game with graphs of functions of two variables.

Definition 3.6. A **trace** of a surface S is the intersection of S with some plane parallel to a coordinate plane. Fix $r, s \in \mathbb{R}$. The intersection of S with either $x = r$ or $y = s$ is a **vertical trace** of S with either $x = r$ or $y = s$. Fix $t \in \mathbb{R}$. The intersection of S with the plane $z = t$ is a **horizontal trace** of S with $z = t$.

In many cases, a trace of a surface is either a curve, a point, or empty. But sometimes it's a collection of curves or points, or both.

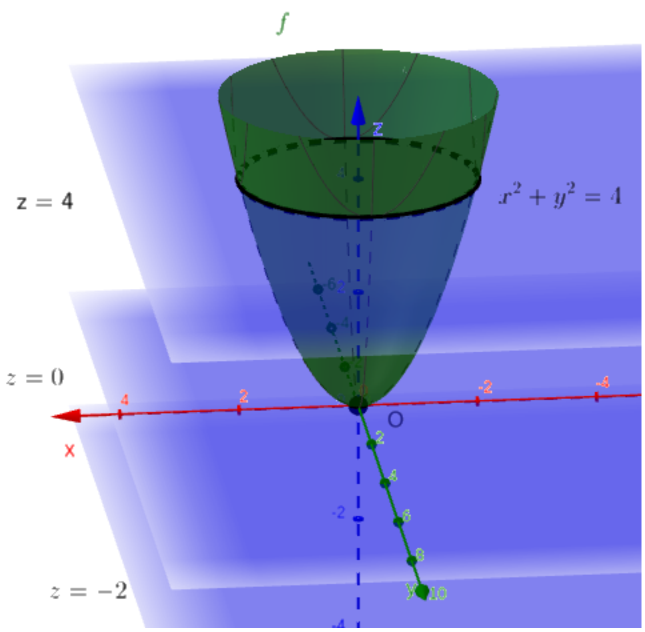
Example 3.7. (a) Suppose $f(x, y) = x^2 + y^2$. The vertical traces of graph f are parabolas. E.g., if $s \in \mathbb{R}$, then the trace of graph f with the plane $y = s$ is the parabola $z = x^2 + s^2$ in the plane $y = s$:



Let's now find the horizontal traces of graph f. If $t > 0$, then the horizontal trace of graph f with the plane $z = t$ is the set of all points $(x, y, f(x, y))$ such that $f(x, y) = t$. Hence, $x^2 + y^2 = t$ and $z = t$, a circle with radius $\sqrt{t}$.

If $t = 0$, then the points which satisfy $x^2 + y^2 = 0$ is just one point, the origin, O.

If $t < 0$, then the trace of graph f with $z = t$ is empty, since $0 \leq x^2 + y^2 = t < 0$ is a contradiction.



graph f is an example of an *elliptic paraboloid*, another quadric surface.

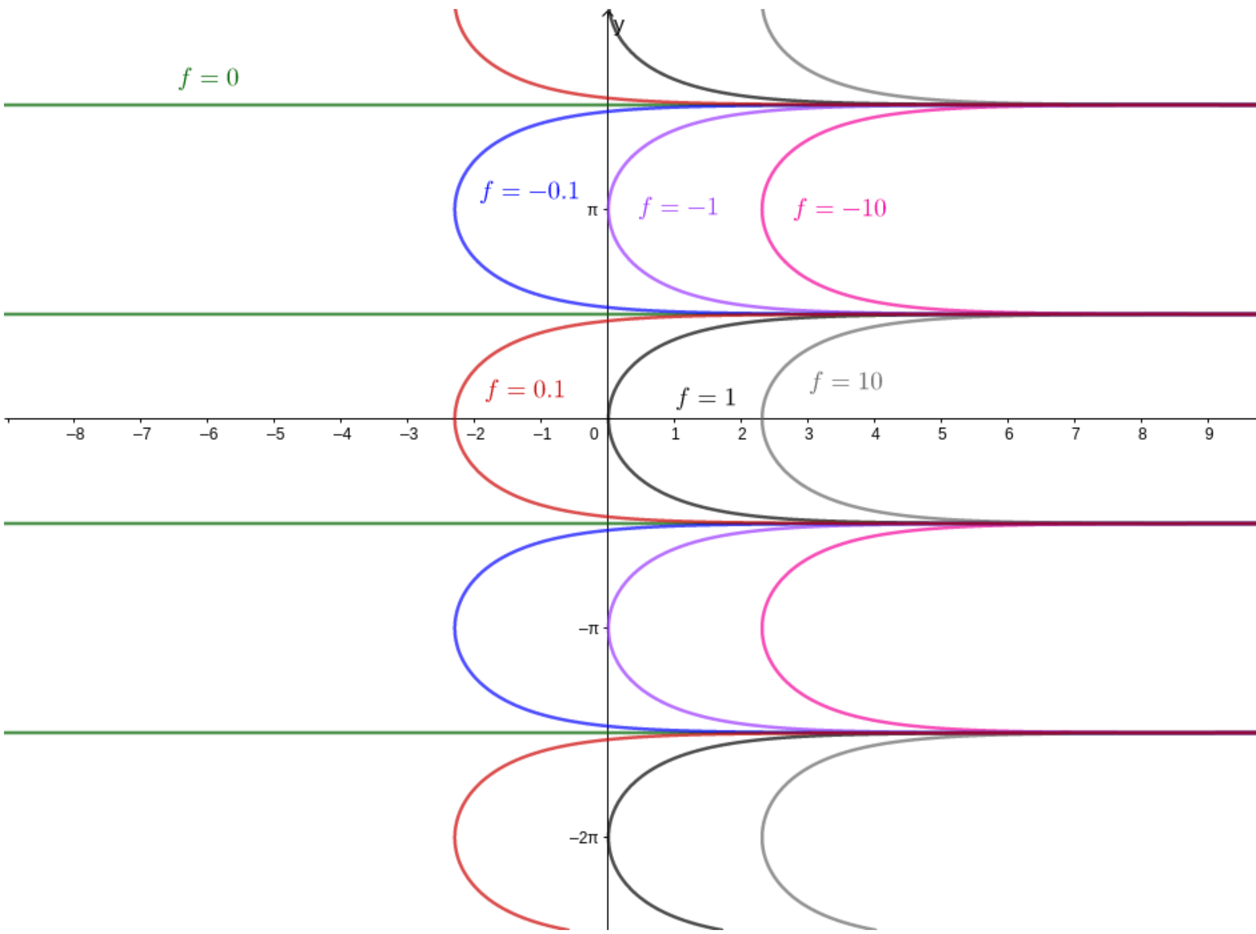
This example hopefully demonstrates that the horizontal trace of a graph of a function of two variables with the plane, say, $z = t$, is the set contained in the plane $z = t$ where the function values take on the constant t. See the definition of a level curve below. $\square$

Definition 3.8. A **level curve** or **contour line** of a function of two variables, f, is the projection of a given horizontal trace onto the xy-plane. E.g., the set of all points $(x, y, 0)$ such that $f(x, y) = k$ for some $k \in \mathbb{R}$. For example, a topographic map has contour lines of constant elevation. See here and here for more discussion and some neat examples. In particular, notice if the difference in the function's value from one level curve to the next is fixed (this common difference called the **contour interval**), then in places where the level curves are closer correspond to places where the function is increasing more (where the graph is steeper), and in places where the level curves are farther apart corresponds to a lack of change (the graph is gentler, flatter).

Example 3.9. (a) Suppose $f(x, y) = e^x \cos(y)$. The domain of this function is certainly the entire plane. Let's try to describe the graph of f by describing its level curves and vertical traces. $f(x, y) = 0$ wherever $\cos(y) = 0$, for any x, since $e^x > 0$:

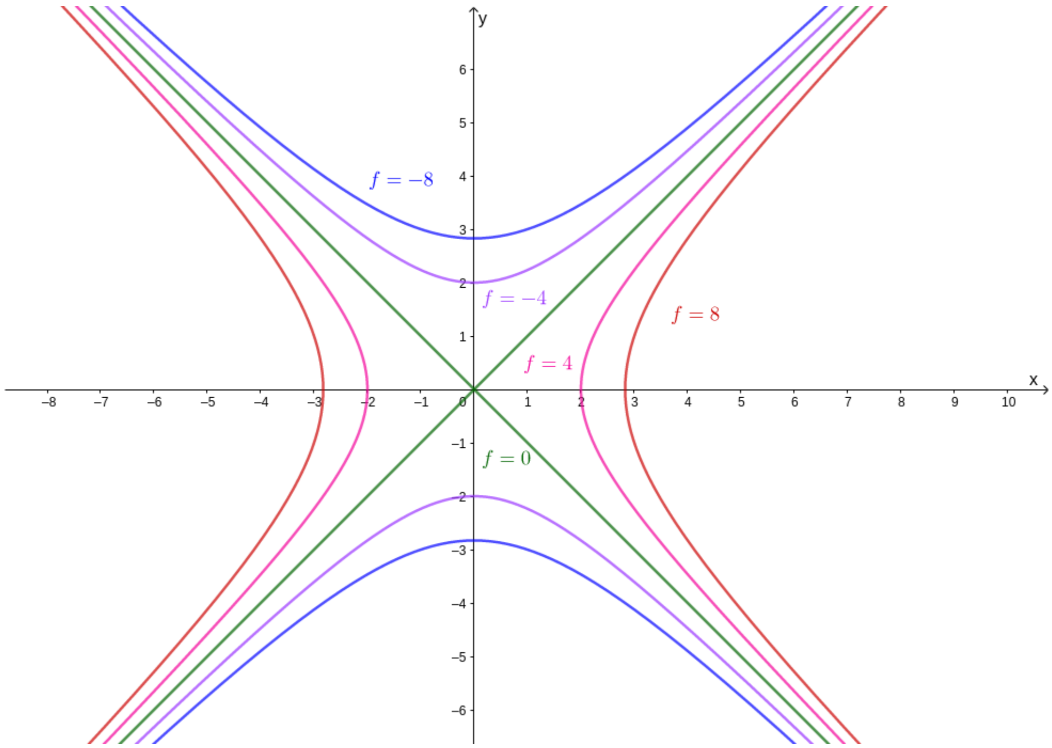
$$\{(x, y, 0) \mid y = \frac{\pi}{2} + \pi n \text{ for some integer } n\}.$$

When $k > 0$, we can argue as follows. The level curve described by $f(x, y) = k$ will again be a collection of curves repeating along the y axis, because of the periodicity of cos. But if $k > 0$, we must have $\cos(y) > 0$, since $e^x > 0$ for all x. $\cos(y) > 0$ when $-\frac{\pi}{2} + 2\pi n < y < \frac{\pi}{2} + 2\pi n$ for some integer n (see unit circle). In any of these intervals, we have $x = \ln(k \sec(y))$. We get the same expression when $k < 0$, except y is in different intervals: $\cos(y) < 0$ when $\frac{\pi}{2} + 2\pi n < y < \frac{3\pi}{2} + 2\pi n$ for some integer n.



The vertical traces are easier to describe. When $x = r$, $f(x, y) = e^r \cos(y)$ is a cosine curve with amplitude e^r. So, the cosine curves get larger or smaller depending on which direction in x one goes. When $y = s$, $f(x, y) = \cos(s)e^x$ is an exponential curve either decreasing or increasing in x depending on the sign of $\cos(s)$. One can see this behavior from the level curves, too. Vary k, r, s here.

(b) Suppose $f(x, y) = x^2 - y^2$. Its level curves are hyperbolas:



Try to sketch these level curves as horizontal traces to get an idea of what graph f looks like, in particular, those that pass through points near the origin. In particular, the function increases in either x direction, while it decreases in either y direction. What are this surface's vertical traces? Its name might give them away. graph f is called a *hyperbolic paraboloid*. Now follow this link to verify your reasoning. □

Definition 3.10. If n is a positive integer, a **function of n variables** is a real-valued function whose domain is a subset of the set of ordered n-tuples, where an n-tuple is an ordered list of n real numbers, e.g., $(x_1, x_2, x_3, \dots, x_n)$ where $x_i \in \mathbb{R}$ for every $i \in \{1, 2, \dots, n\}$. Sometimes we will have a chance to consider f as a function on n-dimensional vectors, where $\mathbb{R}^n := \{\langle x_1, x_2, \dots, x_n\rangle \mid x_i \in \mathbb{R} \text{ for every } i \in \{1, 2, \dots, n\}\}$ is the set of n-dimensional vectors (the set of n-tuples with the structure of vector addition and scalar multiplication). See Definition 1.25.

Example 3.11. (a) $f(x, y, z) = \frac{1}{x^2+y^2+z^2}$ is a function of three variables whose domain is all ordered triples $(x, y, z) \neq (0, 0, 0)$.

(b) $f(x, y, z, w) = \frac{1}{\sqrt{9-x^2-y^2-z^2-w^2}}$ is a function of four variables whose domain is all ordered quadruples (x, y, z, w) such that $x^2 + y^2 + z^2 + w^2 < 9$, the interior of a three dimensional *hypersphere*.

(c) A **level surface** of a function of three variables, f, is the set of all points (x, y, z) satisfying $f(x, y, z) = k$ for some real k. For example, if $f(x, y, z) = x^2+y^2-z^2$, then the level surfaces of f are given by hyperboloids of various sheets, and a double cone if $k = 0$. See Example 2.19 (b)-(d) and the link provided there. □

3.2 Continuity

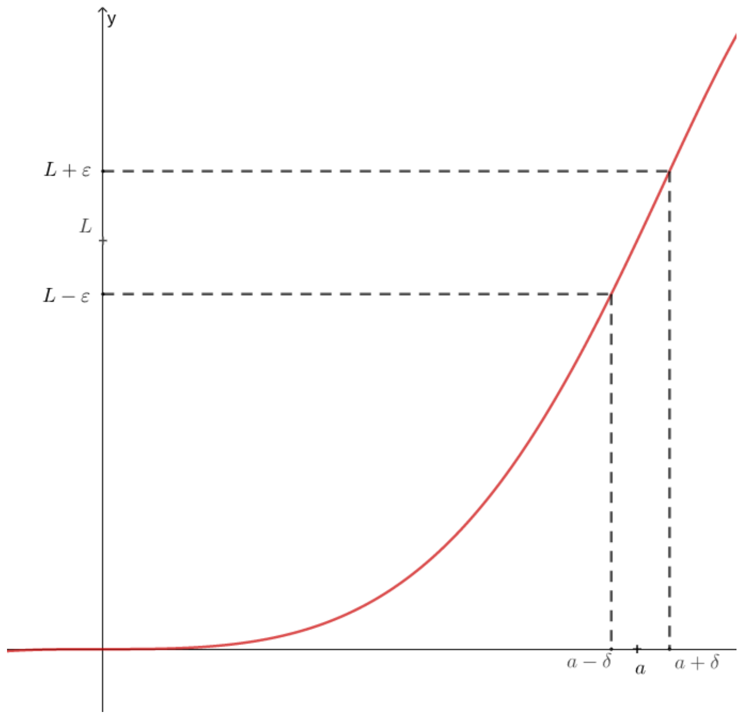
Recall, from single variable calculus, if f is a real-valued function defined on some subset of the real line, we write

$$\lim_{t \to a} f(t) = L$$

if, for any $\varepsilon > 0$, there exists some $\delta > 0$ such that, for any t in the domain of f, if $0 < |t - a| < \delta$, then

$$|f(t) - L| < \varepsilon\,.$$

(we can make $f(t)$ arbitrarily close to L given t is sufficiently close to a but not a itself.)



We use an analogous definition of a limit of a function of two variables. Instead of using the distance formula for the real line (the absolute value), we use the distance formula for points in the plane.

Definition 3.12. Let (a, b) be a point in the domain of a function of two variables, f. The **δ-disk around** (a, b) is the set

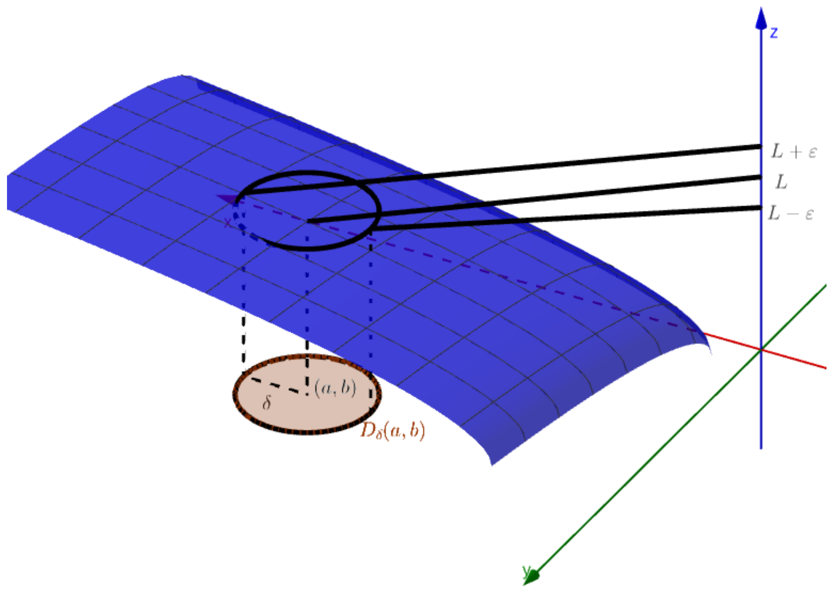
$$D_\delta(a, b) := \{(x, y) \mid \sqrt{(x-a)^2 + (y-a)^2} < \delta\}.$$

Now, we say the **limit of f at** (a, b) **is** L, written

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

provided we can make $f(x, y)$ as close to L as we want, given we make (x, y) sufficiently close to (a, b) but not (a, b) itself. That is, the limit of f at (a, b) is L, if, for any $\varepsilon > 0$, there is some $\delta > 0$ such that, for any (x, y) in the domain of f, if $0 < \sqrt{(x-a)^2 + (y-a)^2} < \delta$, then

$$|f(x, y) - L| < \varepsilon\,.$$



To evaluate limits, see, e.g., Theorem 4.1 and Example 4.8 in [**OS**]. Recall a function of one variable has a jump discontinuity at a point if its left and right hand limits don't coincide there. The situation in several variables is considerably different, because there are now infinitely many ways to approach a point. (See exercise 53 and 55 in the problem set for more precise notions.)

Example 3.13. Determine whether the following limits exist.

(i) [**S** Exercise 14.2.9]

$$\lim_{(x,y)\to(0,0)} \frac{x^4-4y^2}{x^2+2y^2}$$

If we approach $(0,0)$ along the path $x=0$, we obtain

$$\lim_{(x,y)\to(0,0)} \frac{x^4-4y^2}{x^2+2y^2} = \lim_{y\to 0} \frac{0-4y^2}{0+2y^2} = -2,$$

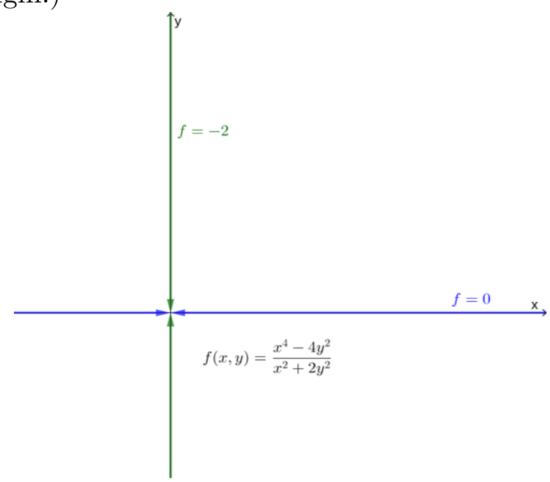
while along the path $y=0$, we obtain

$$\lim_{(x,y)\to(0,0)} \frac{x^4-4y^2}{x^2+2y^2} = \lim_{x\to 0} \frac{x^4-0}{x^2+0} = 0.$$

Since, along two different paths, we obtain different limits, we must conclude

$$\lim_{(x,y)\to(0,0)} \frac{x^4 - 4y^2}{x^2 + 2y^2}$$

does not exist. (See the contrapositive statement of exercise 55 in the problem set. Also, graph this function using technology, and zoom in to the origin.)



(ii) [**S** Exercise 14.2.16]

$$\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4 + y^4}$$

Along the paths $x = 0, y = 0, y = x$, the limit evaluates to 0. Consider $\alpha > 0$, $k \in \mathbb{R}$ and the path $y = kx^{\alpha}$. Then the limit becomes

$$\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4 + y^4} = \lim_{x\to 0} \frac{k^4 x}{x^{4(1-\alpha)} + k^4} = 0.$$

This is possibly decent evidence that the limit might be zero. But now we have to prove it. Notice, for any (x, y), $y^4 \le x^4 + y^4$, since $x^4 \ge 0$, so that

$$\left| \frac{xy^4}{x^4 + y^4} \right| \le |x|.$$

Hence, given $\varepsilon > 0$, we can take $\delta = \varepsilon > 0$, so that, whenever (x, y) satisfies $0 < \sqrt{x^2 + y^2} < \delta$, we have (since $|x| \le \sqrt{x^2 + y^2} < \delta = \varepsilon$)

$$\left| \frac{xy^4}{x^4 + y^4} \right| < \varepsilon\,.$$

That is,

$$\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4+y^4} = 0.$$

(iii) [**S** Exercise 14.2.18]

$$\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4+y^8}$$

Along all paths $y = kx$ and the path $x = 0$, the limit evaluates to 0. But if $y = \sqrt{x}$, then

$$\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4+y^8} = \lim_{x\to 0^+} \frac{1}{2x} = \infty.$$

Hence, the limit

$$\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4+y^8}$$

does not exist. We see that along any straight line passing through the origin, this function approaches zero, yet the limit doesn't exist. This is indeed a pathological example. □

Definition 3.14. We say a function, f, of two variables is **continuous** at a point (a, b) in its domain if

(i) $f(a, b) \in \mathbb{R}$.

(ii) $\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$.

E.g., If we define

$$f(x,y) = \begin{cases} \frac{xy^4}{x^4+y^4} & \text{for } (x,y) \neq (0,0), \\ 0 & \text{for } (x,y) = (0,0). \end{cases},$$

then f is continuous at the origin by Example 3.13 (ii).

3.3 Partial derivatives

Definition 3.15. If f is a function of two variables with domain D, suppose $(a, b) \in D$ is such that there is some δ-disk around (a, b) entirely contained in D. We can then define a new function of one variable $g(x) = f(x, b)$ with domain $D_\delta(a, b)$. If $g'(a) = \lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h}$ exists, we write

$$\frac{\partial f}{\partial x}(a, b) := f_x(a, b) := \lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h} \qquad \big(= g'(a)\big)$$

and call this real number the **partial derivative of f with respect to x.**

Similarly, the limit

$$\frac{\partial f}{\partial y}(a, b) := f_y(a, b) := \lim_{h\to 0} \frac{f(a, b+h) - f(a, b)}{h},$$

if it exists, is called the **partial derivative of f with respect to y.**

In practice, when taking the partial derivative of f with respect to x, one treats y as a constant and differentiates with respect to x. And vice versa.

Example 3.16. If $f(x, y) = x^3 + \frac{\cos(y)}{x} + e^{2x^2y}$, then

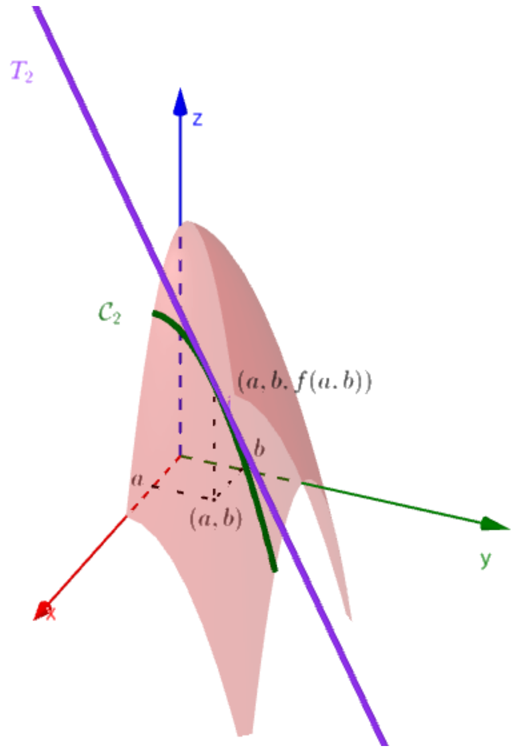
$$f_x(x, y) = 3x^2 - \frac{\cos(y)}{x^2} + 4xye^{2x^2y},$$

and

$$f_y(x, y) = -\frac{\sin(y)}{x} + 2x^2e^{2x^2y}.$$

□

The geometric interpretation of $\frac{\partial f}{\partial x}(a, b)$ and $\frac{\partial f}{\partial y}(a, b)$ is as follows. For example, to interpret $\frac{\partial f}{\partial y}(a, b)$, fixing $x = a$ corresponds to considering the trace of graph f with the plane $x = a$. Suppose we name this intersection curve $\mathcal{C}_2$. That is, $\mathcal{C}_2 = \{(a, y, f(a, y)) \mid y \in \text{domain } f\}$. A parametrization for $\mathcal{C}_2$ is then, evidently, $\mathbf{c}(y) = \langle a, y, f(a, y)\rangle$. So, the tangent line, T_2 to $\mathcal{C}_2$ at $(a, b, f(a, b))$ is given by the parametrization $t\mathbf{c}'(b) + \mathbf{c}(b) = t\left\langle 0, 1, \frac{\partial f}{\partial y}(a, b)\right\rangle + \langle a, b, f(a, b)\rangle$. That is, $\frac{\partial f}{\partial y}(a, b)$ is the slope of the tangent line to the trace of graph f with the plane $x = a$ at $(a, b, f(a, b))$.



Remark 3.17. (a) Partial derivatives for functions of more than two variables are defined similarly: if f is a function of n variables, then, for any $i = 1, 2, \ldots, n$, its partial derivative with respect to x_i at $(a_1, a_2, \ldots, a_n)$ is defined to be

$$\frac{\partial f}{\partial x_i}(a_1, a_2, \ldots, a_n) := \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

if the limit on the right hand side exists.

(b) If, for example, f_x exists in a δ-disk, then we can consider it as a function of two variables, and hence consider its continuity, and whether or not its partial derivatives exist, just as a function of one variable's derivative. Similarly for f_y. More on this later when we talk about higher order partial derivatives.

3.4 The tangent plane and differentiability

(Informal discussion) The tangent line L to a differentiable curve $\mathcal{C}$ at a point P is the best linear approximation to $\mathcal{C}$ at P in the sense that they both pass through P and their tangent vectors agree at P (for a given parametrization of $\mathcal{C}$). (So, near P, $\mathcal{C}$ resembles L and resembles any other line less.)

The tangent *plane* to a differentiable *surface* S at a point P, however defined, should be the best linear approximation to S at P. In this case, what do we mean by best? What does it mean for a surface to be differentiable at a point? Is such a notion necessary for the existence of a tangent plane?

A curve is differentiable at a point if its derivative exists there. This then guarantees the existence of the tangent line. We'll see below that the situation in more than one variable is more subtle. In particular, the existence of partial derivatives only tells part of the story.

Example 3.18. Let

$$f(x,y) = \begin{cases} -\frac{2xy^4}{x^4+y^4} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

We saw, in the example below Definition 3.14, that f is continuous at the origin. Also, a quick application of the definition shows both $f_x(0,0)$ and $f_y(0,0)$ exist and equal 0. E.g.,

$$f_x(0,0) = \frac{f(h,0) - f(0,0)}{h} = \frac{0-0}{h} = 0.$$

Hence, both of the tangent lines to the trace of graph f with the planes $y = 0$ and $x = 0$, respectively, lie in the xy-plane.

However, consider the intersection of graph f with the plane $y = x$. This space curve has parametrization $\mathbf{c}(x) = \langle x, x, -x \rangle$ and hence has as a tangent line parametrization $\mathbf{l}(t) = t\,\langle 1, 1, -1 \rangle$. In particular, this line does not lie in the xy-plane! So, we have three tangent lines passing through the origin, but they don't lie on the same plane. Maybe it's safe to say the tangent plane to graph f doesn't exist at the origin because there's not a reasonable choice for one.

Here's the graph of f, with vertical traces $x = 0$, $y = 0$ and the intersection curve of graph f with the plane $y = x$ highlighted, which also happen to be the tangent lines to these curves. Notice at $(0,0,0)$, this surface is saddle-like, but unlike the surface described in Example 3.9 (b), it has a corner at its saddle point. It would be reasonable if we excluded such surfaces (those with corners) from the class of differentiable ones. □

In these notes we might have a chance to discuss differentiable surfaces in general (at least in the way we defined what a surface is), but it's not likely. For now, we'll focus on differentiable functions, and say the graph of a function is differentiable at a point if the function itself is differentiable there.

Definition 3.19. We say a function of two variables f with domain D is **differentiable at a point**, $P = (x_0, y_0) \in D$, if there exists a δ-disk around P entirely contained in D, and, for any $(x, y) \in D_\delta(P)$, we can write

$$f(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + E(x, y),$$

where the **error term**, E, satisfies

$$\lim_{(x,y)\to(x_0,y_0)} \frac{E(x, y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0.$$

Example 3.20. (a) The function f defined in Example 3.18 is not differentiable at the origin. This can be seen as follows. Since $f(0, 0) = f_x(0, 0) = f_y(0, 0) = 0$, we have $f(x, y) = E(x, y)$. Hence, if f were differentiable at the origin, we'd have

$$\lim_{(x,y)\to(0,0)} -\frac{2xy^4}{(x^4 + y^4)\sqrt{x^2 + y^2}} = 0.$$

However, this is not true, since, in particular, if $y = x$, then

$$\lim_{(x,y)\to(0,0)} -\frac{2xy^4}{(x^4 + y^4)\sqrt{x^2 + y^2}} = \lim_{x\to 0} -\frac{x^5}{x^4|x|} = \pm 1 \neq 0.$$

(b) The function

$$g(x, y) = \begin{cases} \frac{2xy}{\sqrt{x^2+y^2}} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

can be seen to be continuous at the origin. However, it is not differentiable there. Again, $g(0, 0) = g_x(0, 0) = g_y(0, 0) = 0$, so that $g(x, y) = E(x, y)$. Now, since, along $y = x$, we have

$$\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2 + y^2} = \lim_{x\to 0} 1 = 1 \neq 0,$$

we conclude g is not differentiable at $(0, 0)$. □

As we see in the above example, continuity at a point does not imply differentiability there, just as with a single variable. However, just as with a single variable, differentiability does imply continuity.

Theorem 3.21 (**OS** Theorem 4.6)**.** *If a function of two variables, f, is differentiable at a point, P, then f is continuous at P.*

Also, in practice it's hard to use this definition to verify differentiability. Luckily, an easier condition on the partial derivatives guarantees this regularity.

Theorem 3.22 (**OS** Theorem 4.7)**.** *If the partial derivatives f_x and f_y of a function of two variables exist in a small δ-disk around P and are continuous at P, then f is differentiable at P.*

We demonstrated the function in Example 3.20 (a) is not differentiable at the origin. In Example 3.18, we saw that same function's graph has three non-coplanar tangent lines at the origin. In fact, this is not a coincidence given the following definition and theorem.

Definition 3.23. Let S be a surface and P a point on S. Suppose the tangent line, L, to any given curve on S passing through P exists. If all such L lie on the same plane, $\mathcal{P}$, we call $\mathcal{P}$ the **tangent plane** to S at P.

This definition is justified, at least for graphs, by the following

Theorem 3.24. *A function of two variables f is differentiable at a point P if and only if the tangent plane to $\operatorname{graph} f$ exists at P.*

This theorem says the analytic notion of differentiability in Definition 3.19 and the geometric one in 3.6 coincide.

To further discuss linear approximations and tangent planes, it will be useful to obtain a new tool. In fact, this tool is interesting in its own right.

3.5 The chain rule

Recall, if f and g are functions of one variable, g is differentiable at a, and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at t and $(f \circ g)'(a) = f'(g(a))g'(a)$. A similar result holds for functions of more than one variable.

Definition 3.25. Let f be a function of n variables. The **gradient** of f at $P = (a_1, a_2, \dots, a_n)$ is the n dimensional vector

$$\nabla f(P) := \left\langle \frac{\partial f}{\partial x_1}(P), \frac{\partial f}{\partial x_n}(P), \dots, \frac{\partial f}{\partial x_n}(P) \right\rangle$$

given that the partial derivatives of f exist at P. Quite a bit more on this later.

Theorem 3.26. *The chain rule. Suppose a function of two variables f is differentiable at a point (a, b). Further, suppose x, y are both functions of one variable, t, defined in a common interval, and are differentiable at t_0, where $a = x(t_0)$ and $b = y(t_0)$. Then if we define $\mathbf{c} = \langle x, y \rangle$, then $z = f \circ \mathbf{c}$ is differentiable at t_0, and*

$$\begin{aligned}\frac{dz}{dt}(t_0) &= \frac{\partial f}{\partial x}(\mathbf{c}(t_0))\frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(\mathbf{c}(t_0))\frac{dy}{dt}(t_0) \\ &= \nabla f(\mathbf{c}(t_0)) \cdot \mathbf{c}'(t_0)\end{aligned}$$

This then implies, more concisely,

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

as functions.

But writing $\frac{dz}{dt}(t_0) = \nabla f(\mathbf{c}(t_0)) \cdot \mathbf{c}'(t_0)$ is concise too, and shows that this rule is a direct generalization of the chain rule in one variable.

We also have a more general version.

Theorem 3.27. *The chain rule. Suppose a function of two variables f is differentiable at a point (a, b). Further, suppose x, y are both functions of two variables s and t, defined in a common δ-disk, and are differentiable at (s_0, t_0), where $a = x(s_0, t_0)$ and $b = y(s_0, t_0)$. If we define $\mathbf{x} := \langle x, y \rangle$, then $z = f \circ \mathbf{x}$ is differentiable at (s_0, t_0), and*

$$\begin{aligned}\frac{\partial z}{\partial s}(s_0, t_0) &= \frac{\partial f}{\partial x}(\mathbf{x}(s_0, t_0))\frac{\partial x}{\partial s}(s_0, t_0) + \frac{\partial f}{\partial y}(\mathbf{x}(s_0, t_0))\frac{\partial y}{\partial s}(s_0, t_0) \\ &= \nabla f(\mathbf{x}(s_0, t_0)) \cdot \frac{\partial \mathbf{x}}{\partial s}(s_0, t_0)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial z}{\partial t}(s_0, t_0) &= \frac{\partial f}{\partial x}(\mathbf{x}(s_0, t_0))\frac{\partial x}{\partial t}(s_0, t_0) + \frac{\partial x}{\partial y}(\mathbf{x}(s_0, t_0))\frac{\partial y}{\partial t}(s_0, t_0) \\ &= \nabla f(\mathbf{x}(s_0, t_0)) \cdot \frac{\partial \mathbf{x}}{\partial t}(s_0, t_0)\end{aligned}$$

where, e.g., $\frac{\partial \mathbf{x}}{\partial s}(s_0,t_0) := \left\langle \frac{\partial x}{\partial s}(s_0,t_0), \frac{\partial y}{\partial s}(s_0,t_0) \right\rangle$.

Notice how Theorem 3.26 is a special case of Theorem 3.27. Again, Theorem 3.27 implies,

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

as functions. Similarly for $\frac{\partial z}{\partial t}$.

Remark 3.28. If any linear algebra terminology is known to the reader, the two equations in Theorem 3.27 can be written more concisely as

$$\nabla z(s_0,t_0) = \nabla f(\mathbf{x}(s_0,t_0)) \frac{\partial(x,y)}{\partial(s,t)}, \tag{3.1}$$

where, in (3.1),

$$\frac{\partial(x,y)}{\partial(s,t)} = \begin{bmatrix} \frac{\partial x}{\partial t}(s_0,t_0) & \frac{\partial x}{\partial s}(s_0,t_0) \\ \frac{\partial y}{\partial t}(s_0,t_0) & \frac{\partial y}{\partial s}(s_0,t_0) \end{bmatrix}$$

is a 2×2 matrix (called the **Jacobian** of $\mathbf{x}$ at (s_0,t_0)), the gradient vectors are row vectors, and the product on the right hand side is matrix multiplication.

The most general version of the chain rule in these notes is the following

Theorem 3.29. *[**OS** Theorem 4.10] Generalized chain rule. If* $w = f(x_1,x_2,\ldots,x_n)$ *is a differentiable function of* n *variables, and* $x_1,x_2,\ldots,x_n$ *are differentiable functions of* m *variables,* $t_1,t_2,\ldots,t_m$*, then, for any* $i \in \{1,2,\ldots,m\}$*,*

$$\frac{\partial w}{\partial t_i} = \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_i}.$$

See lecture and/or [**OS** section 4.5] for examples.

3.6 The tangent plane and linear approximation

We'll now derive the equation for the tangent plane to a graph of a function of two variables at a point. In fact, we'll find a slightly more general expression.

Suppose we have a level surface described by $F(x,y,z) = k$, whose tangent plane exists at a point, $P = (x_0,y_0,z_0)$. Call this plane T. Let $S = \{(x,y,z) \mid F(x,y,z) = k\}$ be this level surface.

Consider any curve on S passing through P. Suppose we parametrize this curve by $\mathbf{c}(t) = \langle x(t), y(t), z(t) \rangle$ such that $\mathbf{c}(0) = \mathbf{OP}$. Then $\mathbf{c}$ satisfies

$F(\mathbf{c}(t)) = k$ for all t in the domain of $\mathbf{c}$. That is, $(F \circ \mathbf{c})(t) = k$, which implies $(F \circ \mathbf{c})'(0) = 0$. By the chain rule (Theorem 3.29), we have

$$\nabla F(\mathbf{c}(0)) \cdot \mathbf{c}'(0) = 0. \tag{3.2}$$

By the definition of the tangent plane to S at P, $\mathbf{c}'(0)$ lies on T. Hence, if $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$, then $\nabla F(x_0, y_0, z_0)$ is a normal vector to T! (See the discussion at the end of Example 2.15).

So, an equation for the tangent plane to a level surface at a point $P = (x_0, y_0, z_0)$ is

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0.$$

Example 3.30. We can now find a nice description for the tangent plane to the graph of a function of two variables, f. Suppose the tangent plane to the graph of f exists at $(x_0, y_0, f(x_0, y_0))$. Let $F(x, y, z) = z - f(x, y)$. Then $\operatorname{graph} f = \{(x, y, z) \mid F(x, y, z) = 0\}$. That is, $\operatorname{graph} f$ is a level surface of a particular function of three variables. In this case, $F_x = -f_x$, $F_y = -f_y$, $F_z = 1$. Hence, the tangent plane to $\operatorname{graph} f$ at $(x_0, y_0, f(x_0, y_0))$ is described by the equation

$$-f_x(x_0, y_0)(x - x_0) - f_y(x_0, y_0)(y - y_0) + z - f(x_0, y_0) = 0,$$

which is equivalent to

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0), \tag{3.3}$$

which we can also write more concisely as

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (\mathbf{x} - \mathbf{x}_0), \tag{3.4}$$

where $\mathbf{x} = x\mathbf{i} + y\mathbf{j}, \mathbf{x}_0 = x_0\mathbf{i} + y_0\mathbf{j}$. □

Definition 3.31. Let f be a function of two variables, differentiable at (x_0, y_0). The **linear approximation** of f at (x_0, y_0) is the function of two variables given by

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

That is, the graph of L is the tangent plane $\operatorname{graph} f$ at $(x_0, y_0, f(x_0, y_0))$. This makes sense because of Theorem 3.24.

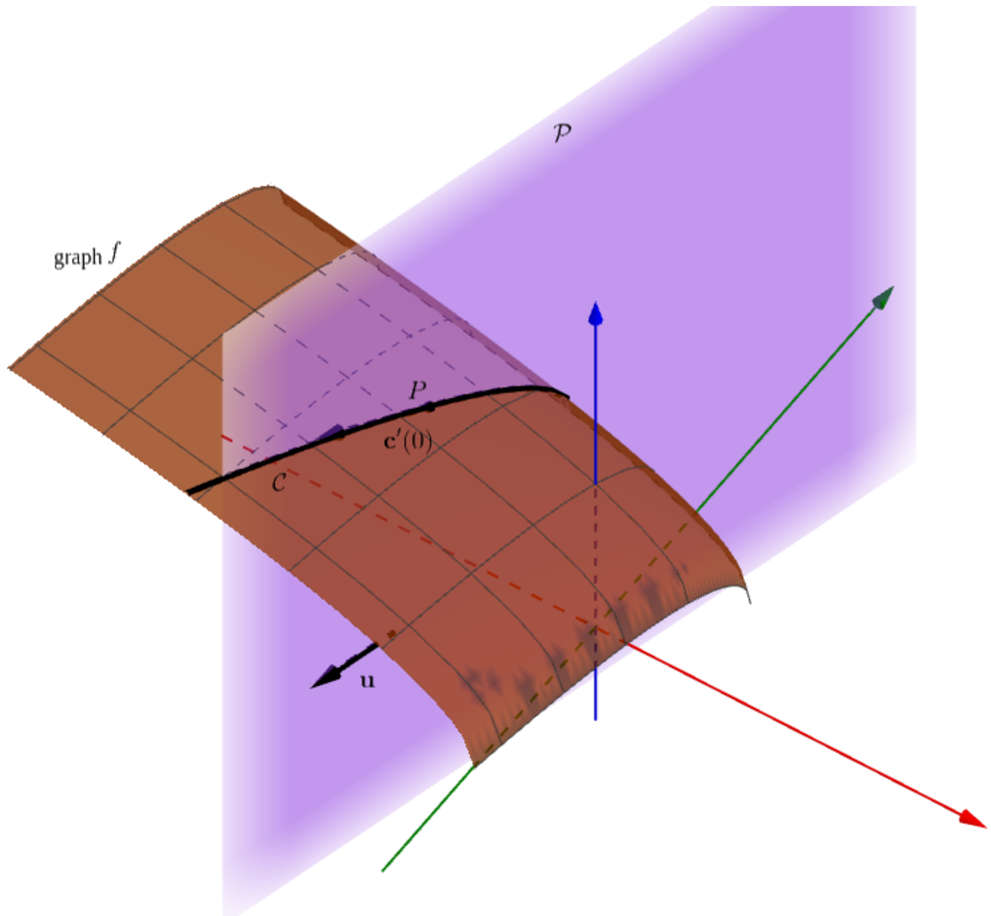
Note, we have $f(x, y) - L(x, y) = E(x, y)$, so that, near (x_0, y_0), $f(x, y) \approx L(x, y)$.

Example 3.32. Suppose we want to estimate $f(0.1, 1.1)$, where $f(x, y) = e^x/y$. We can use the linear approximation of f at $(0, 1)$ to do so. Notice $f(0, 1) = 1$, $f_x(0, 1) = 1$, $f_y(0, 1) = -1$. So, $L(x, y) = 1 + (x - 0) - (y - 1) = 2 + x - y$. Hence, $L(0.1, 1.1) = 1$, so that $f(0.1, 1.1) \approx 1$. □

3.7 Directional derivatives and the gradient vector

Directional derivatives allow us to find the rate of change of a function in any direction, not just coordinate ones.

Definition 3.33. Let f be a function of two variables, differentiable at (x_0, y_0). Fix a unit vector, $\mathbf{u}$ ($\|\mathbf{u}\| = 1$), which lies on the xy-plane. Let $\mathcal{P}$ be the vertical plane in the direction of $\mathbf{u}$, passing through $P = (x_0, y_0, f(x_0, y_0))$. (What's an equation for this plane?) Then the curve, $\mathcal{C}$, parametrized by $\mathbf{c}(s) = s\mathbf{u} + \mathbf{x}_0 + f(s\mathbf{u} + \mathbf{x}_0)\mathbf{k}$ is the intersection of the graph of f with the plane $\mathcal{P}$, where $\mathbf{x}_0 = x_0\mathbf{i} + y_0\mathbf{j}$. The tangent vector to $\mathbf{c}$ at 0 is thus $\mathbf{c}'(0) = \mathbf{u} + (\nabla f(\mathbf{x}_0) \cdot \mathbf{u})\,\mathbf{k}$, where the last expression was found using the chain rule (Theorem 3.26). So, $\nabla f \cdot \mathbf{u}$ is the slope of the tangent line to the intersection of graph f with $\mathcal{P}$.



We say

$$D_{\mathbf{u}}f(x_0, y_0) := \nabla f(x_0, y_0) \cdot \mathbf{u}$$

is the **directional derivative** of f at (x_0, y_0) in the direction of $\mathbf{u}$.

Note that $D_{\mathbf{i}}f = f_x$ and $D_{\mathbf{j}}f = f_y$.

In some sense, if one knows the directional derivatives of a differentiable function, then one knows its rate of change in any direction. Geometrically, this corresponds to the slope of any tangent line to a curve on the graph of a function is given by some directional derivative. We use unit vectors since then $D_{\mathbf{u}}f$ is the rate of change of f per unit change in the direction of $\mathbf{u}$.

This discussion also works in higher dimensions, but we lose the picture.

Theorem 3.34. *Suppose f is a differentiable function of two or three variables. Fix $\mathbf{x}$ and suppose $\nabla f(\mathbf{x}) \neq 0$. The maximum value of $D_{\mathbf{u}}f(\mathbf{x})$ for* *varying unit vectors* $\mathbf{u}$ *is* $\|\nabla f(\mathbf{x})\|$ *and it occurs when* $\mathbf{u} = \frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|}$

Proof. We have

$$D_{\mathbf{u}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} = \|\nabla f(\mathbf{x})\| \|\mathbf{u}\| \cos\theta = \|\nabla f(\mathbf{x})\| \cos\theta,$$

where θ is the angle between $\nabla f(\mathbf{x})$ and $\mathbf{u}$. cos attains its max when $\theta = 0$, i.e., when $\mathbf{u} = \frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|}$. □

Now, from Equation (3.2), the gradient of a function at a point is normal to its level set passing through that point. Notice this fact, along with the above theorem, imply that the curve of greatest change in f intersects every level curve of f perpendicularly.

Example 3.35. (a) Let's find the direction from $P = (1, -1)$ in which $g(x, y) = \frac{x^2}{2} - y^2$ decreases the most.

(b) Let $\mathcal{C}$ be the curve passing through P, orthogonal to every level curve of g. Two curves are orthogonal if their tangent vectors are orthogonal at their point(s) of intersection. In this case, this means direction vectors of tangent lines to $\mathcal{C}$ are parallel to the gradient of g at that point. That is, we can use the direction found in (a) to estimate the intersection of $\mathcal{C}$ with the curve $g = -1$.

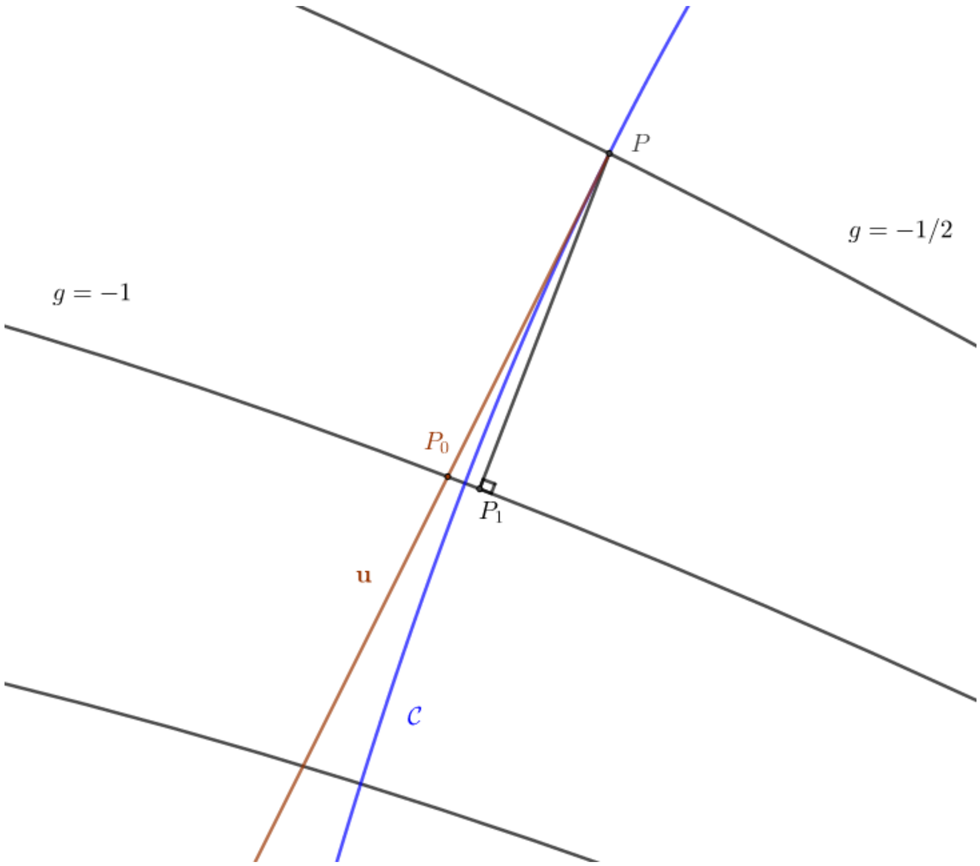
Solution: (a) The direction of greatest decrease of g from P can be seen, from the proof of the above theorem, to be $\mathbf{u} = -\frac{\nabla g(P)}{\|\nabla g(P)\|}$. We have $\nabla g(x, y) = \langle x, -2y\rangle$. Hence,

$$\mathbf{u} = -\frac{1}{\sqrt{5}} \langle 1, 2\rangle .$$

(b) Now, a parametrization of the tangent line to $\mathcal{C}$ at P is then $\mathbf{l}(t) = t\mathbf{u} + \mathbf{OP} = \left\langle -\frac{t}{\sqrt{5}} + 1, -\frac{2t}{\sqrt{5}} - 1\right\rangle$. We can estimate the intersection of $\mathcal{C}$ with $g = 1$ by solving $g(\mathbf{l}(t)) = -1$, which ends up quadratic in t:

$$\frac{\left(1 - \frac{t}{\sqrt{5}}\right)^2}{2} - \left(1 + \frac{2t}{\sqrt{5}}\right)^2 = 1.$$

Taking the positive root: $t_0 = \frac{1}{7}\sqrt{5}(4\sqrt{2} - 5)$ (since $g(P) = -1/2 > -1$, so we want to travel in the direction of $\mathbf{u}$), we find that $P_0 = \left\langle -\frac{t_0}{\sqrt{5}} + 1, -\frac{2t_0}{\sqrt{5}} - 1\right\rangle$ lies on $g = -1$ and the tangent line to $\mathcal{C}$ at P.



Notice the intersection of $\mathcal{C}$ with $g = -1$ is not the same point as P_1 (see figure above, correcting a point made in lecture), the point on $g = -1$ closest to P. (See Example 3.44 below). □

We now look at a major application of the differential calculus in two and three variables.

3.8 Optimization

Definition 3.36. Suppose f is defined in a δ-disk around $P = (x_0, y_0)$. If $f(x, y) \leq f(x_0, y_0)$ for all $(x, y) \in D_\delta(P)$, then we say f **has a** or **attains a local maximum** at (x_0, y_0) and $f(P)$ is a **local maximum value** of f. Similarly, if $f(x, y) \geq f(x_0, y_0)$ for all $(x, y) \in D_\delta(P)$, then we say f **has a local minimum** at (x_0, y_0) and $f(P)$ is a **local minimum value** of f.

(x_0, y_0) is a **global maximum/minimum** if the δ-disk above is replaced with the domain of f.

Theorem 3.37 (**OS** Theorem 4.16)**.** *Fermat's Theorem for functions of two variables. If f is a function of two variables defined in a δ-disk around (x_0, y_0), $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exist, and f has a local extremum(max/min) at (x_0, y_0), then $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$. That is,*

$$\nabla f(x_0, y_0) = 0.$$

The idea behind this result is intuitive. If (x_0, y_0) is a local extrema but $\nabla f(x_0, y_0) \neq 0$, then going in the coordinate direction from (x_0, y_0) where the partial of f doesn't vanish would cause either an increase or decrease in f, contradicting the extreme nature of $f(x_0, y_0)$.

Definition 3.38. (x_0, y_0) is a **critical point** of a function of two variables, f, if one of the following hold:
(i) $\nabla f(x_0, y_0) = 0$
(ii) Either $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ does not exist.

Recall from [**OS** pages 375 - 377] the definition of higher order partials and Clairaut's Theorem for mixed partials: If f_{xy} and f_{yx} are continuous on some disk, then $f_{xy} = f_{yx}$ there.

Theorem 3.39. *The Second Derivative Test. Suppose the second partials of f are continuous on a δ-disk around (x_0, y_0) and that (x_0, y_0) is a critical point of f (so that $\nabla f(x_0, y_0) = 0$ in this case). Define*

$$D := D(x_0, y_0) := f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - (f_{xy}(x_0, y_0))^2,$$

the discriminant of f at (x_0, y_0).
(a) If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then $f(x_0, y_0)$ is a local minimum value.
(b) If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then $f(x_0, y_0)$ is a local maximum value.
(c) If $D < 0$, then $f(x_0, y_0)$ is not a local extremum value. Here, (x_0, y_0) is called a saddle point of graph f.
(d) If $D = 0$, then the test is inconclusive. f could have some local extrema at (x_0, y_0), or (x_0, y_0) could be a saddle point.

The proof of this theorem relies on approximating f up to second order by Taylor's Theorem. (which is why D is called the discriminant). You can find a proof with a quick web search of "proof of the second derivative test".

Example 3.40. See [**OS** Example 4.39]. Maybe an original one will be added here.

Remark 3.41. (a) Here we remark another difference between one and several variables. It is possible for a function of two or more variables to attain more than one local max without possessing a local min. This can't happen in one variable. For an example, try graphing the function $f(x, y) = 10x^2y - 5x^2 - 4y^2 - x^4 - 2y^4$ or see its graph here. This example is from [**S** page 1002 Example 4]. How many saddle points does graph f have?

(b) From Fermat's Theorem and the Extreme Value Theorem (see problem set Exercise 56 or [**OS** Theorem 4.18 pg 445]), we now have a systematic way of finding the global extrema of a continuous function of two variables, f, defined on a closed and bounded domain, D:

Step 1: Find values of f at critical points in the interior of D.

Step 2: Find the maximum and minimum of f on the boundary of D.

Step 3: The largest from steps 1,2 is the global maximum value of f, while the smallest is its global minimum.

Example 3.42. See [**OS** Example 4.40]. Maybe an original one will be added here.

3.8.1 Lagrange multipliers

The method of Lagrange multipliers gives us a concise way of finding extrema of a function of two or three variables, f, when the domain of f is restricted to some level set of another function, g. The idea is the following. Suppose f and g are both functions of two variables and we restrict the domain of f to a level curve of g, described by, say, $g = k$. Suppose (x_0, y_0) is a point on this level curve where f attains an extreme value, say $f(x_0, y_0) = m$. Then the curve $g = k$ must be tangent (just touch) $f = m$ at (x_0, y_0), since otherwise there's a direction on $g = k$ that either increases or decreases f at (x_0, y_0).

See [**OS** Figure 4.60] and figure 1 here. In particular, the normal lines of $g = k$ and $f = m$ must coincide. So, from the discussion at the beginning of Section 3.6, if $\lambda\nabla g(x_0, y_0) \neq 0$, we must have $\nabla f(x_0, y_0) = \lambda\nabla g(x_0, y_0)$ for some $\lambda \in \mathbb{R}$, called a *Lagrange multiplier*. A similar argument can be made for functions of three variables as well.

Theorem 3.43. *Method of Lagrange Multipliers. Let f and g be differentiable functions of three variables. To find the extrema of f (assuming they exist) subject to the constraint $g = k$ (where $\nabla g(x, y, z) \neq 0$ when $g(x, y, z) = k$), it's enough to find all (x, y, z, λ) such that*

(1) $\nabla f(x, y, z) = \lambda\nabla g(x, y, z)$
(2) $g(x, y, z) = k$.

Notice (1) and (2) describe 4 unknowns within 4 equations. See the proof of [**OS** Theorem 4.20].

Example 3.44. Referring back to Example 3.35, we can find the point on $g = -1$ closest to the point $P = (1, -1)$ using the method of Lagrange multipliers.

Here we want to minimize the function $f(x, y) = (x - 1)^2 + (y + 1)^2$, which is the distance between (x, y) and P squared. We want to find all (x, y, λ) such that

$\nabla f(x, y) = \lambda\nabla g(x, y)$
$g(x, y) = -1$.

That is, all (x, y, λ) such that

(1) $2(x - 1) = \lambda x$
(2) $2(y + 1) = -2\lambda y$
(3) $\frac{x^2}{2} - y^2 = -1$.

Notice first (2) implies $y \neq 0$ since otherwise we'd have $2 = 0$, a contradiction. So (2) also implies $\lambda = -1 - \frac{1}{y}$. This implies, with (1), that $2(x - 1) = -x - \frac{x}{y}$, which further implies

$$y = -\frac{x}{3x - 2}.$$

Using this with (3) gives us

$$\frac{x^2}{2} - \left(\frac{x}{3x - 2}\right)^2 = -1.$$

That is,

$$9x^4 - 12x^3 + 20x^2 - 24x + 8 = 0.$$

This quartic polynomial has two real roots, $x \approx 0.50745$ and $x \approx 0.92463$, corresponding then to $y \approx 1.0629$ and $y \approx -1.19478$, respectively, using the expression for y in terms of x above. These decimal representations have closed form solutions that the reader can find with their favorite computer algebra system. It turns out, that if P_1 denotes the point on $g = -1$ closest to P, then $P_1 \approx (0.92463, -1.19478)$, which also makes sense by considering that $(0.50745, 1.0629)$ is closer to the upper component of the hyperbola $g = -1$. See the figure in Example 3.35. The distance from P to $g = -1$ is thus $\sqrt{f(P_1)} \approx 0.208854$.

Notice, though, this does not imply that the other point obtained, $\approx (0.50745, 1.0629)$ is the point on $g = -1$ farthest away from P. In fact, $g = -1$ is unbounded, so there are points on this curve that are arbitrarily far away from P.

This illustrates the importance of the hypothesis of the method. In Theorem 3.43, it is *assumed* that extrema of f exist on $g = k$. But in this problem, we didn't establish any conditions on f or g that guarantee the existence of such points. See the extreme value theorem or here for a sufficient condition.

□

Example 3.45 (**S** page 1013)**.** Let's try to use this method to maximize the volume $V = xyz$ of a box with an open top, subject to a fixed surface area:

$$2xz + 2yz + xy = 12.$$

Where x is the length, y the width, and z the height of the box.

Let $g(x, y, z) = 2xz + 2yz + xy$ and $f(x, y, z) = V = xyz$. Then we want to find all (x, y, z, λ) such that

$$(1) \quad yz = \lambda(2z + y)$$

$$(2) \quad xz = \lambda(2z + x)$$

$$(3) \quad xy = \lambda(2x + 2y)$$

$$(4) \quad 2xz + 2yz + xy = 12$$

We solve (1)-(3) for λ, to obtain

$$(5) \qquad \lambda = \frac{yz}{2z+y}$$

$$(6) \qquad \lambda = \frac{xz}{2z+x}$$

$$(7) \qquad \lambda = \frac{xy}{2x+2y}$$

Notice $z > 0$, since otherwise $z = 0$, so $V = 0$, which certainly isn't a maximum. Hence, (5) and (6) imply

$$\frac{y}{2z+y} = \frac{x}{2z+x},$$

which implies, seen after some rearranging, that

$$y = x.$$

That is, the base should be a square.

Now, $y = x$, along with (6) and (7), imply

$$\frac{x^2}{4x} = \frac{xz}{2z+x}.$$

Hence,

$$(8) \qquad 2xz = x^2.$$

Now, $y = x$ along with (4) implies

$$4xz + x^2 = 12,$$

which along with (8), gives

$$3x^2 = 12.$$

We finally arrive at $x = y = 2$, since $x, y > 0$. Using (8) again, we get $z = 1$.

So, through this method we end up with one point:

$$(2, 2, 1).$$

At this point, $V = 4$. We can also check this point satisfies the constraint.

□

This example has a similar problem as the last example. For one thing, since we only have one point, we don't know if this point is a maximum or a minimum. Worse, though, as in the above example, is that we have no reason that any extrema of f restricted to g exist. The reason, again, is that the surface described by $g = 12$ is not bounded, so we can't rely on the extreme value theorem.

However, in this problem, using the constraint, we can reduce V to a function of two variables and apply the second derivative test to check that, indeed, $V = xyz$ has a maximum at $(2, 2, 1)$ when $2xz + 2yz + xy = 12$.

Let's try an example where we have all of the tools to successfully apply this method.

Example 3.46 (**S** exercise 14.8.23)**.** Find the global extrema of $f(x, y) = e^{-xy}$ on $D = \{(x, y) \mid x^2 + 3y^2 \leq 1\}$.

Notice f is continuous on D and D is closed and bounded. Hence, we can follow the steps in Remark 3.41 (b).

Step 1. Find the critical points in the interior of D, which is

$$D^\circ = \{(x, y) \mid x^2 + 3y^2 < 1\}.$$

We have

$$f_x = -ye^{-xy}, f_y = -xe^{-xy}.$$

Hence, $f_x = f_y = 0$ if and only if $x = y = 0$. $(0, 0) \in D^\circ$. We conclude the only critical point of f in the interior is the origin and $f(0, 0) = 1$.

Step 2. Now we find the extrema of f subject to $x^2 + 3y^2 = 1$, which describes the boundary of D.

Let $g(x, y) = x^2 + 3y^2$.

We want to find all (x, y, λ) such that

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$
$$g(x, y) = 1.$$

That is,

$$(1) \quad -ye^{-xy} = 2\lambda x$$
$$(1) \quad -xe^{-xy} = 6\lambda y$$
$$(3) \quad x^2 + 3y^2 = 1$$

Multiplying (1) by $3y$ and (2) by x, we find

$$-3y^2e^{-xy} = 6\lambda xy$$

$$-x^2e^{-xy} = 6\lambda xy.$$

Hence, $3y^2 = x^2$, so that, by (3), $2x^2 = 1$. That is, $x = \pm\frac{1}{\sqrt{2}}$. Hence, since $3y^2 = x^2$, we have that $y = \pm\frac{1}{\sqrt{6}}$. Thus, we have found four points:

$$\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{6}}\right).$$

We have $f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{6}}\right) = f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}\right) = e^{-\frac{1}{\sqrt{12}}}$. And $f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}\right) = f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{6}}\right) = e^{\frac{1}{\sqrt{12}}}$. Now notice, the boundary of D is an ellipse, which is closed and bounded in the plane (its boundary is empty). So, f attains extrema there. The method of Lagrange multipliers gives us the candidates for where these extrema occur. Hence, we have

Step 3. We conclude, since $e^{-\frac{1}{\sqrt{12}}} < 1 = e^0 < e^{\frac{1}{\sqrt{12}}}$, that f attains a global maximum value of

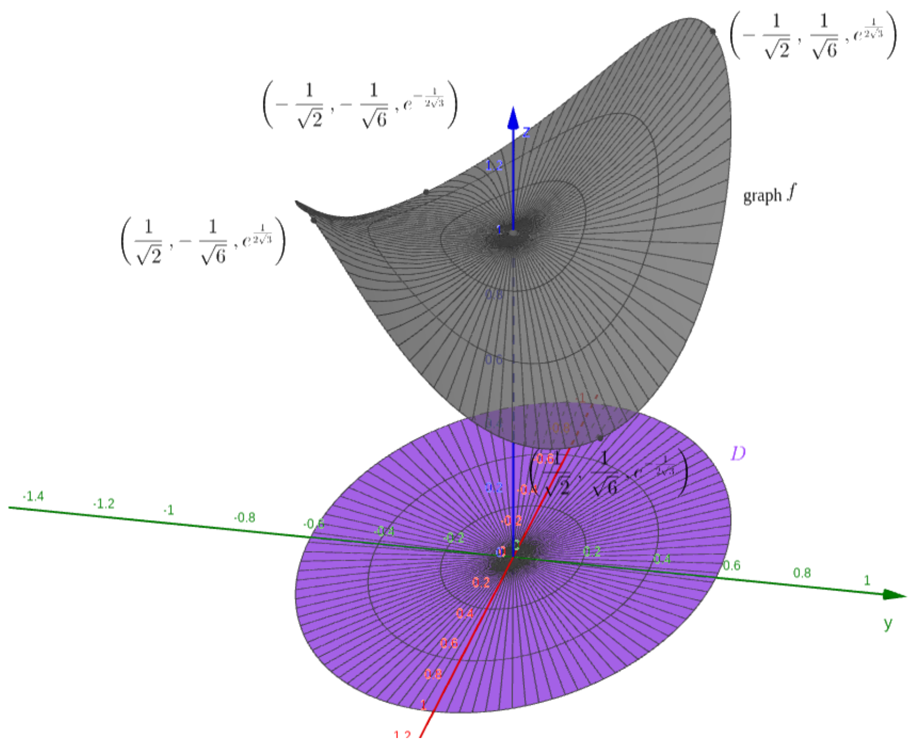
$$e^{\frac{1}{\sqrt{12}}}$$

at the points $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}\right)$ and $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{6}}\right)$ on D. And f attains a global minimum value of

$$e^{-\frac{1}{\sqrt{12}}}$$

at $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{6}}\right)$ and $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}\right)$ on D.

In fact, a quick application of the second derivative test of f at $(0, 0)$ reveals that the origin is in fact a saddle point of graph f, which can also be seen from the figure below.



$\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}, e^{\frac{1}{2\sqrt{3}}}\right)$
$\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{6}}, e^{-\frac{1}{2\sqrt{3}}}\right)$
$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{6}}, e^{\frac{1}{2\sqrt{3}}}\right)$
graph f
$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}, e^{-\frac{1}{2\sqrt{3}}}\right)$
D
y

□

Bibliography

[**OS**] OpenStax, Calculus Volume 3. OpenStax CNX. May 30, 2018 http://cnx.org/contents/a31cd793-2162-4e9e-acb5-6e6bbd76a5fa@2.103.

[**S**] Stewart, James. Multivariable Calculus. 8th ed., Cengage Learning, 2016.

The first four figures are based off of this answer: https://tex.stackexchange.com/a/164162 using pgfplots and tikz. The rest are made using the graphing and 3d graphics calculators here: www.geogebra.org.