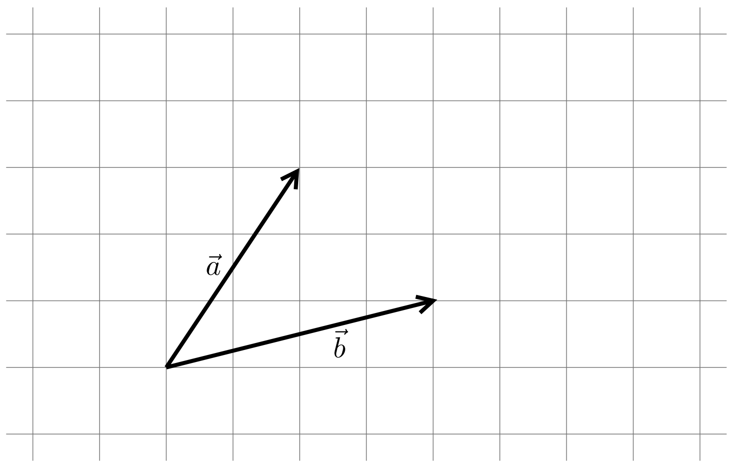
Homework

Math 247 – Honors Calculus 3

Fall 2015 – Jochen Denzler

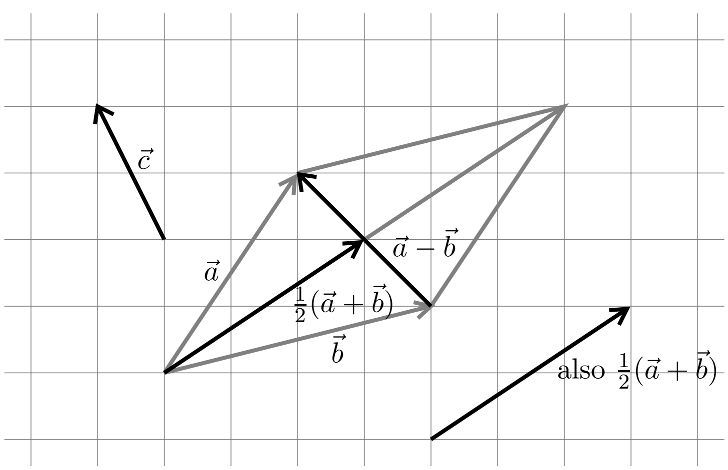
Hwk #1:

The geometric vectors $\vec{a}$ and $\vec{b}$ are as given in the figure below. The coordinate lines drawn are at unit distances apart. The vector $\vec{c}$ is given in coordinates: $\vec{c} = [-1\,,\ 2]^T$.



(a) Draw $\vec{a} - \vec{b}$ and $\frac{1}{2}(\vec{a} + \vec{b})$ into the figure.
(b) Draw $\vec{c} = [-1\,,\ 2]^T$ into the same figure.
(c) Find the coordinates of $\vec{a}$ and $\vec{b}$, and calculate their dot product.
(d) Find the angle between $\vec{a}$ and $\vec{b}$ to a numerical precision of $1/100^{\text{th}}$ of a degree. Since you cannot read off the coordinates from a picture to sufficient precision, I'll tell you that the coordinates of $\vec{a}$ and $\vec{b}$ are indeed intended to be precise integers.

Solution:



$\vec{a} = [2\,,\ 3]^T$

$\vec{b} = [4\,,\ 1]^T$

$\vec{a} \cdot \vec{b} = 2 \cdot 4 + 3 \cdot 1 = 11$

$$\cos \angle(\vec{a}, \vec{b}) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|\,\|\vec{b}\|} = \frac{11}{\sqrt{2^2 + 3^2}\,\sqrt{4^2 + 1^2}} = \frac{11}{\sqrt{13 \cdot 17}}$$

Pocket calculator: $\varphi = \arccos \frac{11}{\sqrt{221}} = 0.73782\text{rad} = 42.27^\circ$

Note that there is no origin in this figure, and none is needed. All vectors could have been placed elsewhere, as has been done for illustration by putting a copy of $\frac{1}{2}(\vec{a} + \vec{b})$ in a second place.

Hwk #2:

(a) Show that $2\|\vec{a}\|^2 + 2\|\vec{b}\|^2 = \|\vec{a} - \vec{b}\|^2 + \|\vec{a} + \vec{b}\|^2$. Draw a figure for illustration, and see why this formula is called the parallelogram identity.

(b) The dot product can be reconstructed from the norm: Indeed, show that $\vec{x} \cdot \vec{y} = \frac{1}{4}(\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2)$.

Solution: Its easier to start evaluating these equalities from the right hand side:

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

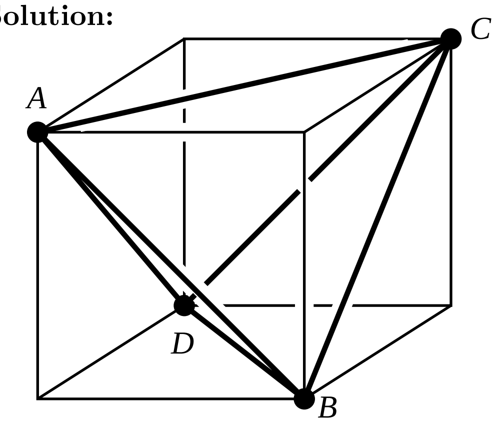
$$\begin{aligned}
\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2 &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) - (\vec{x} - \vec{y}) \cdot (\vec{x} - \vec{y}) \\
&= (\vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{x} + \vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y}) - (\vec{x} \cdot \vec{x} - \vec{y} \cdot \vec{x} - \vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y}) \\
&= 2\vec{y} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} \\
&= 4\vec{x} \cdot \vec{y}
\end{aligned}$$

Hwk #3:

Let A, B, C, D be the vertices of a regular tetrahedron (a polyhedron whose four faces are congruent equilateral triangles). Let O be the center of this tetrahedron. Find the angle AOB. Interpretation: Chemists are interested in this angle b/c methane has a carbon atom in the center and a hydrogen atom at each of the vertices of the tetrahedron.

Hint: Take a cube, whose center is the origin, with the axes of a coordinate system parallel to the sides of the cube; choose four of its eight vertices in a checkerboard manner: if one vertex of the cube is chosen, then the immediately adjacent ones are not and vice versa. The chosen vertices form the corners of a tetrahedron. Draw a figure. It is easy to calculate (eg) the dot product $\vec{OA} \cdot \vec{OB}$ in terms of coordinates. The formula for this same dot product in terms of norms and angles can then be used to find the angle.

Solution:



Assume the sides to be of length 2 and the origin O in the center of the cube, the x axis horizontal, the z axis vertical in the drawing plane, and the y axis perpendicular into the drawing plane. Then the coordinates are as follows:
$A = (-1, -1, 1)$, hence $\vec{a} := \vec{OA} = [-1, -1, 1]^T$
$B = (1, -1, -1)$, hence $\vec{b} := \vec{OB} = [1, -1, -1]^T$
$C = (1, 1, 1)$
$D = (-1, -1, 1)$

Let φ be the angle AOB, which is the angle between $\vec{a}$ and $\vec{b}$. Then $\vec{a} \cdot \vec{b} = \|\vec{a}\| \, \|\vec{b}\| \, \cos\varphi$. Specifically, $-1 = \sqrt{3}\,\sqrt{3}\, \cos\varphi$, hence $\varphi = \arccos(-1/3) \approx 1.9106\text{rad} \approx 109.47°$.

Hwk #4:

(a) By copycating the proof for the Cauchy-Schwarz inequality for vectors in $\mathbb{R}^n$, proof that for any two continuous functions f, g on the interval $[a,b]$, the inequality

$$\left|\int_a^b f(x)g(x)dx\right| \le \sqrt{\int_a^b f(x)^2\,dx}\,\sqrt{\int_a^b g(x)^2\,dx}$$

is true. (This inequality is called the Cauchy-Schwarz inequality for functions. It should be viewed as analogous to the CS inequality for vectors: $|\vec{f}\cdot\vec{g}| \le \|\vec{f}\|\,\|\vec{g}\|$)

(b) To begin appreciating the benefit of the inequality, find out what it tells you specifically about the integral $\int_{\pi/2}^{\pi} \frac{\sin x}{x}\,dx$, an integral you will not be able to do by means of anti-derivatives. Use the CS inequality first for $f(x) = \sin x$ and $g(x) = \frac{1}{x}$, then for $f(x) = \sqrt{\frac{\sin x}{x}}$, $g(x) = \sqrt{x\,\sin x}$ to get a 2-sided estimate for the 'difficult' integral. (This technique doesn't always give numerically as good estimates as in this particular example.)

Note: The analogy between vectors and functions that is exploited here is studied more generally and systematically in a linear algebra course under the headings 'Abstract real vector spaces' and 'Inner product spaces'

Solution: (a) If g is the constant 0 on $[a,b]$, the inequality is obviously true. We now assume that g is not identically 0, and therefore $\int_a^b < g(x)^2\,dx > 0$. Consider the quantity $\int_a^b (f(x)+tg(x))^2\,dx$, which is ≥ 0 for any real number t. We will later choose t conveniently. Expanding the square, we obtain

$$0 \le \int_a^b f(x)^2\,dx + 2t\int_a^b f(x)g(x)\,dx + t^2\int_a^b g(x)^2\,dx$$

If we now let $t := -\int_a^b f(x)g(x)\,dx / \int_a^b g(x)^2\,dx$ (which is possible since $\int_a^b g(x)^2\,dx > 0$), we obtain

$$0 \le \int_a^b f(x)^2\,dx - \frac{(\int_a^b f(x)g(x)\,dx)^2}{\int_a^b g(x)^2\,dx}$$

Clearing the denominator, moving the negative term over, and taking the square root proves the claim.

(b) For $f(x) = \sin x$ and $g(x) = \frac{1}{x}$, we obtain

$$\left|\int_{\pi/2}^{\pi} \frac{\sin x}{x}\,dx\right| \le \sqrt{\int_{\pi/2}^{\pi} \sin^2 x\,dx}\,\sqrt{\int_{\pi/2}^{\pi} \frac{dx}{x^2}}$$

With

$$\int_{\pi/2}^{\pi} \sin^2 x\,dx = \frac{1}{2}\left[x - \sin x\,\cos x\right]_{\pi/2}^{\pi} = \frac{\pi}{4}$$

and

$$\int_{\pi/2}^{\pi} \frac{dx}{x^2} = \frac{1}{\pi}$$

we conclude $\int_{\pi/2}^{\pi} \frac{\sin x}{x}\,dx \le \frac{1}{2}$.

Next, with $f(x) = \sqrt{\frac{\sin x}{x}}$ and $g(x) = \sqrt{x\,\sin x}$, we conclude

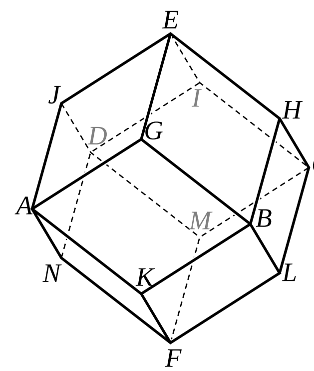
$$\left|\int_{\pi/2}^{\pi} \sin x\,dx\right| \le \sqrt{\int_{\pi/2}^{\pi} \frac{\sin x}{x}\,dx}\,\sqrt{\int_{\pi/2}^{\pi} x\sin x\,dx}$$

Using $\int_{\pi/2}^{\pi} x \sin x \, dx = \pi - 1$, and $\int_{\pi/2}^{\pi} \sin x \, dx = 1$, we conclude after squaring that $\int_{\pi/2}^{\pi} \frac{\sin x}{x} \, dx \geq 1/(\pi - 1)$.

Conclusion: $0.4669 \leq \int_{\pi/2}^{\pi} \frac{\sin x}{x} \, dx \leq 0.5$.

Hwk #5:

A rhombus is a quadrangle whose sides are all of the same length. A rhombododekahedron is a polyhedron with 12 faces, all of which are congruent rhombi. At each vertex, either four rhombi meet with their acute angles, or else, three rhombi meet with their obtuse angles. See the figure.
This does not work with an arbitrary rhombus. If you make the acute angle smaller (thus making the obtuse angle larger), the 'crown' above the zigzag $AGBHCIDJA$ becomes skinnier and taller, and the obtuse angle of the rhombus will become too large to fit below the crown as angle AGB.



Your job is to find the correct angle for a rhombus that is fit to build a rhombododekahedron. (No calculus here; just training your vector geometry and spatial vision a bit more.)

Hint: First choose a convenient cartesian coordinate system. Then set up equations describing that certain lengths are equal, to get coordinates of other points needed, then you can calculate the desired angle using the dot product.

Solution: Put the origin in the center of the square $ABCD$ and the x axis parallel to AB, the y axis parallel to BC, and the z axis passing through E and F. With a convenient choice of the unit length, we have $A = (-1, -1, 0)$, $B = (1, -1, 0)$, $C = (1, 1, 0)$, $D = (-1, 1, 0)$.

Then $E = (0, 0, e)$ and $F = (0, 0, -e)$ with e yet to be determined. (If you see that $EBFD$ should be a square as well, you know e without calculation, but I proceed to explain it assuming you don't see this.) The congruence of the rhombi requires that $\vec{EB}$ has the same length as $\vec{AB}$, namely 2. So $1^2 + 1^2 + e^2 = 2^2 + 0^2 + 0^2$. Hence $e = \sqrt{2}$.

Next $G = (0, -1, g)$ with g yet to be determined. Since the length of $\vec{AG}$ must be the same as the length of $\vec{EG}$, we conclude $1^2 + 0^2 + g^2 = 0^2 + 1^2 + (g - \sqrt{2})^2$, and therefore $g = \sqrt{2}/2$.

Now for the angle φ between the vectors $\vec{AG}$ and $\vec{AK}$, we have

$$\cos\varphi = \vec{AG} \cdot \vec{AK} / \|\vec{AG}\| \, \|\vec{AK}\| = \begin{bmatrix} 1 \\ 0 \\ \sqrt{2}/2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -\sqrt{2}/2 \end{bmatrix} / (1^2 + 0^2 + (\sqrt{2}/2)^2) = \frac{1/2}{3/2} = \frac{1}{3} \, .$$

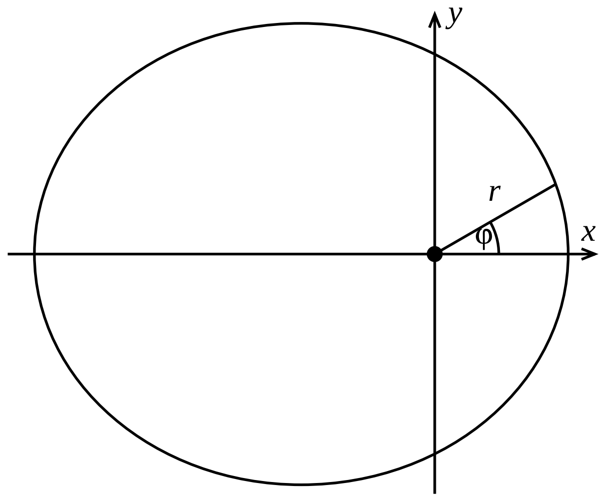
Solution 2: A different solution, choosing another coordinate system: Let's put the origin of the coordiante system again in the center of the square $ABCD$, but this time let's have the axes diagonally: the x-axis goes through DB, the y-axis through AC. So, with a convenient unit length (different than the one before) we have $A = (0, -1, 0)$, $C = (0, 1, 0)$ and $D = (-1, 0, 0)$, $B = (1, 0, 0)$. Similarly, since the z axis would now be FE, we have $F = (0, 0, -1)$, $E = (0, 0, 1)$. With G being verticaly above the midpoint of AB, its coordinates are $G = (\frac{1}{2}, -\frac{1}{2}, g)$. For the length of $\vec{AG}$ to equal the length of $\vec{EG}$, we need $(\frac{1}{2})^2 + (\frac{1}{2})^2 + g^2 = (\frac{1}{2})^2 + (-\frac{1}{2})^2 + (1 - g)^2$. This implies that $g = \frac{1}{2}$. Of course

$K = (\frac{1}{2}, -\frac{1}{2}, -g)$. Again, we can get for the angle φ between the vectors $\vec{AG}$ and $\vec{AK}$ that

$$\cos\varphi = \vec{AG}\cdot\vec{AK}/\|\vec{AG}\|\,\|\vec{AK}\| = \begin{bmatrix} 1/2 \\ 1/2 \\ g \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \\ -g \end{bmatrix} \Big/ \Big((\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 + g^2\Big) = (\tfrac{1}{2} - g^2)/(\tfrac{1}{2} + g^2) = \tfrac{1}{4}/\tfrac{3}{4} = \tfrac{1}{3}$$

Hwk #6:

In astronomy, it is convenient to describe an ellipse (like, e.g., the orbit of the earth around the sun) in polar coordinates (with the sun at the origin, and the positive x axis through the perihelion (the point on the orbit of the earth that is closest to the sun). See figure (not to scale).



In polar coordinates, the orbit is given by $r = r_0/(1 + \varepsilon\cos\varphi)$, where $\varepsilon \in \,]0, 1[$ is called the eccentricity (a measure how much the ellipse deviates from circular shape). Obtain the equation in cartesian coordinates (x, y). The answer should be in the form (you fill in the '?').

$$\frac{(x-?)^2}{?^2} + \frac{y^2}{?^2} = 1$$

Solution:

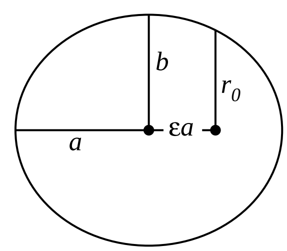
The connection between polar and cartesian coordinates is given by $x = r\cos\varphi$, $y = r\sin\varphi$. Therefore $\cos\varphi = x/r$ and $r^2 = x^2 + y^2$.

We write $r = r_0/(1 + \varepsilon\cos\varphi)$ as $r + r\varepsilon\cos\varphi = r_0$, or equivalently, $r = r_0 - \varepsilon x$. Squaring produces $x^2 + y^2 = (r_0 - \varepsilon x)^2$. Gathering like terms gives us $x^2(1 - \varepsilon^2) + 2r_0\varepsilon x + y^2 = r_0^2$.

After completing the square for x, we get $(1 - \varepsilon^2)(x + \frac{r_0\varepsilon}{1-\varepsilon^2})^2 + y^2 = r_0^2 + r_0^2\frac{\varepsilon^2}{1-\varepsilon^2} = r_0^2/(1 - \varepsilon^2)$.

This can be written as

$$\frac{(x - x_0)^2}{a^2} + \frac{y^2}{b^2} = 1$$



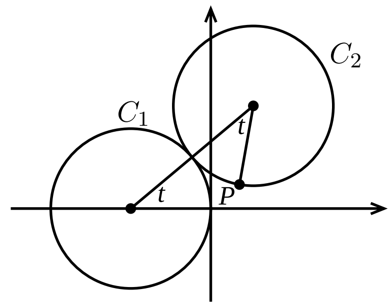
where $x_0 = -r_0\varepsilon/(1 - \varepsilon^2)$, $a = r_0/(1 - \varepsilon^2)$, $b = r_0/\sqrt{1 - \varepsilon^2}$. $(x_0, 0)$ is the center of the ellipse, a is the major semi-axis, and b is the minor semi-axis.

Comment: The point in the allipse that was used as origin of the polar coordinate system is called a *focus* of the ellipse. The focus of the ellipse is distance εa away from the center of the ellipse. This is why ε is called the eccentricity (or excentricity).

Hwk #7:

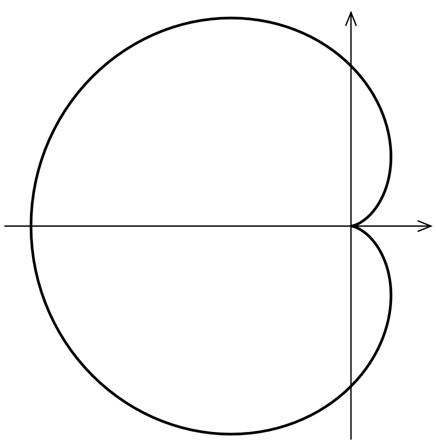
The cardioid is most easily described in terms of polar coordinates: $r = 2(1 - \cos\varphi)$. Its name comes from the Greek word for 'heart', and you'll see why when you graph this curve. So graph it carefully, but don't just steal the graph from a Valentine's card, that would be too corn(er)y!

A circle C_1 of radius 1 sits stationary with center $(-1, 0)$. Another circle C_2 of radius 1 touches it from the right, in the origin. This circle is soon to roll along the fixed circle C_1 without sliding. A point P is marked on the circle C_2. Initially it is the point where both circles touch. As C_2 rolls along C_1, the point P traces out a curve in the stationary plane. Use t for the angle (measured on C_1) of the point of contact of both circles. Give a vector valued function $t \mapsto \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ that describes the position of P as a function of t.



Show that the curve traced out by P is a cardioid.

Solution:



The figure shows a graph of the cardioid, obtained from the formula $r = 2(1 - \cos\varphi)$.

Next we study the curve traced out by P: The center O_1 of circle C_1 has coordinates $(-1, 0)$. The vector from there to the center O_2 of circle C_2 is $[2\cos t\,,\ 2\sin t]^T$. Note that the angle O_1O_2P is also t, because C_2 rolls along C_1 without sliding. This is in addition to the angle t which the vector $\vec{O_2P}$ subtends with the horizontal. This means the vector from O_2 to P has coordinates $[-\cos 2t\,,\ -\sin 2t]^T$. We obtain the following representation of the curve traced out by P:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} -1 + 2\cos t - \cos 2t \\ 2\sin t - \sin 2t \end{bmatrix}$$

In order to show that these two curves coincide, we rewrite the last formula, using the double-angle trig formulas:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 2\cos t\,(1 - \cos t) \\ 2\sin t\,(1 - \cos t) \end{bmatrix}$$

With these formulas, it is clear that the angle φ just coincides with t, since $\tan\varphi = y/x = \tan t$; and that $r = 2(1 - \cos t)$.

Hwk #8:

Given a curve described in parametric form $t \mapsto \vec{x}(t)$, in the plane or in space, we may pretend that the parameter t represents a time and that $\vec{x}(t)$ is the position vector at 'time' t. Then the velocity is $\vec{x}'(t)$, and the speed is $\|\vec{x}'(t)\|$. The length of the curve between parameters t_0 and t_1 is $\int_{t_0}^{t_1} \|\vec{x}'(t)\|\, dt$.

Using this insight, calculate the length (perimeter) of the cardioid.

Solution: We have seen that the cardioid can be parametrized as

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 2(1-\cos t)\cos t \\ 2(1-\cos t)\sin t \end{bmatrix} = \begin{bmatrix} -1+2\cos t-\cos 2t \\ 2\sin t-\sin 2t \end{bmatrix}.$$

Differentiating, we get

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} -2\sin t+2\sin 2t \\ 2\cos t-2\cos 2t \end{bmatrix}$$

and hence

$$\begin{aligned} \left\| \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} \right\|^2 &= (-2\sin t+2\sin 2t)^2+(2\cos t-2\cos 2t)^2 \\ &= 4+4-8(\cos t\ \cos 2t+\sin t\ \sin 2t) = 8-8\cos t = 16\sin^2(t/2) \end{aligned}$$

Therefore the perimeter of the cardioid is

$$L = \int_0^{2\pi} 4|\sin(t/2)|\,dt = 8\int_0^{\pi} \sin s\,ds = 16$$

Hwk #9:

I love to collect airline miles. Let's assume I fly from Atlanta (ATL) to Frankfurt, Germany (FRA). How many miles is the shortest distance? I look up the following info on a map: ATL is at 84.4° western longitude and 33.65° northern latitude. FRA is at 8.6° eastern longitude and 50.1° northern latitude. The radius of the earth is 3975 mi.

Transforming from spherical to cartesian coordinates (the equator plane is the xy plane, with the Greenwich meridian going through the x axis); and using the dot product again, I can calculate the number of miles for this trip (along the shortest route, which is the arc of a circle centered at the center of the earth and connecting from ATL to FRA).

(Note: It is of course more expedient to use symbols $\lambda_{1,2}$ and $\varphi_{1,2}$ for the coordinates first, and postpone plugging in numbers until the end.)

Solution:

Location 1 (ATL) has coordinates $\begin{bmatrix} R\cos\varphi_1\cos\lambda_1 \\ R\cos\varphi_1\sin\lambda_1 \\ R\sin\varphi_1 \end{bmatrix} =: \vec{v}_1$.

Location 2 (FRA) has coordinates $\begin{bmatrix} R\cos\varphi_2\cos\lambda_2 \\ R\cos\varphi_2\sin\lambda_2 \\ R\sin\varphi_2 \end{bmatrix} =: \vec{v}_2$.

(The points are viewed as vectors from the origin to said points.)

For the angle δ between these two vectors (which determines the distance $R\delta$ along the great circle), it holds

$$\begin{aligned} \cos\delta = \frac{\vec{v}_1\cdot\vec{v}_2}{\|\vec{v}_1\|\,\|\vec{v}_2\|} &= \cos\varphi_1\ \cos\varphi_2\,(\cos\lambda_1\ \cos\lambda_2+\sin\lambda_1\ \sin\lambda_2)+\sin\varphi_1\ \sin\varphi_2 \\ &= \cos\varphi_1\ \cos\varphi_2\ \cos(\lambda_1-\lambda_2)+\sin\varphi_1\ \sin\varphi_2 \end{aligned}$$

Specifically, with $\lambda_1=-84.4°$, $\lambda_2=+8.6°$, $\varphi_1=33.65°$, $\varphi_2=50.1°$, and $R=3975$, we obtain $R\delta = 3975\arccos(0.41578) = 4539$. ($\delta$ needed in radian of course for the validity of the formula $R\delta$.)

Hwk #10:

Calculate the curvature of the helix (spiral staircase) given by $\vec{x}(t) = [r\cos t\,,\ r\sin t\,,\ ht]^T$.

Solution: There are two methods: Either way we calculate $\vec{x}\,'(t) = [-r\sin t\,,\ r\cos t\,,\ h]^T$. Therefore $\frac{ds}{dt} = \|\vec{x}\,'(t)\| = \sqrt{r^2 + h^2}$.

Consequently, the unit tangent vector is $\vec{T}(t) = [-r\cos t\,,\ r\cos t\,,\ h]^T/\sqrt{r^2+h^2}$. We get $\kappa = \|d\vec{T}/ds\| = \|\vec{T}'(t)\|/\|\vec{x}\,'(t)\| = r/(r^2+h^2)$.

Alternatively, using the formula $\kappa = \|\vec{x}\,'(t) \times \vec{x}\,''(t)\|/\|\vec{x}\,'(t)\|^3$, we find

$$\vec{x}\,'(t) \times \vec{x}\,''(t) = \begin{bmatrix} -r\sin t \\ r\cos t \\ h \end{bmatrix} \times \begin{bmatrix} -r\cos t \\ -r\sin t \\ 0 \end{bmatrix} = \begin{bmatrix} rh\sin t \\ -rh\cos t \\ r^2 \end{bmatrix}$$

which has norm $r\sqrt{r^2+h^2}$.

Again we find $\kappa = r/(r^2+h^2)$.

Hwk #11:

Let $\vec{u} = [2,1,3]^T$, $\vec{v} = [-1,0,4]^T$, $\vec{w} = [2,-1,-3]^T$. Calculate $\vec{v}\times\vec{w}$, $\vec{u}\times(\vec{v}\times\vec{w})$, $\vec{u}\times\vec{v}$, $(\vec{u}\times\vec{v})\times\vec{w}$.

With these same vectors from the previous problem, calculate $\vec{u}\cdot(\vec{v}\times\vec{w})$ and $\vec{w}\cdot(\vec{u}\times\vec{v})$.

Solution:

$$\vec{v}\times\vec{w} = \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix} \times \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 0(-3) - (-1)4 \\ 4\cdot 2 - (-1)(-3) \\ (-1)(-1) - 0\cdot 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}$$

$$\vec{u}\times(\vec{v}\times\vec{w}) = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \times \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1\cdot 1 - 3\cdot 5 \\ 3\cdot 4 - 2\cdot 1 \\ 2\cdot 5 - 1\cdot 4 \end{bmatrix} = \begin{bmatrix} -14 \\ 10 \\ 6 \end{bmatrix}$$

$$\vec{u}\times\vec{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \times \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 - 3\cdot 0 \\ 3(-1) - 2\cdot 4 \\ 2\cdot 0 - 1(-1) \end{bmatrix} = \begin{bmatrix} 4 \\ -11 \\ 1 \end{bmatrix}$$

$$(\vec{u}\times\vec{v})\times\vec{w} = \begin{bmatrix} 4 \\ -11 \\ 1 \end{bmatrix} \times \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} (-11)(-3) - 1(-1) \\ 1\cdot 2 - 4(-3) \\ 4(-1) - (-11)2 \end{bmatrix} = \begin{bmatrix} 34 \\ 14 \\ 18 \end{bmatrix}$$

(If you find it still difficult to safely remember the pattern of which to multiply with which, a quick check towards correctness of your calculation is to confirm that the dot product of the result with either factor is 0.)

Specifically our example has shown that $\vec{u}\times(\vec{v}\times\vec{w})$ and $(\vec{u}\times\vec{v})\times\vec{w}$ are different!

Now we calculate

$$\vec{u}\cdot(\vec{v}\times\vec{w}) = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix} = 2\cdot 4 + 1\cdot 5 + 3\cdot 1 = 16$$

$$\vec{w}\cdot(\vec{u}\times\vec{v}) = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -11 \\ 1 \end{bmatrix} = 2\cdot 4 + (-1)(-11) + (-3)\cdot 1 = 16$$

As they should, these two are equal.

Hwk #12:

Find the area of the triangle whose vertices are the points $A(1,1,3)$, $B(-2,3,0)$, $C(1,1,-2)$.

Solution: $\vec{AB} = [-3,\ 2,\ -3]^T$ and $\vec{AC} = [0,\ 0,\ -5]^T$. The area of the triangle is half the area of the parallelogram spanned by these two vectors, and this parallelogram area is the norm of the cross product of the two vectors. So the area is

$$\frac{1}{2}\left\| \begin{bmatrix} 10 \\ 15 \\ 0 \end{bmatrix} \right\| = \frac{1}{2}\sqrt{325} = \frac{5}{2}\sqrt{13}$$

Note: You could have done the same kind of calculation (eg) with $\vec{BA}$ and $\vec{BC}$. Different route, but the result would have to be the same. Try it out if you wish.

Hwk #13:

Given the vectors $\vec{u} = [u_1, u_2, u_3]^T$ and $\vec{v} = [v_1, v_2, v_3]^T$ in space, I have defined their cross product $\vec{u}\times\vec{v}$ to be the vector $\vec{w} = [u_2v_3 - u_3v_2\ ,\ u_3v_1 - u_1v_3\ ,\ u_1v_2 - u_2v_1]^T$. Show that indeed, $\|\vec{w}\|^2 = \|\vec{u}\|^2\|\vec{v}\|^2(1-\cos^2\varphi)$, where φ is the angle between $\vec{u}$ and $\vec{v}$.

Solution: The right side is $\|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2$. Now let's evaluate either side in components:

$$\begin{aligned}\|\vec{w}\|^2 &= (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2 = \\ &= u_2^2v_3^2 + u_3^2v_2^2 + u_3^2v_1^2 + u_1^2v_3^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2u_3v_2v_3 - 2u_1u_3v_1v_3 - 2u_1u_2v_1v_2\end{aligned}$$

$$\begin{aligned}\|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2 &= (u_1^2+u_2^2+u_3^2)(v_1^2+v_2^2+v_3^2) - (u_1v_1+u_2v_2+u_3v_3)^2 \\ &= u_2^2v_3^2 + u_3^2v_2^2 + u_3^2v_1^2 + u_1^2v_3^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2u_3v_2v_3 - 2u_1u_3v_1v_3 - 2u_1u_2v_1v_2\end{aligned}$$

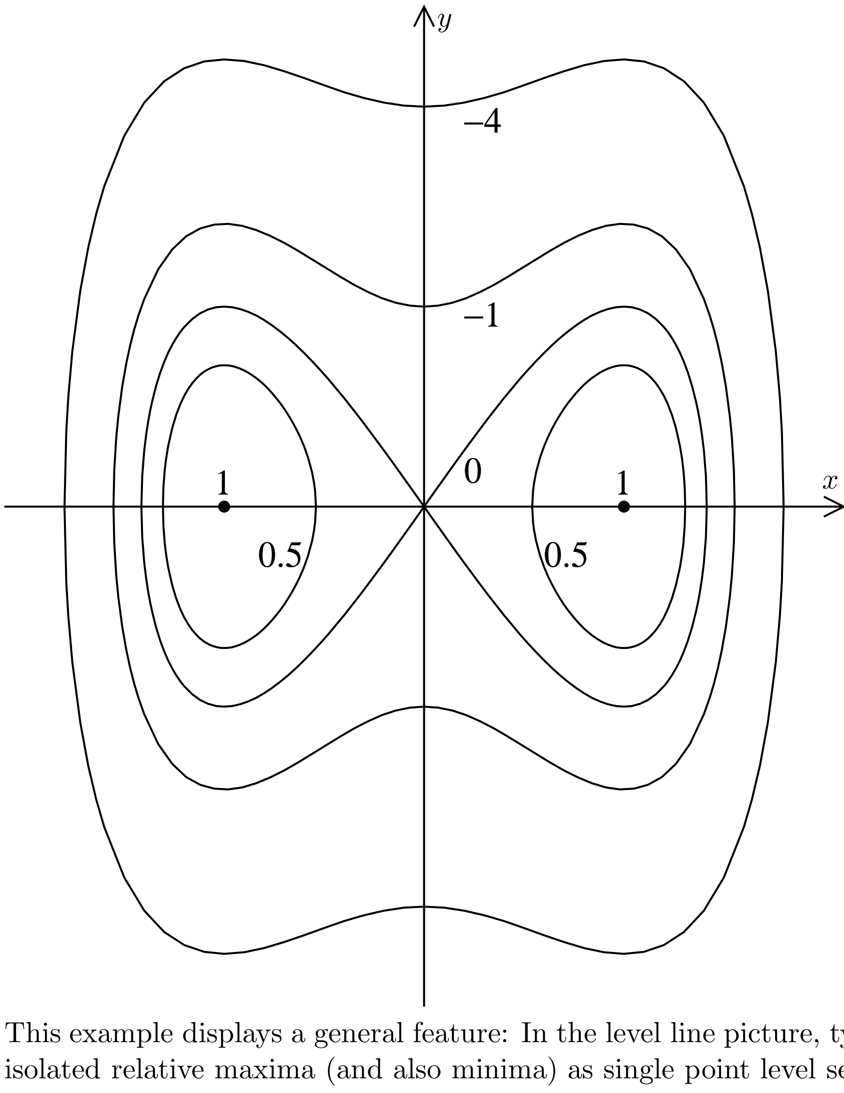
Hwk #14: *See next page. (Order swapped for page break layout reasons)*

Hwk #15:

> Try to understand the level sets of the function f given by $f(x,y) = 2x^2 - x^4 - y^2$. In particular use single-variable calculus and simple algebraic reasoning to find maxima of f. Make sure to sketch at least five level sets $\{(x,y) \mid f(x,y) = c\}$. Namely, for $c \in \{1, \frac{1}{2}, 0, -1, -4\}$. You may also find it useful to sketch graphs of a few single variable functions $g(x) := f(x, y_0)$ for various y_0. Try to avoid using technology that does 'multi-variable graphs', but feel free to enlist the help of technology for single variable graphs if this helps. The skill you are to train here is to piece single variable info together to get a multi-variable picture.

Solution:

As far as the maximum is concerned, $f(x,y) \le f(x,0) \le f(\pm 1, 0) = 1$. Here, the single variable maximum of $g(x) := f(x,0) = 2x^2 - x^4$ can be found by studying the derivative $g'(x) = 4(x - x^3)$ as usual; or else we can argue that $2x^2 = 2\sqrt{1 \cdot x^4} \le 1 + x^4$ by the agm inequality.



The level set for level 1 consists of only the two points $(\pm 1, 0)$. The level set for level z consists of two sv function graphs $y = \pm\sqrt{-z + 2x^2 - x^4}$. For $0 < z < 1$, the term under the root is non-negative on two intervals, namely for $x^2 \in [1 - \sqrt{1-z}, 1 + \sqrt{1-z}]$. These level sets are in the shape of two (slightly deformed) circles, each surrounding a maximum. For $z = 0$, the two circles merge into a figure-8 curve; for $z < 0$, the term under te square root is non-negative on a single interval centered at 0, and the level curves consist of a single closed curve.
Regarding the graph of f, the origin looks like a *saddle*. Looking only into the x direction, it looks like a minimum, but looking into the y direction, it looks like a maximum.

This example displays a general feature: In the level line picture, typical saddles show up as crossings, isolated relative maxima (and also minima) as single point level sets.

Hwk #14:

> Draw level curves of the function f given by $f(x,y) := |x| + |y|$, and describe the graph $z = f(x,y)$.
>
> Same question for $g(x,y) = \sqrt{x^2 + y^2}$.

Solution: The level curves for f are squares whose vertices are on the x and y axes respectively. The graph of f is a four-sided pyramid 'standing on its tip'.

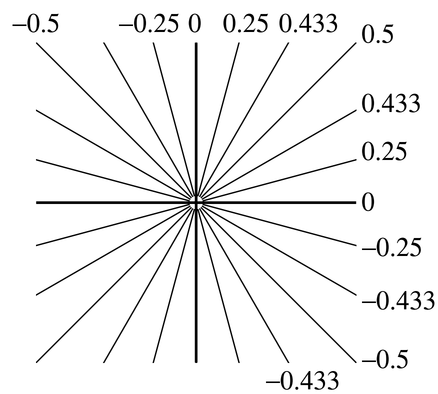
The level curves of g are circles, and the graph of g is a cone 'standing on its tip'.

Hwk #16:

> Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,y) = \frac{xy}{x^2+y^2}$ for $(x,y) \neq (0,0)$ and $f(0,0) = 0$. By using polar cordinates, draw the level curves of this function. As explained in class, this function is *not* continuous at $(0,0)$. Show that nevertheless, all the single variable functions g and h given by $g(x) := f(x,y)$ for any choice of y, and $h(y) := f(x,y)$ for any choice of x, are continuous. For any fixed k, find the limit $\lim_{x\to 0} f(x,kx)$.

Solution:
We let $x = r\cos\varphi$ and $y = r\sin\varphi$ and define $\tilde{f}(r,\varphi) := f(r\cos\varphi, r\sin\varphi)$. (Note that in comparison, physicists may prefer to re-use the symbol f for the new function, with the names of the arguments being used to resolve the ambiguity. – I am using the mathematicians' convention here. Refer to the introductory notes (pg 1) for this issue)

Then $\tilde{f}(r,\varphi) := \cos\varphi\,\sin\varphi = \frac{1}{2}\sin(2\varphi)$. The level curves are straight lines 'through' (but omitting) the origin.
Obviously, the limits $\lim_{x\to 0} f(x,kx) = k/(1+k^2)$ and $\lim_{y\to 0} f(0,y) = 0$ exist.
For fixed $x \neq 0$, the sv function $h : y \mapsto \frac{xy}{x^2+y^2}$ is continuous, b/c it is a rational function with non-vanishing denominator. For $x = 0$, we are considering the constant function $h : y \mapsto 0$ (defined for every y including 0, because of the stipulation $f(0,0) = 0$). This h is trivially continuous. The analogous arguments apply to the functions $g : x \mapsto \frac{xy}{x^2+y^2}$ for fixed y and (in case $y = 0$) to $g : x \mapsto 0$.



Hwk #17:

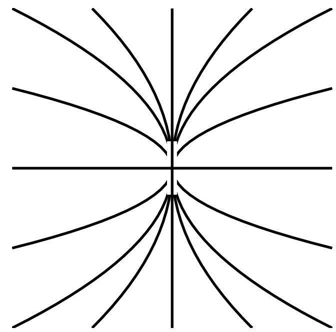
> Cranking the previous example up a notch, consider the function f given by $f(x,y) := \frac{x^2y^4}{x^4+y^8}$ for $(x,y) \neq (0,0)$ and $f(0,0) = 0$.
>
> Show that each of the radial limits $\lim_{x\to 0} f(x,kx)$ and $\lim_{y\to 0} f(0,y)$ equal $0 = f(0,0)$, but that, nevertheless, $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist (and hence f is not continuous in the origin). Draw level lines of f to get insight into this function.
>
> Can you describe a 'curve of approach' in the (x,y) plane along which the single variable limit exists, but is different from 0? That is, can you find $x(t)$ and $y(t)$ such that $\lim_{t\to\infty} x(t) = 0$ and $\lim_{t\to\infty} y(t) = 0$, but $\lim_{t\to\infty} f(x(t),y(t)) \neq 0$.
>
> Can you also describe a curve of approach for which the limit does not exist?

Solution:

$$\lim_{x\to 0} f(x,kx) = \lim_{x\to 0}\frac{k^4x^6}{x^4+k^8x^8} = \lim_{x\to 0}\frac{k^4x^2}{1+k^8x^4} = 0$$

because the last expression under the limit is continuous. Moreover, $\lim_{y\to 0} f(0,y) = \lim_{y\to 0}(0\cdot y^4/y^8) = 0$. (Note that during the $\lim_{y\to 0}$, by definition, y is *not* 0.)

The level lines are parabolas $y^2 = kx$, for level $k^2/(1+k^4)$. For instance, if we approach the origin on such parabolas, we get different limits: $\lim_{y\to 0} f(y^2,y) = \frac{1}{2}$; or $\lim_{y\to 0} f(\frac{1}{2}y^2,y) = \frac{4}{17}$. — (The parametrization for the curve of approach suggested in the problem isn't the most convenient for these parabolas, but we could consider $(x(t),y(t)) = (1/t^2,1/t)$ for the parabola $x = y^2$.)



The limit $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist, because in every ball around the origin, f takes on (among others), in particular the values 0 and $\frac{1}{2}$, so no number L can be within distance $\varepsilon := 0.1$ from both 0 and $\frac{1}{2}$.

To take a curve of approach for which the limit does not exist, we may want to take some kind of spiral around the origin, that keeps encountering all level sets in turn. For instance, we might consider trying $(x(t),y(t)) := (\frac{1}{t}\cos t, \frac{1}{t}\sin t) \to (0,0)$ as $t\to\infty$, and we would get

$$f(x(t),y(t)) = \frac{\cos^2 t\,\sin^4 t}{t^2\cos^4 t + t^{-2}\sin^8 t}$$

This would work, but is technically a bit awkward to write up. We would need to select a sequence of t_k where $\cos^4 t$ is small enough to compensate for the large coefficient t^2, but not exactly 0, to prevent the numerator from being 0. We would specifically argue (using the intermediate value theorem) that there is a sequence $t_k\to\infty$ along which $t^2\cos^4 t = t^{-2}\sin^8 t$, and that for these t_k, from the equality case of the arithmetic and geometric mean inequality, the denominator equals $2\cos^2 t\sin^4 t$.

It is a bit easier to modify the curve of approach to model the different weights of x and y in the formula for f. So we can take $(x(t),y(t)) = (t^{-2}\cos^2 t\ \mathrm{sign}(\cos t), t^{-1}\sin t)$, which is still a spiral but squeezing much more in x direction than in y direction.

Then

$$f(x(t),y(t)) = \frac{\cos^4 t\,\sin^4 t}{\cos^8 t + \sin^8 t}$$

and clearly no limit exists as $t\to\infty$ (take $t_k = \frac{\pi}{4} + k\pi$ and $t'_k = k\pi$ as respective subsequences).

Hwk #18:

Does $\lim_{(x,y)\to(0,0)} y\sin\frac{1}{x}$ exist, and if so, what is its value? Explain.

Solution: The limit exists and is 0 by the squeeze theorem. However, note that the domain of $y\sin\frac{1}{x}$ does not contain an entire punctured disc around the origin. So, due to the restriction on the approach direction, this limit is more akin to the single variable limits like $\lim_{x\to 0+}[\ldots]$, but obviously we cannot maintain specific notations for all the possible limitations on the approach geometry that could occur in multi-variable limits, and this is why we have put the domain hypothesis in the definition of the limit.

Hwk #19:

Consider the functions f and g given by $f(x,y) := x^2y + e^{xy} + y^3$ and $g(x,y) := \arctan \frac{y}{x}$. Calculate the following:

$$(a)\quad \frac{\partial f(x,y)}{\partial x}\,, \qquad (b)\quad \frac{\partial f(x,y)}{\partial y}\,, \qquad (c)\quad \frac{\partial g(x,y)}{\partial x}\,, \qquad (d)\quad \frac{\partial g(x,y)}{\partial y}\,,$$

Also calculate the following:

$$(a')\ \frac{\partial}{\partial y}\left(\frac{\partial f(x,y)}{\partial x}\right), \quad (b')\ \frac{\partial}{\partial x}\left(\frac{\partial f(x,y)}{\partial y}\right), \quad (c')\ \frac{\partial}{\partial y}\left(\frac{\partial g(x,y)}{\partial x}\right), \quad (d')\ \frac{\partial}{\partial x}\left(\frac{\partial g(x,y)}{\partial y}\right)$$

Compare (a') with (b') and (c') with (d').

Solution:

$$\frac{\partial f(x,y)}{\partial x} = 2xy + ye^{xy}\,, \qquad \frac{\partial f(x,y)}{\partial y} = x^2 + xe^{xy} + 3y^2$$

$$\frac{\partial g(x,y)}{\partial x} = \frac{-y/x^2}{1+(\frac{y}{x})^2} = \frac{-y}{x^2+y^2}\,, \qquad \frac{\partial g(x,y)}{\partial y} = \frac{1/x}{1+(\frac{y}{x})^2} = \frac{x}{x^2+y^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f(x,y)}{\partial x}\right) = \frac{\partial}{\partial y}\left(2xy + ye^{xy}\right) = 2x + e^{xy} + yxe^{xy}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f(x,y)}{\partial y}\right) = \frac{\partial}{\partial x}\left(x^2 + xe^{xy} + 3y^2\right) = 2x + e^{xy} + xye^{xy}$$

So the mixed partial derivatives in either order coincide.

$$\frac{\partial}{\partial y}\left(\frac{\partial g(x,y)}{\partial x}\right) = \frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right) = \frac{-1(x^2+y^2)-(-y)(2y)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial g(x,y)}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) = \frac{1(x^2+y^2)-(x)(2x)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

Again, the mixed partial derivatives in either order coincide.

Hwk #20:

More about the ellipse: Given the points $F_\pm = (\pm e, 0)$ in the plane (where e is some positive real number, not to be confused with the Euler number $2.718\ldots$), and a number $a > e$. Show that the set of those points $P = (x,y)$ in the plane that satisfy the condition $\|\vec{PF_+}\| + \|\vec{PF_-}\| = 2a$ is an ellipse $x^2/a^2 + y^2/b^2 = 1$. How does b relate to e and a? What is the eccentricity ε of the ellipse?

Solution:

The vectors $\vec{PF}_+$ and $\vec{PF}_-$ are $-\begin{bmatrix} x-e \\ y \end{bmatrix}$ and $-\begin{bmatrix} x+e \\ y \end{bmatrix}$ respectively. Therefore the condition on the norms reads as

$$\sqrt{(x+e)^2+y^2} + \sqrt{(x-e)^2+y^2} = 2a$$

Squaring the equation and isolating the square root that remains from the mixed term yields

$$2\sqrt{(x+e)^2+y^2}\,\sqrt{(x-e)^2+y^2} = 4a^2 - \Big((x+e)^2+y^2\Big) - \Big((x-e)^2+y^2\Big)$$

Simplifying and squaring again yields

$$\Big((x+e)^2+y^2\Big)\Big((x-e)^2+y^2\Big)=\Big(2a^2-(x^2+e^2)-y^2\Big)^2$$

As we expand both sides, we benefit from a lot of cancellations:

$$\begin{aligned}(x^2-e^2)^2+y^2(2x^2+2e^2)+y^4 &= 4a^4+(x^2+e^2)^2+y^4-4a^2(x^2+e^2)-4a^2y^2+2(x^2+e^2)y^2\\ (x^2-e^2)^2-(x^2+e^2)^2 &= 4a^4-4a^2(x^2+e^2)-4a^2y^2\\ -x^2e^2 &= a^2(a^2-x^2-e^2-y^2)\\ x^2(a^2-e^2)+y^2a^2 &= a^2(a^2-e^2)\\ \frac{x^2}{a^2}+\frac{y^2}{a^2-e^2} &= 1\end{aligned}$$

So we have obtained the desired equation, with $b^2 := a^2-e^2$. — Comparing with Hwk #6, we quote:

$$a=\frac{r_0}{1-\varepsilon^2}\ ,\quad b=\frac{r_0}{\sqrt{1-\varepsilon^2}}$$

and therefore $b^2/a^2 = 1-\varepsilon^2$. Since $b^2=a^2-e^2$, we obtain $\varepsilon = e/a$. Quoting $-x_0=\varepsilon a$ from the solution of #6, we see that $|x_0|=e$. This observation identifies the coordinate origin from #6 with the focus F_+ in the present problem.

Hwk #21:

> Reconsider the function f from Problem #17: $f(x,y):=\frac{x^2y^4}{x^4+y^8}$ for $(x,y)\neq(0,0)$ and $f(0,0)=0$. Write it in polar coordinates: $g(r,\varphi):=f(r\cos\varphi,r\sin\varphi)$. The partial derivative $\partial g(r,\varphi)/\partial r$ at $r=0$ is a directional derivative of f (at the origin). Show that all directional derivatives at the origin vanish, so the graph has a horizontal tangent in each direction. Nevertheless, f is not even continuous at the origin.
>
> Plot, for some choice of fixed φ (other than an integer multiple of $\pi/2$) the graph of the single variable function $g(\cdot,\varphi): r\mapsto g(r,\varphi)$. Include information about the precise location of the maximum of this function.

Solution:

$$g(r,\varphi)=f(r\cos\varphi,r\sin\varphi)=\frac{r^2\cos^2\varphi\,\sin^4\varphi}{\cos^4\varphi+r^4\sin^8\varphi}$$

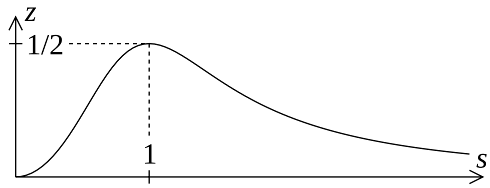
Now,

$$\frac{\partial}{\partial r}g(r,\varphi)\mid_{r=0}=\left.\frac{2r\cos^2\varphi\,\sin^4\varphi\,(\cos^4\varphi+r^4\sin^8\varphi)-r^2\cos^2\varphi\,\sin^4\varphi\,4r^3\sin^8\varphi}{(\cos^4\varphi+r^4\sin^8\varphi)^2}\right|_{r=0}=0$$

provided $\cos\varphi\neq0$. In the case where $\cos\varphi=0$, we have the function $r\mapsto g(r,\varphi)$ constant 0, and the conclusion still holds.

We know from #17 that the maximum of g is $\frac{1}{2}$, and that it occurs when $r=|\cos\varphi|/\sin^2\varphi$. This can be seen by setting the r-derivative 0.

If we let $s:=r\sin^2\varphi/|\cos\varphi|$, then we see that $g(r,\varphi)=s^2/(1+s^4)$, so all radial graphs arise by stretching of the s axis from one graph: $z=s^2/(1+s^4)$:



The maximum of $g(\cdot,\varphi)$ is at $r = |\cos\varphi|/\sin^2\varphi$, with value $\frac{1}{2}$. As $\varphi \to 0$, the location of this maximum moves to ∞, and near the origin, we just see the minimum. – In contrast, as $\varphi \to \pi/2$, the maximum gets closer and closer to the origin. For $\varphi = \pi/2$ exactly, the radial function is 0. This function arises as a limit from the 'decaying tail' of the graph.

Hwk #22:

Sketch level lines for the function $f(x,y) := x^3 - 3xy^2$. Choose levels 4, 1, 0, −1, −4. The most convenient way to do this is to use polar coordinates again. Look for a trig formula involving multiple angles that fits the situation (you'd likely not have memorized this formula to recognize it at first sight, that's why I say you should look for it).

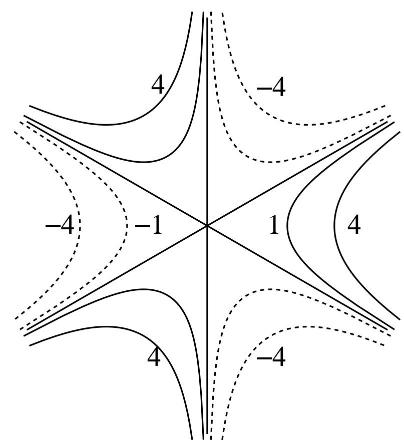
This function is hand-picked to display a rare pattern in the level lines picture

Describe the graph of f in topographer's terms: where are the hills and the valleys? The point $(0,0)$ is said to feature a *monkey saddle* of this function f.

Solution:

$$g(r,\varphi) = f(r\cos\varphi, r\sin\varphi) = r^3(\cos^3\varphi - 3\cos\varphi\sin^2\varphi) = r^3\cos 3\varphi$$

The level curves of level h can be described in polar coordinates as $r = h^{1/3}(\cos 3\varphi)^{-1/3}$ if $h \neq 0$. The level curves for level $h = 0$ are straight lines through the origin, determined by the condition $\cos 3\varphi = 0$. As $\varphi \to \pi/6$, or any other value that makes $\cos 3\varphi$ vanish, the r coordinate on the level line goes to infinity.
This is called a monkey saddle because the monkey can sit in it facing east, with his tail hanging down west, and the legs in southeast and northeast direction.



Hwk #23:

Find $Df(x,y)$ for $f(x,y) = x^y$, where $x > 0$.

Solution: $\partial f(x,y)/\partial x = yx^{y-1}$ (like the derivatives of x^2, x^3, etc.).

$\partial f(x,y)/\partial y = x^y \ln x$ (like derivatives of $2^y = e^{y \ln 2}$ and the like.

Therefore $Df(x,y) = [yx^{y-1} \;\; x^y \ln x]$.

Note: *Don't go on auto-pilot from Calc2 or DiffEq: y is* NOT *a function of x here. It is the context (which functions, which dependences, are we talking about) and not some magic symbol manipulation that governs when the chain rule must be applied.*

Hwk #24:

Consider the following vector valued multi-variable functions:

$$f : (r, \varphi) \mapsto \begin{bmatrix} r\cos\varphi \\ r\sin\varphi \end{bmatrix}$$

defined for $\{(r,\varphi) \mid r > 0\,,\; \varphi \in \mathbb{R}\}$; and

$$g : (r, \vartheta, \phi) \mapsto \begin{bmatrix} r\sin\vartheta\,\cos\phi \\ r\sin\vartheta\,\sin\phi \\ r\cos\vartheta \end{bmatrix}$$

defined for $r > 0$, $\vartheta \in \,]0,\pi[$, $\phi \in \mathbb{R}$.

Find $Df(r,\varphi)$ and $Dg(r,\vartheta,\phi)$ by calculating the necessary partial derivatives and observing they are continuous.

Solution:

$$Df(r,\varphi) = \begin{bmatrix} \partial f_1(r,\varphi)/\partial r & \partial f_1(r,\varphi)/\partial\varphi \\ \partial f_2(r,\varphi)/\partial r & \partial f_2(r,\varphi)/\partial\varphi \end{bmatrix} = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}$$

$$Dg(r,\vartheta,\phi) = \begin{bmatrix} \frac{\partial g_1(r,\vartheta,\phi)}{\partial r} & \frac{\partial g_1(r,\vartheta,\phi)}{\partial\vartheta} & \frac{\partial g_1(r,\vartheta,\phi)}{\partial\phi} \\ \frac{\partial g_2(r,\vartheta,\phi)}{\partial r} & \frac{\partial g_2(r,\vartheta,\phi)}{\partial\vartheta} & \frac{\partial g_2(r,\vartheta,\phi)}{\partial\phi} \\ \frac{\partial g_3(r,\vartheta,\phi)}{\partial r} & \frac{\partial g_3(r,\vartheta,\phi)}{\partial\vartheta} & \frac{\partial g_3(r,\vartheta,\phi)}{\partial\phi} \end{bmatrix} = \begin{bmatrix} \sin\vartheta\,\cos\phi & r\cos\vartheta\,\cos\phi & -r\sin\vartheta\,\sin\phi \\ \sin\vartheta\,\sin\phi & r\cos\vartheta\,\sin\phi & r\sin\vartheta\,\cos\phi \\ \cos\vartheta & -r\sin\vartheta & 0 \end{bmatrix}$$

Hwk #25:

Show that for $f(\vec{x}) := \|\vec{x}\|$ we have $\nabla f(\vec{x}) = \vec{x}/\|\vec{x}\|$ at $\vec{x} \neq \vec{0}$.

Solution: Noticing that $\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots}$, we find $\partial\|\vec{x}\|/\partial x_1 = 2x_1/(2\sqrt{x_1^2 + x_2^2 + \ldots})$ and similar for the other variables. So $\partial\|\vec{x}\|/\partial x_i = x_i/\|\vec{x}\|$. Putting the components in a vector, the claim is immediate.

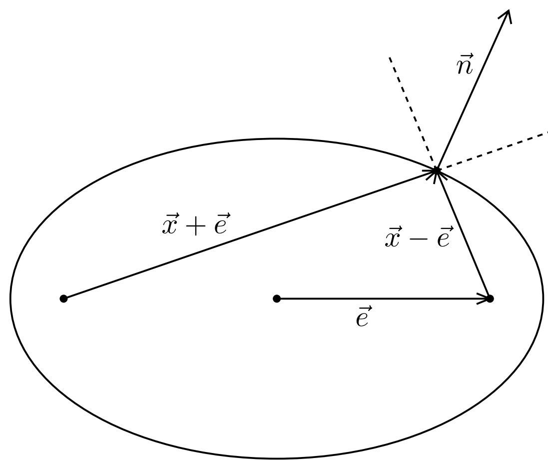
Also note that the function $\vec{x} \mapsto \|\vec{x}\|$ *is* differentiable outside the origin because the partials are continuous there.

Hwk #26:

Consider the vector $\vec{e}$ and any (non-empty) level set of the function $g(\vec{x}) := \|\vec{x}-\vec{e}\|+\|\vec{x}+\vec{e}\|$. We know from #13 that the level set is an ellipse. In this problem we show the reflection property of the ellipse: A ray emanating from one focus $\vec{e}$ and reflected in the ellipse will pass through the other focus $-\vec{e}$. The reflection law in physics says that the incoming ray has the same angle with the normal to a curve as the reflected ray.

Prove the reflection property of the ellipse by checking the angle between the normal to the ellipse and the ray or reflected ray respectively. You can do everything without coordinates or components, just using vector notation.

Solution:



It is easy to write down a normal vector to the ellipse at point $\vec{x}$: namely, $\nabla g(\vec{x})$ is orthogonal to the ellipse, because gradients are normal to the level lines of a function. We calculate (using the previous problem)

$$\vec{n} := \nabla g(\vec{x}) = \frac{\vec{x}-\vec{e}}{\|\vec{x}-\vec{e}\|} + \frac{\vec{x}+\vec{e}}{\|\vec{x}+\vec{e}\|}$$

Note: a use of the chain rule is implicit in this calculation, however the shift $x_i \mapsto x_i \pm e_i$ has inner dervative 1. Or, if you want to use the MV chain rule, you should work with the untransposed Df first: Take $f : \vec{x} \mapsto \vec{x} - \vec{e}$. Then $Df(\vec{x}) = I$, the unit matrix, for every $\vec{x}$. (Remember I is a matrix with 1's on the diagonal and 0's everywhere else). With $g_0(\vec{y}) := \|\vec{y}\|$, we find

$$D(g_0 \circ f)(\vec{x}) = Dg_0(f(\vec{x}))Df(\vec{x}) = (\vec{y}^T/\|\vec{y}\|)_{\vec{y}=f(\vec{x})=\vec{x}-\vec{e}} \cdot I = (\vec{x}-\vec{e})^T/\|\vec{x}-\vec{e}\|$$

Returning to our main calculation, we want to show that the angle between $\vec{x}-\vec{e}$ and $\vec{n}$ is the same as the angle between $\vec{x}+\vec{e}$ and $\vec{n}$. We check equality for the cosines of these angles, using the dot product:

$$\frac{(\vec{x}-\vec{e})\cdot\vec{n}}{\|\vec{x}-\vec{e}\|\,\|\vec{n}\|} \stackrel{?}{=} \frac{(\vec{x}+\vec{e})\cdot\vec{n}}{\|\vec{x}+\vec{e}\|\,\|\vec{n}\|}$$

Equivalently, we have to check

$$\frac{(\vec{x}-\vec{e})\cdot\left(\frac{\vec{x}-\vec{e}}{\|\vec{x}-\vec{e}\|}+\frac{\vec{x}+\vec{e}}{\|\vec{x}+\vec{e}\|}\right)}{\|\vec{x}-\vec{e}\|} \stackrel{?}{=} \frac{(\vec{x}+\vec{e})\cdot\left(\frac{\vec{x}-\vec{e}}{\|\vec{x}-\vec{e}\|}+\frac{\vec{x}+\vec{e}}{\|\vec{x}+\vec{e}\|}\right)}{\|\vec{x}+\vec{e}\|}$$

Simplifying (namely multiply out the dot products in the numerator distributively and separate the large fraction), both sides equal

$$1 + \frac{(\vec{x} - \vec{e}) \cdot (\vec{x} + \vec{e})}{\|\vec{x} - \vec{e}\| \, \|\vec{x} + \vec{e}\|}$$

and the equality of both sides proves the reflection property.

COMMENT: The previous solution does not use any coordinates, but simply works with the vectors geometrically. This more abstract approach is something I hope to get you accustomed to. It is often capable of giving slick calculations. However, you can also work with coordinates. This gives the following

VARIANT SOLUTION: In coordinates, the ellipse is the level set of the function $f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ for level 1. (The x-axis passes through the foci $F_\pm$, with the origin halfway between them. and the y-axis vertical.) Refer to Pblm #12. The coordinates of the Foci $F_\pm$ are $(\pm e, 0)$ respectively, with $e^2 + b^2 = a^2$. The normal vector $\vec{n}$ is $\nabla f(x, y) = [\frac{2x}{a^2}, \frac{2y}{b^2}]^T$. The vector $\vec{x} + \vec{e}$ is $[x + e\,,\ y]^T$, and the vector $\vec{x} - \vec{e}$ is $[x - e\,,\ y]$. The same calculation as before can now be carried out in components:

$$\frac{(x-e)2x/a^2 + y\,2y/b^2}{\sqrt{(x-e)^2 + y^2}\,\|\vec{n}\|} = \frac{(\vec{x} - \vec{e}) \cdot \vec{n}}{\|\vec{x} - \vec{e}\| \, \|\vec{n}\|} \stackrel{?}{=} \frac{(\vec{x} + \vec{e}) \cdot \vec{n}}{\|\vec{x} + \vec{e}\| \, \|\vec{n}\|} = \frac{(x+e)2x/a^2 + y\,2y/b^2}{\sqrt{(x+e)^2 + y^2}\,\|\vec{n}\|}$$

where I have wisely refrained from evaluating $\|\vec{n}\|$ because it gets cancelled anyways. After a bit of simplification, we have to check if indeed

$$\frac{(x-e)x/a^2 + y^2/b^2}{\sqrt{(x-e)^2 + y^2}} \stackrel{?}{=} \frac{(x+e)x/a^2 + y^2/b^2}{\sqrt{(x+e)^2 + y^2}}$$

Squaring and cross-multiplying, this is equivalent to

$$[(x+e)^2 + y^2]\left(\frac{(x-e)x}{a^2} + \frac{y^2}{b^2}\right)^2 - [(x-e)^2 + y^2]\left(\frac{(x+e)x}{a^2} + \frac{y^2}{b^2}\right)^2 \stackrel{?}{=} 0$$

I hope you are already thinking that this will be a mess, and the vector calculation above is preferrable, but now that we're here, let's carry it through with gusto. Sorting out the terms in the brackets first, the left side equals

$$[x^2 + e^2 + y^2]\left(\frac{(x-e)x}{a^2} - \frac{(x+e)x}{a^2}\right)\left(\frac{2x^2}{a^2} + \frac{2y^2}{b^2}\right) + 2xe\left(\left(\frac{(x-e)x}{a^2} + \frac{y^2}{b^2}\right)^2 + \left(\frac{(x+e)x}{a^2} + \frac{y^2}{b^2}\right)^2\right)$$

It's probably time to remember that $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ to simplify this mess further, and then we simplify our expression to

$$-(x^2 + e^2 + y^2)\frac{2ex}{a^2}\,2 + 2ex\left(\left(1 - \frac{ex}{a^2}\right)^2 + \left(1 + \frac{ex}{a^2}\right)^2\right)$$

We still have to show that this expression is 0, and so, by multiplying with $a^2/(2ex)$, this is equivalent to checking whether

$$-2a^2(x^2 + e^2 + y^2) + (a^2 - ex)^2 + (a^2 + ex)^2 \stackrel{?}{=} 0$$

Indeed, replacing e^2 by $a^2 - b^2$, the lhs can be expanded to

$$-2a^2x^2 - 2a^2(a^2 - b^2) - 2a^2y^2 + 2a^4 + 2(a^2 - b^2)x^2 = 2a^2b^2 - 2a^2y^2 - 2b^2x^2 = 2a^2b^2\left(1 - \frac{y^2}{b^2} - \frac{x^2}{a^2}\right) = 0$$

Hwk #27:

The multi-variable chain rule says: $D(f\circ g)(\vec{p}) = Df(g(\vec{p}))Dg(\vec{p})$. Here is one specific example for which I ask you to caclulate all quantities involved in this equation and check the equality, all by explicit calculation.

$$f(x,y,z) := xyz^2 + (y-x)/(1+z^2), \qquad g(\vartheta,\phi) = \begin{bmatrix} \sin\vartheta\,\cos\phi \\ \sin\vartheta\,\sin\phi \\ \cos\vartheta \end{bmatrix}.$$

(Motivation for this example: think of $f\circ g$ as a function on the unit sphere.)

Solution:
A brief comment ahead: In this problem the $\vec{p}$ stands for $\left[\begin{smallmatrix}\vartheta\\ \phi\end{smallmatrix}\right]$. If ϑ and ϕ are indeed angle coordinates on the sphere, as suggested, then $\vec{p}$ is a non-geometric vector, artificially converted from (ϑ,ϕ). However, if ϑ and ϕ happen to be just weirdly chosen names for cartesian coordinates in a plane, then $\vec{p}$ is a geometric vector. Whichever may be the case, it does not affect the validity of the calculations.

$$(f\circ g)(\vartheta,\phi) = \sin^2\vartheta\cos^2\vartheta\sin\phi\cos\phi + \frac{\sin\vartheta(\sin\phi-\cos\phi)}{1+\cos^2\vartheta}$$

Now let's calculate all ingredients for the chain rule formula:

$$D(f\circ g)(\vartheta,\phi) = [T_{11} \quad T_{12}]$$

with

$$T_{11} = (2\sin\vartheta\cos^3\vartheta - 2\sin^3\vartheta\cos\vartheta)\sin\phi\cos\phi + \frac{\cos\vartheta(\sin\phi-\cos\phi)}{1+\cos^2\vartheta} + \frac{2\cos\vartheta\sin^2\vartheta(\sin\phi-\cos\phi)}{(1+\cos^2\vartheta)^2}$$

$$T_{12} = \sin^2\vartheta\cos^2\vartheta(\cos^2\phi-\sin^2\phi) + \frac{\sin\vartheta(\cos\phi+\sin\phi)}{1+\cos^2\vartheta}$$

$$Df(x,y,z) = [S_{11} \quad S_{12} \quad S_{13}]$$

with

$$S_{11} = yz^2 - \frac{1}{1+z^2} \qquad S_{12} = xz^2 + \frac{1}{1+z^2} \qquad S_{13} = 2xyz - \frac{2z(y-x)}{(1+z^2)^2}$$

$$Df(g(\vartheta,\phi)) = Df(\sin\vartheta\cos\phi, \sin\vartheta\sin\varphi, \cos\vartheta) = \left[\tilde{S}_{11} \quad \tilde{S}_{12} \quad \tilde{S}_{13}\right]$$

with

$$\tilde{S}_{11} = \sin\vartheta\cos^2\vartheta\sin\phi - \frac{1}{1+\cos^2\vartheta}$$

$$\tilde{S}_{12} = \sin\vartheta\cos^2\vartheta\cos\phi + \frac{1}{1+\cos^2\vartheta}$$

$$\tilde{S}_{13} = 2\sin^2\vartheta\cos\vartheta\cos\phi\sin\phi - \frac{2\cos\vartheta\sin\vartheta(\sin\phi-\cos\phi)}{(1+\cos^2\vartheta)^2}$$

$$Dg(\vartheta,\phi) = \begin{bmatrix} \cos\vartheta\,\cos\phi & -\sin\vartheta\,\sin\phi \\ \cos\vartheta\,\sin\phi & \sin\vartheta\,\cos\phi \\ -\sin\vartheta & 0 \end{bmatrix}$$

The claim of the chain rule, which we have to check here, is:

$$[T_{11} \quad T_{12}] = \left[\tilde{S}_{11} \quad \tilde{S}_{12} \quad \tilde{S}_{13}\right] \begin{bmatrix} \cos\vartheta\,\cos\phi & -\sin\vartheta\,\sin\phi \\ \cos\vartheta\,\sin\phi & \sin\vartheta\,\cos\phi \\ -\sin\vartheta & 0 \end{bmatrix}$$

In other words, the claim is

$$\begin{aligned} T_{11} &= (\tilde{S}_{11}\cos\phi + \tilde{S}_{12}\sin\phi)\,\cos\vartheta - \tilde{S}_{13}\sin\vartheta \qquad \text{and} \\ T_{12} &= (-\tilde{S}_{11}\sin\phi + \tilde{S}_{12}\cos\phi)\sin\vartheta \end{aligned}$$

and this is immediate to check.

Hwk #28:

> Suppose that a piece of a level curve $g(x, y) = c$, near some point (x_0, y_0), where c is some constant, can be written as the graph of a function f: $y = f(x)$. Express the slope $f'(x_0)$ of this graph in terms of partial derivatives of g. Write your result both in mathematical notation and in physicists' notation with differential quotients (with z for the output variable of g and dy/dx expressed in terms of $\partial z/\partial x$ and $\partial z/\partial y$). *If you think a certain minus sign in your result looks weird, you're right: it does look weird, but it's still correct! Or rather I hope so, I haven't seen your solution; all I say is that the correct solution may have some weird looking detail.*

Solution: If $y = f(x)$ solves the equation $g(x, y) = c$ for y, then $g(x, f(x)) = c$, i.e., the composite function on the left hand side is actually the constant function. Taking the derivative with respect to x, we get

$$\frac{\partial g}{\partial x}(x, f(x)) + \frac{\partial g}{\partial y}(x, f(x))f'(x) = 0$$

Note on notation: The first term in this expression looks a bit funny. In line with previous notation, we have written the partial derivative of g with respect to its first variable as $\partial g/\partial x$, because the default name for the first variable of this function was x. However, now that the arguments we plug into g are $(x, f(x))$ and therefore the variable x shows up in both variables, the notation justly raises some eyebrows, despite it being common usage. So far there is no ambiguity. The entire expression left of the equal sign, which is the (single-variable) derivative of the composite function $x \mapsto g(x, f(x))$ would be called $\frac{d}{dx}g(x, f(x))$. Nevertheless this instance hints at some possible trouble when we use default variable *names* to denote with respect to which *slot* of a multi-variable function we are taking a partial derivative. The cleaner notation would be $\partial_1 g$ instead of $\partial g/\partial x$. The index 1 tells unambiguously (and independent of the name of the variable) that we are taking a partial derivative with respect to the first variable of the function g. And once $\partial_1 g$ is constructed, we plug x in the first slot, and $f(x)$ in the second slot.

We only need to solve for $f'(x)$ yet:

$$f'(x) = -\frac{\partial g(x, f(x))/\partial x}{\partial g(x, f(x))/\partial y} = -\frac{\partial_1 g(x, f(x))}{\partial_2 g(x, f(x))}$$

Note on notation: I have just aggravated the 'default variable name identifies slot' problem in the first version, whereas the second version cleans the problem up. The way how I aggravated the problem in the first version flows out of typographic convenience: As I changed the fraction $\frac{\cdot}{\cdot}$ into $\cdot/\cdot$ to avoid towering fractions, I was forced to put the arguments $(x, f(x))$ from 'behind the formal fraction' $\frac{\partial g}{\partial x}(x, f(x))$ 'into the numerator of the formal fraction' $\frac{\partial g(x,f(x))}{\partial x}$. The first notation $\frac{\partial g}{\partial x}(x, f(x))$ at least indicates that there is a function g of which we take a certain partial derivative $\frac{\partial g}{\partial x}$ into which we then plug in some arguments. So it is manifest that the x in $\frac{\partial g}{\partial x}$ represents a slot and not a variable (b/c none is plugged in there yet). In contrast, the second notation $\frac{\partial g(x,f(x))}{\partial x}$ seems to indicate (and wrongly so) that there is an expression $g(x, f(x))$ obtained from plugging in certain x dependent quantities into the function g, and that we then take the derivative of this expression with respect to x. But this in NOT what we are doing here, and the only indication to this effect that is left is the curly ∂. For if we were really taking the derivative of $g(x, f(x))$ with respect to x, this would be a single variable derivative (as there is no other variable around any more). Likewise, if $\partial g(x, f(x))/\partial y$ were really referring to a derivative of the expression $g(x, f(x))$ with respect to y, the result would be 0, b/c there is no y in this expression any more. But in reality, this is not what is intended, and the second notation with ∂_2 and ∂_1 represents the intent so much more clearly.

Now let's look how the same formula looks in the physicist's notation: z depends on two quantities x and y. If we fix z, then y must become dependent on x (or vice versa). Our formula is written as

$$\frac{dy}{dx} = -\frac{\partial z/\partial x}{\partial z/\partial y}$$

If you think in terms of formally cancelling partial differentials (but you shouldn't!), then this looks paradoxical, b/c the right hand side would become $-\frac{\partial y}{\partial x}$ and the $-$ sign doesn't seem to belong here. The physicist's version of the formula looks pretty neat and concise, but note that a lot of unspoken information is present in the context but is not represented in the formula, and it is only within this context that the formula can be interpreted meaningfully. On the left side, in $\frac{dy}{dx}$, we can *view* y as a function of x, by looking at a level set *on which* z *is constant.* In the partials on the right side, we have $\frac{\partial z}{\partial x}$ *with the implied understanding that* y *is constant* and then $\frac{\partial z}{\partial y}$ *with the implied understanding that* x *is held constant.* If you were to formally cancel partials on the right hand side, you would change or disable the implied contextual inferences: $\frac{\partial y}{\partial x}$ has no meaning at all! – The curly-∂'s always hide some information that must be implied by the context (namely which variable does *not* change?), and formal algebra with them that is unaware or negligent of this context (as formal cancellation in $\frac{\partial z/\partial x}{\partial z/\partial y}$ would be), is a recipe for errors.

Variant Solution: You could also do the following: $\nabla g(x, y) = \begin{bmatrix} \partial g(x,y)/\partial x \\ \partial g(x,y)/\partial y \end{bmatrix}$ is orthogonal to the level line $g(x, y) = c$. To get a vector that is *tangent* to this level line, we need to find a vector that is orthogonal to $\nabla g(x, y)$. For instance $\begin{bmatrix} -\partial g(x,y)/\partial y \\ \partial g(x,y)/\partial x \end{bmatrix}$. (Any multiple of this vector would also qualify of course.) The slope of this tangent is $\Delta y/\Delta x = -\frac{\partial g(x,y)/\partial x}{\partial g(x,y)/\partial y}$. Since this slope if $f'(x)$, we conclude $f'(x) = -\frac{\partial g(x,y)/\partial x}{\partial g(x,y)/\partial y}$.

Hwk #29:

(For warmup only: this is predominantly single variable calculus)
Consider $f(x) := \int_0^\infty e^{-xt} \frac{\sin t}{t}\, dt$. You will not find an immediate antiderivative by which to evaluate this integral. Nevertheless, calculate $f'(x)$. [The Math 447 expert tells you that moving the x-derivative under the t-integral is legitimate here.] You should be able to evaluate the integral that you obtain for $f'(x)$.

What do you think $\lim_{x\to\infty} f(x)$ is? [The Math 447 expert tells you that in this example it is legitimate to move $\lim_{x\to\infty}$ past the integral sign.]

Finally, knowing $f'(x)$, $\lim_{x\to\infty} f(x)$ and the fundamental theorem of calculus, find $f(x)$. Specifically, what is $\int_0^\infty \frac{\sin t}{t}\, dt$?

The late physics Nobel prize winner Richard Feynman reports in his memoirs how, as a student, he got the reputation of being an integration wizard, because he was familiar with this 'differentiation under the integral sign technique', which his peers hadn't learned.

Solution:

$$f'(x) = \int_0^\infty \frac{\partial}{\partial x} e^{-xt} \frac{\sin t}{t}\, dt = -\int_0^\infty t e^{-xt} \frac{\sin t}{t}\, dt$$

We live on the cancelled t and evaluate the integral (e.g.) by a sequence of two integrations by parts (assuming $x > 0$). Here I derive the exponentials and integrate the trigs; the other way around would work as well.

$$\int_0^\infty e^{-xt} \sin t \, dt = [-e^{-xt} \cos t]_0^\infty - x \int_0^\infty e^{-xt} \cos t \, dt = 1 - x[e^{-xt} \sin t]_0^\infty - x^2 \int_0^\infty e^{-xt} \sin t \, dt$$

By solving for the unknown integral, which was the negative of $f'(x)$, we find

$$f'(x) = -\frac{1}{1+x^2} \; .$$

Variant: Remember: if you accept complex exponentials and Euler's formula, there is another way to evaluate the same integral:

$$\int_0^\infty e^{-xt} \sin t \, dt = \frac{1}{2i} \int_0^\infty e^{-xt}(e^{it} - e^{-it}) \, dt = \frac{1}{2i} \left[\frac{1}{-x+i} e^{-xt} e^{it} - \frac{1}{-x-i} e^{-xt} e^{-it} \right]_0^\infty$$
$$= \frac{1}{2i} \left(\frac{1}{x-i} - \frac{1}{x+i} \right) = \frac{1}{x^2+1}$$

If we trust the M447 expert about moving the limit under the integral, we conclude

$$\lim_{x \to \infty} f(x) = \int_0^\infty \lim_{x \to \infty} e^{-xt} \frac{\sin t}{t} \, dt = \int_0^\infty 0 \frac{\sin t}{t} \, dt = 0$$

The limit under the integral is 0 for every $t > 0$ (and we trust that the integral doesn't bother about the single exception $t = 0$; this is part of the M447 expertise on which we need to rely here for lack of deeper abstract theory).

Now we can integrate over x and argue

$$f(x_1) - f(x_0) = \int_{x_0}^{x_1} f'(x) \, dx = -\int_{x_0}^{x_1} \frac{dx}{x^2+1} = -\arctan x_1 + \arctan x_0$$

As $x_1 \to \infty$, we get $0 - f(x_0) = \arctan x_0 - \frac{\pi}{2}$. Writing x for x_0 and letting $x_0 \to 0+$ we obtain

$$\lim_{x \to 0+} f(x) = \frac{\pi}{2} - \arctan 0 = \frac{\pi}{2}$$

Note that I have been careful in writing $\lim_{x \to 0+}$ rather than plugging in $x = 0$, because our calculation of $f'(x)$ had relied on x being positive. The integral from which we calculated $f'(x)$ does NOT converge for $x = 0$! So there is really another technical issue for which we need the $M447$ expert (and it is not easy). We need to know by means of the M447 expert's theorem toolkit, that f, as defined by the integral formula, is indeed continuous (from the right) at $x = 0$. (Only) then can we conclude that $\int_0^\infty \frac{\sin t}{t} \, dt = f(0) = \frac{\pi}{2}$.

I promise you that all these theoretical gaps can be filled in, but urge you to be aware that the need to fill in the necessary theory is genuine. This type of problem is to be your motivation when you study more advanced analysis courses. On the other hand, you will see this type of calculation abundantly in physics. When it goes smoothly, physicists will not bother to justify the theoretical legitimacy. But in cases when the calculation fails (producing a wrong result because the hypotheses for legitimacy of moving limits and/or derivatives past an integral sign are not fulfilled), there is usually also an intuitive physical reason to explain why such unworried calculation is not advisable.

Hwk #30:

Use the multi-variable chain rule to determine $f'(x)$, when $f(x) := \int_0^x \frac{\sin(xt)}{t}\,dt$.

Analogous question for $g(x) := \int_{x/2}^{2x} \frac{e^{xt}}{t}\,dt$.

Again, we rely on the Math 447 expert, who tells us that it is legitimate to move derivatives past the integral sign in this example.

Solution: This time, x occurs in two places in the formula for $f(x)$. The MV chain rule, written in terms of partials, tells us to consider each location separately and apply a partial (single variable) derivative (as in the single variable chain rule), and then to add the results obtained for each separate location. More formally, we consider

$$F(u,v) := \int_0^u \frac{\sin vt}{t}\,dt$$

and we substitute $u = x$ and $v = x$ to get $f(x) = F(x,x)$. So we have

$$f'(x) = (\partial_1 F)(x,x)\frac{\partial u}{\partial x} + (\partial_2 F)(x,x)\frac{\partial v}{\partial x}$$

For $\partial_1 F$, we use the fundamental theorem (derivative of an antiderivative) to get $(\partial_1 F)(u,v) = \frac{\sin vu}{u}$. For $\partial_2 F$, we use differentiation under the integral sign (with permission from the M447 expert again given specifically for the situation of this problem, not as a blank cheque!) and get $(\partial_2 F)(u,v) = \int_0^u \frac{t\cos vt}{t}\,dt = [\frac{1}{v}\sin vt]_{t=0}^{t=u} = \frac{\sin vu}{v}$. Putting it all together (with $\partial u/\partial x = 1 = \partial v/\partial x$ because $u = x$ and $v = x$), we get

$$f'(x) = \frac{\sin(x^2)}{x} + \frac{\sin(x^2)}{x} = 2\frac{\sin(x^2)}{x}\ .$$

The same works for g, in principle: We define $G(u,v,w) := \int_w^u \frac{e^{vt}}{t}\,dt$ and let $w = x/2$, $u = 2x$, and $v = x$. So $g(x) = G(2x, x, \frac{x}{2})$. Note that derivatives with respect to the *lower* limit of integration get a minus sign from the fundamental theorem, and that we have inner derivatives $\frac{\partial w}{\partial x} = \frac{1}{2}$ and $\frac{\partial u}{\partial x} = 2$ this time.

$$g'(x) = 2\frac{e^{x\cdot 2x}}{2x} + \int_{x/2}^{2x} \frac{te^{xt}}{t}\,dt - \frac{1}{2}\frac{e^{x\cdot x/2}}{x/2} = \frac{e^{2x^2}}{x} + \left[\frac{e^{xt}}{x}\right]_{t=x/2}^{t=2x} - \frac{e^{x^2/2}}{x} = \frac{2e^{2x^2}}{x} - \frac{2e^{x^2/2}}{x}$$

Hwk #31:

A quantity w depends on the coordinates x, y, z in 3-space as follows: $w = x^2 + y^2 + xyz$ (1). We study w especially on the plane given by $z = x + 2y$. Then we have there $w = x^2 + y^2 + xy(x + 2y) = x^2 + y^2 + x^2y + 2xy^2$ (2).

Now we calculate $\frac{\partial w}{\partial x}$ from (1): $\frac{\partial w}{\partial x} = 2x + yz$. On the plane, this simplifies to $\frac{\partial w}{\partial x} = 2x + y(x + 2y) = 2x + xy + 2y^2$.

Calculating $\frac{\partial w}{\partial x}$ on the plane directly from (2), we get $\frac{\partial w}{\partial x} = 2x + y(x+2y) = 2x + 2xy + 2y^2$. We clearly have a discrepancy by a term xy. What is wrong? Clear up the confusion. (This requires some text as well as formulas.)

Solution: When we calculate $\frac{\partial w}{\partial x}$ from (1), we are taking a partial derivative with respect to z of a three-variable function f given by $f(x, y, z) = x^2 + y^2 + xyz$. In this partial derivative, both y and z are treated as constant. The partial derivative $\frac{\partial w}{\partial x}$ is a directional derivative in direction $[1\,,\ 0\,,\ 0]^T$, which is a direction that goes off the plane $z = x + 2y$, even if we later evaluate this derivative in a point on that plane.

In contrast, when we substitute first $z = x + 2y$ and then take the partial derivative $\frac{\partial w}{\partial x}$ form (2), we are taking a partial derivative of a two-variable function g given by $g(x, y) = x^2 + y^2 + xy(x + 2y)$. In this partial derivative, y is still constant and x still varies, as before; but z is now not constant, but also varies because $z = x + 2y$. For f, this is now a directional derivative in direction $[1\,,\ 0\,,\ 1]^T$, which is a direction lying *in* the plane $z = x + 2y$.

The notation $\frac{\partial w}{\partial x}$ is ambiguous in this context, because the variable w can refer to two different functions f versus g. Unless the ambiguity is resolved by an explaining text or context, the notation $\frac{\partial w}{\partial x}$ should be avoided in such a situation. Rather the result according to part (1) should be written as $\partial_1 f(x, y, x + 2y)$, and the result at part (2) should be written as $\partial_1 g(x, y)$.

We then have

$$(*)\qquad \begin{aligned} \frac{\partial}{\partial x} f(x, y, x + 2y) &= \partial_1 f(x, y, x + 2y)\frac{\partial x}{\partial x} + \partial_2 f(x, y, x + 2y)\frac{\partial y}{\partial x} + \partial_3 f(x, y, x + 2y)\frac{\partial (x + 2y)}{\partial x} \\ &= \partial_1 f(x, y, x + 2y) + \partial_3 f(x, y, x + 2y) \end{aligned}$$

and

$$(**)\qquad \frac{\partial}{\partial x} g(x, y) = \partial_1 g(x, y)$$

Both of these results are equal and correspond to calculation (2), whereas calculation (1) gave just the first term of the sum arising in $(*)$.

Hwk #32:

This example is taken from *I. Rosenholtz, L. Smylie: "The only Critical Point in Town" Test*, Mathematics Magazine **58**(1985), 149–150.

Show that the function

$$g : (x, y) \mapsto y^2 + 3(y + e^x - 1)^2 + 2(y + e^x - 1)^3\ ,\ \mathbb{R}^2 \to \mathbb{R}$$

has exactly one critical point, and that this point is a relative mimimum.

Furthermore explain why this point is NOT an absolute minimum.

Solution: It is convenient to use the abbreviation $u := y + e^x - 1$.

$$\frac{\partial g(x, y)}{\partial x} = 6ue^x + 6u^2e^x \qquad \frac{\partial g(x, y)}{\partial y} = 2y + 6u + 6u^2$$

For both to vanish we need $y = 0$ and $u + u^2 = 0$, which latter means $u = 0$ or $u = -1$. Since $y = 0$, $u = e^x - 1$, and this cannot equal -1. So $u = 0$, and this means $x = 0$.

Therefore the only critical point is $(x, y) = (0, 0)$. Let's calculate the Hessian:

$$Hg(x, y) = \begin{bmatrix} (6u + 6u^2)e^x + (6 + 12u)e^{2x} & (6 + 12u)e^x \\ (6 + 12u)e^x & 8 + 12u \end{bmatrix}$$

$$Hg(0,0) = \begin{bmatrix} 6 & 6 \\ 6 & 8 \end{bmatrix}$$

This matrix is positive definite, by the Hurwitz test: $h_{11} = 6 > 0$ and $h_{11}h_{22} - h_{12}^2 = 12 > 0$. Therefore the origin is a relative minimum of g. $g(0,0) = 0$. But $g(0,y) = 4y^2 + 2y^3 \to -\infty$ as $y \to -\infty$. The function is unbounded below and does not have an absolute minimum.

Comment: This example shows two things. Firstly, it does not suffice to check the values of a function at the only candidates for a relative minimum in order to determine a global minimum, *unless* the existence of an absolute minimum is established beforehand.

But there is a second lesson hidden in this example. An intuitively plausible argument would go like this: "If I am hiking in a landskape and standing in a local minimum, but there are points out there with lower elevation than my present location, then it must be possible, in principle, to reach them by a walk that goes through some pass (mathematically a saddle point). So, if a local minimum is not a global minimum, there should be another critical point, a saddle point, somewhere." This argument does actually have some merit. However, the saddle point could have run away to infinity. In our example the 2-variable function $(u,y) \mapsto y^2 + 3u^2 + 2u^3$ has a local minimum at $(u,y) = (0,0)$ and a saddle point at $(u,y) = (-1,0)$. But $u \to -1$ corresponds to $x \to -\infty$ when $u = y + e^x - 1$ and $y = 0$.

In advanced applications, this principle 'there should be a saddle point, if only we can make sure that it hasn't run off to infinity' is a powerful tool in solving partial differential equations.

Hwk #33:

> This example is taken from Marsden-Tromba: Show that the function f given by $f(x,y) = (y - 3x^2)(y - x^2)$ has a critical point in the origin, which is neither a relative minimum nor a relative maximum. What kind of '***'definite is the Hessian?
>
> Show also that all single-variable radial functions $t \mapsto f(t\cos\phi, t\sin\phi)$ have a relative minimum at $t = 0$.

Solution: For fixed ϕ, let $g(t) := f(t\cos\phi, t\sin\phi) = t^2\sin^2\phi - 4t^3\sin\phi\cos^2\phi + 3t^4\cos^4\phi$. Then $g'(0) = 0$ and $g''(0) = 2\sin^2\phi$. For $\sin\phi \neq 0$, we can argue that g has a local minimum at 0 because the first derivative vanishes and the second derivative is positive there. For $\sin\phi = 0$, we have $\cos^2\phi = 1$ and $g(t) = 3t^4$, and we again have a local minimum at 0, albeit a 'degenerate' one that cannot be detected by the second derivative test. .

Now clearly $f(0,0) = 0$, but there are both positive and negative values in any neighbourhood of $(0,0)$. For instance $f(x, 2x^2) = -x^4 < 0$. Let's calculate the Hessian from $f(x,y) = 3x^4 - 4x^2y + y^2$: $f_{xx}(x,y) = 36x^2 - 8y$, $f_{xy}(x,y) = -8x$, $f_{yy}(x,y) = 2$. So $Hf(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$. This matrix is positive semidefinite, but not positive definite.

Note: Similar as with continuity and differentiability, we learn here for the local minimu property that the MV version cannot be captured by having the single variable version in all directions. However, a slightly strengthened version of the single variable local minimality (namely that the sufficient condition of having positive second derivative is satisfied) in all directions does suffice to prove local minimality in the multi-variable context for C^2 functions.

Hwk #34:

> *This example is geometrically appealing, but alas calculationally lengthy. This is why I give you the intermediate steps and hints to navigate you through. Ideally, it should be*

done with the help of symbolic algebra software, and you are welcome to use this tool, if available.

We want to find a shortest connection between two plane curves, namely $y = x^2 + 2$ and $y = \frac{1}{2}(x-1)^2$. A precise plot is attached. Choose points $P = (a, a^2+2)$ on the first parabola and $Q = (b, \frac{1}{2}(b-1)^2)$ on the second and minimize the square of the distance. Determine all critical points and classify them. Does one of them probide a *global* minimum? Why?

Hint 1: While it is possible to take one of the equations for a critical point and solve it for b via by means of the quadratic formula, and then plug in the result in the other equation, this is tedious. It is more straightforward to take successively linear combinations of the two equations with the strategy of first eliminating b^3, then b^2, then b, until one polynomial equation in a remains.

Hint 2: After an obvious factorization of this polynomial equation, an easy solution $a = 1$ can be guessed, and when this is factored off, a 4th order polynomial remains that can be factored into two quadratics with integer coefficients; indeed one factor is $a^2 + 2a + 3$.

Solution: We take the distance squared between a point $P = (a, a^2 + 2)$ on the first parabola and a point $Q = (b, \frac{1}{2}(b-1)^2)$ on the second parabola and get the function

$$f(a,b) := (a-b)^2 + \Big(a^2 + 2 - \frac{1}{2}(b-1)^2\Big)^2 .$$

We obtain

$$\partial f(a,b)/\partial a = 2(a-b) + 4a\Big(a^2 + 2 - \frac{1}{2}(b-1)^2\Big) = 0$$
$$\partial f(a,b)/\partial b = 2(b-a) - 2(b-1)\Big(a^2 + 2 - \frac{1}{2}(b-1)^2\Big) = 0$$

So we have to solve the system of two polynomial equations

$$-b + 2a^3 + 4a - ab^2 + 2ab = 0 \quad (1)$$
$$2a^2 - 2a + 3 - 2a^2 b + b^3 - 3b^2 + b = 0 \quad (2)$$

Following the hint, we take a times (2) plus b times (1) to get rid of b^3. We obtain

$$2a^3 - 2a^2 + 3a + 5ab - (a+1)b^2 = 0 \quad (3)$$

Now we add $-(a+1)$ times (1) to a times (3) to eliminate b^2. We obtain

$$-4a^3 - a^2 - 4a + b(3a^2 - a + 1) = 0 \quad (4)$$

We solve this for b and plug it back into (1):

$$2a^3 + 4a + (2a-1)\frac{4a^3 + a^2 + 4a}{3a^2 - a + 1} - a\left(\frac{4a^3 + a^2 + 4a}{3a^2 - a + 1}\right)^2 = 0$$

Clearing denominators and expanding, we get

$$2a^7 + 4a^6 + 3a^5 - 5a^4 - 7a^3 + 3a^2 = 0 \quad (5)$$

We can factor off $a = 0$ twice and guess the solution $a = 1$, thus factoring off $(a-1)$. After long division of polynomials, we get

$$0 = 2a^4 + 6a^3 + 9a^2 + 4a - 3 = (2a^2 + 2a - 1)(a^2 + 2a + 3)$$

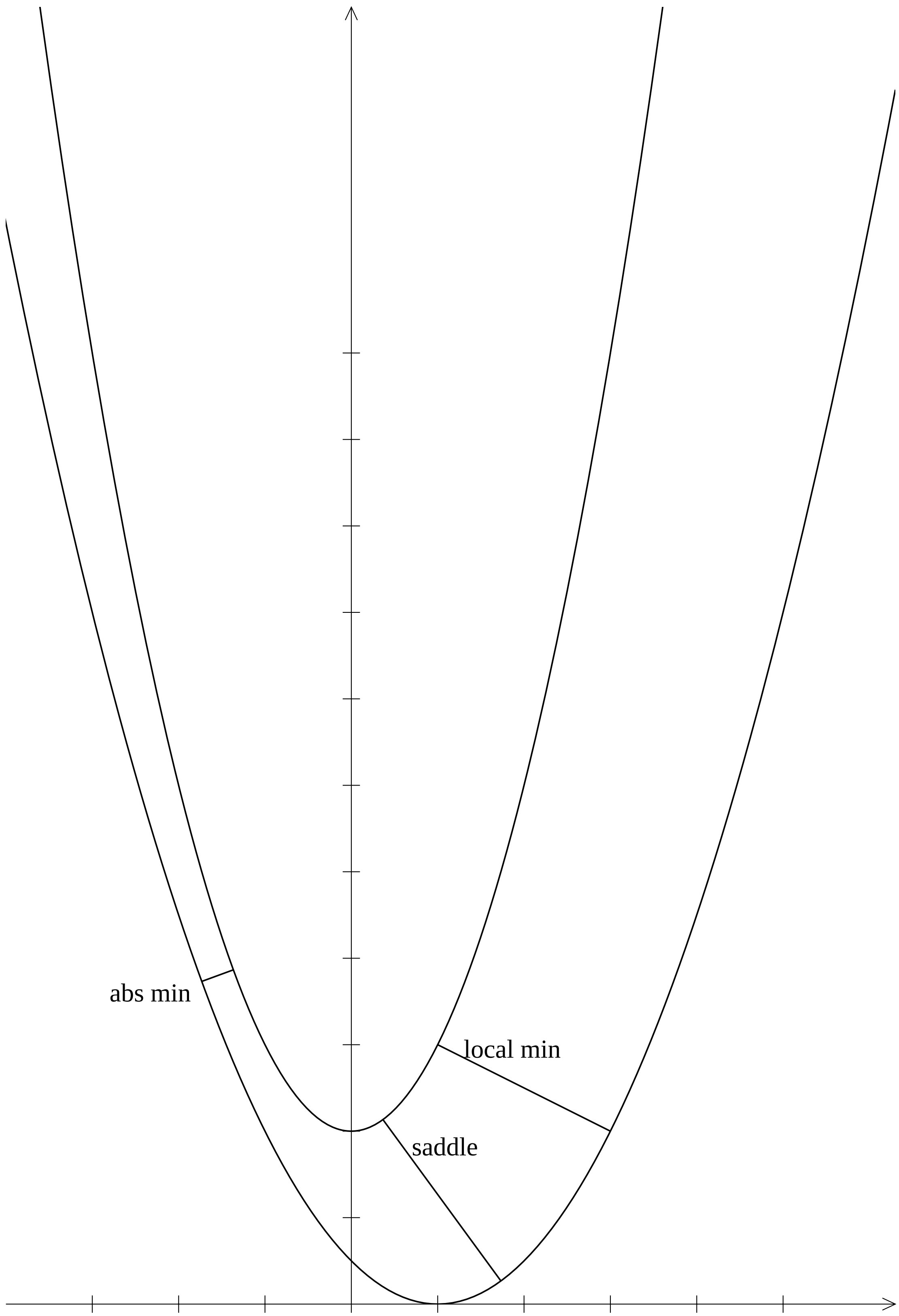


Figure 1: Figure for Hwk # 34

This factorization, short of the hint given, would best be found by means of a symbolic algebra software. $a^2+2a+3=0$ doesn't have real solutions; $2a^2+2a-1=0$ can be solved by the quadratic formula. We have thus found the following solutions of (5):

$$a_0 = 0\ , \quad a_1 = 1\ , \quad a_{2,3} = \frac{1}{2}\Big(-1 \pm \sqrt{3}\Big)$$

We can get b form a by means of (4), but initially, all of these are mere *consequences* of (1),(2); they are not *equivalent* to (1),(2). We must plug them back into the original equations. Clearly $a_0 = 0$ violates (2). For $a \neq 0$, (1),(2) $\iff$ (1),(3) $\iff$ (1),(4) $\iff$ (4),(5). The last equivalence uses $3a^2 - a + 1 \neq 0$. So we have found

$$\begin{array}{lll}
(a_1, b_1) = (1, 3) & P_1 = (1, 3) & Q_1 = (3, 2) \\[1ex]
(a_2, b_2) = \left(\frac{1}{2}\Big(-1+\sqrt{3}\Big), \sqrt{3}\right) & P_2 = \left(\frac{1}{2}\Big(-1+\sqrt{3}\Big), 3 - \frac{1}{2}\sqrt{3}\right) & Q_2 = \Big(-\sqrt{3}, 2-\sqrt{3}\Big) \\[1ex]
(a_3, b_3) = \left(\frac{1}{2}\Big(-1-\sqrt{3}\Big), -\sqrt{3}\right) & P_3 = \left(\frac{1}{2}\Big(-1-\sqrt{3}\Big), 3 + \frac{1}{2}\sqrt{3}\right) & Q_3 = \Big(\sqrt{3}, 2+\sqrt{3}\Big)
\end{array}$$

Numerical values are:

$$\begin{array}{ll|l}
P_1 = (1, 3) & Q_1 = (3, 2) & |P_1Q_1| = 2.236 \\[1ex]
P_2 = (0.366, 2.134) & Q_2 = (1.732, 0.268) & |P_2Q_2| = 2.313 \\[1ex]
P_3 = (-1.366, 3.866) & Q_3 = (-1.732, 3.732) & |P_3Q_3| = 0.390
\end{array}$$

The Hessian of f is

$$Hf(a,b) = \begin{bmatrix} 8 + 12a^2 + 4b - 2b^2 & -2 + 4a - 4ab \\ -2 + 4a - 4ab & 1 - 2a^2 - 6b + 3b^2 \end{bmatrix}$$

$$Hf(a_1, b_1) = \begin{bmatrix} 14 & -10 \\ -10 & 8 \end{bmatrix} \qquad \text{pos def}$$

$$Hf(a_2, b_2) = \begin{bmatrix} 14 - 2\sqrt{3} & -10 + 4\sqrt{3} \\ -10 + 4\sqrt{3} & 8 - 5\sqrt{3} \end{bmatrix} \approx \begin{bmatrix} 10.536 & -3.072 \\ -3.072 & -0.660 \end{bmatrix} \qquad \text{indefinite}$$

$$Hf(a_3, b_3) = \begin{bmatrix} 14 + 2\sqrt{3} & -10 - 4\sqrt{3} \\ -10 - 4\sqrt{3} & 8 + 5\sqrt{3} \end{bmatrix} \approx \begin{bmatrix} 17.464 & -16.928 \\ -16.928 & 16.660 \end{bmatrix} \qquad \text{pos def}$$

So (a_2, b_2) is a saddle point, the other two are local minima. The one with the smaller distance, namely (a_3, b_3) can be accepted as a global minimum, PROVIDED we know apriori that a global minimum EXISTS.

We cannot argue directly, because $\mathbb{R} \times \mathbb{R} \ni (a, b)$ is not bounded. However, we have the additional feature that $f(a, b) \to \infty$ as either a or b goes to infinity. For instance, when asking whether a global minimum exists, we can a-priori neglect all (a, b) for which $f(a, b) > 100$, because at some points f has smaller values, e.g., $f(0, 1) = 5 \ll 100$. So we only need to consider those (a, b) where $|a - b| \leq 10$, because otherwise $f(a, b) > 10^2 + 0$. But if $|a - b| \leq 10$, we have $a^2 + 2 - \frac{1}{2}(b-1)^2 = \frac{1}{2}a^2 + \frac{1}{2}[a^2 - (b-1)^2] + 2 \geq \frac{1}{2}a^2 + 2 - \frac{1}{2}|a - b + 1|\,|a + b - 1| \geq \frac{1}{2}a^2 + 2 - \frac{11}{2}(2|a| + 11)$. For $|a|$ sufficiently large, this will again be > 10, making $f(a, b) > 0 + 10^2$. So we need to consider only the absolute minimum on some closed and bounded set $\{(a, b) : |a - b| \leq 10\ ,\ |a| \leq C\}$, and an abolute minimum exists on this set. Outside ths set, the values of f are larger, so we do have an absolute minimum on $\mathbb{R}^2$.

Hwk #35:

Suppose in the following matrices, the starred entries are not known. Which of the five possibilities 'positive definite', 'positive semidefinite (but not definite)', 'negative definite', 'negative semidefinite (but not definite)', 'indefinite' remains a possibility, based on knowledge only of the known entries?

$$(a) \begin{bmatrix} 3 & * \\ * & * \end{bmatrix} \qquad (b) \begin{bmatrix} * & * \\ * & -5 \end{bmatrix} \qquad (c) \begin{bmatrix} * & 6 \\ 6 & * \end{bmatrix} \qquad (d) \begin{bmatrix} 3 & * \\ * & -1 \end{bmatrix} \qquad (e) \begin{bmatrix} 0 & 1 & * \\ 1 & * & * \\ * & * & * \end{bmatrix}$$

Consider the following: While the Hurwitz test was worded in a way to calculate determinants starting from the top, the order in which the variables are listed (and thus determine entries of the matrix) is not essential for definiteness of a matrix; so you could use the determinants in the Hurwitz test starting at any diagonal element and then calculating 2×2, 3×3, etc. determinants, adding any one variable (row and column) at a time.

Solution: (a) this matrix could be positive (semi-)definite or indefinite, but not negative (semi-) definite, because the vector $[1\,,\ 0]^T$ makes the quadratic form positive.

(b) this matrix could be negative (semi-)definite or indefinite, but not positive (semi-)definite, because the vector $[0\,,\ -1]^T$ makes the quadratic form negative.

(c) no conclusion can be drawn: If one diagonal element is positive and one negative, then the matrix is indefinite. If both diagonal elements are larger than 6, the matrix is positive definite by Gershgorin's test. If both diagonal elements are below -6, the matrix is negative definite. If both diagonal elements are equal 6 (or equal -6), the matrix is positive (or negative) semidefinite by explicit writing down of the quadratic form. (This does not exhaust all possibilities that the $*$ could stand for, but it already represents all cases for definiteness properties of the matrix.)

(d) this matrix is indefinite, regardless of the off-diagonal elements.

(e) this matrix is indefinite. The determinant of the 2×2 submatrix $\begin{bmatrix} 0 & 1 \\ 1 & * \end{bmatrix}$ is -1, so the matrix cannot be positive definite (and the 0 in the corner also confirms this). Similarly, the negative of this matrix cannot be positive definite either. – We do not have a Hurwitz test for semi-definiteness, so we cannot rule this case out by determinants. But let's just calculate the quadratic form with the vector $[1\,,\ t\,,\ 0]^T$. We get $0 \cdot 1^2 + 2 \cdot 1 \cdot t + * \cdot t^2 = 2t + *t^2$. Whatever the value of $*$, this expression is positive for t positive and sufficiently small, but is negative for t negative of sufficiently small absolute value.

Hwk #36:

Find the absolute minimum and absolute maximum of $x^2 + (y-1)^2 + z^2 - xyz$ on the ball $x^2 + y^2 + z^2 \leq 3^2$. *Hint: For the boundary consideration, use the xz plane as equator plane for the spherical coordinates, to benefit from the symmetry of the problem. Otherwise formulas get obnoxiously messy.*

Solution: Since this is a continuous function on a bounded and closed set, we know that an absolute minimum and maximum *exist*. We do not need to calculate Hessians to check for local minima or maxima, because they are not asked. The absolute extrema can be found by selecting from critical points in the interior and critical points on the boundary.

Let's first study any interior critical points of the function f given by $f(x,y,z) := x^2+(y-1)^2+z^2-xyz$: For the gradient to vanish, we need

$$\begin{aligned}(\partial_1 f)(x,y,z) &= 2x - yz = 0\\ (\partial_2 f)(x,y,z) &= 2(y-1) - xz = 0\\ (\partial_3 f)(x,y,z) &= 2z - xy = 0\end{aligned}$$

Combining the first and third condition, we get $2x^2 = 2z^2 = xyz$. So we conclude either (1) $x = z$, $y = 2$ or (2) $x = -z$, $y = -2$, or (3) $x = z = 0$. Plugging each case into the 2nd equation we get: two solutions $(x,y,z) = (\pm\sqrt{2}, 2, \pm\sqrt{2})$ from (1). Two solutions $(x,y,z) = (\pm\sqrt{6}, -2, \mp\sqrt{6})$ from (2). One solution $(x,y,z) = (0,1,0)$ from (3). Since we have only been drawing conclusions from this system, we need to check that all these five solutions do satisfy all three equations. They do. The solutions from (2) are not in the ball and may therefore be discarded.

We note $f(\pm\sqrt{2}, 2, \pm\sqrt{2}) = 1$ and $f(0,1,0) = 0$.

We parametrize the boundary by spherical coordinates: $y = 3\cos\vartheta$, $x = 3\sin\vartheta\,\cos\varphi$, $z = 3\sin\vartheta\,\sin\varphi$. While this is originally meant for $\varphi \in [0, 2\pi]$ and $\vartheta \in [0,\pi]$, we can extend the coordinates by 'wrapping over' to $\varphi, \vartheta \in \mathbb{R}$. This reflects the fact that the sphere does not have a boundary. We only need to look for interior critical points of the 2-variable function

$$\begin{aligned}g(\vartheta,\varphi) &:= f(3\sin\vartheta\,\cos\varphi, 3\cos\vartheta, 3\sin\vartheta\,\sin\varphi)\\ &= 9\sin^2\vartheta + (3\cos\vartheta - 1)^2 - 27\cos\vartheta\,\sin^2\vartheta\,\sin\varphi\,\cos\varphi\\ &= 10 - 6\cos\vartheta - \tfrac{27}{2}\cos\vartheta\,\sin^2\vartheta\,\sin 2\varphi\end{aligned}$$

The vanishing of g_φ requires $\cos\vartheta\,\sin^2\vartheta\,\cos(2\varphi) = 0$. This means $\vartheta \in \{0, \frac{\pi}{2}, \pi\}$ or $\varphi \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$.

The vanishing of g_ϑ requires $6\sin\vartheta + \frac{27}{2}\sin^3\vartheta\,\sin 2\varphi - 27\sin\vartheta\,\cos^2\vartheta\,\sin 2\varphi = 0$.

$\vartheta \in \{0, \pi\}$ satisfy this condition and give rise to boundary critical points $f(0,3,0) = 4$ and $f(0,-3,0) = 16$.

$\vartheta = \frac{\pi}{2}$ (and hence $y = 0$) still needs $\sin 2\varphi = -\frac{4}{9}$ for g_ϑ to vanish. We do not need to pursue these critical points further because $g(\frac{\pi}{2}, \varphi) \equiv 10$, so they would be neither absolute minima nor absolute maxima. *Note: Lest this seem confusing: the function g is constant on the 'equator' $y = 0$, $x^2 + z^2 = 9$. But not all points on the equator are critical points, because criticality also depends on how the function changes off the equator.*

$\varphi \in \{\frac{\pi}{4}, \frac{5\pi}{4}\}$ still needs $6\sin\vartheta + \frac{27}{2}\sin^3\vartheta - 27\sin\vartheta\,\cos^2\vartheta = 0$ for g_ϑ to vanish. Apart from retrieving $\vartheta \in \{0,\pi\}$, which has been discussed already, we have to solve $6 + \frac{27}{2}(1 - \cos^2\vartheta) - 27\cos^2\vartheta = 0$, or $\cos\vartheta = \pm(\frac{13}{27})^{1/2}$. Then $\sin^2\vartheta = \frac{14}{27}$ and $g = 10 \mp 13(\frac{13}{27})^{1/2}$.

Similarly, $\varphi \in \{\frac{3\pi}{4}, \frac{7\pi}{4}\}$ still needs $6\sin\vartheta - \frac{27}{2}\sin^3\vartheta + 27\sin\vartheta\,\cos^2\vartheta = 0$ for g_ϑ to vanish. Apart from retrieving old critical points, we have to solve $6 - \frac{27}{2}(1 - \cos^2\vartheta) + 27\cos^2\vartheta = 0$, or $\cos\vartheta = \pm(\frac{5}{27})^{1/2}$. Then $\sin^2\vartheta = \frac{22}{27}$ and $g = 10 \pm 5(\frac{5}{27})^{1/2}$, values that do not qualify for a global extremum.

Concludingly, we have found the global minimum 0 at $(x,y,z) = (0,1,0)$ and the global maximum $10 + 13(\frac{13}{27})^{1/2} \approx 19.0206$ on the boundary at $(x,y,z) = (\pm(\frac{7}{3})^{1/2}, -(\frac{13}{3})^{1/2}, \pm(\frac{7}{3})^{1/2})$.

Hwk #37:

> Redo the boundary part of the calculations from the previous problem using Lagrange multipliers.

Solution: Minimizing/maximizing the expression $f(x, y, z) := x^2 + (y-1)^2 + z^2 - xyz$ on the level set $h(x, y, z) = x^2 + y^2 + z^2 - 9 = 0$ leads to the following necessary conditions for a local extremum:

$$\begin{aligned} 2x - yz &= 2\lambda x \quad (1) \\ 2(y-1) - xz &= 2\lambda y \quad (2) \\ 2z - xy &= 2\lambda z \quad (3) \\ x^2 + y^2 + z^2 &= 9 \quad (4) \end{aligned}$$

Guided by the symmetry between x and z, we seek to simplify this system by calculating $x\cdot(1) - z\cdot(3)$. After slight rearrangement, this gives $2(1-\lambda)(x^2 - z^2) = 0$. So either $x = \pm z$ or $\lambda = 1$. We also subtract $(1) - (3)$ directly and get $(2+y)(x-z) = 2\lambda(x-z)$. So if $x \neq z$, we need $y = 2\lambda - 2$.

The case $\lambda = 1$ leads to $yz = 0 = xy$, $xz = -2$. So since x, z cannot vanish, we need $y = 0$. The combination $xz = -2$, $x^2 + z^2 = 9$ gives four solutions (the ones on the equator in the previous problem), and for all of them, the value of f is 10. We still have to consider the cases $x = z$ and $x = -z$.

Let's now look at the case $x = z$. Either $x = z = 0$ and hence $y = \pm 3$ with values 4 or 16 for f, or else we can cancel x from (1) and get $2\lambda = 2 - y$. Plugging this into (2) we get $2(y-1) - x^2 = (2-y)y$. And (4) becomes $2x^2 + y^2 = 9$. These togehter are equivalent to $x^2 = \frac{7}{3}$, $y^2 = \frac{13}{3}$.

In the case $x = -z$ we may assume $x \neq z$ (hence $x, z \neq 0$), b/c $x = z$ has been studied already. So we have then $2\lambda = y + 2$ and this case leads to $-2 + x^2 = y^2$ from (2) and $2x^2 + y^2 = 9$ from (4). Hence $x^2 = \frac{11}{3}$ and $y^2 = \frac{5}{3}$. The values of f corresponding to this case are $10 \pm \frac{5}{3}(\frac{5}{3})^{1/2}$.

Hwk #38:

> This problem gives you the most celebrated use of Lagrange multipliers, but it requires some intrduction to appreciate it. (The calculations aren't bad at all.)
>
> A famous task in linear algebra and matrix theory is to find *eigenvalues* of a given matrix. If A is a square matrix and you can find a *non-zero* vector v such that Av is actually a multiple of v, i.e., λv where λ is a number, then we call λ an eigenvalue of the matrix A (and v an eigenvector). For instance, the matrix $A = \begin{bmatrix} 3 & 2 \\ -3 & -4 \end{bmatrix}$ has 2 as an eigenvalue and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ as a corresponding eigenvector, because
>
> $$\begin{bmatrix} 3 & 2 \\ -3 & -4 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$
>
> It also has -3 as an eigenvalue with $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ as an eigenvector. Of course multiples of eignevectors are again eigenvectors, e.g., if $Av = 2v$ then also $A(7v) = 2(7v)$. — In the example, there are only these two numbers $\lambda_1 = 2$ and $\lambda_2 = -3$ that are eigenvalues. If you try to find $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ solving $Av = \lambda v$ for any other λ you will only get the solution $v_1 = v_2 = 0$, i.e., only the zero vector. (Try it, just to gain familiarity with the notions.)

This problem is about eigenvalues of *symmetric* matrices. They play a role in studying definiteness of symmetric matrices. In physics, they are key concepts in describing rotating motions of rigid bodies. To every body, there is associated a symmetric 3×3 matrix called its 'tensor of inertia', whose eigenvectors point in the directions of such axes about which the body can rotate without wobbling (i.e., in a balanced way). The eigenvalues are called the moments of inertia about these axes.

To every symmetric $n \times n$ matrix A we associate the quadratic function $f(x) := x^T Ax$ where $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$. We try to minimize or maximize $f(x)$ under the constraint $x^T x = 1$ (i.e., for x on the unit sphere).

(a) Write out $f(x)$ in components x_i for a 3×3 matrix A whose entries are called a_{ij}. Explain why a global maximum and a lobal minimum of $f(x)$ on the sphere are a-priori guaranteed to exist.

(b) Use the Lagrange multiplier method to set up equations satisfied by the x providing a minimum or a maximum. (You may have written all these in components; but now make sure to rewrite the whole stuff again in matrix and vector form.) While you are not asked to actually solve for x (that would be very tedious, involving a cubic equation for λ), I ask you to express the value of f at the minimum and maximum in terms of the Lagrange multiplier. [Be aware that when finding the max vs the min, x and λ will typically refer to numerically different quantities in these two cases.]

You have just proved that every symmetric 3×3 matrix has (at least) two real eigenvalues. (Actually, if A is a multiple of the identity matrix, these two eigenvalues coincide.) And with jut a bit more writing, the same can be done for symmetric $n \times n$ matrices.

The method can be cranked up a bit, by throwing in further constraints, to prove that every symmetric $n \times n$ matrix has n real eigenvalues (some of which may coincide). This may well be among the most important pieces of insight in undergraduate mathematics, and it's a pity that it often falls between the cracks of separating Calc 3 and Linear Algebra into independent courses of the curriculum.

(c) Show, in a very brief calculation: If A is positive definite, then all its eigenvalues are positive. If A is positve semidefinte, then all of its eigenvalues are ≥ 0. FYI: The converse is also true; so indeed a symmetric matrix is positve definite (resp. emidefinite) IF AND ONLY IF all of its eigenvalues are positive (resp. non-negative). This statement is sponsored by the above proof (a), (b) and some extra dose of linear algebra. It is the launch pad for proving the Hurwitz and Gershgorin tests I gave you before.

Solution: (a) Using the symmetry of the matrix already, we get

$$\begin{aligned} f(x_1, x_2, x_3) &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \end{aligned}$$

Being a polynomial, this function is in particular continuous, and we constrain it to the sphere $x_1^2 + x_2^2 + x_3^2 = 1$, which is a closed and bounded set. Therefore a minimum and a maximum of f on the sphere exist.

(b) The Lagrange multiplier method says that at a constrained minimum, and at a constrained maximum, there exists a $\lambda \in \mathbb{R}$ such that $\nabla f = \lambda \nabla g$ (where $g(x) := x^T x - 1 = 0$ is the constraint), provided ∇g does not vanish on the set $g = 0$. We can easily see $\nabla g(x) = 2x$ (and this clearly does not vanish on the constraining set $x^T x = 1$). Let's calculate $\nabla f(x)$:

$$\nabla f(x) = \begin{bmatrix} \partial_1 f(x) \\ \partial_2 f(x) \\ \partial_3 f(x) \end{bmatrix} = \begin{bmatrix} 2a_{11}x_1 + 2a_{12}x_2 + 2a_{13}x_3 \\ 2a_{12}x_1 + 2a_{22}x_2 + 2a_{23}x_3 \\ 2a_{13}x_1 + 2a_{23}x_2 + 2a_{33}x_3 \end{bmatrix} = 2Ax$$

We conclude there exist $\lambda \in \mathbb{R}$ and a nonzero vector $x \in \mathbb{R}^3$ such that $2Ax = 2\lambda x$, hence λ i an eigenvalue and x is an eigenvector. (We know more about x than that it is non-zero; actually $\|x\| = 1$.)

Since $Ax = \lambda x$ at the extremum, we conclude $x^T Ax = \lambda x^T x = \lambda$, so the eigenvalue is actually the value of the constrianed maximum / minimum of f. Unless these two coincide (in which case $x^T Ax$ would have to be constant on the sphere, i.e., A would have to be a multiple of the unit matrix I), we indeed have found two different solutions λ (one for the min and one for the max).

(c) If A has an eigenvalue λ (with eigenvector $x \neq 0$), i.e., if $Ax = \lambda x$, then $x^T Ax = \lambda \|x\|^2$. For a positive definite matrix A, this expression must be positive for *all* vectors $x \neq 0$, in particular for the eigenvector x Hence $\lambda > 0$. Similarly we can argue for positive semidefinite.

Hwk #39:

> Think of the task of finding the absolute maximum of $x^2 + \frac{1}{2}y^2 + y^4 - xy$ on the set S given by $(x-1)^2 + |y + y^3| \leq 5$. The purpose of this problem is NOT that you would actually do calculations to FIND the maximum (which would require numerical methods). Rather, in preparation for such a search. I want you to use the Hessian to conclude that the maximum exists and is on the BOUNDARY of the set S.
>
> *The message here is: While modest problems can already lead to prohibitively complicated calculations that may need numerical tools, simple analytic arguments may still be able significantly to reduce the amount of labor in a numerical search.*

Solution: First we note that an absolute maximum exists, b/c we have a continuous expression on a bounded and closed set. We argue that any absolute maximum in this case cannt be in the interior because the Hessian is not negative semidefinite anywhere. (We do not attempt to solve the equations from vanishing of the gradient, even though we might of course have tried this, too.

The Hessian is $\begin{bmatrix} 2 & -1 \\ -1 & 1 + 12y^2 \end{bmatrix}$. It is clearly positive definite everywhere by the Hurwitz test. So if a critical point were to be found in the interior of S, it would not be an absolute maximum (but a relative minimum at least; possibly an absolute minimum).

Hwk #40:

> You may or may not have seen the following formula (called Heron's formula): The area of a triangle with sides a, b, c is $\sqrt{s(s-a)(s-b)(s-c)}$ where s is the semiperimeter $\frac{1}{2}(a+b+c)$. Show that among all triangles with a given perimeter $2s = a + b + c$, the area takes an absolute minimum exactly for the equilateral triangle. (Explain first why an absolute minimum exists before calclulating it.)

Solution: Since $a, b, c \geq 0$ and $a + b + c = 2s$, clearly, the admissible choices of a, b, c lie in a bounded set. The set is also closed (because it is given by equations and non-strict inequalities for continuous expressions). The function $A = \sqrt{s(s-a)(s-b)(s-c)}$ is continuous, so it does take on global minima and maxima.

The boundary (when one of a, b, c is 0) does not qualify for maxima, because the area must then be 0. This is geometrically expected, but can also be seen from the formula as follows: If $a = 0$ then $s(s-a)(s-b)(s-c) = \frac{s}{8}(b+c-a)(c+a-b)(a+b-c) = -\frac{s}{8}(b+c)(c-b)^2 = -\frac{s^2}{4}(c-b)^2$ which can only be ≥ 0 (to have a square root) if $b = c$, and then the expression is 0, hence the area is 0. Similarly, when $b = 0$ or $c = 0$, the area is 0. Since there are triangles with positive area available, the boundary cases where one of a, b, c is 0, cannot be absolute maxima. There are other boundary points when e.g., $a = b + c$, which makes one factor under the square root 0. But they, too, give rise to area 0.

So the absolute maximum (or maxima, if several) must be in the interior.

Now we can search for critical points by means of Lagrange multipliers, and the absolute maximum (or maxima) must be among the critical points. Rather than minimizing the positive quantity A, we minimize A^2, which gives the same solutions (when $A \geq 0$, then A is largest exactly when A^2 is largest).

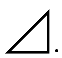
So we maximize $s(s-a)(s-b)(s-c)$ under the constraint $a + b + c - 2s = 0$. We get the equations

$$\begin{aligned}
\tfrac{\partial}{\partial a}s(s-a)(s-b)(s-c) &= \lambda\tfrac{\partial}{\partial a}(a+b+c-2s) \quad \text{i.e.,} \quad & -s(s-b)(s-c) &= \lambda \\
\tfrac{\partial}{\partial b}s(s-a)(s-b)(s-c) &= \lambda\tfrac{\partial}{\partial b}(a+b+c-2s) \quad \text{i.e.,} \quad & -s(s-a)(s-c) &= \lambda \\
\tfrac{\partial}{\partial c}s(s-a)(s-b)(s-c) &= \lambda\tfrac{\partial}{\partial c}(a+b+c-2s) \quad \text{i.e.,} \quad & -s(s-a)(s-b) &= \lambda
\end{aligned}$$

Since we may neglect cases where $s-a$ or $s-b$ or $s-c$ is 0 (as this gives area 0 and thus certainly not a maximum), we can divide the equations on the right pairwise: For instance dividing $-s(s-b)(s-c) = \lambda$ by $-s(s-a)(s-c) = \lambda$ immediately gives $\frac{s-b}{s-a} = 1$, hence $a = b$. Dividing the 2nd by the 3rd equation gives $b = c$. So the only possible maximum can happen when $a = b = c$, an equilateral triangle. Since $a + b + c = 2s$, this requires $a = b = c = \frac{2}{3}s$. Then $A^2 = s(s-a)(s-b)(s-c) = \frac{3}{2}a(\frac{1}{2}a)^3 = \frac{3}{16}a^4$.

Hwk #41:

> Let T be the set $\{(x, y) \mid 0 \leq x \leq \pi,\ 0 \leq y \leq x\}$. Draw a figure of this set. Then evaluate the integral $I := \int_T \sin x \, \sin y \, d(x, y)$ in two ways: as iterated integral in either order.
>
> *Hint: Make sure you get the limits of integration right. If any of your calculations leaves a dangling x or y in the result you sure haven't gotten the limits right. This alert applies to all MV integral problems.*

Solution: T is a triangle: . We can integrate over y first: for each fixed x, we note that y runs from 0 to x. These y integrals occur for x from 0 to π. So,

$$I = \int_0^\pi \Big(\int_0^x \sin x \, \sin y \, dy\Big) \, dx = \int_0^\pi (1 - \cos x) \sin x \, dx = 2 \ .$$

Alternatively, we can integrate over x first. Then, for each fixed y, the x-integral extends from y to π. So we have

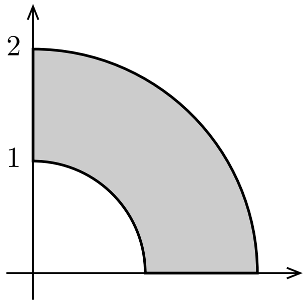
$$I = \int_0^\pi \Big(\int_y^\pi \sin x \, \sin y \, dx\Big) \, dy = \int_0^\pi (1 + \cos y) \sin y \, dy = 2 \ .$$

Hwk #42:

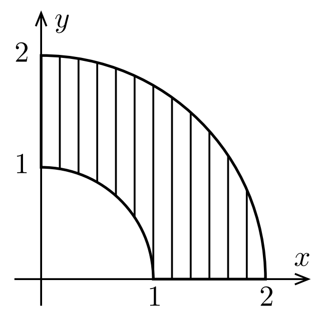
Let A be the set $\{(x,y) \mid 1 \le x^2+y^2 \le 4,\ x,y \ge 0\}$. Draw a figure of this set. Then evaluate the integral $I := \int_A x^2y\,d(x,y)$ in two ways: one version using cartesian coordinates, and one using polar coordinates.

Using cartesian coordinates here is a bit dumb, admittedly. But I am asking that you do it anyways, to see the comparison with polar coordinates, and as a training to deal with the limits of integration correctly. Note that one order of integration in cartesian coordinates is easier to calculate than the other. Can you see which, and why?

Solution:



The domain is a quarter of an annulus as depicted on the left. Integration in cartesian coordinates requires splitting the outer integral into two, because the limits of integration in the inner integral are given by a piecewise function. We'll carry out both orders of integral in cartesian coordinates, for illustration purposes. But using $\int \dots dy$ as the inner integral is easier, because the antiderivative $y^2/2$ makes the square roots disappear from the outer integral.



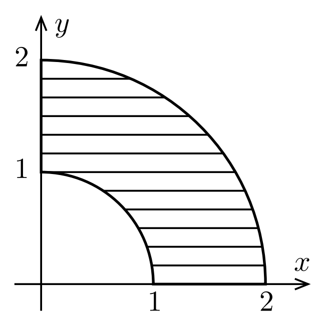
First we consider the integration with y as the inner integral. If $0 \le x \le 1$, the y integral extends between the two circular arcs, namely from $y = \sqrt{1-x^2}$ to $y = \sqrt{4-x^2}$. If $1 \le x \le 2$, the y integral extends from $y = 0$ to $y = \sqrt{4-x^2}$. So the outer integral needs to be split into $\int_0^1 \dots dx + \int_1^2 \dots dx$.

$$I = \int_0^1 \left(\int_{\sqrt{1-x^2}}^{\sqrt{4-x^2}} x^2y\,dy \right) dx + \int_1^2 \left(\int_0^{\sqrt{4-x^2}} x^2y\,dy \right) dx$$

This evaluates to

$$I = \int_0^1 x^2 \left[\frac{y^2}{2}\right]_{y=\sqrt{1-x^2}}^{y=\sqrt{4-x^2}} dx + \int_1^2 x^2 \left[\frac{y^2}{2}\right]_{y=0}^{y=\sqrt{4-x^2}} dx = \int_0^1 \frac{3}{2}x^2\,dx + \int_1^2 \frac{1}{2}x^2(4-x^2)\,dx = \frac{1}{2} + \frac{14}{3} - \frac{31}{10}$$

Doing the fractions we get $I = \frac{31}{15}$.



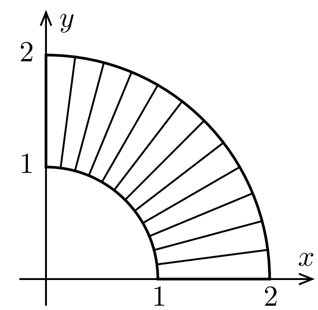
Next we consider the integration with x as the inner integral. If $0 \le y \le 1$, the x integral extends between the two circular arcs, namely from $x = \sqrt{1-y^2}$ to $x = \sqrt{4-y^2}$. If $1 \le y \le 2$, the x integral extends from $x = 0$ to $x = \sqrt{4-y^2}$. So the outer integral needs to be split into $\int_0^1 \dots dy + \int_1^2 \dots dy$.

$$I = \int_0^1 \left(\int_{\sqrt{1-y^2}}^{\sqrt{4-y^2}} x^2y\,dx \right) dy + \int_1^2 \left(\int_0^{\sqrt{4-y^2}} x^2y\,dx \right) dy$$

This evaluates to

$$
\begin{aligned}
I &= \int_0^1 y \left[\frac{x^3}{3}\right]_{x=\sqrt{1-y^2}}^{x=\sqrt{4-y^2}} dy + \int_1^2 y \left[\frac{x^3}{3}\right]_{x=0}^{x=\sqrt{4-y^2}} dy \\
&= \int_0^1 \frac{y}{3}\left((4-y^2)^{3/2} - (1-y^2)^{3/2}\right) dy + \int_1^2 \frac{y}{3}(4-y^2)^{3/2} dy \\
&= -\frac{1}{15}\left[(4-y^2)^{5/2} - (1-y^2)^{5/2}\right]_0^1 - \frac{1}{15}\left[(4-y^2)^{5/2}\right]_1^2 = \frac{4^{5/2}}{15} - \frac{1}{15} = \frac{31}{15}
\end{aligned}
$$

Finally, we evaluate the integral in polar coordinates, using $x = r\cos\varphi$, $y = r\sin\varphi$ and $d(x,y) = r\,dr\,d\varphi$. DON'T FORGET THE r IN THE DIFFERENTIAL! The limits of integration are much easier now: $1 \le r \le 2$ and $0 \le \varphi \le \frac{\pi}{2}$. Order of integration is inessential: since the integrand is a product of an r function and a φ function, both single variable integrals can be evaluated independently of each other.



$$I = \int_0^{\pi/2} \int_1^2 (r^3 \cos^2\varphi \, \sin\varphi)\, r\, dr\, d\varphi$$

We can move the trigs in front of the r-integral. Then the remaining r integral is a constant, not depending on φ, and it can therefore be moved out of the φ integral. Look carefully at this procedure: it will occur often, and at first glance it may look like a "product rule $\int fg = (\int f)(\int g)$". NOT SO! The essence is that the integrand is a product of functions each of which depends on a *different* variable, and that the limits of each integration do not depend on the other variables. It is ONLY then that you can use this trick. If you want to have a name for this rule (there doesn't seem to be an official name), call it the 'TENSOR PRODUCT RULE'.

$$I = \int_0^{\pi/2} \left(\cos^2\varphi\,\sin\varphi \int_1^2 r^4\, dr\right) d\varphi = \left(\int_1^2 r^4\, dr\right)\left(\int_0^{\pi/2} \cos^2\varphi\,\sin\varphi\, d\varphi\right) = \frac{31}{5}\cdot\frac{1}{3}[-\cos^3\varphi]_0^{\pi/2} = \frac{31}{15}$$