General Information: This assignment contains written and/or programming tasks. Combine all the answers to the written tasks in a single PDF document, named `{lastname}-written.pdf`. You can also scan or take pictures of (readable) handwritten papers. JPEG/PNG image files are accepted in this case and they should be named `{exercisenumber}-{lastname}-written.{jpeg/png}`. Make sure that we can follow the manual calculations. Do not combine too many small steps into one. The programming tasks have to be solved in *Julia* and the source code files have to be submitted using the following naming scheme: `{exercisenumber}-{lastname}.jl`.

(1) (1 Point) Compute the arc length of the following functions. **Note** that the arc length for polar coordinates is defined as $L = \int ds$ with $ds = \sqrt{r^2 + (\frac{dr}{d\theta})^2}\, d\theta$

a) $r(\theta) = 1 - cos\theta$ for $0 \leq \theta < \pi$ and sketch the plot of $r(\theta)$ for $0 \leq \theta < 2\pi$.
Hint: $cos(2a) = 2cos^2(a) - 1$.
SOLUTION: We are looking for the arc lenght one half of a cardioid. Considering it's easier and cleaner, we will use polar coordinates.

$$r' = sin(\theta)$$
$$r^2 + (r')^2 = (1 - cos(\theta))^2 + sin(\theta)^2 = 2 - 2cos(\theta)$$

using double angle formula:

$$cos^2(\frac{\theta}{2}) = \frac{1}{2}(1 + cos(\theta))$$
$$r^2 + (r')^2 = 4 - 4cos^2(\frac{\theta}{2})$$
$$\int_0^\pi \sqrt{r^2 + (r')^2} = \int_0^\pi \sqrt{4 - 4cos^2(\frac{\theta}{2})}d\theta$$
$$= \int_0^\pi \sqrt{4sin^2(\frac{\theta}{2})}d\theta$$
$$= \int_0^\pi 2\sqrt{sin^2(\frac{\theta}{2})}$$
$$= 4$$

b) $\gamma(\theta) = (2(\theta + sin\theta), 2(1 + cos\theta))$ for $0 \leq \theta < 2\pi$
SOLUTION: We are solving for the arc length of one full arch of the cycloid with $r = 2$. The arc length for a cycloid is given as $8r$ so 16 in this example. To arrive at this we consider the general formula of cycloids $\gamma(\theta) = (r(\theta + sin\theta), r(1 + cos\theta))$ from which follows:

$$\int_0^{2\pi} \sqrt{(\frac{dx}{d\theta})^2 + (\frac{dy}{d\theta})^2 d\theta}$$

with:

$$\frac{dx}{d\theta} = r(cos(\theta) + 1) \quad \frac{dy}{d\theta} = -rsin(\theta)$$
$$(\frac{dx}{d\theta})^2 + (\frac{dy}{d\theta})^2 = 2r^2(1 + cos(\theta))$$

with: $cos\frac{\theta}{2} = \pm\sqrt{\frac{1+cos(\theta)}{2}}$ and $(1 + cos\theta) = 2cos^2(\frac{\theta}{2})$

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{4r^2cos^2(\frac{\theta}{2})}d\theta \\ &= \int_0^{2\pi} 2rcos(\frac{\theta}{2})d\theta \\ &= 8r = 16 \end{aligned}$$

c) $x(t) = t^{\frac{3}{2}}$, $y(t) = (4-t)^{\frac{3}{2}}$ for $0 \leq t \leq 8$
SOLUTION: We have $ds = \sqrt{dx^2 + dy^2} = \sqrt{(\frac{dx}{dt})^2dt^2 + (\frac{dy}{dt})^2dt^2} = \sqrt{x'(t)^2 + y'(t)^2}dt$. Thus,

$$\begin{aligned} \int_0^8 \sqrt{(\frac{3}{2}t^{\frac{1}{2}})^2 + (\frac{3}{2}(-(4-t)^{\frac{1}{2}}))^2}dt &= \int_0^8 \sqrt{\frac{9}{4}t + \frac{9}{4}(4-t)}dt = \frac{3}{2}\int_0^8 \sqrt{4}dt \\ &= \frac{3}{2}\int_0^8 2dt = 24 \end{aligned}$$

(2) (1 Point) Definiteness of matrices: Check the following matrices on their definiteness with respect to free variables:

a) rotation matrix $\varphi = \frac{2\pi}{6}$

$$rot(\varphi) = \begin{bmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{bmatrix}$$

SOLUTION:

$$rot(\frac{2\pi}{6}) = \begin{pmatrix} cos(\frac{2\pi}{6}) & -sin(\frac{2\pi}{6}) \\ sin(\frac{2\pi}{6}) & cos(\frac{2\pi}{6}) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

The matrix is not symmetric and therefore definiteness does not hold.

b) shearing matrix, $x \in \mathbb{R}$

$$shear(x) = \begin{bmatrix} 2 & x \\ 0 & 2 \end{bmatrix}$$

SOLUTION: To get a symmetric matrix x has to be 0:

$$shear(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow (x, y)^T \cdot \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 + y^2$$

So for $x = 0$ the matrix is positive definite. For all other values definiteness does not hold.

c) scaling matrix, $x \in \mathbb{R}$

$$scale(x) = \begin{bmatrix} 0.5 & 0 \\ 0 & x \end{bmatrix}$$

SOLUTION: To get a symmetric matrix x can take any value in $\mathbb{R}$. Check for definiteness:

$$(y_1, y_2)^T \begin{pmatrix} 0.5 & 0 \\ 0 & x \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_1^2 + x \cdot y_2^2$$

We can see, that if $x > 0$ the matrix is positive definite. If $x = 0$ or $x < 0$ then a vector $(y_1, y_2)^T$ like $(0, 0.5)^T$ for which the linear combination returns 0 as well yields an indefinite matrix.

d) (0.5 Points) Consider some arbitrary high-dimensional non-convex function. Are local minima/maxima encountered less frequently than saddle points? **Hint**: Think in terms of the eigenvalues of the Hessian matrix.

SOLUTION: Saddle points are more common compared to local minima. Actually, the ratio of saddle points to local minima grows exponentially as the dimension n of the function increases. In a saddle point, the Hessian matrix has both positive and negative eigenvalues. On the other hand, a point being a local minima implies that the eigenvalues of the Hessian matrix need to be all positive or all negative. If the sign of an eigenvalue is determined by a coin toss then this means that all n tosses need to have the same outcome (heads or tails) which is a much more rare event compared to the one of having eigenvalues with a different sign. In more detail, the probability of getting all heads or tails after n drops is 2^{1-n} and the probability of everything else is $1 - 2^{1-n}$. Thus the ratio of the probabilities of saddle points to local minima or maxima is $\frac{1-2^{1-n}}{2^{1-n}}$.

(3) (1.5 Points) For each of the following surfaces state whether they are given in implicit, explicit, or parametric form and compute the two missing representations.

a) $2z = 4x + 10y - 15$
SOLUTION: The function is given in explicit form and represents a plane. The corresponding implicit form is $4x+10y-2z-15 = 0$ where $F(x, y, z) = 4x+10y-2z-15 = 0$. To find a parameterization, we need to find two vectors parallel to the plane and a point on the plane. A point on the plane is $\boldsymbol{c} = (0, 0, -7.5)$. The vector normal to the plane is $n = (4, 10, -2)$. Thus we need we search for 2 vector $\boldsymbol{a} = (a_1, a_2, a_3)$ and $\boldsymbol{b} = (b_1, b_2, b_3)$ such that $\boldsymbol{a} \cdot \boldsymbol{n} = 0$ and $\boldsymbol{b} \cdot \boldsymbol{n} = 0$. The equations that we need to satisfy are:

$$4a_1 + 10a_2 - 2a_3 = 0$$
$$4b_1 + 10b_2 - 2b_3 = 0$$

We are free to choose any two parameters and determine the third one so that the equation is satisfied. For example for $a_1 = a_2 = 1$ and $b_1 = 1, b_2 = 0$ we get $a_3 = 7.5$ and $b_3 = 2$. Thus, the plane in parametric form is given by $\boldsymbol{x} = \boldsymbol{c} + s\boldsymbol{a} + \boldsymbol{b}$.

b) $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})(x, y, z)^\mathsf{T} + 5 = 0$
SOLUTION: The expression represents a plane in implicit form, namely $\boldsymbol{n}^T\boldsymbol{x} + c = 0$, with $\boldsymbol{n} = ((1/\sqrt(3), 1/\sqrt(3), 1/\sqrt(3)))$, the normal of the plane, and $c = 5$. The implicit form can be also expressed as $x + y + z + 5\sqrt{3} = 0$, thus the explicit form is $z = -x - y - 5\sqrt{3}$. For the parametric form again we need a point on the plane, which is given by $\boldsymbol{c} = (0, 0, -5\sqrt{3})$ and two vectors parallel to the plane $\boldsymbol{a}$ and $\boldsymbol{b}$ such that $\boldsymbol{a} \cdot \boldsymbol{n} = 0$ and $\boldsymbol{b} \cdot \boldsymbol{n} = 0$. The equations that we need to satisfy are:

$$a_1 + a_2 + a_3 = 0$$
$$b_1 + b_2 + b_3 = 0$$

We are free to choose any two parameters and determine the third one so that the equation is satisfied. For example for $a_1 = a_2 = 1$ and $b_1 = 1, b_2 = 0$ we get $a_3 = -2$ and $b_3 = -1$. Thus, the plane in parametric form is given by $\boldsymbol{x} = \boldsymbol{c} + s\boldsymbol{a} + \boldsymbol{b}$.

c) $(x-6)^2 + (y-3)^2 + z^2 = 25$
SOLUTION: The function is given in implicit form, namely $F(x,y,z) = 0$, where $F(x,y,z) = (x-6)^2 + (y-3)^2 + z^2 - 25 = 0$ and it represents a sphere centered at $(6,3,0)$ with radius $\rho = 5$. For the explicit form we actually get two parts, the upper half of the sphere given by $z = \sqrt{25-(x-6)^2-(y-3)^2}$ and the lower half given by $z = -\sqrt{25-(x-6)^2-(y-3)^2}$. For the parametric form we use:

$$\begin{aligned} x &= 6 + 5\sin\theta\cos\varphi \\ y &= 3 + 5\sin\theta\sin\varphi \\ z &= 5\cos\theta \end{aligned}$$

with $0 \leq \theta \leq \pi$ and $0 \leq \varphi < 2\pi$.

(4) (2 Points) Consider the following functions $f, g\colon \mathbb{R}^2 \to \mathbb{R}^2$ and $h\colon \mathbb{R}^3 \to \mathbb{R}^3$ of several real variables

a) $f(x) = (x_1^2, x_2^2)^\intercal$,
b) $g(x) = (x_1^3x_2^2 - x_1x_2^3 - 1, x_1^2 - x_1x_2^3 - 4)^\intercal$,
c) $h(x) = (2x_1^2 - \cos(x_2x_3) - \frac{3}{2}, 4x_1^2 - 420x_2^3 + 4x_3 - 1, 20x_3 + \exp(-x_1x_2) + 10)^\intercal$,

and compute the Jacobian matrices by hand.

Familiarize yourself with the backslash operator \ in *Julia*. Implement the multivariate Newton method and the three Jacobian matrices in the provided template `multivariate_newton.jl`. Do not explicitly invert the Jacobian matrices, cf. lecture notes. Test your code by running `runtests.jl` to see whether your solution is correct.

```
# Snippet for backslash operator.
A = [-4 -1; 2 2]  # 2x2 matrix
b = [-3; 0]  # right-hand side
x = A\b  # solving the system of linear equations
A*x == b
```

(0.5 Points) Which conditions need to be met in order to use Newton's Method? Could you use another optimization method that we discussed in the lecture? If so, describe the differences between both approaches.

SOLUTION: Function should ideally be differentiable once (twice if we use Hessian), and the Jacobian needs to be non-singular (otherwise it's zero and there is nothing to optimize). Newton's method could further encounter convergence issues if the selected starting value is too far from the root/desired extremum.
We could use gradient descent for slope-based root finding (and other optimization issues). Instead of only updating w.r.t. the Jacobian, we'd use the derivative of our objective function weighted by a learning rate. Thus, gradient descent (GD) is parametric and requires hyperparameter tuning. Convergence is not guaranteed for either method.

(5) (1 Point) Given the vector field $F = [y, -x]^T$ compute the work W done on a particle moving clockwise around $r(t) = (\cos(t) + 3, sin(t) - 3)^T$ for $0 \leq t < 2\pi$. What would happen if we change the orientation of the particle? Re-parameterize its curve and calculate the result.

SOLUTION: For clockwise we can reparameterize to $r(t) = \begin{bmatrix} cos(t) + 3 \\ -sin(t) - 3 \end{bmatrix}$. We then apply $W = \int_C F \cdot dr = \int_a^b F(r(t)) \cdot r'(t)dt$ as discussed during the lecture and substitude:

$$F(r(t)) = \begin{bmatrix} -sin(t) - 3 \\ cos(t) + 3 \end{bmatrix} \qquad r'(t) = \begin{bmatrix} -sin(t) \\ -cos(t) \end{bmatrix}$$

Yielding:

$$\begin{aligned} W &= \int_0^{2\pi} \begin{bmatrix} -sin(t) - 3 \\ -cos(t) - 3 \end{bmatrix} \cdot \begin{bmatrix} -sin(t) \\ -cos(t) \end{bmatrix} dt \\ &= \int_0^{2\pi} sin^2(t) - 3sin(t) + cos^2(t) + 3cos(t)dt \\ &= \int_0^{2\pi} 1 + 3cos(t) - 3sin(t))dt \\ &= (t + 3sin(t) + 3cos(t))|_0^{2\pi} \\ &= 2\pi \approx 6.2832 \end{aligned}$$

When considering movements in a counter-clockwise direct we need to re-parameterize r(t). Since clock-wise movement along a curve is given by: $x = r \cdot cos(kt)$ and $y = r \cdot sin(kt)$, we obtain: $r(t) = \begin{bmatrix} cos(t) + 3 \\ sin(t) - 3 \end{bmatrix}$. We could re-do all of the above calculations or multiply the line integral with -1 to reverse the circle.