ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Visual Computing

AS 2018

Prof. M. Pollefeys
(Computer Vision Lab)
Prof. S. Coros (Computational Robotics Lab)

Y. Aksoy, J. Song, E. Aksan
V. Megaro, J. Bern, M. Geilinger

Final Exam

06 February 2019

First and Last name: ____________________

ETH number: ____________________

Signature: ____________________

General Remarks

- At first, please check that your exam questionnaire is complete (there are 34 pages total).
- Remove all material from your desk which is not allowed by examination regulations.
- Fill in your first and last name and your ETH number and sign the exam. Place your student ID in front of you.
- You have 3 hours for the exam. There are 9 questions, where you can earn a total of 180 points. You don't have to score all points to obtain a very good grade.
- Answer each question directly on the exam sheets. There is enough space left for you to fully answer the questions. If, for whatever reason, your answer is on the back of the sheet (or on an additional piece of paper), state it clearly on the space reserved for the answer.
- Please do not use a pencil or red color pen to write your answers.
- You may provide at most one valid answer per question. Invalid solutions must be cancelled out clearly.

	Topic	Max. Points	Points Achieved	Visum
1	Filtering	21		
2	Miscellaneous	10		
3	PCA	16		
4	Optical Flow	20		
5	Fourier Transform	23		
6	Geometry and Rendering	30		
7	Transformations	30		
8	Animation	22		
9	Optimization	8		
Total		180		

Grade: ....

Question 1: Filters and Image Features (21 pts.)

a) What is the purpose of applying a blur filter before edge or corner detection? **1 pt.**

b) Briefly explain what a separable kernel is and why it is favorable. **2 pts.**

c) What could be the purpose of the kernel given below? What is the effect of convolving an image with this kernel? Explain whether it is a separable kernel or not. **4 pts.**

1	0	-1
2	0	-2
1	0	-1

d) We apply the following filtering operation on an image I:

$$I'(x, y) = \sum_{i,j \in \mathcal{N}(x,y)} f(i, j) I(x + i, y + j)$$

What is the effect if f is a Gaussian kernel? **1 pt.**

e) Bilateral filtering makes the following modification:

$$I'(x,y) = \frac{1}{Z} \sum_{i,j \in \mathcal{N}(x,y)} f(i,j) g(I(x,y) - I(x+i, y+j)) I(x+i, y+j)$$

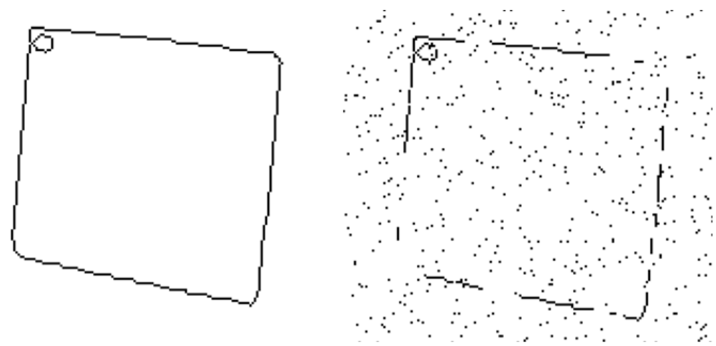
where $g : \mathbb{R} \to \mathbb{R}$ is a Gaussian on the intensity difference and Z is the normalization factor.

i) Explain what this modification implies. **1 pt.**

ii) Write down the normalization factor Z such that the weights in the weighted combination of pixels sum to 1. **1 pt.**

iii) Explain if bilateral filtering is linear and shift-invariant. **2 pts.**

f) Explain how the Hough Transform detects straight lines in an image using the polar parameterization (θ, ρ). **3 pts.**



g) We apply Hough Transform to detect 4 straight lines on the left image below. How does degrading the original image as shown in the right image affect the result? **2 pts.**

h) Explain if Harris corner detector is invariant to intensity shift $I(x, y) \to I(x, y) + a$ for some constant a. **1 pt.**

i) Explain if Harris corner detector is invariant to **intensity scale** $I(x, y) \to bI(x, y)$ for some constant b. **1 pt.**

j) When may temporal redundancy reduction for video compression be ineffective?. **2 pts.**

Question 2: Miscellaneous (10 pts)

a) Typically, why are the colors green and blue chosen for chromakeying? **1 pts.**

b) List three potential problems that may arise when doing background subtraction using a chromakeying approach with a solid colored background (e.g. blue or green screen). Justify each answer. **3 pts.**

c) Suppose you want to capture a video of an actor in front of a static background and you need to perform background subtraction to further process your footage. Assume that the camera is not moving and you cannot cover the surroundings with colored sheets. 1) What else can you do? 2) What if the filming location is inside the Zurich Hauptbahnhof where you cannot stop people from passing by the field of view of your camera? **2 pts.**

d) What are the two input arguments for taking morphological operations? How to determine the output of the morphological operation given these two input arguments? **2 pts.**

e) Name two functions(objectives) of erosion operation. **2 pts.**

Question 3: PCA (16 pts.)

a) Show that the variance of the projected data on a principal component is equivalent to the corresponding eigenvalue of the original covariance matrix. **4 pts.**

b) We apply PCA on two datasets A, B $\in \mathbb{R}^2$. The normalized eigenvalues of dataset A and B are $[0.95, 0.05]$ and $[0.55, 0.45]$, respectively. Illustrate the distribution of dataset A and B in separate plots and explain the relationship between the shape of dataset distribution and eigenvalues. **4 pts.**

c) The PCA algorithm is trained on dataset F of faces and is then used to compress photographs from the dataset C of chairs. Why does not it work well despite "packing the most energy in its first J coefficients"? **2 pts.**

d) You are given a dataset of 1000 images with size 50×50. Calculate the maximum number of principal components that allows us to save the dataset by using at least 1052500 unit space. You should consider all information required to decompress an image or compress a new one. We assume that each array entry or number costs 1 unit space. For example, the original dataset takes $1000 \times 50 \times 50 = 2500000$ unit space. **6 pts.**

Question 4: Optical flow (20 pts.)

a) The Lucas-Kanade algorithm can be used to estimate the optical flow of an image sequence. Nevertheless, it doesn't work for all cases. State the 3 assumptions that have to be fulfilled so that this method works **3 pts.**

b) The Lucas-Kanade method provides the following formulation

$$\begin{pmatrix} \sum \mathrm{I}_x^2 & \sum \mathrm{I}_x\mathrm{I}_y \\ \sum \mathrm{I}_x\mathrm{I}_y & \sum \mathrm{I}_y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{pmatrix} \sum \mathrm{I}_x\mathrm{I}_t \\ \sum \mathrm{I}_y\mathrm{I}_t \end{pmatrix}$$

Where (u, v) is the displacement of a pixel within an image patch.

Mathematically derive the above result using the 3 basic assumptions of this method. Indicate these 3 assumptions during derivation and also point out what is the aperture problem. **8 pts.**

c) Some problems may arise while applying the previous method:

i) Sometimes the solution found is not unique. Under what condition does this happen and why? Give an example and a mathematical explanation. **3 pts.**

ii) In some cases the movement of the objects on the images is too big. What can be done to solve this problem and still be able to apply the Lucas-Kanade algorithm? Explain the process with iterative refinement. **3 pts.**

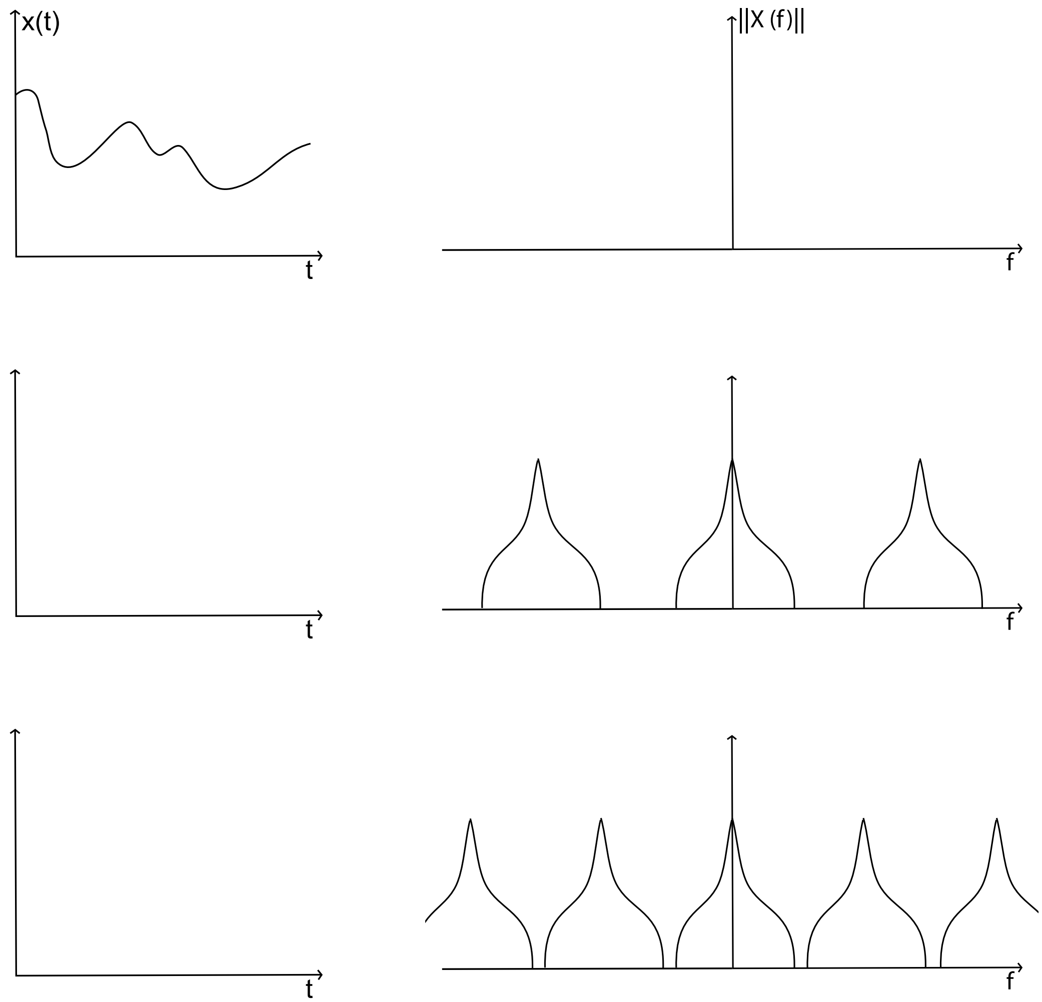
d) i) Explain the general idea of how optical flow information can be used for video compression. **1 pts.**

ii) What image artifacts can appear in video compression if you compute frame n + 1 by applying forward propagation to frame n using the optical flow computed from image n to n + 1. **2 pts.**

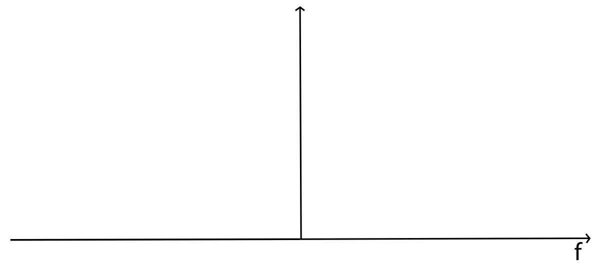
Question 5: Fourier transform (23 pts.)

a) Briefly state the convolution theorem in one or two sentences. **5 pts.**

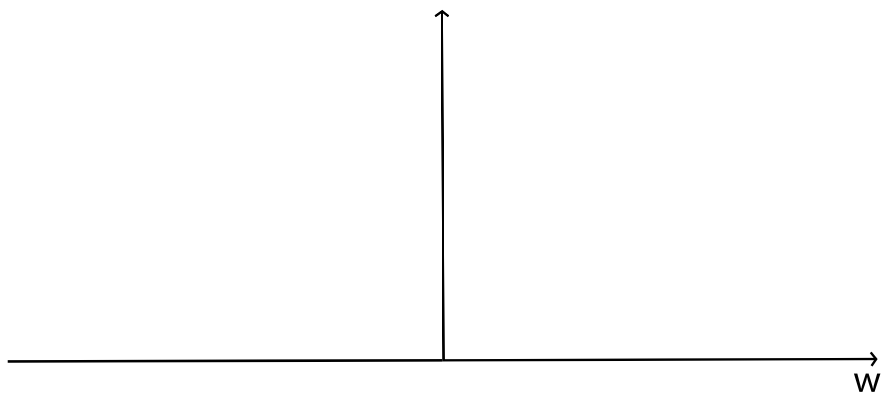
b) Consider the continuous-time signal and the Fourier transforms of two signals that were sampled from the original below. Roughly fill in the missing diagrams. **9 pts.**



c) Assume that the signal in the previous part was sampled at a frequency that is below the Nyquist rate. Roughly draw the sampled signal and its Fourier transform below. **5 pts.**



d) Briefly explain what the `fftshift` function we use to visualize the Fourier domain representations in MATLAB does using the 1-D diagram for the frequency domain below. **4 pts.**



Question 6: Geometry and Rendering (30 pts.)

a) **Geometric Representations**

Given the function

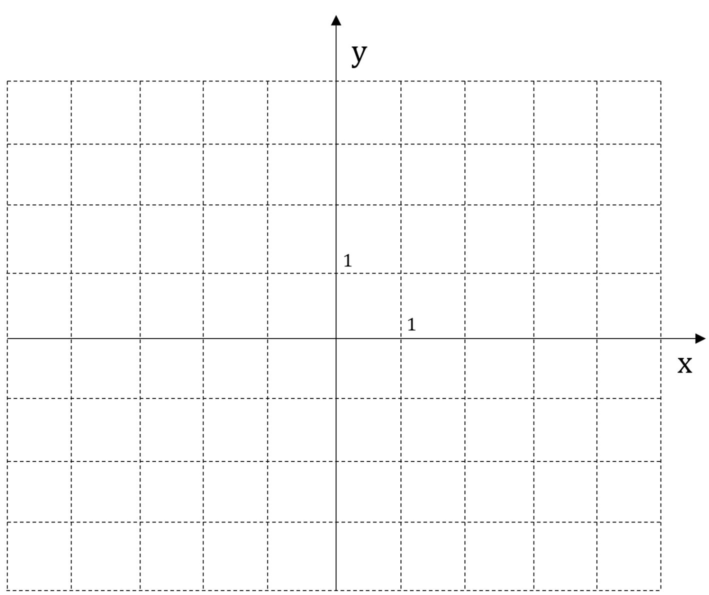
$$f(x, y) = \max\left(\frac{|x|}{2}, |y|\right),$$

$f(x, y) = 1$ describes the boundary of a two-dimensional geometric shape.

i) What type of geometric representation is this? **2 pts.**

ii) Give two advantages and two disadvantages of this type of geometric representation. **3 pts.**

iii) What shape does $f(x, y) = 1$ describe? Draw it in the coordinate system below! **4 pts.**



iv) If $f(x, y) = a$, how does the choice of a influence the resulting shape? **2 pts.**

b) **Rendering and the Graphics Pipeline**

i) In the most general term, what are barycentric coordinates good for? **2 pts.**

ii) With barycentric coordinates we can compute - for example - the normal $\mathbf{n}_x$ of a triangle at point $\mathbf{x}$ with

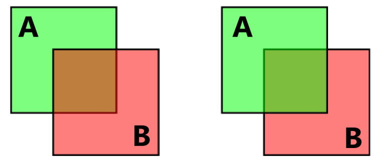
$$\mathbf{n}_x = \alpha\mathbf{n}_a + \beta\mathbf{n}_b + \gamma\mathbf{n}_c$$

where $\mathbf{n}_a$, $\mathbf{n}_b$ and $\mathbf{n}_c$ are the normals at points $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.
Compute the scalars α, β and γ.
Use the notation A_{uvw} for the area of a triangle formed by points $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$. **3 pts.**

iii) Name one method to avoid aliasing using textures. **2 pts.**

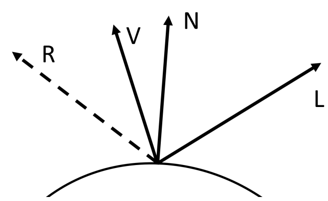
iv) In the graphics pipeline, what is the depth-buffer used for? **2 pts.**

v) In the figure below, two transparent rectangles with transparencies α_A and α_B are overlaid. Below is the *incomplete* equation to compute the composite color C. Complete the equation, such that it computes the composite color of the **right** image below. **3 pts.**



$$C = \qquad \alpha_B B + \qquad \alpha_A A$$

vi) In many reflection models, e.g. the Phong reflection model, the reflection vector $\mathbf{R}$ is used, which is the direction that a perfectly reflected ray of light would take from a point on the surface.
Given the normal vector $\mathbf{N}$, the vector pointing towards the light source $\mathbf{L}$ and the vector pointing towards the camera $\mathbf{V}$, derive the formula to compute the reflection vector $\mathbf{R}$. **4 pts.**



vii) The Phong reflection model describes the way surface reflects light by the combination of three reflection terms. What are these three terms called? And which of these terms uses the reflection vector $\mathbf{R}$? **3 pts.**

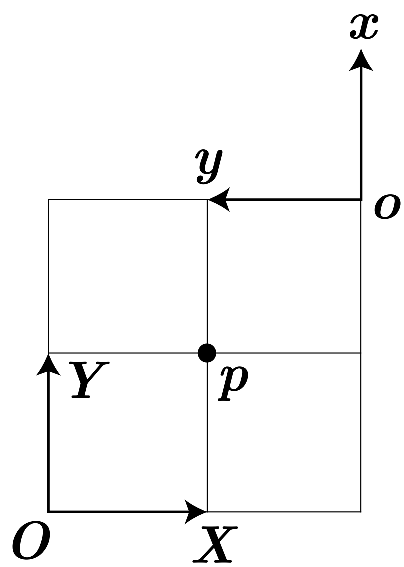
Question 7: Coordinate Systems and Transformations (30 pts.)

a) Coordinate System Basics

This is a 2D problem in homogeneous coordinates.

The 2D Cartesian point $\begin{pmatrix} x \\ y \end{pmatrix}$ is written as a homogeneous point $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$.

The 2D Cartesian vector $\begin{pmatrix} x \\ y \end{pmatrix}$ is written as a homogeneous vector $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$.



The world coordinates are defined by axes $\boldsymbol{X} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\boldsymbol{Y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and origin $\boldsymbol{O} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

The object coordinates are defined by axes $\boldsymbol{x}, \boldsymbol{y}$ and origin $\boldsymbol{o}$.

The grid cells are unit length.

i) Write $\boldsymbol{p}_{\text{world}}$ ($\boldsymbol{p}$ in world coordinates) as a homogeneous point. **1 pt.**

ii) Write $\boldsymbol{p}_{\text{object}}$ ($\boldsymbol{p}$ in object coordinates) as a homogeneous point. **1 pt.**

iii) Write $\boldsymbol{x}$ (in world coordinates) as a homogeneous vector. **1 pt.**

iv) Write $\boldsymbol{y}$ (in world coordinates) as a homogeneous vector. **1 pt.**

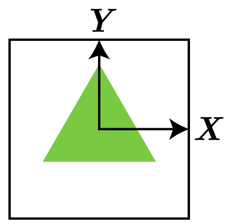
v) Write $\boldsymbol{o}$ (in world coordinates) as a homogeneous point. **1 pt.**

vi) Write $\boldsymbol{p}_{\text{world}}$ as a function of $\boldsymbol{p}_{\text{object}}$ and the 3x3 matrix $\boldsymbol{F} = \begin{pmatrix}\boldsymbol{x} & \boldsymbol{y} & \boldsymbol{o}\end{pmatrix}$.
Substitute in your answers to the (i)-(v) and show the equation holds. **2 pts.**

vii) Say $\boldsymbol{F} = \boldsymbol{T}\boldsymbol{R}$, where $\boldsymbol{R}$ is a rotation matrix, and $\boldsymbol{T}$ is a translation matrix.
Write down valid choices for $\boldsymbol{R}$ and $\boldsymbol{T}$ (using numbers). **4 pts.**

viii) (Challenge) Say $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{T}$, where $\boldsymbol{R}$ is a rotation matrix, and $\boldsymbol{T}$ is a translation matrix.
Write down valid choices for $\boldsymbol{R}$ and $\boldsymbol{T}$ (using numbers). **4 pts.**
HINT: $\boldsymbol{F}$ maps $\boldsymbol{X} \to \boldsymbol{x}$, $\boldsymbol{Y} \to \boldsymbol{y}$, and $\boldsymbol{O} \to \boldsymbol{o}$.

b) Transformation Families

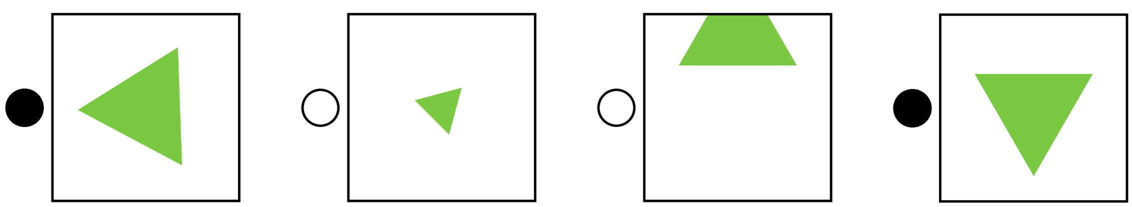


This is the original triangle. The world origin is at the center of the triangle, and the world axes are as drawn as unit vectors. The box around the triangle is **your screen**.

For each question in this section I will give you a parameterized transformation $\boldsymbol{F}_\theta$ and four pictures of triangles **clipped to the screen**. Consider each of the four pictures individually. If a picture is the result of applying $\boldsymbol{F}_\theta$ to the original triangle **for some** θ then **fill in the circle next to it**.

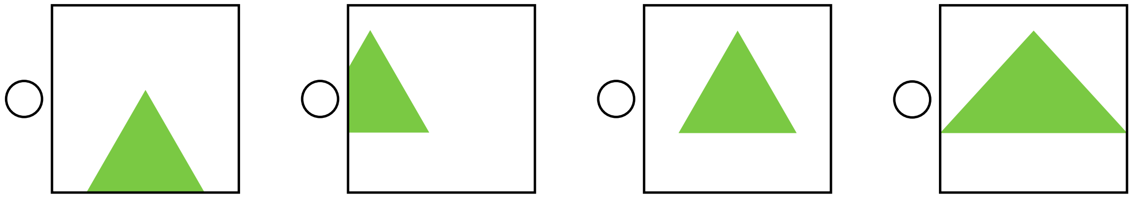
NOTE: To be clear, if the vertices of the original triangle are $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$, then choose a new triangle if and only if its vertices are $\{\boldsymbol{F}_\theta \boldsymbol{p}_1, \boldsymbol{F}_\theta \boldsymbol{p}_2, \boldsymbol{F}_\theta \boldsymbol{p}_3\}$ for some θ.
NOTE: I'm not trying to trick you with the pictures. If a picture looks like the original triangle moved to the right, then it is the original triangle moved to the right.

EXAMPLE: $\boldsymbol{F}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$

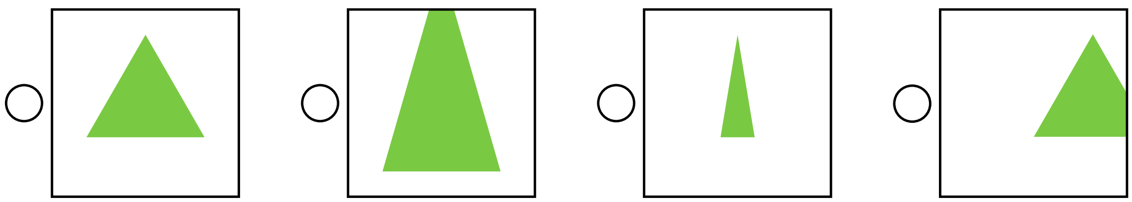


Explanation: This is a rotation by θ. I fill in the circles next to the two pictures that could be made by rotating the original triangle about the origin by some angle θ.
The other two pictures cannot possibly be made by $\boldsymbol{F}_\theta$ so I leave their circles blank.

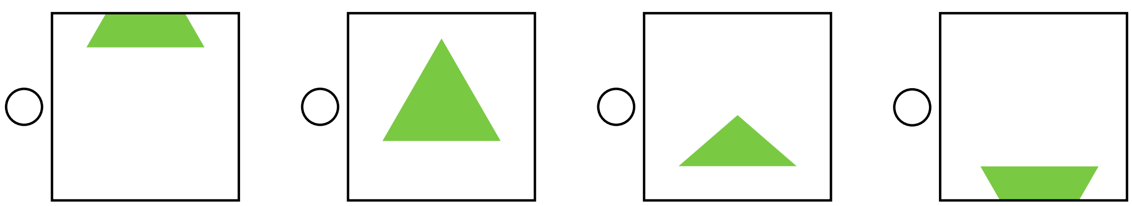
i) $\boldsymbol{F}_\theta = \begin{pmatrix} 1 & 0 & \theta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **4 pts.**



ii) $\boldsymbol{F}_\theta = \begin{pmatrix} \theta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **4 pts.**



iii) (Challenge) $\boldsymbol{F}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \theta & \theta \\ 0 & 0 & 1 \end{pmatrix}$ **4 pts.**



c) Random Questions

i) Do rotations commute in 2D? **1 pt.**

ii) Do translations commute in 7D? **1 pt.**

iii)

$$\boldsymbol{R} = \begin{pmatrix} \cos 1^\circ & -\sin 1^\circ & 0 \\ \sin 1^\circ & \cos 1^\circ & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\prod_{i=1}^{N} \boldsymbol{R} = \underbrace{\boldsymbol{R} * \boldsymbol{R} * \cdots * \boldsymbol{R}}_{N \text{ copies}} = \boldsymbol{I}$$

NOTE: $\boldsymbol{I}$ is the identity matrix.

What is a possible value for N, where $N > 0$? **1 pt.**

Question 8: Animation (22 pts.)

a) **Keyframing**

i) A widely used technique to create animations is called keyframing. Describe the idea behind this technique (maximum two sentences). **1 pt.**

.

ii) A property that can be keyframed is for example the position of an object. Name two other properties of a scene that can be keyframed. **1 pt.**

.

b) **Interpolation**

i) Splines are a great tool to interpolate data. The idea is to create a piece-wise polynomial function to fit data at specific locations. The most commonly used polynomial is of 3rd degree. Explain why this is preferred to linear polynomials and to higher degree polynomials. **2 pts.**

.

ii) Figure 3 shows a 2D plane, a pair of 2D points ($p_0 = [1.0, 0.0]^T$, $p_1 = [2.0, 1.0]^T$), and their respective 2D tangents ($m_0 = [0.0, 1.0]^T$, $m_1 = [0.0, 1.0]^T$). Your task is to use the Cubic Hermite Spline interpolation scheme to compute the 2D point at $t = 0.5$ and then draw the curve between p_0 and p_1. An approximate curve is fine, just hit the critical points and make sure the overall shape is correct.
The general formulation for the Cubic Hermite Spline interpolation is:

$$\mathbf{p}(t) = (2t^3 - 3t^2 + 1)\mathbf{p_0} + (t^3 - 2t^2 + t)\mathbf{m_0} + (-2t^3 + 3t^2)\mathbf{p_1} + (t^3 - t^2)\mathbf{m_1}. \quad (1)$$

5 pts.

.

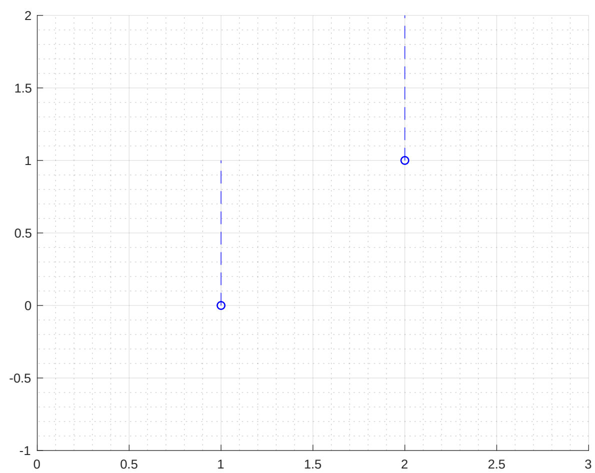


Figure 1: Data points and their tangents.

iii) While creating an animation for a 1D parameter, you decide to use a quadratic polynomial for interpolation. You recall that a parabola has the form:

$$y = at^2 + bt + c. \quad (2)$$

You are also aware that to fully define a parabola you need three points. Therefore, you decide to use three control points: y_0, $y_{0.5}$, and y_1 which are defined at $t = 0$, $t = 0.5$, and $t = 1$, respectively. After solving a linear system of equations, you figure out that the values of a, b, and c:

$$a = 2y_0 - 4y_{0.5} + 2y_1 \quad (3)$$

$$b = -3y_0 + 4y_{0.5} - y_1 \quad (4)$$

$$c = y_0. \quad (5)$$

First, write down the three equations that helped you find a, b, and c. Then, bring your quadratic spline in the following form (i.e. replace the dots between parenthesis):

$$y(t) = (...)y_0 + (...)y_{0.5} + (...)y_1 \tag{6}$$

6 pts.

.

c) **Numerical Integration**

i) In order to advance a physics based animation, we need to perform numerical integration. We have seen: Forward Euler, Backward Euler, and Symplectic Euler. Write down one pro and one con for each of these methods. **3 pts.**

.

ii) Figure 4 shows the image that is used as texture for a cart's wheel (note: the image is completely symmetric). The wheel is represented by a 3D (rigid) disc and its flat surface is textured with figure 4. The disc center as well as its boundaries are aligned with the image. The z-axis is perpendicular to the disc flat surface. In red you can see some measurements: the radius of the wheel is $1m$ and the angle between the spokes of the wheel is $45°$.

The current disc's linear velocity is $v = [0, 0, 0]^T$, whereas its angular velocity is $\omega = [0, 0, \frac{5}{2}\pi]^T$. By using the Forward Euler integration scheme, what is the smallest (positive) time step Δt for which the wheel would appear to be still, even if you are evolving the simulation? **4 pts.**

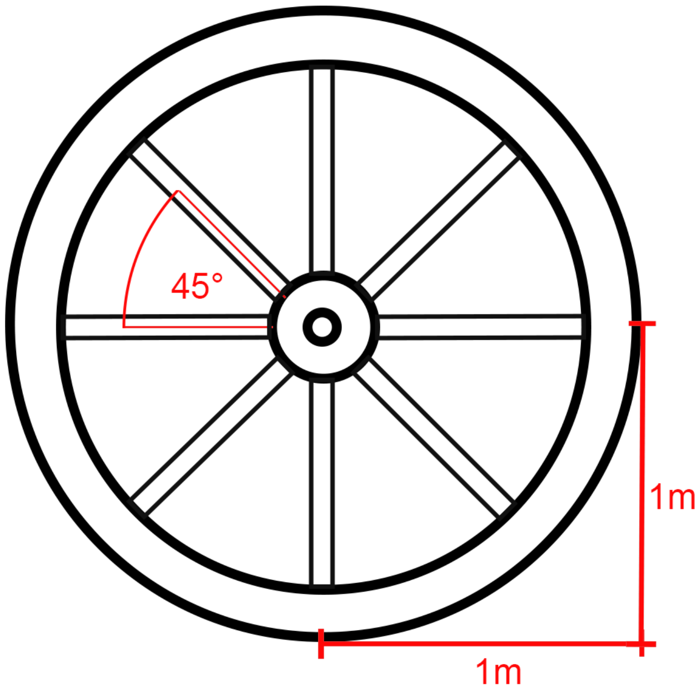


Figure 2: Orthographic view of the wheel. In black: wheel image. In red: measurements (not part of the actual image).

.

Question 9: Optimization (8 pts.)

a) **Local Minima**

i) We want to perform an unconstrained continuous optimization on a function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$. To achieve such a task, we are interested in finding a local minima. List the two optimality conditions which make $\mathbf{x}^*$ the solution to local minima. **2 pts.**

.

ii) We are now interested in maximizing $f(\mathbf{x})$ by minimizing an auxiliary function $g(\mathbf{x})$. Write a function $g(\mathbf{x})$, for which the following equivalence holds:

$$\max_{\mathbf{x}} f(\mathbf{x}) = \min_{\mathbf{x}} g(\mathbf{x}) \tag{7}$$

1 pt.

.

b) **Gradient Descent**

i) We want to perform the first two steps of the *Gradient Descent* algorithm to minimize the function $f(\mathbf{x}) = f(x, y) = \frac{1}{3}x^3 + \frac{1}{2}y^2$. First, write down the general update rule to compute the approximate solution $\mathbf{x}_{k+1}$ given the current approximation $\mathbf{x}_k$. Then, starting at $\mathbf{x}_0 = [1.0, 1.0]^T$, perform the first two steps of Gradient Descent to find out the value of $\mathbf{x}_2$ (for the "step control" variable, use $\tau = \frac{1}{2}$). **5 pts.**

.