

#### BEng, BSc, MEng and MMath Degree Examinations 2022-23

**Department** Computer Science

**Title Computer Vision & Graphics** 

TIME ALLOWED THREE hours (Recommended time to complete TWO hours)

Papers late by up to 30 minutes will be subject to a 5 mark penalty; papers later than 30 minutes will receive 0 marks.

The time allowed includes the time to download the paper and to upload the answers.

**Word Limit** For questions requiring an explanation, two to three sentences should usually be sufficient. A single (textual) answer must not be longer than 100 words (a short paragraph).

Allocation of Marks: Section marks are given in the questions, Justify your answers and, where questions require mathematical calculations, show your working and intermediary steps as appropriate.

#### **Instructions:**

Students must answer all questions.

Submit your answers to the Department's Teaching Portal as a single PDF file.

Use black-on-white only, unless otherwise instructed.

If you have urgent queries regarding a suspected error in the exam, inform (cs exams@york ac.uk) with enough time for a response to be considered and made within the first hour of the start of the exam. Corrections or clarifications will **NOT** be announced after the first hour of the exam.

If a question is unclear and no correction or clarification has been issued, then answer the question as best you can and note the assumptions you have made to allow you to proceed. Inform  $\langle cs\text{-exams@york.ac.uk} \rangle$  about any suspected errors on the paper immediately after you submit.

#### **Note on Academic Integrity**

We are treating this online examination as a time-limited open assessment, and you are therefore permitted to refer to written and online materials to aid you in your answers. However, you must ensure that the work you submit is entirely your own, and for the whole time the assessment is live you must not:

- communicate with other students on the topic of this assessment.
- communicate with departmental staff on the topic of the assessment (other than to highlight an error or issue with the assessment which needs amendment or clarification).
- seek assistance with the assessment from academic support services, such as the Writing and Language Skills Centre or Maths Skills Centre, or from Disability Services (unless you have been recommended an exam support worker in a Student Support Plan).
- seek advice or contribution from any other third party including proofreaders, friends, or family members.

We expect, and trust, that all our students will seek to maintain the integrity of the assessment, and of their award, through ensuring that these instructions are strictly followed. Where evidence of academic misconduct is evident this will be addressed in line with the Academic Misconduct Policy and if proven be penalised in line with the appropriate penalty table. Given the nature of these assessments, any collusion identified will normally be treated as cheating/breach of assessment regulations and penalised using the appropriate penalty table (see AM3.3. of the Guide to Assessment).

- 1 (16 marks) Perspective projection
- (i) [3 marks] The intrinsic matrix of a camera C is given by

$$\mathbf{K} = \left( \begin{array}{ccc} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{array} \right)$$

What do the symbols f,  $c_x$  and  $c_y$  represent in terms of camera C?

- (ii) [13 marks] The intrinsic parameters of camera C are f=300,  $c_x=320$ ,  $c_y=240$ . The extrinsic pose parameters of camera C are given by the matrix  $\mathbf R$  and the vector  $\mathbf t$ .
  - (a) [2 marks] Let  $\mathbf{x}_i$  be an image point in homogeneous coordinates. What is the meaning of the term homogeneous in this context?
  - (b) [3 marks] What do the symbols  $\mathbf{R}$  and  $\mathbf{t}$  represent?
  - (c) [8 marks] Let  $\mathbf{x}_i' = \mathbf{K}^{-1}\mathbf{x}_i$  where both  $\mathbf{x}_i$  and  $\mathbf{x}_i'$  are in homogeneous image coordinates. The x-component of  $\mathbf{x}_i'$  is represented by  $x_i'$  and the y-component by  $y_i'$ . In the special case when  $\mathbf{R}$  is the identity matrix, the vector  $\mathbf{t}$  should satisfy the equation  $\mathbf{A}\mathbf{t} = \mathbf{y}$  where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & x_1 \\ 0 & 1 & -y_1 \\ \vdots & \vdots \\ 1 & 0 & -x_m \\ 0 & 1 & -y_m \end{pmatrix}, \mathbf{y} \neq \begin{pmatrix} x_1w_1 - u_1 \\ y_1w_1 - v_1 \\ \vdots \\ x_m'w_m - u_m \\ y_m'w_m - v_m \end{pmatrix}$$
 (2)

Table 1 gives three image points and their corresponding locations in world coordinates.

,	Point:	#1	#2	#3	
	x	364	395	353	
	y	261	296	290	
	u	-1	1	-1	
	v	-1	1	1	
	w	10	12	14	

Table 1: Table of image points and their corresponding world points

Estimate the vector t.

### 2 (15 marks) Digital cameras and colour spaces

A 12-patch colour target, uniformly lit by a D65 illuminant, is measured using a tele-colorimeter. The tele-colorimeter measurements, represented as xyY values, are reported in table 2, where each column refers to a different patch. The first row reports the x colour coordinate, the second row corresponds to y and the third row to Y:

Table 2: Reference xyY values

P#01 P#02 P#03 P#04 P#05 P#06 P#07 P#08 P#09 P#10	D#111 D#12
0.31 0.31 0.31 0.19 0.37 0.45 0.53 0.30 0.18 0.38	0.20
0.32 0.33 0.33 0.26 0.25 0.48 0.33 0.48 0.14 0.49	0.31 0.17
0.32     0.33     0.33     0.26     0.25     0.48     0.33     0.48     0.14     0.49       0.03     0.20     0.95     0.21     0.20     0.61     0.13     0.25     0.07     0.45	0.19 0.12

A photograph of the same colour target has been taken with a digital camera (12-bit ADC). The de-bayered digital values are given in table 3, where the first row refers to the red channel, the second row to the green channel and the last row to the blue channel. As in table 2, each column refers to a specific patch out of the 12 in the colour target.

Table 3: Acquired values (12-bit)

P#01	P#02	P#03	P#04	P#05	P#06	P#07	P#08	F#09	P#10	P#11	P#12
81	515	2483	305	790	1835	<b>57</b> L	431	145	1057	854	266
128	804	3890	1173	754	1912	327	995	471	1616	558	746
110	680	3260	1314	896	657	236	431 <b>7</b> 995 490	818	656	468	1145

- (i) [10 marks] Compute the matrix M required to map camera data to linear sRGB, using the tele-colorimeter data as ground truth. Describe all the steps required to get to the matrix from the given input data.
- (ii) [5 marks] Apply the colour correction matrix to the camera data. Encode the colours as 16-bit linear RGB values. Which colour corrected value shows the largest distance with respect to the corresponding reference value, if computing the distance within the linear sRGB cube?

- 3 (14 marks) Binocular Stereo
- (i) [4 marks] Two images are taken of the same scene with a stereo camera. The *fundamental* matrix relating the two images is given by

$$\mathbf{F} = \left( \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{array} \right)$$

A point in the left image is given by x and the corresponding point in the right image is given by x'. By considering  $x'^T F x$ , or otherwise, show that the image pair is rectified.

- (ii) [10 marks] The cameras are fronto-parallel with a translation between optical centres of 100mm in the x-direction and 0 in the y- and z-directions. The optical centre and focal length of both cameras are (500,500) pixels and 35mm respectively, and the sensor has 100 pixels per mm.
  - (a) [6 marks] If a point (x, y) in the left image corresponds to a point (x', y) in the right image, then show that

$$x' - x = \frac{\alpha}{w} \tag{4}$$

where w is the depth of the point in camera coordinates. Give the value of  $\alpha$ .

(b) [4 marks] Point (150,675) in the left image and (325,675) in the right image are found to be matching in the stereo image pair. Find the 3D position of the scene point, in the coordinate system of the left camera.

## 4 (18 marks) Reflectance models

In this question you will compute the value of some analytical Bidirectional Reflectance Distribution Function (BRDF) models for a given set of parameters. All vectors are expressed in the local reference frame.

- (i) [10 marks] Compute the value of the Oren-Nayar BRDF for the following sets of parameters. For each point, also write down the components of the halfway vector to three significant figures:
  - Viewing vector:  $\mathbf{v}_r: (\theta_r=\pi/4, \phi_r=3\pi/2)$ ; Lighting direction:  $\mathbf{v}_i: (\theta_r=\pi/4, \phi_r=\pi/2)$ ; diffuse albedo = 0.6;  $\sigma=\pi/6$ ;
  - Viewing vector:  $\mathbf{v}_r = [-0.936, -0.341, 0.087]$ ; Lighting direction:  $\mathbf{v}_i = [-0.168, -0.951, 0.259]$ ; diffuse albedo = 0.9;  $\sigma = 2\pi/9$ .
- (ii) [8 marks] Consider a flat teflon sample (refractive index  $\eta=1.35$ ). Given the viewing vector  $\mathbf{v}_r: (\theta_r=13\pi/36,\phi_r=3\pi/2)$ , compare the value of the Fresnel reflectance using the Schlick's approximation (parameterised in terms of difference angle) with the value given by the Cook-Torrance approximation, for the following viewing directions:
  - $\mathbf{v}_i: (\theta_i = \frac{2\pi}{9}, \phi_i = \frac{2\pi}{3});$
  - $\mathbf{v}_i = [-0.492, 0.853, 0.174].$

- 5 (17 marks) Volumetric representations and Machine Learning
- (i) [4 marks] Explain the concept of a *signed distance function* (SDF) and describe how it can be used in neural networks to learn about shapes.
- (ii) [4 marks] An SDF representing a 3D object is given by

$$s(u, v, w) = 2uv + 2uw - 14$$

The point (2,3,a) lies on the surface of the object. Calculate the value of a and the unit surface normal of the object at this point.

- (iii) [4 marks] You are given the task of training a deep neural network to learn the SDF for a shape. The network is represented by  $f(\mathbf{x};\Theta)$  where  $\mathbf{x}$  is a 3D position and  $\Theta$  contains the network parameters. You are given a set of pairs  $(\mathbf{x}_i,d)$  containing a 3D point and the signed distance to the closest point on the surface. Propose a cost function for training the network.
- (iv) [5 marks] You are given a set of images of the object instead of point clouds. Explain, in general terms, how you might use the concept of self-supervision to learn the SDF, and what extra elements you would need to add.

## 6 (20 marks) Ray tracing

In this question, you will trace the path of a ray within a scene. The only objects contained in the scene are a transparent slab and a sphere, both with smooth surfaces. The coordinates of the vertices of the slab are given in table 4; the sphere has radius  $s_r=5$ , and the coordinates of its center c are: c(10,5,-3). Figure 1 shows the coordinate system and a view of the scene on the x-y plane.

Table 4: Coordinates of the vertices of the transparent slab

							-	
	v#01	v#02	v#03	v#04	v#05	v#06	v#07	v#08
Х	-10	-10	-10	-10	-6	-6	-6	-6
У	-4	-4	2	2	-4	-4	2	2
Z	-2	2	2	-2	-2	2		-2

- (i) [12 marks] Consider the ray with origin  $p_o(-14, -4, 0)$  and direction  $\mathbf{d} = [0.853, 0.523, 0]$ . Assume the slab is made of a transparent material with index Of Refraction (IOR) equal to 1.81. Compute the coordinates of the intersection points between the ray and the sphere.
- (ii) [8 marks] Consider the ray with origin  $p_o(-12, -4.6, 0)$  and direction  $\mathbf{d} = [0.820, 0.572, 0]$ . Assume the slab is made of a transparent material with Index Of Refraction (IOR) equal to 1.31. Compute the coordinates of the intersection points between the ray and the sphere.

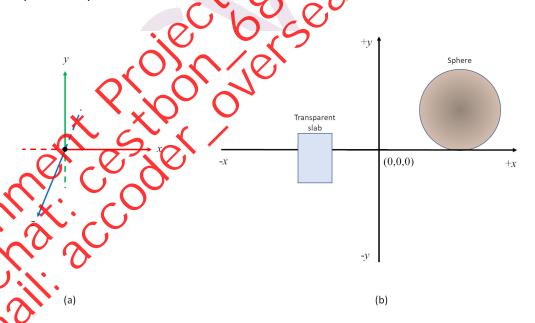


Figure 1: Coordinate system (a) and orthographic view of the scene in the x-y plane (b).

# **End of examination paper**



A Signification of the Property of the Propert

# University of York Department of Computer Science

Degree Examinations 2022-23

**Computer Vision & Graphics** 

**Marking Notes** 

Question 1 (16 marks) Perspective projection

Part (i) [3 marks]

f=sd is the focal length of the camera in pixels (consisting of the physical focal length d and the pixels per metre s.  $(c_z,c_y)$  is the position of the optical centre in the image.

Part (ii) [13 marks]

Sub-part (a) [2 marks]

We construct homogeneous coordinates by adding an extract co-ordinate which is nominally 1. This allows us to convert all operations (rotation, translation and projection) into linear operations implemented by matrix multiplications. The non-linear element of project is then carried out by a final scaling of the homogeneous coordinates so that the final value is restored to 1.

Sub-part (b) [3 marks]

 ${f R}$  is a rotation matrix which rotates world coordinate axes into the same orientation as the camera axes.  ${f t}$  is a translation which moves the origin of the world coordinates onto the camera coordinates origin

Sub-part (c) [8 marks]

Then we should calculate  $f = (A^T A)^{-1} A^T f$  for the least-squares estimate

$$\begin{array}{c}
3.00 \\
1.98 \\
3.95
\end{array}$$

## Question 2 (15 marks) Digital cameras and colour spaces

#### Part (i) [10 marks]

• Step 1 (4 marks): convert the CIE xyY values in table 2 to CIE XYZ. The missing X and Z values can be retrieved using the following equations, as seen in the lectures:

$$\begin{split} & \mathsf{X} = \frac{xY}{y}; \\ & \mathsf{Z} = \frac{(1-x-y)Y}{y}. \end{split}$$

The conversion from CIE XYZ to linear sRGB is done by multiplying the XYZ value by the standard XYZ to RGB matrix:

$$\mathbf{XYZ} - \mathbf{to} - \mathbf{RGB} = \begin{bmatrix} 3.240 & -1.537 & -0.499 \\ -0.969 & 1.876 & 0.042 \\ 0.056 & -0.204 & 1.057 \end{bmatrix}$$

Clamp resulting values as required (data must be in the range [1-1]). The resulting values are reported below, in matrix B.

- Step 2 (1 mark): normalise the acquired camera data, knowing that the ADC is 12-bit. Resulting values are reported below in matrix A
- Step 3 (5 marks): find the colour correction matrix M as the least squares solution of the linear system  $A \cdot M \in B$ .

0.031

0.027

In step 3, the matrices A and B are given by:

$$\mathbf{A} = \begin{bmatrix} 0.126 & 0.196 & 0.166 \\ 0.606 & 0.950 & 0.796 \\ 0.074 & 0.286 & 0.321 \\ 0.193 & 0.184 & 0.219 \\ 0.448 & 0.467 & 0.160 \\ 0.139 & 0.0799 & 0.058 \\ 0.105 & 0.243 & 0.120 \\ 0.035 & 0.115 & 0.200 \\ 0.258 & 0.395 & 0.160 \\ 0.209 & 0.136 & 0.114 \\ 0.065 & 0.182 & 0.280 \end{bmatrix}$$

0.020

$$\mathbf{B} = \begin{bmatrix} 0.031 & 0.030 & 0.032 \\ 0.193 & 0.202 & 0.200 \\ 0.915 & 0.960 & 0.952 \\ 0 & 0.264 & 0.435 \\ 0.500 & 0.100 & 0.297 \\ 0.871 & 0.594 & 0.001 \\ 0.449 & 0.044 & 0.043 \\ 0.065 & 0.322 & 0.079 \\ 0.015 & 0.058 & 0.350 \\ 0.380 & 0.511 & 0.054 \\ 0.574 & 0.083 & 0.120 \\ 0.051 & 0.107 & 0.454 \end{bmatrix}$$

3.011

The resulting  $3 \times 3$  matrix  $\mathbf M$  is:

$$\mathbf{M} = \begin{bmatrix} 3.476 & -1.571 & 0.422 \\ -0.393 & 1.889 & -0.754 \\ -0.091 & -0.616 & 1.998 \end{bmatrix}$$

## Part (ii) [5 marks]

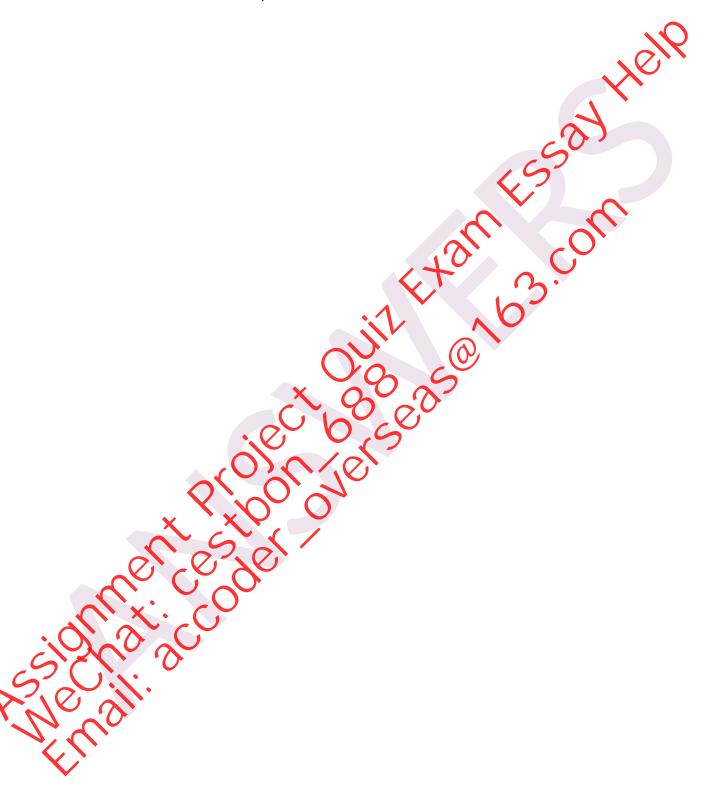
• Step 1 (3 marks): Correct the acquired colours using M and clamp values as required. The resulting colours are:



Step 2: multiply the values by  $2^{16}-1$  and cast to uint16;

- Step 3: do the same as in step 2 with ground truth linear sRGB (calculated in the previous part of the question, matrix B);
  - Step 4 (2 marks): compute the euclidean distance between the reference and mapped rgb values, for corresponding patches.

The patch with the largest distance between corrected colour - ground truth colour is the 7th one. 1 mark deducted if steps 2 and 3 are omitted.



Question 3 (14 marks) Binocular Stereo

Part (i) [4 marks]

If the images are rectified, then epipolar lines should be scanlines. So if  $\mathbf{x}=(u,v,1)$  then  $x'=(u+\Delta,v,1)$ . By definition of the Fundamental matrix,  $\mathbf{x}'TF\mathbf{x}=0$ , so we need to put in the given values to verify this equation holds. Alternatively, we can compute  $F\mathbf{x}$  to find the epipole and observe that it is the (homogenous) equation of a horizonal line.

Part (ii) [10 marks]

Sub-part (a) [6 marks]

Using the basic projection equation

$$x = dsu/w + c_x$$

To obtain the right camera projection, we apply the translation to right camera coordinates  $u \to u + 100$ .

$$x' = ds(u + 100)/w + c_a$$

Subtracting

$$x' - x = -ds(u + 100) / + c_x - dsu/w + c_x = 100 ds/w$$

as required with  $\alpha=100ds$ .

Sub-part (b) [4 marks]

Using the given equation,

$$-x = 175 = 100 ds/$$

 $d=35,\ s=100$  so we can find w=100 (  $35\times 100$ ) 175=2000mm. Now we use the projection equation to find the other coordinates

$$u \neq \frac{w(x - c_x)}{ds} = -200mm$$

$$v = \frac{w(y - c_y)}{ds} = 100mm$$

# Question 4 (18 marks) Reflectance models

## Part (i) [10 marks]

• BRDF: 0.148;  $h_x = 0$ ,  $h_y = 0$ ,  $h_z = 1$ ;

• BRDF: 0.606;  $h_x = -0.636$ ,  $h_y = -0.745$ ,  $h_z = 0.200$ .

Marks breakdown: 5 marks for each bullet point.

## Part (ii) [8 marks]

• Schlick  $\mathcal{F} = 0.028$ , Cook-Torrance  $\mathcal{F} = 0.051$ ;

• Schlick  $\mathcal{F}=0.106$ , Cook-Torrance  $\mathcal{F}=0.106$ .

Marks breakdown: 4 marks for each bullet point.

Question 5 (17 marks) Volumetric representations and Machine Learning

Part (i) [4 marks]

A signed distance function  $s(\mathbf{x})$  is a function of a spatial point which returns the distance from the surface of the object at that point. The distance is signed so that (for example) distances are positive outside the surface and negative inside the surface. From the definition  $s(\mathbf{x}) = 0$  on the surface, so the surface of the object is defined by the zero level set of the function. We can use this idea in a neural network as an NN is a general function which can be learnty given data we can learn a neural network which reproduces the SDF.

Part (ii) [4 marks]

Point position:

$$2uv + 2uw - 14 = 2 \times 2 \times 3 + 2 \times 2 \times w - 14 = 12 + 4w - 14 = 0$$

$$w = (14 - 12)/4 = 1/2$$

Surface normal:

$$\mathbf{n} = \begin{pmatrix} \frac{\partial F}{\partial u} \\ \frac{\partial F}{\partial v} \\ \frac{\partial F}{\partial w} \end{pmatrix} = \begin{pmatrix} 2v + 2w \\ 2u \\ 2u \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ 4 \end{pmatrix}$$

Finally, it should be a unit vector, so

$$\hat{\mathbf{n}} = 0.49$$

Part (iii) [4 marks]

There are a number of possible answers here, the key feature is to obtain an appropriate estimate of the SDF without under fitting. DeepSDF for example only fits the SDF within  $\delta$  of the surface using  $\mathrm{Clamp}(x,\delta) = x$  if  $-\delta < x < \delta$ , 0 otherwise. Then the cost function is the error in the local gradient

$$\ell \in \mathbf{Clamp}(f(\mathbf{x}, \Theta), \delta) - \operatorname{clamp}(d, \delta)|$$

Part (iv) [5 marks]

Self-supervision, in this case, means using the know properties of image formation to provide a supervision signal for the network. We supervise the network by ensuring that the output of a render of the SDF is the same as the images we are given. We require a differentiable rendered R(f,R) which takes the output of the SDF network and renders an image in pose P. We also therefore need to know the pose P. The cost function is the difference between the output of the renderer and the actual image. Since all elements are differentiable, we can train the SDF network with backpropagation.

#### Question 6 (20 marks) Ray tracing

#### Part (i) [12 marks]

• From the vertices of the transparent slab it's trivial to observe that the side facing the ray origin has normal direction  $\mathbf{n}_s = [-1,0,0]$  (in the scene coordinate system). Since we are interested in the angle of incidence with respect to the surface of the slab, we need to refer to the local reference frame, in which the normal direction is always [1,0,0].

The angle of incidence  $\theta_i$  can be computed as:  $\theta_i = \arccos(-\overrightarrow{d} \cdot -\overrightarrow{n_s}) \neq 0.54$  rad

• We need to compute the point  $p_s$  in which the ray hits the surface of the slab. Given the coordinates of the vertices of the slab, the ray can only hit its surface at x=-10. We could use trigonometry or the ray-plane intersection method. Trigonometric solution:

$$\Delta x = -10 - (-14) = 4;$$
  
 $\Delta y = \Delta x * \tan \theta_i = 2.448.$   
 $\Delta z = 0;$   
 $p_s = p_m + (\Delta x, \Delta y, \Delta z) = (-10, -1.548, 0)$ 

• Refracted ray, interface air-slab, trigonometric solution:  $\theta_{air-slab} = \arcsin\left(ior_{air}/ior_{slab} * \sin t_i\right) = 0.2925 \text{rad.}$  The thickness of the slab is 4 (from the vertices), so  $\Delta t_i = 4$ 

$$\Delta y = \Delta x * \tan \theta_{air-slab} = 1.205$$
  
 $\Delta z = 0;$ 

The ray will hit the other side of the slab at coordinates:

$$p_{s2} = p_s + (\Delta x, \Delta y, \Delta x) = (-6, -0.343.0)$$

• Refracted ray, interface slab air: Trivially,  $\theta_{stab}$   $q_{ir} = \theta_i$  Therefore, the new ray to trace has origin  $p_{s2}$  and direction  $\overrightarrow{r}$ .

ullet We can now compute the ray sphere intersection, knowing  $\overrightarrow{d}$  ,  $p_{s2}, s_r$  and c :

$$\begin{array}{l} a = \begin{bmatrix} d & \parallel = 1; \\ b = 2*(r_x*(p_{s2_x} + c_x) + r_y*(p_{s2_y} - c_y) + r_z*(p_{s2_z} - c_z)) = -32.885; \\ \neq (p_{s2_x} - c_x)^2 + (p_{s2_y} - c_y)^2 + (p_{s2_z} - c_z)^2 - s_r^2 = 268.546; \\ \Delta = b - 4ac = 7.216. \end{array}$$

Since  $\Delta>0$  , the ray intersects the sphere in two points,  $p_1$  and  $p_2$ :

$$p_1 = (-b)/(2a = 15.099;$$

$$p_1 = p_{s2} + t_1 \overrightarrow{d} = (6.880, 7.554, 0);$$

$$t_2 = (-b + \sqrt{\Delta})/2a = 17.785;$$

$$p_2 = p_{s2} + t_2 \overrightarrow{d} = (9.171, 8.959, 0);$$

Part (ii) [8 marks]

We can apply the same method used in the first part of the exercise. Intermediate results:

• Point  $p_s$  in which the ray hits the surface of the slab:

 $\Delta x = -2;$ 

 $\Delta y = 1.396.$ 

 $\Delta z = 0$ ;

 $p_s = p_m + (\Delta x, \Delta y, \Delta z) = (-10, -3.205, 0).$ 

• Refracted ray, interface air-slab, trigonometric solution:

 $\theta_{air-slab} = 0.452 \text{rad}.$ 

 $\Delta x = 4$ ;

 $\Delta y = 1.942;$ 

 $\Delta z = 0;$ 

The ray will hit the other side of the slab at coordinates:

 $p_{s2} = p_s + (\Delta x, \Delta y, \Delta z) = (-6, -1.263, 0).$ 

• Refracted ray, interface slab-air:

 $\theta_{slab-air} = \theta_i$ . Therefore, the new ray to trace has origin  $p_s$  and direction  $\overrightarrow{r}$ 

• We can now compute the ray-sphere intersection, knowing  $\overrightarrow{r}$ ,  $p_{s2}$ ,  $s_{r}$  and  $s_{r}$ 

a = 1;

b = -33.411;

c = 279.222;

 $\Delta = b^2 - 4ac = -0.575.$ 

Since  $\Delta < 0$ , in this case there is no intersection, the ray does not hit the sphere.

**End of examination paper**