Phase Space Ray Tracing for Illumination Optics

Carmela Filosa

Cover art:
Photography:

A catalogue record is available from the Eindhoven University of Technology Library

ISBN: 978-90-386-3972-7

title

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een
commissie aangewezen door het College voor
Promoties, in het openbaar te verdedigen

door

Carmela Filosa

geboren te Torre del greco, Italië

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.
1e promotor:	prof.dr. W.L. IJzerman
copromotor:	dr. J.H.M. ten Thije Boonkkamp
leden:	

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Contents

1 Introduction **3**
1.1 Motivation . . . 3
1.2 Methods and results . . . 3
1.3 Content of this thesis . . . 3

2 Illumination optics **5**
2.1 Radiometric and photometric variables . . . 5
2.2 Reflection and refraction law . . . 10
2.3 Fresnel's equations . . . 11

3 Ray tracing **19**
3.1 Ray tracing for two-dimensional optical systems . . . 19
3.2 Monte Carlo ray tracing . . . 21
3.3 Quasi-Monte Carlo method . . . 24

4 Ray tracing on phase space **27**
4.1 Phase space concept . . . 27
4.2 The edge-ray principle . . . 27
4.3 Phase space ray tracing . . . 27

5 Two different appoaches to compute the boundaries in target phase space **29**
5.1 The α-shapes approach . . . 29
5.2 The two-faceted cup . . . 32
5.3 Results for a TIR collimator . . . 32
5.4 The triangulation refinement approach . . . 32
5.5 The two-faceted cup . . . 32
5.6 Results for a TIR collimator . . . 32
5.7 Results for a Parabolic reflector . . . 32
5.8 Results for the Compound Parabolic Concentrator (CPC) . . . 32

6 The inverse ray mapping method: analytic approach **33**
6.1 Explanation of the method . . . 33
6.2 The two-faceted cup . . . 33

6.3 The multi faceted cup . . . 33
6.4 Results for the two-faceted cup . . . 33
6.5 Results for the multi-faceted cup . . . 33
6.6 Discussions . . . 33

7 The extended ray mapping method 35
7.1 Explanation of the method . . . 35
7.2 Bisection procedure . . . 35
7.3 Results for a parabolic reflector . . . 35
7.4 Results for two different kind of TIR-collimators . . . 35

8 Extended ray mapping method to systems with Fresnel reflection 37

9 Discussion and conclusions 39

Acknowledgments 41

Bibliography 43

List of symbols

$\mathrm{d}\Omega$	Solid angle
θ	Angle between the direction of the solid angle and the normal $\boldsymbol{\nu}$
θ_i	Angle between the incident ray and the normal $\boldsymbol{\nu}$
θ_r	Angle between the reflected ray and the normal $\boldsymbol{\nu}$
θ_t	Angle between the transmitted ray and the normal $\boldsymbol{\nu}$
ϕ	Angle that the ray located on line $i$forms with respect to the optical axis

2

Chapter 1

Introduction

1.1 Motivation

1.2 Methods and results

1.3 Content of this thesis

Chapter 2

Illumination optics

This chapter provides some concepts of illumination optics used in this thesis. We start explaining the difference between radiometry and photometry. In particular, we focus on the photometric variables, defining them both in three and two dimensions. The reflection and refraction laws and the phenomenon of total internal reflection are explained. The last paragraph of the chapter gives a brief introduction to Fresnel reflection.

2.1 Radiometric and photometric variables

Radiometry is concerned with the measurement of electromagnetic radiation across the entire electromagnetic spectrum. Photometry is the subfield of radiometry that takes into account only the portion of the electromagnetic spectrum corresponding to the visible light, [1]. Radiometry deals with radiometric quantities. An important radiometric quantity is the radiant flux Φ_{r} (unit watt, [W]) which is the total energy emitted from a source or received by a target per unit time:

$$\Phi_{\mathrm{r}} = \frac{\mathrm{d}Q}{\mathrm{d}T}\,, \tag{2.1.1}$$

where Q is the energy and T the time.

In illumination optics the measurement of light is given in terms of the impression that it gives on the human eye. Therefore, illumination optics deals with photometric variables. The most important photometric variables are defined as follows using the same notation adopted by Chaves in [2]. The luminous flux Φ (unit lumen, [lm]) is defined as the perceived power of light by the human eye. The radiant and the luminous flux are related by the luminous efficacy function, unit [lm/W], which tells us how many lumen there are for each Watt of power at a given wavelength. The luminous efficacy reaches its maximum at a wavelength of 555 nm where it is equal to 683 lm/W. We may normalize the luminous efficacy function with its maximum value of 683. This normalized function is the dimensionless luminosity function $\bar{y}(\lambda)$ shown in Figure 2.1 where λ is the wavelength.

The luminous flux corresponding to one Watt of radiation power at any wavelength is given by the product of 683 lm/W and the luminosity function at the same wavelength,

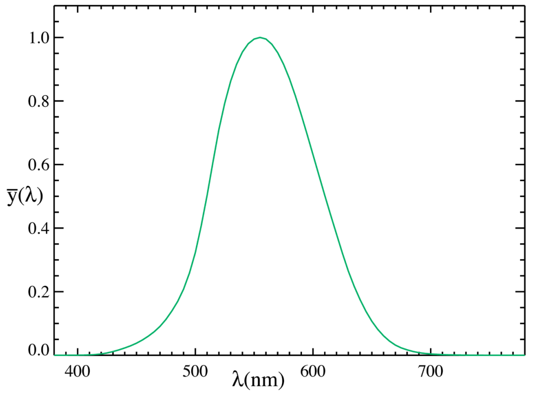

Figure 2.1: Luminosity function $\bar{y}(\lambda)$: relation between the eye's sensitivity and the wavelength of light. The luminosity function is dimensionless, [3].

i.e. $683\,\bar{y}(\lambda)$. Hence, the total luminous flux Φ has unit lumen [lm] and it is defined as:

$$\Phi = 683 \int_0^\infty \Psi_{\mathrm{r}}(\lambda)\bar{y}(\lambda)\mathrm{d}\lambda \tag{2.1.2}$$

where $\Psi_{\mathrm{r}}(\lambda)$ is the power in watt per unit wavelength (unit [W/m]).

A beam of light can be described as a collection of parallel light rays, where a light ray can be interpreted as a path along which the energy travels. The luminous flux $\mathrm{d}\Phi$ incident on a surface is called illuminance E $\left(\text{unit } \left[\mathrm{lm/m^2}\right]\right)$ and is defined as:

$$E = \frac{\mathrm{d}\Phi}{\mathrm{d}A}\,, \tag{2.1.3}$$

where $\mathrm{d}A$ is an infinitesimal area receiving radiation. The density of light emitted by a point source in a given direction is determined by the solid angle. The solid angle on a given direction is defined by the infinitesimal surface area $\mathrm{d}S$ of a sphere subtended by the radius of that sphere and by the rays emitted by the center on that direction, [4]. Indicating with r the radius of the sphere, the infinitesimal solid angle $\mathrm{d}\Omega$ defined by $\mathrm{d}S$ is given by:

$$\mathrm{d}\Omega = \frac{\mathrm{d}S}{r^2}\,. \tag{2.1.4}$$

The solid angle on the entire sphere is $\Omega = 4\pi$, its unit is the steradian [sr] and it is usually defined on a unit sphere. The luminous intensity I (unit candela (cd), [cd = lm/sr]) is defined as the luminous flux $\mathrm{d}\Phi$ per solid angle $\mathrm{d}\Omega$ and is given by:

$$I = \frac{\mathrm{d}\Phi}{\mathrm{d}\Omega}\,. \tag{2.1.5}$$

The luminance L $\left(\text{unit } \left[\mathrm{cd/m^2}\right]\right)$ is the luminous flux per unit solid angle $\mathrm{d}\Omega$ and per unit projected area $\cos\theta \mathrm{d}A$ where θ is the angle that the normal $\boldsymbol{\nu}$ to the area $\mathrm{d}A$ makes with the direction of the solid angle $\mathrm{d}\Omega$, as shown in Figure 2.2. L is given by:

$$L = \frac{\mathrm{d}\Phi}{\cos\theta \mathrm{d}A \mathrm{d}\Omega}\,. \tag{2.1.6}$$

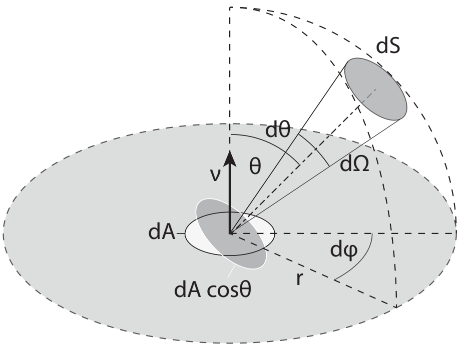

Figure 2.2: Solid angle $\mathrm{d}\Omega$ in a direction making an angle θ with the normal to the area $\mathrm{d}A$.

Note that from (2.1.5) and (2.1.6) we can derive a relation between the intensity and the luminance. The intensity I emitted by the infinitesimal area $\mathrm{d}A$ is given by:

$$I = \frac{\mathrm{d}\Phi}{\mathrm{d}\Omega} = L(\boldsymbol{x}, \theta) \cos\theta \mathrm{d}A\,. \tag{2.1.7}$$

When the luminance is uniform over a finite area A, the luminous intensity emitted in the direction θ is equal to:

$$I(\theta) = L(\theta) A \cos\theta\,. \tag{2.1.8}$$

Thus, when $L(\boldsymbol{x}, \theta)$ does not depend on the position and the direction (i.e. $L(\boldsymbol{x}, \theta) = L$), we obtain Lambert's cosine law:

$$I(\theta) = I_0 \cos\theta\,. \tag{2.1.9}$$

where $I_0 = I(0) = LA$.

Finally, the étendue U (unit [m sr]) describes the ability of a source to emitt light or the capability of an optical system to receive light, [5]. The quantity $\mathrm{d}U$ is defined as:

$$\mathrm{d}U = n^2 \cos\theta \mathrm{d}A \mathrm{d}\Omega. \tag{2.1.10}$$

where n is the index of refraction of the medium in which the surface A is immersed. In optics the étendue is considered to be a volume in phase space (or an area for two-dimensional systems). This concept will be clarified in Chapter 4 in which we treat the phase space in more detail. An important property of the étendue is that it is conserved within an optical system in absence of absorption. We now show, using the approach of Chaves in [2], how conservation of this quantity can be derived. Consider a light ray emitted from an infinitesimal area $\mathrm{d}A_1$ to the area $\mathrm{d}A_2$. Suppose that the centers of $\mathrm{d}A_1$ and $\mathrm{d}A_2$ are located at a distance d to each other, see Figure 2.3. Indicating with $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ the normals to the surfaces $\mathrm{d}A_1$ and $\mathrm{d}A_2$, respectively and with θ_1 and θ_2 the angles that the central ray forms with $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$, respectively, the flux $\mathrm{d}\Phi_1$ passing through $\mathrm{d}A_2$ coming from $\mathrm{d}A_1$ and the corresponding solid angle

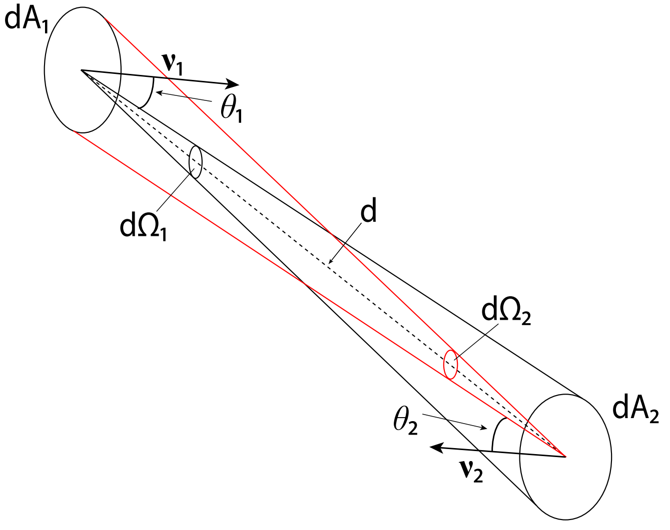

Figure 2.3: $\mathrm{d}A_1$ and $\mathrm{d}A_2$ are two surfaces with normals ν_1 and ν_2, respectively. Their centers are located at a distance d. θ_1 and θ_2 are the angles made by the central ray with the normals ν_1 and ν_2, respectively.

$\mathrm{d}\Omega_1$ are defined as:

$$\begin{aligned} \mathrm{d}\Phi_1 &= L\cos\theta_1 \mathrm{d}A_1 \mathrm{d}\Omega_1, \\ \mathrm{d}\Omega_1 &= \frac{\mathrm{d}A_2\cos(\theta_2)}{d^2}\,. \end{aligned} \tag{2.1.11}$$

Similarly, the flux $\mathrm{d}\Phi_2$ passing through $\mathrm{d}A_1$ coming from $\mathrm{d}A_2$ is equal to:

$$\begin{aligned} \mathrm{d}\Phi_2 &= L\cos\theta_2 \mathrm{d}A_2 \mathrm{d}\Omega_2 \\ \mathrm{d}\Omega_2 &= \frac{\mathrm{d}A_1\cos\theta_1}{d^2}\,. \end{aligned} \tag{2.1.12}$$

Then from Eq. (2.1.10) we obtain the following relations:

$$\begin{aligned} \mathrm{d}U_1 &= n^2 \mathrm{d}A_1\cos\theta_1 \mathrm{d}\Omega_1 = \frac{n^2 \mathrm{d}A_1\cos\theta_1 \mathrm{d}A_2\cos\theta_2}{d^2}, \\ \mathrm{d}U_2 &= n^2 \mathrm{d}A_2\cos\theta_2 \mathrm{d}\Omega_2 = \frac{n^2 \mathrm{d}A_2\cos\theta_2 \mathrm{d}A_1\cos\theta_1`}{d^2} \end{aligned} \tag{2.1.13}$$

for $\mathrm{d}A_1$ and $\mathrm{d}A_2$, respectively. From the previous equations we can conclude that $\mathrm{d}U_1 = \mathrm{d}U_2$ and therefore the étendue $\mathrm{d}U$ is conserved along a beam of light. Since also the flux through the areas $\mathrm{d}A_1$ and $\mathrm{d}A_2$ is conserved, the following relation holds:

$$L := n^2\frac{\mathrm{d}\Phi}{\mathrm{d}U} = constant\,. \tag{2.1.14}$$

In the optical systems we will consider in this work, the source and the target are located in the same medium (air) with $n = 1$, so the luminance L equals the basic luminance $L^* = L/n^2$ at the source and the target of the system.

In this thesis we consider two-dimensional optical systems. Hence, the definitions of the photometric parameters have to be given in two dimensions. An infinitesimal line segment of length da that emits a light beam and the ray that makes an angle θ with the normal $\boldsymbol{\nu}$ are considered, see Fig. 2.4.

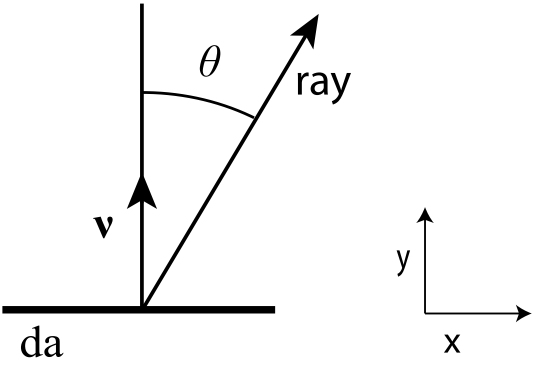

Figure 2.4: Ray emitted by an infinitesimal line segment da that makes an angle θ with respect to the line normal $\boldsymbol{\nu}$.

The two-dimensional illuminance $\big($unit $[\mathrm{lm/m}]\big)$ denotes the luminous flux falling on an infinitesimal line segment of length da and it is given by:

$$E = \frac{\mathrm{d}\Phi}{\mathrm{d}a}\ . \tag{2.1.15}$$

The luminous intensity $\big($unit $[\mathrm{lm/rad}]\big)$ is the luminous flux per angle dθ:

$$I = \frac{\mathrm{d}\Phi}{\mathrm{d}\theta}\ . \tag{2.1.16}$$

The two-dimensional luminance $\big($unit $[\mathrm{lm/(rad \cdot m)}]\big)$ is given by:

$$L = \frac{\mathrm{d}\Phi}{\cos\theta\,\mathrm{d}a\,\mathrm{d}\theta}. \tag{2.1.17}$$

Thus the following relation holds:

$$I = L(x,\theta)\cos\theta\,\mathrm{d}a \tag{2.1.18}$$

where x is a certain position at the light source da. Finally, the étendue dU (unit $[\mathrm{m \cdot rad}]$) in two dimensions is given by:

$$\mathrm{d}U = n\cos\theta \mathrm{d}a\,\mathrm{d}\theta. \tag{2.1.19}$$

In order to determine the light distribution on a surface and to compute the photometric variables on that surface, we need to understand how the light emitted from the source propagates. In the field of geometric optics the light propagation is described by light rays. The propagation of a light ray traveling through different media is determined by the reflection and refraction law. In the following we introduce these two laws and we explain the total internal reflection phenomenon.

2.2 Reflection and refraction law

A light ray is described by a position vector $\boldsymbol{x}$ on a surface and a direction vector $\boldsymbol{t}$ and can be parameterized by the arc length s. Light rays travel in a homogeneous medium along straight lines, once they hit a reflective surface their direction changes. Denoting with $\boldsymbol{t}_{\mathrm{i}}$ the direction of the incident ray and with $\boldsymbol{\nu}$ the unit normal to the surface at the location of incidence, the direction $\boldsymbol{t}_{\mathrm{r}}$ of the reflected ray is given by:

$$\boldsymbol{t}_{\mathrm{r}} = \boldsymbol{t}_{\mathrm{i}} - 2(\boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{\nu})\boldsymbol{\nu}\,, \tag{2.2.1}$$

where the vectors $\boldsymbol{t}_{\mathrm{i}}$ and $\boldsymbol{\nu}$ are unit vectors and $\boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{\nu}$ indicates the scalar product between $\boldsymbol{t}_{\mathrm{i}}$ and $\boldsymbol{\nu}$. From Eq. (2.2.1) it follows that the vector $\boldsymbol{t}_{\mathrm{r}}$ is a unit vector too, indeed considering the scalar product $(\boldsymbol{t}_{\mathrm{r}}, \boldsymbol{t}_{\mathrm{r}})$ we conclude:

$$\boldsymbol{t}_{\mathrm{r}} \cdot \boldsymbol{t}_{\mathrm{r}} = \boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{t}_{\mathrm{i}} - 4(\boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{\nu})(\boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{\nu}) + 4(\boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{\nu})^2(\boldsymbol{\nu} \cdot \boldsymbol{\nu}) = 1. \tag{2.2.2}$$

The vectors $\boldsymbol{t}_{\mathrm{i}}$, $\boldsymbol{t}_{\mathrm{r}}$ and $\boldsymbol{\nu}$ live all in the same plane. Defining the incident angle θ_{i} and the reflective angle θ_{r} such that θ_{i}, $\theta_{\mathrm{r}} \in [0, \pi/2)$. the reflection law states that $\theta_{\mathrm{i}} = \theta_{\mathrm{r}}$, see Fig. 2.5.

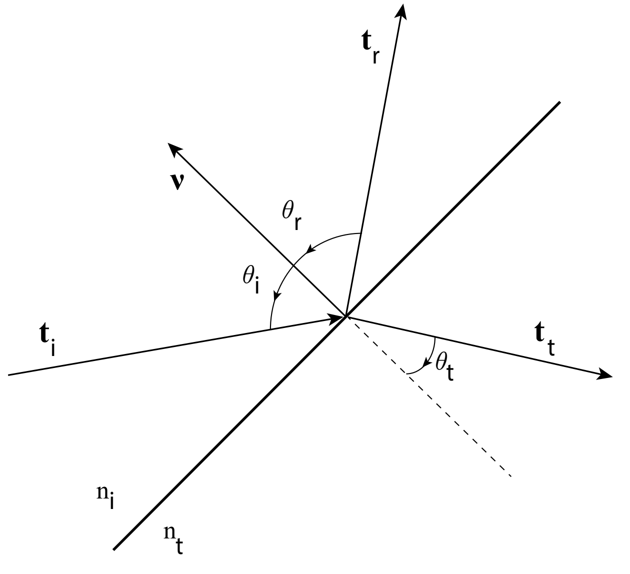

Figure 2.5: Propagation of a ray through two different media with index of refraction n_{i} and n_{t}.

When a ray propagates through two different media, its direction changes according to the law of refraction. Indicating with n_{i} the index of refraction of the medium in which the incident ray travels and with n_{t} the index of refraction of the medium of the transmitted ray, the direction $\boldsymbol{t}_{\mathrm{t}}$ of the transmitted ray is given by:

$$\boldsymbol{t}_{\mathrm{t}} = n_{\mathrm{i,t}}\, \boldsymbol{t}_{\mathrm{i}} + \left[\sqrt{1 - n^2_{\mathrm{i\cdot t}} + n^2_{\mathrm{i,t}}(\boldsymbol{\nu} \cdot \boldsymbol{t}_{\mathrm{i}})^2} - n_{\mathrm{i,t}}(\boldsymbol{\nu} \cdot \boldsymbol{t}_{\mathrm{i}})\right]\boldsymbol{\nu}\,, \tag{2.2.3}$$

where $n_{\mathrm{i,t}} = n_{\mathrm{i}}/n_{\mathrm{t}}$, [2]. Note that in Eq. (2.2.1) the direction of the normal $\boldsymbol{\nu}$ to the surface is not relevant for the computation of the direction of the reflective ray, since:

$$\boldsymbol{t}_{\mathrm{r}} = \boldsymbol{t}_{\mathrm{i}} - 2(\boldsymbol{t}_{\mathrm{i}} \cdot \boldsymbol{\nu})\boldsymbol{\nu} = \boldsymbol{t}_{\mathrm{i}} - 2(\boldsymbol{t}_{\mathrm{i}} \cdot -\boldsymbol{\nu})(-\boldsymbol{\nu}), \tag{2.2.4}$$

however, this is not the case for Eq. (2.2.3), therefore in the latter case we need to specify the direction of $\boldsymbol{\nu}$ which is usually chosen in such a way that the angle that it forms with the incident ray $\boldsymbol{t}_{\mathrm{i}}$ is smaller than or equal to $\pi/2$. Hence, if $(\boldsymbol{t}_{\mathrm{i}}, \boldsymbol{\nu}) \leq 0$ the normal $\boldsymbol{\nu}$ directed inside the same medium in which travels the incident ray is taken as in Fig. 2.5, otherwise the normal $-\boldsymbol{\nu}$ directed inside the same medium in which the transmitted ray will travel has to be considered.

Eq. (2.2.3) is only valid for

$$1 - n_{\mathrm{i,t}}^2 + n_{\mathrm{i,t}}^2(\boldsymbol{\nu} \cdot \boldsymbol{t}_{\mathrm{i}})^2 \geq 0 \tag{2.2.5}$$

which implies that

$$\frac{n_{\mathrm{t}}}{n_{\mathrm{i}}} \geq \sqrt{1 - (\boldsymbol{\nu} \cdot \boldsymbol{t}_{\mathrm{i}})^2} \tag{2.2.6}$$

from which we obtain:

$$n_{\mathrm{t}} \geq n_{\mathrm{i}} \sin\theta_{\mathrm{i}} \,. \tag{2.2.7}$$

The angle θ_{c} for which the equality holds is

$$\theta_{\mathrm{c}} = \arcsin\left(\frac{n_{\mathrm{t}}}{n_{\mathrm{i}}}\right) \tag{2.2.8}$$

and it is called the critical angle, [2]. When the incident angle θ_{i} is exactly equal to the critical angle θ_{c}, the square root in Eq. (2.2.3) is zero and the inner product $(\boldsymbol{t}_{\mathrm{t}}, \boldsymbol{\nu}) = 0$, hence the transmitted ray propagates parallel to the refractive surface. When $\theta_{\mathrm{i}} > \theta_{\mathrm{c}}$ the light ray is no longer refracted but is only reflected by the surface. This phenomenon is called total internal reflection (TIR). When TIR occurs, 100% of light is reflected and there is no loss of energy. Therefore, optical systems designed such that rays are reflected by TIR are very efficient. Lght that hits an ordinary refractive surface can be reflected and refracted. The energy that is reflected and refracted is determined by the Fresnel's coefficients. In the next paragraph an overview of the Fresnel coefficients is given.

2.3 Fresnel's equations

In order to derive Fresnel's equations we need to describe light as an electromagnetic wave. It is therefore useful to study the light propagation from the perspective of electromagnetic theory which gives information about the incident, reflected and transmitted radiant flux density that are denoted with E_i, E_r and E_t, respectively. Any component of the electric field $\boldsymbol{\mathcal{E}}$ can be written as

$$\boldsymbol{\mathcal{E}}(\boldsymbol{x}, T) = \boldsymbol{\mathcal{E}}_0(\boldsymbol{x})e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega T)} \tag{2.3.1}$$

where $\boldsymbol{x}$ is the position vector and T is the time. The amplitude $\boldsymbol{\mathcal{E}}_0(\boldsymbol{x})$ is constant in time and $\omega = \frac{ck}{n}$ is the value of the angular frequency with c the velocity of light and n the index of refraction in which the wave is traveling, which is the ratio of the speed of light c in vacuum and the speed of light v in the material. Note that the angular frequency can be also written as $\omega = vk$, in particular when a wave travels in vacuum $n = 1$ and $\omega = ck$. The vector $\boldsymbol{k}$ has the same direction of the wave and its absolute value $|\boldsymbol{k}| = k = \frac{2\pi}{\lambda}$ is the wave number in vacuum, with λ the wavelength. Similarly, the magnetic field has the form:

$$\boldsymbol{\mathcal{B}}(\boldsymbol{x}, T) = \boldsymbol{\mathcal{B}}_0(\boldsymbol{x})e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega T)}\,. \tag{2.3.2}$$

Light can be seen as an electromagnetic wave, that is an oscillating electric field $\boldsymbol{\mathcal{E}}$ and an oscillating magnetic filed $\boldsymbol{\mathcal{B}}$ which propagates always perpendicular to $\boldsymbol{\mathcal{E}}$. Th eelectric field oscillates perpendicular to the wave propagation. Light is said to be polarized if the direction of the electic field is well defined. When the electric field propagates in different directions we talk about unpolarized light. By convention, we refer to the light's polarization as the direction of the electric field $\boldsymbol{\mathcal{E}}$, [6] with respect to the incident plane that is defined by the incident and reflected rays as is shown in Fig. 2.6.

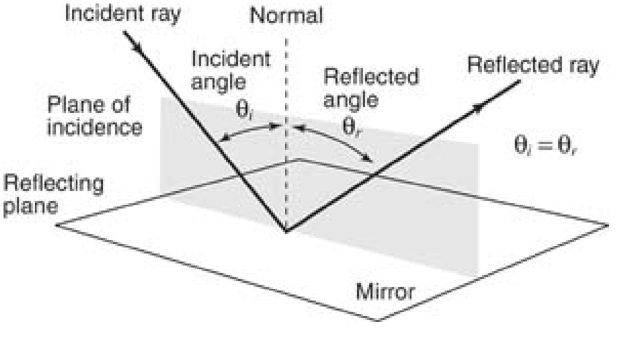

Figure 2.6: Light ray that hits a mirror located on the reflecting plane. The incident and the reflected ray leave in the same plane of the normal to the mirror that is called plane of incident.

In order to derive the Fresnel's coefficients the polarization of light must to be taken into account. Those coefficients are obtained considering Maxwell's equations and the boundary conditions due to the conservation of energy. The details of Fresnel's equations are widely explained in the literature. In the following we provide Fresnel coefficients and we briefly explain their physical interpretation. We refer the reader to [7, 8] for more details. Fresnel's coefficients can also be derived using a different approach that does not involve Maxwell's equations, this method is explained in [9]. The following particular cases of light's polarization need are considered.

1. $\boldsymbol{\mathcal{E}}$ is perpendicular to the plane of incidence (see Fig. 2.7). In this case light is said to be s-polarized.

2. $\boldsymbol{\mathcal{E}}$ is parallel to the plane of incidence (see Fig. 2.8). In this case light is said to be p-polarized.

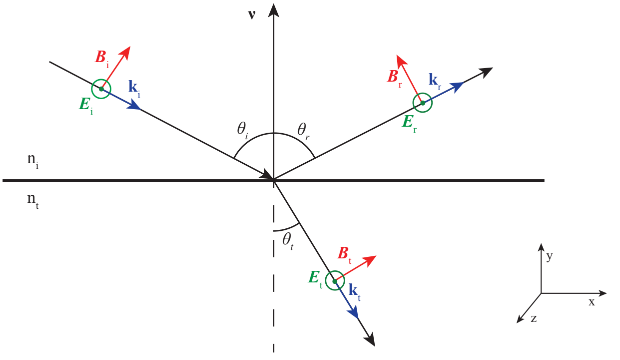

Figure 2.7: Propagation of an electromagnetic wave where $\boldsymbol{\mathcal{E}}$ is perpendicular to the incident plane. The components of $\boldsymbol{\mathcal{E}}$ are is indicated with the green circles. The components of $\boldsymbol{\mathcal{B}}$ are indicated with red arrows.

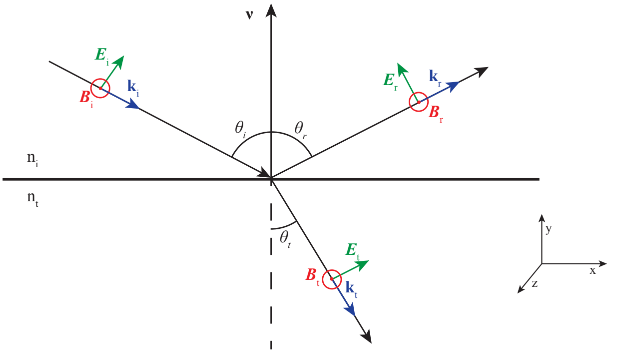

Figure 2.8: Propagation of an electromagnetic wave where $\boldsymbol{\mathcal{E}}$ is parallel to the incident plane. The components of $\boldsymbol{\mathcal{B}}$ are indicated with the red circle. The components of $\boldsymbol{\mathcal{E}}$ are indicated with greed arrows.

Energy conservation gives the boundary conditions of the electromagnetic field at the plane of the interface (which is perpendicular to the incident plane). In the following we derive Fresnel's coefficients for case 1. Similarly, the Fresnel's coefficients can be derived for the second case.

For s-polarized light the tangential components of $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{B}}/\mu$ across the boundary between the two different media must be continuous. The continuity of the tangential component of $\boldsymbol{\mathcal{E}}$ leads to:

$$|\boldsymbol{\mathcal{E}}_{0\mathrm{i}}| + |\boldsymbol{\mathcal{E}}_{0\mathrm{r}}| = |\boldsymbol{\mathcal{E}}_{0\mathrm{t}}|, \tag{2.3.3}$$

while the continuity of the tangential component of $\boldsymbol{\mathcal{B}}/\mu$ gives:

$$-\frac{|\boldsymbol{\mathcal{B}}_{0,\mathrm{i}}|}{\mu_\mathrm{i}}\cos\theta_\mathrm{i} + \frac{|\boldsymbol{\mathcal{B}}_{0,\mathrm{r}}|}{\mu_\mathrm{r}}\cos\theta_\mathrm{r} = -\frac{|\boldsymbol{\mathcal{B}}_{0,\mathrm{t}}|}{\mu_\mathrm{t}}\cos\theta_\mathrm{t}, \tag{2.3.4}$$

where the negative sign in front of $|\boldsymbol{\mathcal{B}}_{0,\mathrm{i}}|$ and $|\boldsymbol{\mathcal{B}}_{0,\mathrm{t}}|$ is due to the convention that a positive direction is considered with increasing x. Since $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{E}}/v$, Eq. (2.3.4) can be written as

$$\frac{1}{\mu_\mathrm{i} v_\mathrm{i}}(|\boldsymbol{\mathcal{E}}_{0,\mathrm{i}}| - |\boldsymbol{\mathcal{E}}_{0,\mathrm{r}}|)\cos\theta_\mathrm{i} = \frac{1}{\mu_\mathrm{t} v_\mathrm{t}}|\boldsymbol{\mathcal{E}}_{0,\mathrm{t}}|\cos\theta_\mathrm{t}, \tag{2.3.5}$$

where we employed the fact that $v_\mathrm{i} = v_\mathrm{r}$, and $\theta_\mathrm{i} = \theta_\mathrm{r}$. Using Eq. (2.3.1) and $n = c/v$, the previous equation becomes:

$$\frac{n_\mathrm{i}}{\mu_\mathrm{i}}(|\boldsymbol{\mathcal{E}}_{0\mathrm{i}}| - |\boldsymbol{\mathcal{E}}_{0\mathrm{r}}|)\cos\theta_\mathrm{i} = \frac{n_\mathrm{t}}{\mu_\mathrm{i}}|\boldsymbol{\mathcal{E}}_{0\mathrm{t}}|\cos\theta_\mathrm{t} \tag{2.3.6}$$

Finally, assuming that $\mu_\mathrm{i} = \mu_\mathrm{t} = \mu_0$ and employing Eq. (2.3.3) we obtain:

$$\begin{aligned} r_s &= \frac{|\boldsymbol{\mathcal{E}}_{0\mathrm{r}}|_s}{|\boldsymbol{\mathcal{E}}_{0\mathrm{i}}|_s} = \frac{n_\mathrm{i}\cos\theta_\mathrm{i} - n_\mathrm{t}\cos\theta_\mathrm{t}}{n_\mathrm{i}\cos\theta_\mathrm{i} + n_\mathrm{t}\cos\theta_\mathrm{t}}, \\ t_s &= \frac{|\boldsymbol{\mathcal{E}}_{0\mathrm{t}}|_s}{|\boldsymbol{\mathcal{E}}_{0\mathrm{i}}|_s} = \frac{2n_\mathrm{i}\cos\theta_\mathrm{i}}{n_\mathrm{i}\cos\theta_\mathrm{i} + n_\mathrm{t}\cos\theta_\mathrm{t}}. \end{aligned} \tag{2.3.7}$$

The coefficients r_s and t_s are amplitude coefficients for the reflected and transmitted light. They are the perpendicular components of r and t for s-polarized light. Using Snell's law, that is $n_\mathrm{i}\sin\theta_\mathrm{i} = n_\mathrm{t}\sin\theta_\mathrm{t}$, the relations for r_s and t_s are simplified as follows:

$$\begin{aligned} r_s &= -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)}, \\ t_s &= -\frac{2\sin\theta_t\cos\theta_i}{\sin(\theta_i + \theta_t)}\,. \end{aligned} \tag{2.3.8}$$

A similar argument for the p-polarized light leads to the calculation of the parallel components r_p and t_p of r and t. In case $\boldsymbol{\mathcal{E}}$ is parallel to the plane of incidence the amplitude coefficients are:

$$\begin{aligned} r_p &= \frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_i\cos\theta_t + n_t\cos\theta_i}, \\ t_p &= \frac{2n_i\cos\theta_i}{n_i\cos\theta_t + n_t\cos\theta_i}\,, \end{aligned} \tag{2.3.9}$$

and their simplified relations are:

$$\begin{aligned} r_p &= \frac{\tan(\theta_i - \theta_t)}{\theta_i + \theta_t}, \\ t_p &= \frac{2\sin\theta_t\cos\theta_i}{\sin(\theta_i + \theta_t)\cos(\theta_i - \theta_t)}. \end{aligned} \tag{2.3.10}$$

Furthermore, it can be checked that

$$\begin{aligned} t_s - r_s &= 1, \\ t_p + r_p &= 1. \end{aligned} \tag{2.3.11}$$

The amplitude coefficients are shown in Fig. 2.9 for the case in which light travels from a less dense to a more dense medium ($n_i < n_t$), that is external reflection. In

Fig. 2.10 the reflection coefficients are shown for the case in which $n_i > n_t$, that is internal reflection. Note from Fig. 2.9 that r_p approaches to 0 when θ_i approaches to θ_p and it gradually decreases reaching -1 for an incident angle $\theta_i = 90°$. The angle θ_p is called Brewster's angle or polarization angle as only the component perpendicular to the incident plane is reflected at that angle and therefore light is perfectly polarized. Similarly, Fig. 2.10 shows that $r_p = 0$ for $\theta_i = \theta_{p\prime}$. It can be show that $\theta_p + \theta_{p\prime} = 90°$. Both r_p and r_s reach 1 when $\theta_i = \theta_c$. θ_c is called the critical angle. Light that hits the incident plane with an incident angle equal to or greater than the critical angle is totally reflected back and no transmitted light is observed. This phenomenon is called total internal reflection.

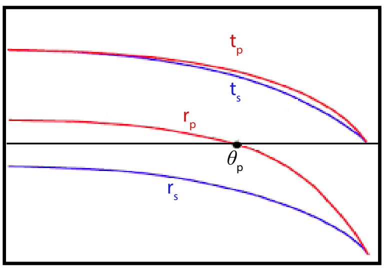

Figure 2.9: Amplitude coefficients of reflection and transmission as a function of the incident angle θ_i in the case of external reflection, i.e. $n_t < n_i$ ($n_t = 1$ and $n_i = 1.5$). θ_p is the polarization angle, [8].

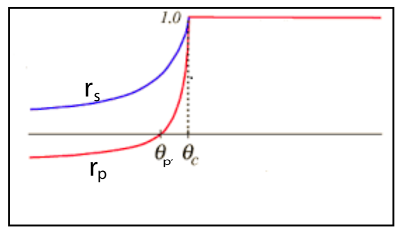

Figure 2.10: Reflection coefficients as a function of the incident angle θ_i in the case of internal reflection, i.e. $n_t > n_i$ ($n_t = 1.5$ and $n_i = 1$). $\theta_{p\prime}$ is the polarization angle and θ_c is the critical angle, [8].

The we introduce the Poynting vector $\boldsymbol{P}$ that defines the energy flux of an electromagnetic field. It is measured in [W/m^2], and it is given by:

$$\boldsymbol{P} = \frac{1}{\mu}\Big(\boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{B}}\Big), \tag{2.3.12}$$

where $\mu = \frac{1}{\varepsilon v^2}$ is the permeability and ε the permittivity of the medium. In the following, the parameters for vacuum are indicated with the subscript 0. All quantities defined in the media of the incident, reflective and transmitted light are indicated with the subscripts i, r and t, respectively. Optical rays are perpendicular to the wave front of an electromagnetic wave and parallel to the Poynting vector, [10]. The irradiance E is defined as the average energy that crosses in unit time a unit area A perpedicular to the direction of the energy flow. Therefore, defining the average of the vector $\boldsymbol{P}$ over the time as:

$$\langle \boldsymbol{P} \rangle_T = \frac{1}{T}\int_0^T \boldsymbol{P}\mathrm{d}T \tag{2.3.13}$$

we can write the irradiance E as:

$$\boldsymbol{E} = \langle \boldsymbol{P} \rangle_T = v\varepsilon|\boldsymbol{\mathcal{E}}|^2 \,. \tag{2.3.14}$$

Considering a beam of light that hits a surface such that an area A is illuminated, the incident, reflected and transmitted beams are $\boldsymbol{E}_{\mathrm{i}} A\cos\theta_{\mathrm{i}}$ $\boldsymbol{E}_{\mathrm{r}} A\cos\theta_{\mathrm{r}}$ and $\boldsymbol{E}_{\mathrm{t}} A\cos\theta_{\mathrm{t}}$, respectively. The reflectance $\mathcal{R}$ is the ratio of the reflected power to the incident power:

$$\mathcal{R} = \frac{|\boldsymbol{E}_{\mathrm{r}}|\cos\theta_r}{|\boldsymbol{E}_{\mathrm{i}}|\cos\theta_{\mathrm{i}}} = \frac{|\boldsymbol{\mathcal{E}}_{0\mathrm{r}}|^2}{|\boldsymbol{\mathcal{E}}_{0\mathrm{i}}|^2} = r^2 \tag{2.3.15}$$

where the second equality holds because $v_{\mathrm{i}} = v_{\mathrm{t}}$, $\varepsilon_{\mathrm{i}} = \varepsilon_{\mathrm{t}}$ and $\theta_{\mathrm{i}} = \theta_{\mathrm{t}}$. Similarly, the transmittance $\mathcal{T}$ is the ratio between the transmitted to the incident power:

$$\mathcal{T} = \frac{|\boldsymbol{E}_{\mathrm{t}}|\cos\theta_{\mathrm{t}}}{|\boldsymbol{E}_{\mathrm{r}}|\cos\theta_{\mathrm{r}}} = \frac{n_{\mathrm{t}}\cos\theta_{\mathrm{t}}}{n_{\mathrm{t}}\cos\theta_{\mathrm{i}}}\frac{|\boldsymbol{\mathcal{E}}_{0\mathrm{t}}|^2}{|\boldsymbol{\mathcal{E}}_{0\mathrm{i}}|^2} = \frac{n_{\mathrm{t}}\cos\theta_{\mathrm{t}}}{n_{\mathrm{t}}\cos\theta_{\mathrm{i}}}t^2\,. \tag{2.3.16}$$

Employing total energy conservation, that is:

$$\boldsymbol{E}_{\mathrm{i}}A\cos\theta_{\mathrm{i}} = \boldsymbol{E}_{\mathrm{r}}A\cos\theta_{\mathrm{r}} + \boldsymbol{E}_{\mathrm{t}}A\cos\theta_{\mathrm{t}}, \tag{2.3.17}$$

we can easily prove that:

$$\mathcal{R} + \mathcal{T} = 1\,. \tag{2.3.18}$$

The parallel and perpendicular components of $\mathcal{R}$ and $\mathcal{T}$ are:

$$\begin{aligned}
\mathcal{R}_p &= r_p^2,\\
\mathcal{T}_p &= \frac{n_t\cos\theta_t}{n_t\cos\theta_i}t_p^2,\\
\mathcal{R}_s &= r_s^2,\\
\mathcal{T}_s &= \frac{n_t\cos\theta_t}{n_t\cos\theta_i}t_s^2.
\end{aligned} \tag{2.3.19}$$

it can be show that

$$\begin{aligned}
\mathcal{R}_s + \mathcal{R}_p &= 1,\\
\mathcal{T}_s + \mathcal{T}_p &= 1\,.
\end{aligned} \tag{2.3.20}$$

For normal incidence, i.e. $\theta_i = 0$, there is no polarization and Eqs. (2.3.19) lead to:

$$\begin{aligned}
\mathcal{R} = \mathcal{R}_p = \mathcal{R}_s &= \left(\frac{n_i - n_t}{n_t + n_i}\right)^2,\\
\mathcal{T} = \mathcal{T}_p = \mathcal{T}_s &= \frac{4n_in_t}{(n_t+n_i)^2}\,.
\end{aligned} \tag{2.3.21}$$

Many common light sources such as sunlight, halogen lighting, LED spotlights, and incandescent bulbs produce unpolarized light. In case of unpolarized light the amount of reflected and transmitted light is given by the average of reflectance $\mathcal{R}$ and transmittance $\mathcal{T}$ calculated considering first p-polarized light and then s-polarization, that is:

$$\begin{aligned}
\mathcal{R} &= \frac{\mathcal{R}_p + \mathcal{R}_s}{2},\\
\mathcal{T} &= \frac{\mathcal{T}_p + \mathcal{T}_s}{2},
\end{aligned} \tag{2.3.22}$$

where $\mathcal{R}_p$, $\mathcal{R}_s$, $\mathcal{T}_p$ and $\mathcal{T}_s$ are given in Eqs. (2.3.19).

With this overview we conclude this chapter. The notions given in Section 2.1 will be used in the entire thesis as our goal is to study the distribution of light at the target of some optical systems. In particular we will focus on the computation of the output intensity distribution. The reflection and refraction laws explained in Section 2.2 are needed to determine how the optical system changes the ray's direction every time that it hits a surfaces (or a line in the two-dimensional case). In Chapters 3, 4, 5, 6 and 7 only systems where the reflection and refraction laws play a role are considered. Systems with Fresnel reflection are treated in the last chapter. The amount of reflected and transmitted light is calculated using the Fresnel's equation (introduced in the last paragraph of this chapter). Since, we restrict ourselves to two-dimensional systems, the value of reflectance and transmittance will be computed using Eqs. (2.3.22).

Chapter 3

Ray tracing

Ray tracing is a geometric problem that describes the transport of light within optical systems. It uses single rays to describe the propagation of light through an optical system. The influence of diffraction on the transport of a ray is neglected and geometrical modeling of an optical system is considered. Generally, the method can be implemented for two or more dimensions and for any optical system. In the following a description of the ray tracing method is given. We restrict ourselves to two-dimensional systems and to the case where only reflection and refraction are taken into account.

3.1 Ray tracing for two-dimensional optical systems

Ray tracing is a procedure that can be applied to non-imaging systems. It consists of tracing each ray, which is considered to be a broken line, through the optical system. Given a Cartesian coordinate system (x, z), a two-dimensional optical system symmetric with respect to the z-axis is defined. Hence, usually the optical axis coincides with the var z-axis. The optical system is formed by a source S, a target T and some optical components i where $i \in \{2, \cdots, \mathrm{Nl}-1\}$ and Nl indicates the number of lines that form the system. S and T are indicated with the indexes 1 and Nl, respectively. The index of refraction of the medium in which line i is located is indicated with n_i Every ray emitted by S (line 1) can hit some optical components $i \in \{2, \cdots, \mathrm{Nl}-1\}$ before reaching T (line Nl). The intersection point of the rays with line i are $(x_i, z_i)_{i=1,\cdots,\mathrm{Nl}}$ and, $\boldsymbol{s}_i = (-\sin\phi_i, \cos\phi_i)$ indicates the direction vector of the rays that leave i, with ϕ_i the angle that the ray forms with respect to the optical axis measured counterclockwise. As we consider only forward rays, the angles $\phi_i \in (-\pi/2, \pi/2)$. Therefore, a ray segment between (x_i, z_i) and (x_j, z_j) with $j \neq i$ is parameterized in real space by:

$$\boldsymbol{r}(s) = \begin{pmatrix} x_i - s\sin\phi_i \\ z_i + s\cos\phi_i \end{pmatrix} \qquad 0 < s \leq s_{\max}\,, \tag{3.1.1}$$

where s denotes the arc-length and $s_{\max}$ is the maximum value that it can assume. Fig. 3.1 shows an example where a single ray is traced inside a very simple optical system, the so-called two-faceted cup. The light source $\mathsf{S}= [-a, a]$ (line 1) and the target $\mathsf{T} = [-b, b]$ (line 4) are two segments normal to the z-axis, where $a = 2$ and

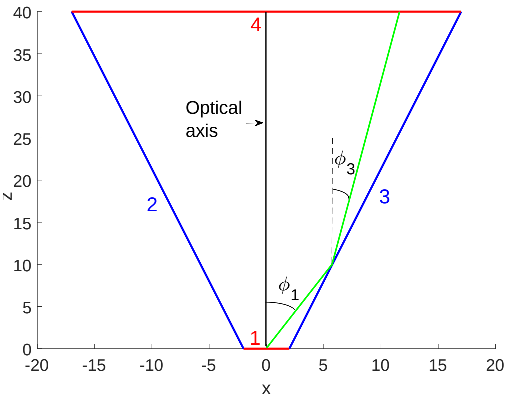

Figure 3.1: Shape of the two-faceted cup. Each line of the system is labeled with a number. The source S$= [-2, 2]$ (line number 1) is located on the x-axis. The target T$= [-17, 17]$ (line 4) is parallel to the source and is located at a height $z = 40$. The left and right reflectors (line 2 and 3) connect the source with the target.

$b = 17$. The left and right reflectors (line 2 and 3) are oblique segments that connect the source and the target. All the optical lines $i \in \{1, \cdots, 4\}$ are located in air, therefore the refractive index $n_i = 1$ for every i.

In order to compute the target photometric variables, we need to know how the optical system influences the direction of the rays when they hit an optical line. To this purpose, the ray tracing procedure is often used in optics. Ray tracing relates the position coordinates (x_1, z_1) and the direction vector $\boldsymbol{s}_1$ of every ray at the source S with the corresponding position $(x_{\mathrm{Nl}}, z_{\mathrm{Nl}})$ and direction $\boldsymbol{s}_{\mathrm{Nl}}$ at the target T. As in the following we will use often the target coordinates of the rays, from now on, to simplify the notation, we write (x, z) instead of $(x_{\mathrm{Nl}}, z_{\mathrm{Nl}})$, ϕ instead of ϕ_{Nl} and $\boldsymbol{s}$ instead of $\boldsymbol{s}_{\mathrm{Nl}}$ for the target coordinates. The ray tracing algorithm can be schematized as follows.

1. Given a ray that leaves S with initial position (x_1, z_1) and initial direction $\boldsymbol{s}_1 = (-\sin\phi_1, \cos\phi_1)$, use Eq. (3.1.1) to implement the ray parametrization $\boldsymbol{r}(s_1)$;

2. Compute the coordinates $(x_k, z_k)_{k=1\cdots,\mathrm{Nl}}$ of the intersection point of the ray with all the lines that it hits;

   a) If the shape of the lines is described by equations with a known analytical description, the intersection points are determined analytically;

   b) If there is no analytic description for the optical lines, the intersections need to be determined using iterative methods;

1. Computing the minimum of the distances between the previous ray's position coordinates (x_1, z_1) and the coordinates (x_k, z_k) for every $k = 1 \cdots, \mathrm{Nl}$ and determine the closest line i that the forward ray encounters;

2. If $i = \mathrm{Nl}$ stop the procedure, the target ray's coordinates (x, z) and $\boldsymbol{s}$ are found.

3. Calculate the normal $\boldsymbol{\nu}_i$ to line i at the point (x_i, z_i);

4. Compute the new ray direction $\boldsymbol{s}_i$ of the ray that leaves line i at the point (x_i, z_i);
   a) If the incident line is a reflective line, $\boldsymbol{s}_i$ is given by Eq. (2.2.1);
   b) If the incident line is a refractive line, $\boldsymbol{s}_i$ is given by Eq. (2.2.3);
5. Restart the procedure from 1. for the ray that leaves line i instead of S. Consider as initial ray position coordinates (x_i, z_i) instead of (x_1, z_1) and as initial ray direction $\boldsymbol{s}_i = (-\sin\phi_i, \cos\phi_i)$ instead of $\boldsymbol{s}_1$.

The procedure explained above is repeated for every ray traced through the system, [11]. There are different ways to implement the ray tracing procedure. An often used method is MC ray tracing which is explained in the next paragraph.

3.2 Monte Carlo ray tracing

Given a function $f : D \subseteq \mathbb{R}^d \mapsto \mathbb{R}$ and a random variable $y \in D$ with probability density function $\rho(y)$, the expected value of f is defines as:

$$E[f(y)] = \int_D f(y)\rho(y)\mathrm{d}y. \tag{3.2.1}$$

In Monte Carlo methods N independent random variables $y_0, \cdots, y_\mathrm{N}$ from the density ρ are considered and, the sum

$$f_\mathrm{N}(y) = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}} f(y_i) \tag{3.2.2}$$

is calculated, [12]. The strong law of large numbers tells us that:

$$\Pr\Big(\lim_{\mathrm{N}\to\infty}\sum_{i=1}^{\mathrm{N}} f(y_i) = E[f(y)]\Big) = 1. \tag{3.2.3}$$

Therefore the following relation holds:

$$E[f(y)] \approx \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}} f(y_i). \tag{3.2.4}$$

MC technique can be combined with the ray tracing procedure in order to compute the light distribution at the target of an optical system. This is called Monte Carlo ray tracing and it is explained in the following. MC ray tracing uses as sample of random variables a set of points the coordinates of which correspond to the position and direction coordinates of every ray at the source of the optical system. Therefore, rays with a random initial position and a random initial direction are traced from S to T within the optical system. In this way a probabilistic interpretation of the output photometric variables is provided. In particular, we are interested in the total target intensity I which is computed as a function of the angular coordinate ϕ. The MC intensity is calculated dividing the target into intervals of the same length, the so-called bins. A partitioning $P_1 : -\pi/2 = \phi_0 < \phi_1 < \cdots < \phi_{\mathrm{Nb}} = \pi/2$ of the interval $[-\pi/2, \pi/2]$ is defined where Nb is the number of bins in P_1. We remark that, with a slight abuse of notation, we indicated the angular coordinates of the rays at the target (line Nl) with ϕ_j instead of $\phi_{\mathrm{Nl},j}$ for every $j \in \{0, \cdots, \mathrm{Nb}\}$. The normalized approximated intensity $I_{\mathrm{MC}}(\phi)$ is a piecewise constant function and its value over the j-th bin is the ratio between the number of rays that fall into that bin $\mathrm{Nr}[\phi_{j-1}, \phi_j)$ and the total number of rays traced $\mathrm{Nr}[-\pi/2, \pi/2]$. Hence, I_{MC} is defined by:

$$I_{\mathrm{MC}}(\phi) = \frac{\mathrm{Nr}[\phi_{j-1}, \phi_j)}{\mathrm{Nr}[-\pi/2, \pi/2]} \qquad \text{for } \phi \in [\phi_{j-1}, \phi_j). \tag{3.2.5}$$

Furthermore, the output intensity is computed from the value of the intensity $I_{\mathrm{MC}}(\phi_{j-1/2})$ along the direction $\phi_{j-1/2} = (\phi_{j-1} + \phi_j)/2$ for every bin $[\phi_{j-1}, \phi_j)_{j=1,\cdots,\mathrm{Nb}}$. The intensity $I_{\mathrm{MC}}(\phi_{j-1/2})$ gives an estimate of the probability that a ray reaches the target with an angle in the j-th interval $[\phi_{j-1}, \phi_j)$ of the partitioning P_1. This probability $\mathrm{P}_{j,\Delta\phi}$ is given by:

$$\mathrm{P}_{j,\Delta\phi} = \Pr(\phi_{j-1} \leq \phi < \phi_j) = \frac{\int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi}{\int_{-\pi/2}^{\pi/2} I(\phi)\mathrm{d}\phi}\,, \tag{3.2.6}$$

where $I(\phi)$ is the output intensity (not normalized) and it is measured in lumen per radian [lm/rad]. Note that $\sum_{j=1}^{\mathrm{Nb}} \mathrm{P}_{j,\Delta\phi} = 1$. Using the mean value theorem for the function $I(\phi)$ continuous in $[\phi_{j-1}, \phi_j]$, the integral at the numerator of the previous equation can be written as:

$$\int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi = \Delta\phi\, I(\phi_{j-1/2}). \tag{3.2.7}$$

Hence, $\mathrm{P}_{j,\Delta\phi}$ is proportional to the size $\Delta\phi = (\phi_{\mathrm{Nb}} - \phi_0)/\mathrm{Nb}$ of the intervals, i.e., inversely proportional to the number of bins Nb of the partitioning P_1. Indicating with $\Phi = \int_{-\pi/2}^{\pi/2} I(\phi)\mathrm{d}\phi$ the total flux (measured in lumen [lm]), the error between the intensity $I(\phi_{j-1/2})$ and the averaged MC intensity $\Phi I_{\mathrm{MC}}(\phi_{j-1/2})/\Delta\phi$ is given by:

$$\begin{aligned} \Big|I(\phi_{j-1/2}) - \frac{\Phi}{\Delta\phi} I_{\mathrm{MC}}(\phi_{j-1/2})\Big| &\leq \\ &\Big|I(\phi_{j-1/2}) - \frac{1}{\Delta\phi}\int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi\Big| + \\ &\frac{1}{\Delta\phi}\Big|\int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi - \Phi\, I_{\mathrm{MC}}(\phi_{j-1/2})\Big|. \end{aligned} \tag{3.2.8}$$

The first term of the right hand side of inequality (3.2.8) gives an estimate of how much the averaged intensity $\frac{1}{\Delta\phi}\int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi$ differs from the exact intensity $I(\phi_{j-1/2})$. This term is due to the discretization of the target and therefore it depends on the number of bins Nb considered. Substituting $I(\phi)$ with its Taylor expansion around the point $\phi_{j-1/2}$ we obtain that this term is proportional to the square of the size of the bins, therefore the following equality holds:

$$\Big|I(\phi_{j-1/2}) - \frac{1}{\Delta\phi}\int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi\Big| = \mathrm{C}_1/\mathrm{Nb}^2 \tag{3.2.9}$$

with $C_1 > 0$ a certain constant.

The second part of the right hand side of inequality (3.2.8) gives an estimate of the MC error and therefore it depends also on the number of rays traced. In order to show how this term decreases as a function of the number of rays traced, we define the random variable $X_j(\phi)$ as the variable that is equal to 1 if the ray with angular coordinate ϕ is inside the interval $[\phi_{j-1}, \phi_j)$ and equal to 0 otherwise,

$$X_j(\phi) = \begin{cases} 1 & \text{if} \quad \phi \in [\phi_{j-1}, \phi_j), \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.10}$$

The Bernoulli trial X_j follows a binomial distribution $B(1, \mathrm{P}_{j,\Delta\phi})$. Considering a sample of Nr rays, the variable $Y_j = \sum_{k=1}^{\mathrm{Nr}} X_j(\phi_k)$ follows a binomial distribution $B(\mathrm{Nr}, \mathrm{P}_{j,\Delta\phi})$, where ϕ_k is the angle that the k-th ray forms with the optical axis. Then, using the de Moivre-Laplace theorem, we conclude that the variable Y_j is approximated by a normal distribution with mean value $E[Y_j] = \mathrm{Nr}\mathrm{P}_{j,\Delta\phi}$ and variance $\sigma^2[Y_j] = \mathrm{Nr}\mathrm{P}_{j,\Delta\phi}(1 - \mathrm{P}_{j,\Delta\phi})$ when a large number of rays is considered, see [13, 14]. Thus, the normalized intensity along the direction $\phi_{j-1/2}$ is given by:

$$I_{\mathrm{MC}}(\phi_{j-1/2}) = \sum_{k=1}^{\mathrm{Nr}} X_j(\phi_k)/\mathrm{Nr}. \tag{3.2.11}$$

The mean value $E[I_{\mathrm{MC}}(\phi_{j-1/2})] = \mathrm{P}_{j,\Delta\phi}$ and the variance $\sigma^2[I_{\mathrm{MC}}(\phi_{j-1/2})] = \mathrm{P}_{j,\Delta\phi}(1 - \mathrm{P}_{j,\Delta\phi})/\mathrm{Nr}$. Note that the standard deviation $\sigma_j := \sigma[I_{\mathrm{MC}}(\phi_{j-1/2})]$ equals:

$$\sigma_j = \sqrt{\mathrm{P}_{j,\Delta\phi}(1 - \mathrm{P}_{j,\Delta\phi})/\mathrm{Nr}} = \frac{\mathrm{C}_2}{\sqrt{\mathrm{NbNr}}}\,, \tag{3.2.12}$$

for some $\mathrm{C}_2 > 0$. σ_j can be used to give an estimate of the difference between the intensity $I_{\mathrm{MC}}(\phi_{j-1/2})$ and its mean value $\mathrm{P}_{j,\Delta\phi}$. Therefore, the second term of the right hand side of relation (3.2.8) becomes:

$$\begin{aligned} \frac{1}{\Delta\phi}\Big| \int_{\phi_{j-1}}^{\phi_j} I(\phi)\mathrm{d}\phi - \Phi\, I_{\mathrm{MC}}(\phi_{j-1/2})\Big| &= \\ \frac{\Phi}{\Delta\phi}\Big|\mathrm{P}_{j,\Delta\phi} - I_{\mathrm{MC}}(\phi_{j-1/2})\Big| &\propto \\ \frac{\Phi}{\Delta\phi}\sigma_j[I_{\mathrm{MC}}(\phi_{j-1/2})] = \mathrm{C}_3\frac{\mathrm{Nb}}{\sqrt{\mathrm{NbNr}}} &= \mathrm{C}_3\sqrt{\frac{\mathrm{Nb}}{\mathrm{Nr}}}\,, \end{aligned} \tag{3.2.13}$$

for some $\mathrm{C}_3 > 0$, where the approximation holds because σ_j gives a measure for the error between $I_{\mathrm{MC}}(\phi_{j-1/2})$ and the probability $\mathrm{P}_{j,\Delta\phi}$, [15]. The second equality follows from Eq. (3.2.12). To conclude, the MC error over the j-th bin is estimated by:

$$\Big|I(\phi_{j-1/2}) - \frac{\Phi}{\Delta\phi}I_{\mathrm{MC}}(\phi_{j-1/2})\Big| = \frac{\mathrm{C}_1}{\mathrm{Nb}^2} + \mathrm{C}_4\sqrt{\frac{\mathrm{Nb}}{\mathrm{Nr}}}, \tag{3.2.14}$$

for $\mathrm{C}_4 > 0$. Considering a fixed number of rays, we obtain that the minimal error is reached when $\mathrm{Nb} \approx \mathrm{Nr}^{1/5}$. Hence, if 10^{10} rays are considered the target has to be divided into 10^2 bins to minimize the MC error. This leads to computational efforts resulting in a very slow procedure.

To give an example we implemented MC ray tracing for the two-faceted cup the profile of which is depicted in Fig. 3.1. A set of random rays with coordinates $(x_k, \phi_k)_{k=1,\cdots,\mathrm{Nr}}$ is generated at the source. We considered for instance $\mathrm{Nr} = 10^4$ rays. The rays distribution on the (x_k, ϕ_k)-plane is shown in Fig. 3.2.

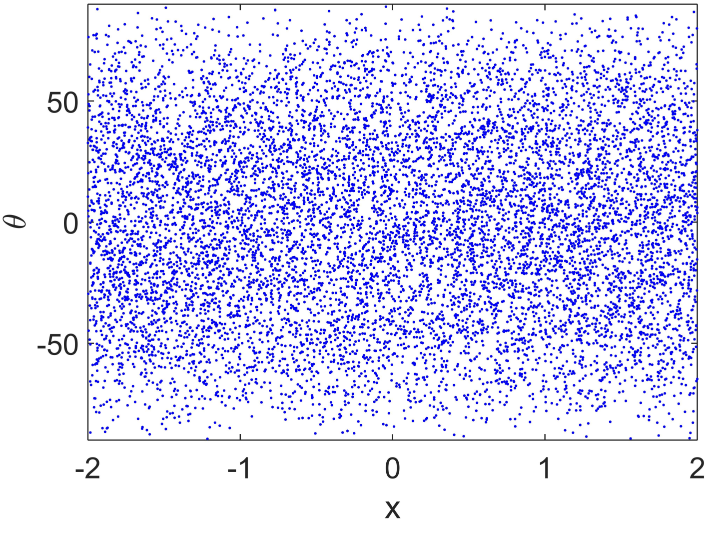

Figure 3.2: Rays at the source of the two-faceted cup with random position coordinate x and random angular coordinates ϕ. 10^4 rays are depicted in this figure.

3.3 Quasi-Monte Carlo method

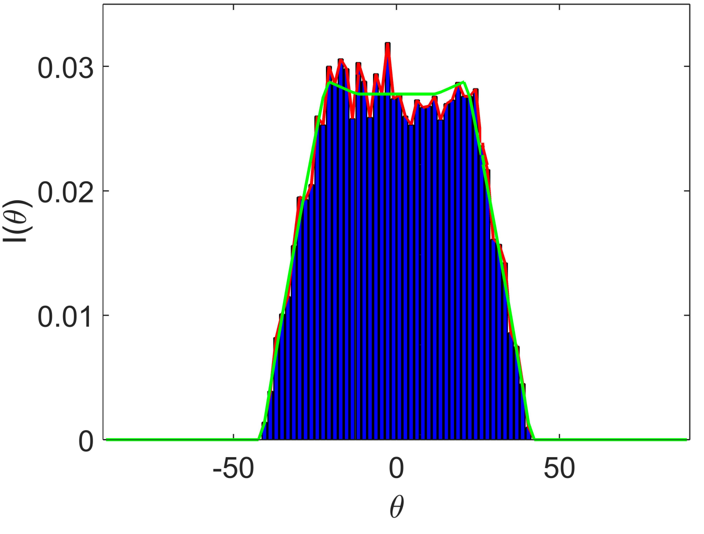

Figure 3.3: Comparison between the averaged normalized MC intensity $\hat{I}_{\mathrm{MC}}$ and the normalized exact intensity $\hat{I}_{\mathrm{exact}}$

.

Chapter 4

Ray tracing on phase space

4.1 Phase space concept

4.2 The edge-ray principle

4.3 Phase space ray tracing

Chapter 5

Two different appoaches to compute the boundaries in target phase space

5.1 The α-shapes approach

Given a finite set $\mathcal{S}$ of points we want to determine the shape formed by these points. α-shapes are geometrical objects which give us a good approximation of the shape of a given point set $\mathcal{S}$. Before giving a formal definition we explain an intuitive interpretation of α-shapes. As mentioned in [16] we can think of an α-shape as a mass of ice-cream with several chocolate pieces. The mass making up the space $\mathbb{R}^3$ and the chocolate pieces are the point set $\mathcal{S}$. Then the aim is to find the shape formed by the chocolate pieces. We can use a spoon with a spherical shape and carve out all parts of the ice-cream without removing the chocolate pieces. We will obtain a shape formed by arcs and points (see figure below for the two-dimensional case). Straightening the arcs to triangles and line segments we have an intuitive description of what is called the α-shape of $\mathcal{S}$. In our example, the parameter α determines of the radius of the carving spoon. If α is equal to 0 the shape degenerates to the point set $\mathcal{S}$. On the other hand, when $\alpha \to \infty$ the α-shape is simple the convex hull. More precisely the process is summarized as follows. Given a point cloud $\mathcal{S}$ we start with a triangulation of it (a possible choice could be the Delaunay triangulation described in the next section). For each triangle we calculate the radius of the circumcircle. If the radius is larger that α the triangle is removed from the shape. The rule of the parameter α is highly significant in this procedure. Hence we have to choose it in such a way to get a better approximation. The choice of the parameter α is closely related to the radius of the circumcircles. A possible strategy is to find the radius of the greater empty circumcircle. Thus α is related to the density of the points. In particular we have:

$$\alpha = C\frac{1}{\Delta}\ , \tag{5.1.1}$$

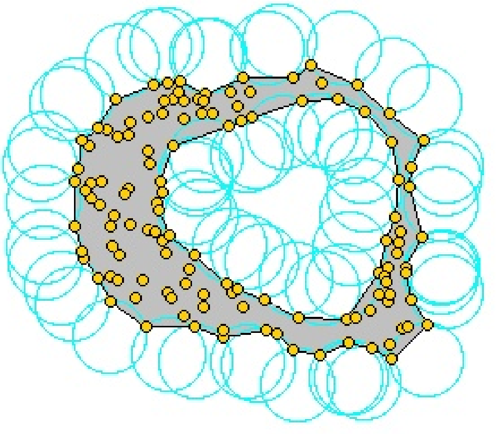

Figure 5.1: Construction of α-shape given a set of points in $\mathbb{R}^2$.

with C a constant that can be determined by a simulation and Δ the density of the point set $\mathcal{S}$ defined as:

$$\Delta = \frac{N}{\text{surface area}} , \tag{5.1.2}$$

where N is the number of points in $\mathcal{S}$ and the surface area is the area inside the boundaries of the region formed by the points cloud. Hence Δ is a constant. As mentioned above to find the α-shape of a point cloud we need a triangulation and a possible choise could be the Delaunay triangulation. As explained in [17] we can see a Delaunay triangulation as the dual of a Voronoi diagram. Let us define a Voronoi diagram in a metric space.

Definition 5.1.1. Let X be a space endowed with a distance d and $\mathcal{S} = \{S_1, \cdots, S_n\}$ a set formed by subsets of X. The Voronoi cell R_k associated with the set S_k where $k \in \{1, \cdots, n\}$ is defined as follows:

$$R_k = \{\mathbf{x} \in X \mid \mathrm{d}(\mathbf{x}, S_k) \leq \mathrm{d}(\mathbf{x}, S_j) \quad \forall j \neq k\} , \tag{5.1.3}$$

where $\mathrm{d}(x, A) = \inf\{\mathrm{d}(x, a) \mid a \in A\}$. The Voronoi diagram is defined as the tuple of the cells $(R_k)_{k\in\{1,\ldots,n\}}$ that are assumed to be disjoint.

The simplest case that we can have is the two-dimensional case that is the case where $X = \mathbb{R}^2$. The tuple $\mathcal{S} = \{1, \cdots, n\} \subset \mathbb{R}^2$ is now a set of points. The Voronoi diagram of $\mathcal{S}$ is a subsection of $\mathbb{R}^2$ such that every other region around a point $p \in \mathcal{S}$ contains all points that are closer to p than to every point in $\mathcal{S}$. A triangulation of the point set $\mathcal{S}$ is a set of edges $\mathcal{E}$ whose extremes are points of $\mathcal{S}$ such that the faces of each triangle are bounded by three edges and any edge that is not in $\mathcal{E}$ intersects one of the existing edges. The Delaunay triangulation is the dual graph of the Voronoi diagram: it consists of vertices (the points in $\mathcal{S}$) and it has an edge between two vertices if the two corresponding faces share an edge.
The Delaunay triangulation triangulates the convex hull of the point set $\mathcal{S}$. Instead,

the α-shape of a point set is formed only by the triangles (taken from the Delaunay triangulation) that satisfy the "α-test" and therefore is a suitable method to reconstruct the surface formed by a point cloud. Even if α-shapes are a powerful tool to reconstruct surfaces, some simulations show that there exist surfaces that are not described well by α-shapes. Indeed for some particular surface there exist no value of α that includes all desired triangles and deletes all undesired triangles. For instance, since the parameter α depends on the density of the point cloud, is intuitively clear that using α-shapes for a non-uniform points set we won't get a good approximation of the surface. Furthermore, the α-shape method doesn't work well when there is a sharp turn or a joint. In this case α-shapes often give a "webbed-foot" appearance at such joints since they improperly connect the adjacent surfaces. Hence a generalization of "classical" α-shapes is required. In the next section a method to solve the "density problem" for two separated and close objects is described. In [18] Teichmann and Capps present "Density-scaled α-shapes". The first step of this method is to make a triangulation of the point cloud. Then the key idea is to compute somehow the point-density of each point and use this to get an approximation of the point density of a triangle. In this way one can reduce the α-value in areas where the triangle's point density (see equation 5.1.6 for the definition) is higher than average in such a way that is possible to obtain a finer level of detail for areas that have an higher density. More precisely, each point $\mathbf{p} \in \mathcal{S}$ has a local point density defined as

$$\delta(\mathbf{p}) = \sum_{\mathbf{q}\in\mathcal{S}} \left(1 - \frac{\mathrm{d}(q,p)}{\lambda}\right) \qquad \forall \mathbf{q} \text{ such that } \mathrm{d}(\mathbf{p},\mathbf{q}) < \lambda\,, \tag{5.1.4}$$

where λ is the constant radius of the local neighborhood and $\mathrm{d}(\mathbf{x},\mathbf{y})$ is the Euclidean distance. When local density is larger than the average, that is when

$$\delta(\mathbf{p}) > \frac{1}{|\mathcal{S}|} \sum_{\mathbf{q}\in\mathcal{S}} \delta(\mathbf{q}) \tag{5.1.5}$$

we know some properties about the region surrounding $\mathbf{p}$. For instance, if the point set is uniformly distributed then it is possible to find areas with a high-density in the case where there are two closely separated surfaces. In point sets of non-uniform distribution, high densities are found when the surface presents a joint discontinuity. The algorithm developed by Teichmann and Capps is structured as follow. After computing density information for each point they make a triangulation of the point set. Then they calculate the average density $\delta(t)$ for each triangle Δ_{abc} defined as:

$$\delta(t) = \frac{\delta(a) + \delta(b) + \delta(c)}{3\mu}\,, \tag{5.1.6}$$

where μ is the global average density of the entire point set $\mathcal{S}$. If $\delta(t)$ is greater than 1 the density of the point cloud is higher. Hence is necessary to define another value of α:

$$\alpha\,' = \frac{\alpha}{\delta(t)^{\sigma}} \tag{5.1.7}$$

where σ is a value that is adjusted by the user. If δ is less than 1 the α-value is not modified. In this way it is possible to have a finer precision on the shape formed by the point set where the density is higher than the average density. Hence it is possible to distinguish two separated objects with different density.

5.2 The two-faceted cup

5.3 Results for a TIR collimator

5.4 The triangulation refinement approach

5.5 The two-faceted cup

5.6 Results for a TIR collimator

5.7 Results for a Parabolic reflector

5.8 Results for the Compound Parabolic Concentrator (CPC)

Chapter 6

The inverse ray mapping method: analytic approach

6.1 Explanation of the method

6.2 The two-faceted cup

6.3 The multi faceted cup

For two-dimensional systems every ray in the PS of a line is given by a two-tuple point. Therefore, the PS of every line is a two-dimensional space. The position coordinate in the PS of line i is the x-coordinate of the intersection point between the ray and the line i. The direction coordinate is the sine of the angle that the ray forms with respect to the normal of the line i multiplied by the index of refraction of the medium in which the ray is located. Let's now introduce some notation before explaining the details of the method. We indicate the PS with $S = Q \times P$, where Q is the set of the position coordinates q and P is the set of the direction coordinates $p = n \sin \tau$ with τ the angle between the ray and the normal $\boldsymbol{\nu}$ of the line and n is the index of refraction of the medium in which the line is located.

6.4 Results for the two-faceted cup

6.5 Results for the multi-faceted cup

6.6 Discussions

Chapter 7

The extended ray mapping method

7.1 Explanation of the method

7.2 Bisection procedure

7.3 Results for a parabolic reflector

7.4 Results for two different kind of TIR-collimators

Chapter 8

Extended ray mapping method to systems with Fresnel reflection

Chapter 9

Discussion and conclusions

Acknowledgments

Bibliography

[1] E. F. Zalewski, “Radiometry and photometry,” *Handbook of optics*, vol. 2, pp. 24–1, 1995.

[2] J. Chaves, *Introduction to nonimaging optics.* CRC press, 2015.

[3] “Luminous efficacy-wikipedia the free encyclopedia,” *https://commons.wikimedia.org/wiki/File:CIE 1931 Luminosity.png media/File:CIE 1931 Luminosity.png.*

[4] A. V. Arecchi, R. J. Koshel, and T. Messadi, “Field guide to illumination,” SPIE, 2007.

[5] H. Zhu and P. Blackborow, “Etendue and optical throughput calculations,” *Energetiq Technology, Inc., Woburn, MA*, 2011.

[6] R. P. Feynman, “Feynman lectures on physics. volume 2: Mainly electromagnetism and matter,” *Reading, Ma.: Addison-Wesley, 1964, edited by Feynman, Richard P.; Leighton, Robert B.; Sands, Matthew*, 1964.

[7] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light.* Elsevier, 2013.

[8] E. Hecht, *Optics.* Parson Addison-Wesley, 2002.

[9] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman lectures on physics, Vol. I: The new millennium edition: mainly mechanics, radiation, and heat*, vol. 1. Basic books, 2011.

[10] P. H. Jones, O. M. Maragò, and G. Volpe, *Optical tweezers: Principles and applications.* Cambridge University Press, 2015.

[11] H. Gross, *Handbook of the Optical Systems*, vol. 1. Wiley-VCH, 2005.

[12] A. B. Owen, “Quasi-monte carlo sampling,” *Monte Carlo Ray Tracing: Siggraph*, vol. 1, pp. 69–88, 2003.

[13] V. M. Zolotarev, *Modern theory of summation of random variables.* Walter de Gruyter, 1997.

[14] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo method*, vol. 10. John Wiley & Sons, 2016.

[15] D. M. Diez, C. D. Barr, and M. Cetinkaya-Rundel, *OpenIntro statistics*. CreateSpace, 2012.

[16] H. Edelsbrunner and E. P. Mücke, "Three-dimensional alpha shapes," *ACM Transactions on Graphics (TOG)*, vol. 13, no. 1, pp. 43–72, 1994.

[17] J. Portegies and P. Lighting, "Fast ray tracing in phase space for optical design," 2013.

[18] M. Teichmann and M. Capps, "Surface reconstruction with anisotropic density-scaled alpha shapes," in *Visualization'98. Proceedings*, pp. 67–72, IEEE, 1998.