METASURFACES AND WAVEGUIDES IN OPTICS

A Dissertation
Submitted to
the Temple University Graduate Board

in Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

by
Luca Pallucchini
May 2019

Examining Committee:
Prof. Cristian Gutiérrez, Advisory Chair, Mathematics
Prof. Shiferaw Berhanu, Mathematics
Prof. Brian Rider, Mathematics
Prof. David Futer, Mathematics (Graduate Chair)
Prof. Peter Riseborough, External Member, Department of Physics, Temple University

ABSTRACT

METASURFACES AND WAVEGUIDES IN OPTICS

Luca Pallucchini

DOCTOR OF PHILOSOPHY

Temple University, May 2019

Professor Cristian E. Gutiérrez, Chair

This thesis analyzes metasurfaces and waveguides in geometric optics. In the first and second chapters, we give a mathematical approach to study metasurfaces. A metasurface is a surface together with a function called phase discontinuity. The phase discontinuity is chosen so that the metasurface produces a desired reflection or refraction job. We give analytical conditions between the curvature of the surface and the set of refracted directions to guarantee the existence of phase discontinuities. The approach contains both the near and far field cases. A starting point is the formulation of a vector Snell's law in the presence of abrupt discontinuities on the interfaces. Also, we derive the equations that the phase discontinuity function must satisfy in order for the metasurface to refract or reflect energy with a prescribed energy pattern, they are Monge-Ampére partial differential equations, and we prove the existence of solutions. In the third chapter, we model energy losses in waveguides. In particular, we give quantitative estimates of the energy internally reflected in case of a straight guide and a circularly curved guide. We give a detailed

ray tracing and internally reflected energy analysis for each striking point on the boundary of the guide.

ACKNOWLEDGEMENTS

First and foremost, I would like to give the deepest and most sincere gratitude possible to my advisor Prof. Cristian E. Gutiérrez. He has always been very patient and enthusiastic and I have found his advice and support invaluable. I feel very fortunate and privileged for being a student of such a great mentor and mathematician. Without his coaching, this research project would have never been completed.

I also owe my special thanks to my thesis committee members Prof. Shiferaw Berhanu, Prof. Brian Rider, Prof. David Futer and Prof. Peter Riseborough.

I would also like to thank all the Temple math faculty who I have encountered in some way or another. My knowledge of mathematics has grown so much in the past years and I owe that to the professors of the many classes I have taken as a graduate student at Temple. In particular, I would like to thank Prof. Gerardo Mendoza and Prof. Irina Mitrea for all the analysis and PDEs I learned from them.

I want to thank my wonderful and supportive wife Giulia and all my friends for being an integral part of my life.

Finally, I would like to thank my parents, my sisters and my entire family for their love and support.

Luca Pallucchini

April 8, 2019

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This dissertation concerns the solution of three problems in Geometric Optics. The first two problems are related to metalenses, and the third one is related to waveguides. Metalenses are ultra thin surfaces which use nano structures to focus light. The shaping of light wave fronts with standard lenses relies on gradual phase changes accumulated along the optical path inside the lens. Metalenses introduce abrupt phase shifts (phase discontinuities) over the scale of the wavelength along the optical path to bend light in unusual ways. The nano structures used are composed of arrays of tiny pillars, rings, and other arrangements of materials, which work together to manipulate light waves as they pass by. The subject of metalenses is a flourishing area of research and one of the nine runners-up for Science's Breakthrough of the Year 2016 [37]. That year, researchers used computer chip-patterning techniques to create the first metamaterial lens, or metalens, that can focus the full spectrum of visible light. A purpose in this dissertation is to give a mathematically rigorous foundation for metalenses, in particular, when it is theoretically possible to find non flat metalenses that bend light in the desired way or that yield prescribed distributions of energy. These questions are potentially important in the applications because metalenses are thinner than a sheet of paper and far lighter than glass, and they could revolutionize everything from microscopes to virtual reality displays to cameras, including the ones in smartphones [37]. The third question considered in this dissertation concerns energy losses in waveguides. A waveguide is a structure that leads the way of electromagnetic or sound waves. Waveguides cannot guide electromagnetic energy around bends without losing power by radiation, and therefore, we study energy losses in circularly curved waveguides.

We next describe more precisely the problems solved in the thesis and how it is organized. The first problem solved is about the existence of nonflat metalenses that bend light in a desired way. This is done in Chapter 2, and the results have appeared in [17]. For classical lens design, a typical problem is to find two surfaces so that the region sandwiched between them and filled with a homogeneous material refracts light in a desired manner. For metalens design, a surface is given and the question is to find a function on the surface -a phase discontinuity- so that the pair, surface together with the phase discontinuity -the metalens-, refracts light in a desired manner. We first derive in Section 2.1 the following generalized Snell's law in the presence of a phase discontinuity using wavefronts. Let n_1, n_2 be the refractive indices of two homogeneous media I and II, respectively. Suppose a surface Γ separates these media, and an incoming light ray in medium I with unit direction vector

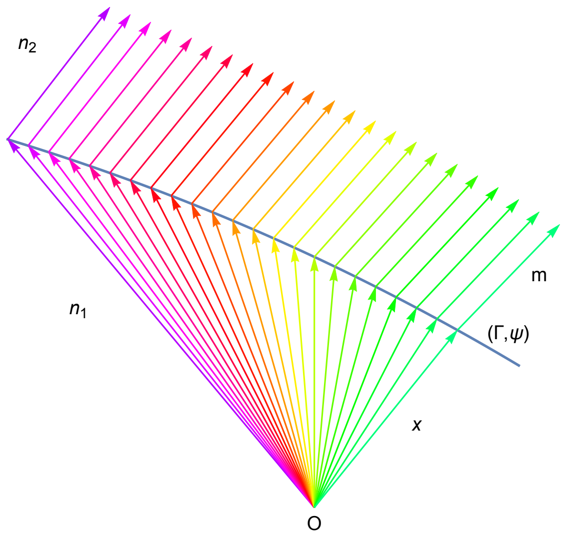

Figure 1.1: Metalens refracting into a fixed direction

x strikes Γ. Assume that there is a real-valued function ψ, the phase discontinuity, defined in a neighborhood of the surface Γ. If ν denotes the unit normal vector to Γ, then the refracted wave vector m satisfies (see (2.6))

$$n_1 x - n_2 m = \lambda \nu + \nabla \psi,$$

where λ satisfies

$$\lambda^2 - [2(n_1 x - \nabla \psi) \cdot \nu] \lambda + |n_1 x - \nabla \psi|^2 - n_2^2 = 0.$$

If ψ is constant, then we recover the classical Snell law in vector form. Using this generalized Snell's law, we introduce a mathematical method to construct phase discontinuities on a given surface, not necessarily flat, so that radiation is steered into a prescribed set of directions. More precisely, given a surface Γ in three dimensional space, we determine when it is possible to have a function ψ defined on a very thin (comparable to the wave length of the radiation) neighborhood of Γ so that radiation emanating from a point source is refracted by the pair (Γ, ψ), surface and function, into a set of directions prescribed in advance. In other words, for a given set of directions where we want to steer the radiation, we discover what kind of surfaces Γ allow the existence of a function ψ so that the pair (Γ, ψ) directs the radiation in the desired way. This leads to ultra thin (not flat) optical components that produce abrupt changes over the scale of the free-space wavelength in the phase. This is in contrast with classical lens design, where the question is to engineer the gradual accumulation of phase delay as the wave propagates in the device, reshaping the scattered wavefront and beam profile at will. In particular, in standard lenses light propagates over distances much larger than the wavelength to shape wavefronts. The existence of phase discontinuity functions is intimately related with the shape of the given surface and the given set of directions. If these two objects satisfy the following condition

$$m_u \cdot r_v = m_v \cdot r_u$$

and the determinant of the matrix

$$-\rho \begin{pmatrix} x_u \cdot x_u & x_u \cdot x_v \\ x_v \cdot x_u & x_v \cdot x_v \end{pmatrix} + \kappa \begin{pmatrix} r_u \cdot m_u & r_u \cdot m_v \\ r_v \cdot m_u & r_v \cdot m_v \end{pmatrix} - B \begin{pmatrix} r_{uu} \cdot \nu & r_{uv} \cdot \nu \\ r_{vu} \cdot \nu & r_{vv} \cdot \nu \end{pmatrix}, \tag{1.1}$$

is not zero, then the existence of the desired phase discontinuity is guaranteed (see Theorem 2.3.1). Here $m(u, v) = (m_1(u, v), m_2(u, v), m_3(u, v))$ is the given C^2 unit field of directions, $r(u, v) = \rho(u, v)\, x(u, v)$ is a parametrization of Γ where $x(u, v)$ are spherical coordinates and $\rho(u, v) > 0$ is the polar radius, $B = (x - \kappa\, m) \cdot \nu$ and $\kappa = \frac{n_2}{n_1}$. Notice that the first and third matrices in (1.1) are respectively the first fundamental form of the 2-sphere, and the second fundamental form of the surface Γ. In particular, given a surface, one can obtain from these conditions what kind of sets of steering directions are allowed. Conversely, given a phase discontinuity and a desired transmission direction, we described the admissible surfaces that are possible (see Theorem 2.4.1).

The second problem solved is about the derivation of the partial differential equations governing light reflection and refraction in metalenses. This is the content of Chapter 3 and the results have appeared in [18]. A reflection problem considered in [41] is to find a perfectly reflecting surface Γ such that rays emitted from the origin with direction $x \in \Omega_1 \subset S^2$ and intensity $f(x)$, after being reflected by Γ cover a prescribed region $\Omega_2 \subset S^2$, and the density of the distribution of the reflected rays is a prescribed function $g(y)$ of the direction y. In [41] it is assumed conservation of energy, i.e.

$$\int_{\Omega_1} f(x)\, d\sigma(x) = \int_{T(\Omega_1)} g(y)\, d\sigma(y),$$

where T indicates the reflection map. The surface Γ is given as a solution of a Monge-Ampére type equation [41]. In our case for metalenses the surface Γ is given and the question consists in finding the phase discontinuity ψ doing a similar reflection job. We answer this question when either rays are emitted in a collimated way from an extended source (see Section 3.1.1) or when rays are emitted from a point source (see Section 3.1.2). For example, in Section 3.1.1 we consider the case in which rays are emitted from an open set of the $x-y$ plane, Γ is a plane parallel to the $x-y$ plane and the rays have direction $e_3 = (0, 0, 1)$. We have found that the phase discontinuity is given by the solution of the following Monge-Ampére type equation

$$\frac{1}{\sqrt{1-\psi_x(x,y)^2-\psi_y(x,y)^2}}|\det D^2\psi| = \frac{f(x)}{g(T(x))}.$$

We then study existence, uniqueness, and smoothness of the solutions. In addition, we also derive similar pdes and solve similar problems in case of refraction (see Sections 3.2.1 and 3.2.2). In Section 3.3 we give a summary of the equations, in all cases considered, when Γ is a plane.

The third problem solved in this dissertation is about energy losses in waveguides which is the contents of Chapter 4. The first mathematical analysis of electromagnetic waves in metal cylindrical structures was performed by Lord Rayleigh in [34]. For sound waves, Lord Rayleigh published a full analysis of propagation modes in his seminal work, "The Theory of Sound"[36]. For a detailed history about the origin of waveguides, we refer the reader to [32]. In this thesis, we give quantitative estimates of the energy internally reflected in the case of a straight guide (Section 4.1) and of a circularly curved guide (Section 4.2). To model this we use the Fresnel formulas and the set up is as follows. Suppose we have two homogeneous media I and II with refractive indices n_1, n_2, respectively, and with $n_1 > n_2$; we set $\kappa = \dfrac{n_2}{n_1}$. Suppose media I and II are separated by a smooth surface S. If an incident wave with unit direction x is traveling within medium I and strikes S at a point P, then the wave splits into two waves: one transmitted into medium II and another internally reflected into medium I. The unit directions of these waves are m_t and m_r, respectively, which are determined by the Snell law (2.2). Therefore, the incident energy E_i carried by the incident wave with direction

x splits into two: the transmitted energy E_t carried by the wave having direction m_t and the internally reflected energy E_r carried by the wave having direction m_r, with $E_i = E_t + E_r$, assuming no losses. The percentages of energy carried by the transmitted and internally reflected waves depends on the incident direction x via the Fresnel formulas, a consequence of Maxwell's equations [5]. With these, we model the losses of energy within a waveguide confined between two parallel surfaces S_1 and S_2. These surfaces are planes in the case of a straight waveguide and circularly curved surfaces for the circularly curved guide; we assume the dielectric within the two surfaces has refractive index n_1, and the cladding, i.e., the material outside, has refractive index n_2 with $n_1 > n_2$. An incident polarized wave will zigzag inside between the two surfaces. Depending on the normal to the surfaces at the striking points, one can calculate the energy transmitted and internally reflected. In other words, we follow the path of the ray and tally the internally reflected energy at each striking point on the boundary of the waveguide. Finally, in Section 4.4 we study asymptotics for the integral that represent the energy internally reflected. In case of a periodic circular guide, as $N \to \infty$, where N is the number of striking points on the outer contour of the guide, we show that the energy internally reflected goes to zero as $\dfrac{C}{N^2}$ where C is a positive constant. This result requires a careful analysis of the integrals that represent the energy internally reflected and it follows from Theorem 4.4.1 and 4.4.3. At the end of the thesis (Section 4.5), we briefly describe some open problems to continue the research in this area.

CHAPTER 2

GENERAL REFRACTION PROBLEMS WITH PHASE DISCONTINUITIES ON NON FLAT METASURFACES

Here, we recall the classical Snell law in vector form. Suppose Γ is a surface in $\mathbb{R}^3$ that separates two media I and II that are homogeneous and isotropic, with refractive indices n_1 and n_2 respectively. If a light ray [1] having direction $x \in S^2$, the unit sphere in $\mathbb{R}^3$, and traveling through medium I strikes Γ at the point P, then this ray is refracted in the direction $m \in S^2$ through medium II according to

[1]Since the refraction angle depends on the frequency of the radiation, we assume that light rays are monochromatic.

the Snell law in vector form:

$$n_1(x \times \nu) = n_2(m \times \nu), \tag{2.1}$$

where ν is the unit normal to the surface to Γ at P pointing towards medium II; see [29, Subsection 4.1]. It is assumed here that $x \cdot \nu \geq 0$. This has several consequences:

(a) the vectors x, m, ν are all on the same plane (called the plane of incidence);

(b) the well known Snell law in scalar form holds:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2,$$

where θ_1 is the angle between x and ν (the angle of incidence), and θ_2 is the angle between m and ν (the angle of refraction).

Equation (2.1) is equivalent to $(n_1 x - n_2 m) \times \nu = 0$, which means that the vector $n_1 x - n_2 m$ is parallel to the normal vector ν. If we set $\kappa = n_2/n_1$, then

$$x - \kappa\, m = \lambda \nu, \tag{2.2}$$

for some $\lambda \in \mathbb{R}$. Notice that (2.2) univocally determines λ. Taking dot products with x and m in (2.2) we get $\lambda = \cos\theta_1 - \kappa \cos\theta_2$, $\cos\theta_1 = x \cdot \nu > 0$, and $\cos\theta_2 = m \cdot \nu = \sqrt{1 - \kappa^{-2}[1 - (x \cdot \nu)^2]}$. In fact, there holds

$$\lambda = x \cdot \nu - \kappa \sqrt{1 - \kappa^{-2}\,(1 - (x \cdot \nu)^2)}. \tag{2.3}$$

The formulation (2.2) is useful to solve refraction problems for lens design, see [12], [13], [14], [15], and [8] for a numerical implementation.

2.1 Derivation of a Vector Snell's Law with phase discontinuity using wavefronts

Let n_1, n_2 be the refractive indices of two homogeneous media I and II, respectively. Suppose a surface Γ separates these media, and an incoming light ray in medium I with wave vector $\mathbf{k}_1$ strikes Γ. Assume that there is a real-valued function ψ, *the phase discontinuity*, defined in a neighborhood of the surface Γ. Notice that ψ must be defined in a neighborhood of Γ because the gradient of ψ will be considered. If ν denotes the unit normal vector to Γ, then the refracted wave vector $\mathbf{k}_2$ satisfies [2, Equation (2)]:

$$\nu \times (\mathbf{k}_2 - \mathbf{k}_1) = \nu \times \nabla\psi. \tag{2.4}$$

We give an alternate formulation and derivation of this result by using wavefronts; our starting point is [16, Section 2.2]. For each t, $\Psi(x, y, z, t) = 0$ denotes a surface in the variables x, y, z that separates the part of the space that is at rest from the part of the space that is disturbed by the electric and magnetic fields. This surface is called a *wave front*, and the light rays are the orthogonal trajectories to the wave fronts at each time t. We assume that $\Psi_t \neq 0$, and so we can solve $\Psi(x, y, z, t) = 0$ in t, obtaining that $\phi(x, y, z) = ct$; so letting t run, the wave fronts are then the level sets of $\phi(x, y, z)$.

Let n_1, n_2, and Γ be as above. An incoming wave front Ψ_1 on medium I strikes the surface Γ and it is then transmitted into a wave front Ψ_2 in medium II (of course, there is also a wave front reflected back). Assuming as before that

$(\Psi_j)_t \neq 0$, $j = 1, 2$, and solving in t, we get that the wave fronts are given by $\phi_j(x, y, z) = ct$ for $j = 1, 2$, respectively. Suppose the surface Γ is parameterized by $x = f(\xi, \eta)$, $y = g(\xi, \eta)$, $z = h(\xi, \eta)$. If there were no phase discontinuity on the surface Γ, then we would have $\phi_1 = \phi_2$ along Γ. But since there is now a phase discontinuity ψ on Γ, we have the following jump condition along Γ:

$$\phi_1(f(\xi, \eta), g(\xi, \eta), h(\xi, \eta)) - \phi_2(f(\xi, \eta), g(\xi, \eta), h(\xi, \eta)) = \psi(f(\xi, \eta), g(\xi, \eta), h(\xi, \eta)).$$

Taking derivatives in ξ and η yields

$$(\nabla\phi_1 - \nabla\phi_2 - \nabla\psi) \cdot (f_\xi, g_\xi, h_\xi) = 0,$$

and

$$(\nabla\phi_1 - \nabla\phi_2 - \nabla\psi) \cdot (f_\eta, g_\eta, h_\eta) = 0.$$

That is, the vector $\nabla\phi_1 - \nabla\phi_2 - \nabla\psi$ must be normal to Γ; as such there exists a real number λ such that

$$\nabla\phi_1 - \nabla\phi_2 - \nabla\psi = \lambda\nu \tag{2.5}$$

where ν is the unit normal to Γ.

Let $\gamma_j(t)$ denote the light rays in medium j having speed v_j, for $j = 1, 2$; i.e., the orthogonal trajectories to ϕ_j. In particular, we have that $\phi_j(\gamma_j(t)) = ct$, and by the chain rule

$$\nabla\phi_j(\gamma_j(t)) \cdot \gamma_j'(t) = c, \quad j = 1, 2$$

If we parameterize the rays so that $|\gamma_j'(t)| = v_j$, then we obtain

$$|\nabla\phi_j(\gamma_j(t))| = \frac{c}{v_j} = n_j, \quad j = 1, 2$$

since $\nabla\phi_j$ is parallel to γ_j'. Letting

$$x = \frac{\nabla\phi_1(\gamma_1(t))}{|\nabla\phi_1(\gamma_1(t))|}, \quad m = \frac{\nabla\phi_2(\gamma_2(t))}{|\nabla\phi_2(\gamma_2(t))|}$$

we obtain from (2.5) the following formula

$$n_1 x - n_2 m = \lambda\nu + \nabla\psi. \tag{2.6}$$

Taking cross products with the unit normal ν in (2.6), we obtain the equivalent formula

$$\nu \times (n_1 x - n_2 m) = \nu \times \nabla\psi. \tag{2.7}$$

Recall that x is the unit direction of the incident ray, m is the unit direction of the refracted ray, ν is the unit outer normal at the incident point on Γ and $\nabla\psi$ is calculated at the incident point. Note that in the case ψ is constant, we recover the classical Snell law in vector form (2.1)[2].

Starting from (2.6), we now calculate λ. Taking dot products in (2.6) and solving for $x \cdot m$ yields

$$x \cdot m = \frac{n_1 - \lambda\, x \cdot \nu - x \cdot \nabla\psi}{n_2}.$$

Next taking dot products in (2.6) with itself, expanding, and substituting $x \cdot m$ from the previous expression, yields that λ satisfies the quadratic equation:

$$\lambda^2 - \left[2(n_1 x - \nabla\psi) \cdot \nu\right] \lambda + |n_1 x - \nabla\psi|^2 - n_2^2 = 0. \tag{2.8}$$

[2]Notice that if ψ =constant, then $n_1\, \nu \times x = n_2\, \nu \times m$. Taking dot product with m yields $n_1\, m \cdot (\nu \times x) = 0$. This means that m is on the plane through the origin having normal $\nu \times x$ which is the plane generated by ν and x. Therefore, ν, x, m are all on the same plane, i.e., the plane of incidence. On the other hand, if ψ is not necessarily constant, then from (2.7) $n_1\, \nu \times x = n_2\, \nu \times m + \nu \times \nabla\psi$. Again taking dot product with m yields $n_1\, m \cdot (\nu \times x) = m \cdot (\nu \times \nabla\psi)$, that is, $m \cdot (\nu \times (n_1\, x - \nabla\psi)) = 0$. That is, now the refracted vector m lies on the plane through the origin and perpendicular to the vector $\nu \times (n_1\, x - \nabla\psi)$ where $\nabla\psi$ is calculated at the point on the surface Γ where the ray with direction x strikes it. This shows that in the general case the refracted vector m is not on the plane generated by ν and x.

Solving for λ yields

$$\lambda = (n_1 x - \nabla\psi)\cdot\nu \pm \sqrt{n_2^2 - \left(|n_1 x - \nabla\psi|^2 - [(n_1 x - \nabla\psi)\cdot\nu]^2\right)}. \tag{2.9}$$

Since λ must be a real number, the quantity under the square root must be non-negative, i.e.,

$$n_2^2 \geq |n_1 x - \nabla\psi|^2 - [(n_1 x - \nabla\psi)\cdot\nu]^2. \tag{2.10}$$

Assuming this for now, it remains to check which sign ($\pm$) to take in (2.9). Dotting (2.6) with ν and using (2.9) yields

$$\begin{aligned} n_1 x\cdot\nu - n_2 m\cdot\nu &= (n_1 x - \nabla\psi)\cdot\nu \\ &\quad \pm \sqrt{n_2^2 - \left(|n_1 x - \nabla\psi|^2 - [(n_1 x - \nabla\psi)\cdot\nu]^2\right)} + \nabla\psi\cdot\nu, \end{aligned}$$

so

$$-n_2 m\cdot\nu = \pm\sqrt{n_2^2 - \left(|n_1 x - \nabla\psi|^2 - [(n_1 x - \nabla\psi)\cdot\nu]^2\right)}.$$

Since $n_2 > 0$ and $m\cdot\nu \geq 0$, we obtain that

$$\lambda = (n_1 x - \nabla\psi)\cdot\nu - \sqrt{n_2^2 - \left(|n_1 x - \nabla\psi|^2 - [(n_1 x - \nabla\psi)\cdot\nu]^2\right)}. \tag{2.11}$$

We next analyze (2.10), which will yield the critical angles. Equation (2.10) is equivalent to

$$\left(\left(x - \frac{\nabla\psi}{n_1}\right)\cdot\nu\right)^2 \geq \left|x - \frac{\nabla\psi}{n_1}\right|^2 - \kappa^2.$$

Thus, if x is such that

$$\left|x - \frac{\nabla\psi}{n_1}\right| \leq \kappa,$$

then (2.10) holds. On the other hand, if

$$\left|x - \frac{\nabla\psi}{n_1}\right| > \kappa$$

then (2.10) holds when either

$$x \cdot \nu \geq \frac{\nabla\psi}{n_1} \cdot \nu + \sqrt{\left|x - \frac{\nabla\psi}{n_1}\right|^2 - \kappa^2} \quad \text{or} \quad x \cdot \nu \leq \frac{\nabla\psi}{n_1} \cdot \nu - \sqrt{\left|x - \frac{\nabla\psi}{n_1}\right|^2 - \kappa^2}.$$

Therefore, the critical angles between x and ν are θ_c with

$$x \cdot \nu = \cos\theta_c = \frac{\nabla\psi}{n_1} \cdot \nu + \sqrt{\left|x - \frac{\nabla\psi}{n_1}\right|^2 - \kappa^2}$$

or

$$x \cdot \nu = \cos\theta_c = \frac{\nabla\psi}{n_1} \cdot \nu - \sqrt{\left|x - \frac{\nabla\psi}{n_1}\right|^2 - \kappa^2}.$$

Remark 2.1.1. *In two dimensions the critical angles are considered in [43]. It is assumed there that the interface* Γ *is the* x*-axis, the region* $y > 0$ *is filled with a material with refractive index* n_1*, and the region* $y < 0$ *with a material with refractive index* n_2*. Also the phase discontinuity satisfies that* $\nabla\psi$ *is constant and is tangential to the interface, i.e.,* $\nabla\psi = (a, 0)$ *with, for example,* $a > 0$*. Therefore, the above calculations applied to this case yield*

$$\cos\theta_c = x \cdot \nu = \sqrt{\left|x - \frac{\nabla\psi}{n_1}\right|^2 - \kappa^2} = \sqrt{1 - \frac{2|\nabla\psi|}{n_1}\cos(\pi/2 - \theta_c) + \frac{|\nabla\psi|^2}{n_1^2} - \kappa^2},$$

where $\kappa = \dfrac{n_2}{n_1}$*. Squaring both sides we obtain*

$$\cos^2\theta_c = 1 - \frac{2|\nabla\psi|}{n_1}\sin\theta_c + \frac{|\nabla\psi|^2}{n_1^2} - \kappa^2,$$

and the critical angles θ_c *are therefore the solutions to the equation*

$$\sin^2\theta_c - \frac{2|\nabla\psi|}{n_1}\sin\theta_c + \frac{|\nabla\psi|^2}{n_1^2} - \kappa^2 = 0,$$

i.e.,

$$\theta_c = \arcsin\left(\frac{|\nabla\psi|}{n_1} \pm \kappa\right),$$

which is in agreement with [43, Formula (3)].

In three dimensions the critical angles are considered in [2]. The interface Γ *is the* $x-y$*-plane, the region* $z>0$ *is filled with a material with refractive index* n_1*, and the region* $z<0$ *with a material with refractive index* n_2*. Also the phase discontinuity is tangential to the interface, i.e.,* $\nabla\psi = \left(\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, 0\right)$ *and without loss of generality we may assume* $x=(0,y,z)$*. Once again, the above calculations applied to this case yield*

$$\cos\theta_c = x\cdot\nu = \sqrt{\left|x-\frac{\nabla\psi}{n_1}\right|^2-\kappa^2} = \sqrt{1-\frac{2}{n_1}\left|\frac{\partial\psi}{\partial y}\right|\cos(\pi/2-\theta_c)+\frac{|\nabla\psi|^2}{n_1^2}-\kappa^2}.$$

Proceeding as before we find

$$\theta_c = \arcsin\left(\frac{1}{n_1}\frac{\partial\psi}{\partial y}\pm\sqrt{\kappa^2-\frac{1}{n_1^2}\left|\frac{\partial\psi}{\partial x}\right|^2}\right),$$

recovering [2, Formula (8)].

Remark 2.1.2. *The reflection case is when* $n_1=n_2$*, so* (2.6) *and* (2.11) *become*

$$x-m = \frac{1}{n_1}\lambda\nu+\frac{\nabla\psi}{n_1},\ \lambda = (n_1 x-\nabla\psi)\cdot\nu+\sqrt{n_1^2-\left(|n_1x-\nabla\psi|^2-[(n_1x-\nabla\psi)\cdot\nu]^2\right)},$$

with x *the unit incident direction,* m *the unit reflected vector,* ν *the unit normal to the interface at the striking point, and* $\nabla\psi$ *at the striking point. Notice that the choice of the plus sign in front of the square root is because for reflection* $m\cdot\nu\leq 0$*.*

2.2 Far field uniformly refracting planar and spherical metalenses

Let Γ be a surface in three dimensional space and V be a vector valued function defined on Γ; $V:\Gamma\to\mathbb{R}^3$. If x is an incident unit direction striking Γ at a point P, and m is the unit refracted direction, then we obtain, dividing by n_1 in the generalized Snell law (2.6), that

$$x - \kappa\, m = \lambda\, \nu(P) + V(P), \tag{2.12}$$

where $\nu(P)$ is the unit outer normal to Γ at P for some $\lambda \in \mathbb{R}$; $\kappa = n_2/n_1$. Suppose rays emanate from the origin and we are given a fixed unit vector m. Our goal is to answer the following two questions. First, given a surface Γ separating media n_1 and n_2, find a field V defined on Γ so that all rays from the origin are refracted into the direction m. The second question is, given a field V defined in a region of $\mathbb{R}^3$, find a separation surface Γ between n_1 and n_2 within that region so that all rays emanating from the origin are refracted into the direction m.

We begin in this section answering the first question when Γ is either a plane or a sphere, surfaces of traditional interest in optics, showing explicit phase discontinuities. For general surfaces, the first question is considered in Section 2.3, even for the more general case of variable m. The second question is answered in Section 2.4.

2.2.1 Case of the plane

Let Γ be the plane $x_1 = a$ in $\mathbb{R}^3$ with $a > 0$. We want to determine a field $V = (V_1, V_2, V_3)$ defined on Γ so that all rays emanating from the origin are refracted into the unit direction $m = (m_1, m_2, m_3)$, with $m_1 > 0$, Figure 2.1. Using spherical coordinates $x(u, v) = (\cos u\, \sin v, \sin u\, \sin v, \cos v)$, $0 \leq u \leq 2\pi, 0 \leq v \leq \pi$,

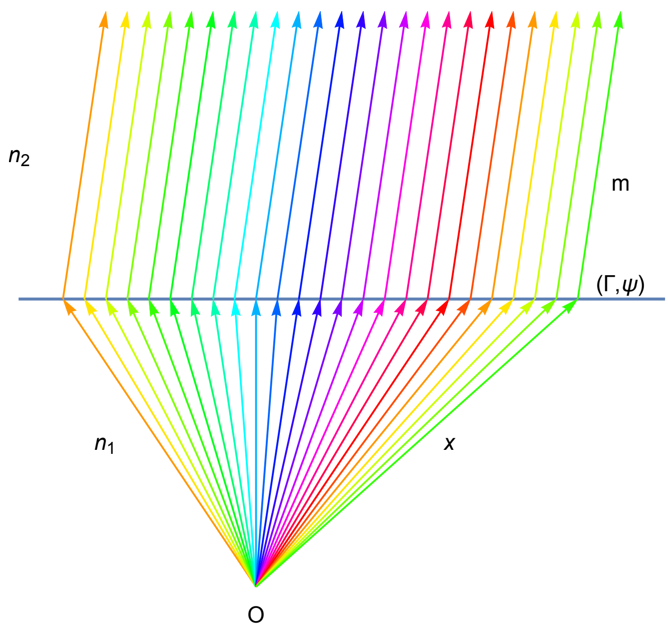

Figure 2.1: Planar metalens

Γ is described parametrically by

$$r(u,v) = \frac{a}{\cos u\, \sin v}\, x(u,v) = a\left(1, \tan u, \frac{1}{\cos u\, \tan v}\right). \tag{2.13}$$

Since the normal to the plane Γ is $\nu = (1,0,0)$, then (2.12) implies that $\sin u\, \sin v - \kappa\, m_2 = V_2(r(u,v))$ and $\cos v - \kappa\, m_3 = V_3(r(u,v))$. Hence V_2 and V_3 are univocally determined. Also, from (2.12) we get

$$V_1(r(u,v)) = \cos u\, \sin v - \kappa\, m_1 - \lambda(u,v). \tag{2.14}$$

Notice also that from (2.11),

$$\lambda = \nu \cdot (x - V) - \sqrt{(\nu \cdot (x - V))^2 - |x - V|^2 + \kappa^2},$$

which in the present case yields

$$\begin{aligned}
\lambda &= \cos u\, \sin v - V_1 - \sqrt{\kappa^2 - (\sin u\, \sin v - V_2)^2 - (\cos v - V_3)^2} \\
&= \cos u\, \sin v - V_1 - \sqrt{\kappa^2 - (\kappa\, m_2)^2 - (\kappa\, m_3)^2} \\
&= \cos u\, \sin v - V_1 - \kappa\, m_1 \qquad \text{since } m_1 > 0 \\
&= \frac{a}{\sqrt{a^2 + x_2^2 + x_3^2}} - V_1(a, x_2, x_3) - \kappa\, m_1.
\end{aligned}$$

This means that in (2.14) each V_1 determines λ and vice-versa.

We now write the field V in rectangular coordinates x_1, x_2, x_3. Since $\sqrt{a^2 + x_2^2 + x_3^2} = \dfrac{a}{\cos u\, \sin v}$, we can write

$$\begin{aligned}
V_2(a, x_2, x_3) &= \frac{x_2}{\sqrt{a^2 + x_2^2 + x_3^2}} - \kappa\, m_2 = \frac{\partial}{\partial x_2}\sqrt{x_1^2 + x_2^2 + x_3^2}\Big|_{x_1 = a} - \kappa\, m_2, \\
V_3(a, x_2, x_3) &= \frac{x_3}{\sqrt{a^2 + x_2^2 + x_3^2}} - \kappa\, m_3 = \frac{\partial}{\partial x_3}\sqrt{x_1^2 + x_2^2 + x_3^2}\Big|_{x_1 = a} - \kappa\, m_3
\end{aligned}$$

$$V_1(a,x_2,x_3) = \frac{a}{\sqrt{a^2+x_2^2+x_3^2}} - \kappa\, m_1 - \lambda = \frac{\partial}{\partial x_1}\sqrt{x_1^2+x_2^2+x_3^2}\Big|_{x_1=a} - \kappa\, m_1 - \lambda,$$

for $-\infty < x_2, x_3 < \infty$. From (2.13) $u = \arctan(x_2/a)$ and $v = \arctan\left(\dfrac{\sqrt{a^2+x_2^2}}{x_3}\right)$, so $\lambda(u,v) = h(x_2,x_3)$. Let $\psi(x_1,x_2,x_3) = \sqrt{x_1^2+x_2^2+x_3^2} - \kappa\, m_1\, x_1 - \kappa\, m_2\, x_2 - \kappa\, m_3\, x_3$. Therefore, if on the plane $x = a$ we give the field

$$V(x_1,x_2,x_3) := \nabla\psi(x_1,x_2,x_3) - h(x_2,x_3)\,\mathbf{i}, \tag{2.15}$$

then resulting metasurface does the desired refraction job. If we want V to be the gradient of a function, then $h(x_2,x_3)\,\mathbf{i}$ must be a gradient, which is only possible when $h(x_2,x_3) = C_0$ a constant; that is, $V = \nabla\left(\psi(x_1,x_2,x_3) - C_0\, x_1\right)$. As a particular case when $m_1 = 1$, $m_2 = m_3 = 0$, and $C_0 = 0$, we obtain the equivalent [42, Formula (2)] (where a different orientation of the coordinates is used) with $x_1 = a = f$. Notice also that if we want V in (2.15) to be tangential to the plane $x_1 = a$, that is, $(\nabla\psi(a,x_2,x_3) - h(x_2,x_3)\,\mathbf{i})\cdot(1,0,0) = 0$, then $h = \dfrac{a}{\sqrt{a^2+x_2^2+x_3^2}} - \kappa\, m_1$.

2.2.2 Case of the sphere

Now, the surface Γ considered is a sphere of radius R centered at the origin, that is, $r(u,v) = R\,x(u,v)$, with $x(u,v)$ spherical coordinates. We denote by $x = x(u,v)$; Figure 2.2. Since Γ is a sphere, the normal $\nu = x$ and from (2.12) we get $(x - \kappa\, m - V)\times x = 0$, so

$$(V + \kappa\, m)\times x = 0. \tag{2.16}$$

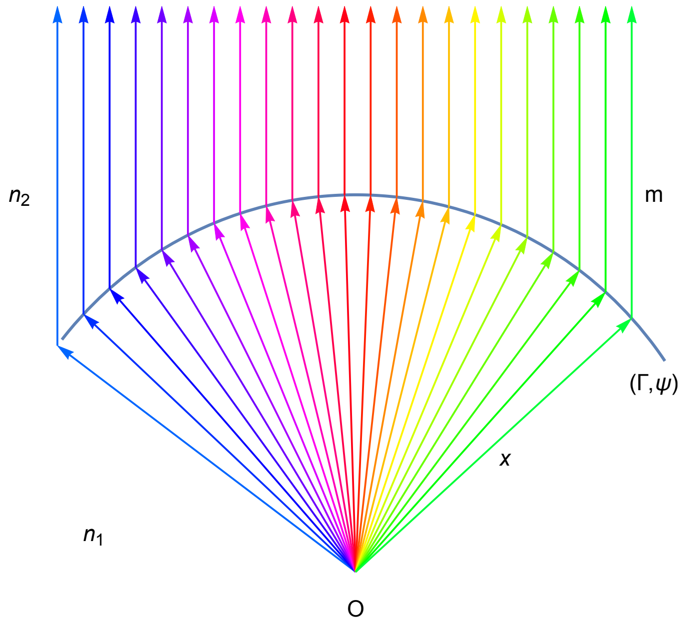

Figure 2.2: Spherical metalens

That is,

$$\begin{bmatrix} x_2 & -x_1 & 0 \\ -x_3 & 0 & x_1 \\ 0 & x_3 & -x_2 \end{bmatrix} \begin{pmatrix} V_1 + \kappa\, m_1 \\ V_2 + \kappa\, m_2 \\ V_3 + \kappa\, m_3 \end{pmatrix} = 0.$$

Notice that $\det \begin{pmatrix} x_2 & -x_1 & 0 \\ -x_3 & 0 & x_1 \\ 0 & x_3 & -x_2 \end{pmatrix} = 0$. Set $W_i = V_i + \kappa\, m_i$, so the system is equivalent to

$$\begin{bmatrix} 0 & 0 & 0 \\ x_2 x_3 & -x_1 x_3 & 0 \\ 0 & x_1 x_3 & -x_1 x_2 \end{bmatrix} \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} = 0.$$

If $x_1 x_2 x_3 \neq 0$, the last matrix has rank two, so the space of solutions has dimension one and the solutions are given by

$$(W_1, W_2, W_3) = \left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1 \right) W_3,$$

with W_3 arbitrary. Therefore,

$$V_1\,(R\, x(u,v)) = \frac{x_1}{x_3}\,(V_3\,(R\, x(u,v)) + \kappa\, m_3) - \kappa\, m_1$$

$$V_2\,(R\, x(u,v)) = \frac{x_2}{x_3}\,(V_3\,(R\, x(u,v)) + \kappa\, m_3) - \kappa\, m_2,$$

with V_3 arbitrary.

Notice that if in (2.16) we take cross product with x, we get

$$\begin{aligned} 0 &= x \times ((V + \kappa\, m) \times x) \\ &= (V + \kappa\, m)\ (x \cdot x) - x\ ((V + \kappa\, m) \cdot x) \end{aligned}$$

$$= V + \kappa\, m - (\kappa\, (m \cdot x) + V \cdot x)\; x.$$

Hence, if we want to pick V tangential to the sphere, we obtain

$$V(R\,x) = -\kappa\, m + \kappa\, (m \cdot x)\, x \text{ with } |x| = 1.$$

V is a field defined on the sphere of radius R. We shall determine a function ψ defined in a neighborhood of the sphere of radius R such that $V(Rx) = \nabla\psi(Rx)|_{|x|=1}$, and satisfying

$$\psi_{x_j}(R\,x) = -\kappa\, m_j + \kappa\, (m \cdot x)\, x_j, \text{ for } |x| = 1, \quad 1 \le j \le 3. \tag{2.17}$$

In fact, we have ($x = x(u, v)$)

$$\frac{\partial\psi(Rx(u,v))}{\partial u} = R \sum_{k=1}^{3} \frac{\partial\psi}{\partial x_k}(Rx(u,v))\, (x_k)_u = R\, (D\psi)(Rx(u,v)) \cdot x_u$$

$$= R\; (-\kappa\, m \cdot x_u + \kappa\, (m \cdot x)\, (x \cdot x_u)) = -\kappa\, R\, (m \cdot x_u) = -\kappa\, R\, \frac{\partial}{\partial u}(m \cdot x),$$

and similarly,

$$\frac{\partial\psi(Rx(u,v))}{\partial v} = -\kappa\, R\, \frac{\partial}{\partial v}(m \cdot x).$$

Integrating the derivative in u yields

$$\psi(Rx(u,v)) = -\kappa\, R\, (m \cdot x) + g(v),$$

and integrating the derivative in v we obtain

$$\psi(Rx(u,v)) = -\kappa\, R\, (m \cdot x(u,v)) + C_1,$$

with C_1 an arbitrary constant. Writing this in rectangular coordinates yields

$$\psi\, (R\, (z_1, z_2, z_3)) = -\kappa\, R\, (m \cdot (z_1, z_2, z_3)) + C_1, \text{ for } |(z_1, z_2, z_3)| = 1.$$

We now define ψ on a neighborhood of $|z|= R$ so that (2.17) holds. Define

$$\psi(z) = -\kappa\, R\,(m \cdot z)\,|z|^{-1}+C_1,\ \text{ for } R-\epsilon < |z|< R+\epsilon. \tag{2.18}$$

We have

$$\nabla\psi(z) = -\kappa\, R\, m\,|z|^{-1}+\kappa\, R\,(m\cdot z)\,z\,|z|^{-3},$$

so for $z = Rx$, with $|x|= 1$, we obtain

$$\nabla\psi(Rx) = -\kappa\, m + \kappa\,(m\cdot x)\,x$$

as desired. Therefore, the phase discontinuity ψ from (2.18) has gradient tangential to the sphere and can be placed on the spherical interface $|z|= R$ so that all rays from the origin are refracted into the fixed direction m.

2.3 Metalenses refracting into a set of variable directions

Suppose $m(u, v) = (m_1(u, v), m_2(u, v), m_3(u, v))$ is a given C^2 unit field of directions, and let Γ be a C^2 surface given parametrically by $r(u, v) = \rho(u, v)\,x(u, v)$ where $x(u, v)$ are spherical coordinates and $\rho(u, v) > 0$ is the polar radius. We want to see when is it possible to have a phase discontinuity ψ on the surface Γ so that each ray from the origin with direction $x(u, v)$ is refracted into the direction $m(u, v)$. The following theorem gives sufficient conditions for such a phase discontinuity to exist.

Theorem 2.3.1. *If a variable field m and a surface Γ given parametrically by $r(u, v) = \rho(u, v)\,x(u, v)$, where $x(u, v)$ are spherical coordinates and $\rho(u, v) > 0$ is* *the polar radius, satisfy the compatibility condition*

$$m_u \cdot r_v = m_v \cdot r_u, \tag{2.19}$$

and the determinant of the matrix

$$\left(-\rho \begin{pmatrix} x_u \cdot x_u & x_u \cdot x_v \\ x_v \cdot x_u & x_v \cdot x_v \end{pmatrix} + \kappa \begin{pmatrix} r_u \cdot m_u & r_u \cdot m_v \\ r_v \cdot m_u & r_v \cdot m_v \end{pmatrix} - B \begin{pmatrix} r_{uu} \cdot \nu & r_{uv} \cdot \nu \\ r_{vu} \cdot \nu & r_{vv} \cdot \nu \end{pmatrix} \right), \tag{2.20}$$

with $B = (x - \kappa\, m) \cdot \nu$, *is not zero at a point* (u_0, v_0), *then there is a neighborhood* U *of the point* $r(u_0, v_0)$ *and a phase discontinuity function* ψ *defined in* U *for the surface* Γ, *with gradient* $\nabla\psi$ *tangential to* Γ, *so that each ray emanating in the direction* $x(u, v)$, *for* (u, v) *in a neighborhood of* (u_0, v_0), *is refracted by the metasurface* (Γ, ψ) *into the direction* $m(u, v)$.

Notice that the first and third matrices in (2.20) are respectively the first fundamental form of the 2-sphere, and the second fundamental form of the surface Γ.

Proof. From (2.12)

$$x(u, v) - \kappa\, m(u, v) - V(r(u, v)) = \lambda\, \nu(r(u, v))$$

so

$$(x - \kappa\, m - V) \times \nu = 0.$$

Taking cross product with ν yields

$$0 = \nu \times ((x - \kappa\, m - V) \times \nu) = (x - \kappa\, m - V)\,(\nu \cdot \nu) - \nu\,((x - \kappa\, m - V) \cdot \nu)\,.$$

If V is tangential to Γ, then $V \cdot \nu = 0$ and so

$$0 = x - \kappa\, m - V - ((x - \kappa\, m) \cdot \nu)\ \nu,$$

that is,

$$V = x - \kappa\, m - ((x - \kappa\, m) \cdot \nu)\ \nu.$$

If $V(r(u,v)) = (\nabla\psi)(r(u,v))$, then

$$\psi_{x_j}(r(u,v)) = x_j(u,v) - \kappa\, m_j(u,v) - ((x(u,v) - \kappa\, m(u,v)) \cdot \nu(r(u,v)))\ \nu_j(r(u,v)).$$

Since $\nu \cdot r_u = \nu \cdot r_v = 0$ and $x \cdot x_u = x \cdot x_v = 0$,

$$\begin{aligned}
&\frac{\partial}{\partial u}\,(\psi(r(u,v))) \\
&= (\nabla\psi)(r(u,v)) \cdot r_u = (x - \kappa\, m) \cdot r_u - ((x - \kappa\, m) \cdot \nu)\ (\nu \cdot r_u) \\
&= (x - \kappa\, m) \cdot r_u = (x - \kappa\, m) \cdot (\rho_u\, x + \rho\, x_u) \\
&= \rho_u\,(x - \kappa\, m) \cdot x + \rho\,(x - \kappa\, m) \cdot x_u \\
&= \rho_u\,(1 - \kappa\, m \cdot x) - \kappa\,\rho\, m \cdot x_u = \rho_u\,(1 - \kappa\, m \cdot x) - \kappa\,\rho\,(m \cdot x)_u + \kappa\,\rho\,(m_u \cdot x) \\
&= \{\rho\,(1 - \kappa\, m \cdot x)\}_u + \kappa\,\rho\,(m_u \cdot x),
\end{aligned}$$

and similarly

$$\frac{\partial}{\partial v}\,(\psi(r(u,v))) = \{\rho\,(1 - \kappa\, m \cdot x)\}_v + \kappa\,\rho\,(m_v \cdot x).$$

Let us now consider the first order system in Φ

$$\begin{cases} \Phi_u &= \kappa\,\rho\,(m_u \cdot x) \\ \Phi_v &= \kappa\,\rho\,(m_v \cdot x). \end{cases} \tag{2.21}$$

where $\Phi(u,v) = \psi(r(u,v)) - \rho\,(1-\kappa m\cdot x)$. Then (2.21) can be written as

$$\nabla\Phi = H(u,v,\Phi), \tag{2.22}$$

where $H(u,v,\Phi) = (\kappa\,\rho\,(m_u\cdot x), \kappa\,\rho\,(m_v\cdot x))$. To solve the system (2.21) we need an initial condition, say $\Phi(u_0,v_0) = \Phi_0$, and use a result from [23, Chapter 6, pp. 117-118], that is, if

$$\frac{\partial H_1}{\partial v}(u,v,\Phi) + \frac{\partial H_1}{\partial \Phi}(u,v,\Phi)H_2(u,v,\Phi) = \frac{\partial H_2}{\partial u}(u,v,\Phi) + \frac{\partial H_2}{\partial \rho}(u,v,\Phi)H_1(u,v,\Phi) \tag{2.23}$$

holds for all (u,v,Φ) in an open set O, then for each $(u_0,v_0,\Phi_0)\in O$ there is neighborhood U of (u_0,v_0) and a unique solution $\Phi(u,v)$ defined for $(u,v)\in U$ solving the system (2.21) and satisfying $\Phi(u_0,v_0)=\Phi_0$. Therefore, if the given set of directions $m(u,v)$ and the surface Γ satisfy

$$m_u\cdot r_v = m_v\cdot r_u, \tag{2.24}$$

that is condition (2.23) for $H(u,v,\Phi) = (\kappa\,\rho\,(m_u\cdot x), \kappa\,\rho\,(m_v\cdot x))$, then there exists Φ solving (2.21). By integration we then obtain that the phase discontinuity ψ satisfies along Γ that

$$\psi(r(u,v)) = \rho\,(1-\kappa\,m\cdot x)+\Phi(u,v) = |r(u,v)|-\kappa\,(m(u,v)\cdot r(u,v))+\Phi(u,v). \tag{2.25}$$

To find the gradient of ψ we need to have ψ defined in a neighborhood of the surface $r(u,v)$ such that (2.25) holds and that its gradient satisfies on $r(u,v)$

$$(\nabla\psi)(r(u,v)) = x-\kappa\,m-((x-\kappa\,m)\cdot\nu)\;\nu. \tag{2.26}$$

Notice that this implies $(\nabla\psi)(r(u,v)) \perp \nu$. To construct the function ψ in a neighborhood of the surface Γ (we will construct it in a neighborhood of each point in Γ), given parametrically by $r(u,v)$, we use the notion of envelope from classical differential geometry; see for example [33, Chapter 5, Section 4] or [7, Chapter 3]. Since the required ψ must satisfy (2.25), consider the surface Γ' given parametrically by

$$P(u,v) = (r(u,v), |r(u,v)| - \kappa\,(m(u,v)\cdot r(u,v)) + \Phi(u,v)) \tag{2.27}$$

in four dimensions. At each point $P(u,v)$, consider the 4-dimensional vector

$$N(u,v) = (x - \kappa\, m - ((x - \kappa\, m)\cdot \nu)\ \nu, -1)\,,$$

where $x = x(u,v)$ and ν is the unit normal to the surface Γ at $r(u,v)$. Next consider the plane Π_{uv} passing through the point $P(u,v)$ and with normal $N(u,v)$, that is, in coordinates x_1, x_2, x_3, x_4, Π_{uv} has equation

$$F(x_1,x_2,x_3,x_4,u,v) := N(u,v)\cdot((x_1,x_2,x_3,x_4) - P(u,v)) = 0. \tag{2.28}$$

Therefore, we have a family of planes Π_{uv} depending on the parameters u, v, and we will let $x_4 = \psi(x_1,x_2,x_3)$ be by definition the envelope to this family of planes. Of course, we need to know under what conditions on $r(u,v)$ and $m(u,v)$ this envelope ψ exists. It will be defined by solving the system of equations

$$\begin{cases} F(x_1,x_2,x_3,x_4,u,v) &= 0 \\ \dfrac{\partial F}{\partial u}(x_1,x_2,x_3,x_4,u,v) &= 0 \\ \dfrac{\partial F}{\partial v}(x_1,x_2,x_3,x_4,u,v) &= 0. \end{cases} \tag{2.29}$$

In fact, let us fix values $u = u_0, v = v_0$, and let $P_0 = P(u_0, v_0) = (p_1, p_2, p_3, p_4)$ be the corresponding value on the surface Γ'; and consider the map

$$G(x_1, x_2, x_3, x_4, u, v) = \left(F(x_1, x_2, x_3, x_4, u, v), \frac{\partial F}{\partial u}(x_1, x_2, x_3, x_4, u, v), \frac{\partial F}{\partial v}(x_1, x_2, x_3, x_4, u, v) \right).$$

The function G has continuous partial derivatives in a neighborhood of the point $(p_1, p_2, p_3, p_4, u_0, v_0)$, and

$$G(p_1, p_2, p_3, p_4, u_0, v_0) = 0.$$

By the implicit function theorem, if the Jacobian determinant

$$\frac{\partial G}{\partial(x_4, u, v)}(p_1, p_2, p_3, p_4, u_0, v_0) = \det \left. \begin{pmatrix} \frac{\partial F}{\partial x_4} & \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial^2 F}{\partial x_4 \partial u} & \frac{\partial^2 F}{\partial u \partial u} & \frac{\partial^2 F}{\partial v \partial u} \\ \frac{\partial^2 F}{\partial x_4 \partial v} & \frac{\partial^2 F}{\partial u \partial v} & \frac{\partial^2 F}{\partial v \partial v} \end{pmatrix} \right|_{(p_1, p_2, p_3, p_4, u_0, v_0)} \neq 0, \tag{2.30}$$

then there are unique differentiable functions g_1, g_2, g_3 in the variables x_1, x_2, x_3 defined in a neighborhood U of (p_1, p_2, p_3) such that $p_4 = g_1(p_1, p_2, p_3)$, $u_0 = g_2(p_1, p_2, p_3)$ and $v_0 = g_3(p_1, p_2, p_3)$ with

$$G(x_1, x_2, x_3, g_1(x_1, x_2, x_3), g_2(x_1, x_2, x_3), g_3(x_1, x_2, x_3)) = 0$$

for all $(x_1, x_2, x_3) \in U$. Therefore, if we let $\psi(x_1, x_2, x_3) = g_1(x_1, x_2, x_3)$ for $(x_1, x_2, x_3) \in U$, then ψ is the function we need, i.e., ψ is by construction defined in a neighborhood of the point $(p_1, p_2, p_3) \in \Gamma$ and satisfies (2.25) and (2.26). We now analyze under what conditions on the surface Γ and m, (2.30) holds. Notice first that since $\partial_{x_4} F = -1$, the matrix inside the determinant in (2.30) equals

$$\begin{pmatrix} 1 & \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ 0 & \frac{\partial^2 F}{\partial u \partial u} & \frac{\partial^2 F}{\partial v \partial u} \\ 0 & \frac{\partial^2 F}{\partial u \partial v} & \frac{\partial^2 F}{\partial v \partial v} \end{pmatrix},$$

and therefore, (2.30) means

$$\det \begin{pmatrix} \frac{\partial^2 F}{\partial u \partial u} & \frac{\partial^2 F}{\partial v \partial u} \\ \frac{\partial^2 F}{\partial u \partial v} & \frac{\partial^2 F}{\partial v \partial v} \end{pmatrix} \neq 0.$$

Let us find what this means in terms of the initial surface Γ and the field m. To simplify the notation let $X = (x_1, x_2, x_3, x_4)$, so we can write (2.28) as

$$F(X, u, v) = N(u, v) \cdot (X - P(u, v)) \,.$$

By calculation

$$\begin{cases} F_u &= N_u \cdot (X - P) - N \cdot P_u \\ F_{uu} &= N_{uu} \cdot (X - P) - 2\, N_u \cdot P_u - N \cdot P_{uu} \\ F_{uv} &= N_{uv} \cdot (X - P) - N_u \cdot P_v - N_v \cdot P_u - N \cdot P_{uv} \\ F_{vv} &= N_{vv} \cdot (X - P) - 2\, N_v \cdot P_v - N \cdot P_{vv}. \end{cases} \tag{2.31}$$

We first show that

$$N \cdot P_u = N \cdot P_v = 0. \tag{2.32}$$

Indeed, we have

$$P(u, v) = \rho(u, v)\ (x, 1 - \kappa\, m \cdot x) + (0, \Phi),$$

so

$$P_u = \rho_u\ (x, 1 - \kappa\, m \cdot x) + \rho\ (x_u, -\kappa\, m \cdot x_u - \kappa\, m_u \cdot x) + (0, \Phi_u) \tag{2.33}$$

$$P_v = \rho_v\ (x, 1 - \kappa\, m \cdot x) + \rho\ (x_v, -\kappa\, m \cdot x_v - \kappa\, m_v \cdot x) + (0, \Phi_v).$$

Hence

$$\begin{aligned} & N \cdot P_u = \\ & \{\rho_u\ (x, 1 - \kappa\, m \cdot x) + \rho\ (x_u, -\kappa\, m \cdot x_u - \kappa\, m_u \cdot x) + (0, \Phi_u)\} \cdot \\ & \qquad (x - \kappa\, m - [(x - \kappa\, m) \cdot \nu]\ \nu, -1) \\ & = (\rho_u\, x + \rho\, x_u) \cdot (x - \kappa\, m - [(x - \kappa\, m) \cdot \nu]\, \nu) - \rho_u\ (1 - \kappa\, m \cdot x) \\ & \qquad + \rho\, (\, \kappa\, m \cdot x_u + \kappa\, m_u \cdot x) - \Phi_u \\ & = \rho_u - \rho_u\, \kappa\, x \cdot m - \rho\, \kappa\, x_u \cdot m - \rho_u\ + \rho_u\, \kappa\, m \cdot x + \rho\, \kappa\, m \cdot x_u \\ & \qquad + \rho\, \kappa\, m_u \cdot x - \rho\, \kappa\, m_u \cdot x \\ & = 0, \end{aligned}$$

since $(\rho_u\, x + \rho\, x_u) \cdot \nu = r_u \cdot \nu = 0$ and $x_u \cdot x = 0$. The same calculation with P_v instead of P_u yields the second identity in (2.32).

Next, differentiating (2.32) with respect to u and v yields

$$N \cdot P_{uu} = -N_u \cdot P_u, \qquad N \cdot P_{uv} = -N_u \cdot P_v = -N_v \cdot P_u, \qquad N \cdot P_{vv} = -N_v \cdot P_v,$$

since $P_{uv} = P_{vu}$. Hence letting $X = P$ in (2.33) yields

$$F_{uu} = -N_u \cdot P_u, \qquad F_{uv} = -N_v \cdot P_u = -N_u \cdot P_v, \qquad F_{vv} = -N_v \cdot P_v.$$

Now let us calculate these dot products. First set

$$B = (x - \kappa\, m) \cdot \nu,$$

and write

$$\begin{aligned}
& N_u \cdot P_u \\
& = \{\rho_u\ (x, 1 - \kappa\, m \cdot x) + \rho\ (x_u, -\kappa\, m \cdot x_u) + (0, \Phi_u)\} \cdot \\
& \qquad \{x_u - \kappa\, m_u - [(x - \kappa\, m) \cdot \nu]_u\ \nu - [(x - \kappa\, m) \cdot \nu]\ \nu_u, 0\} \\
& = (\rho_u\, x + \rho\, x_u) \cdot (x_u - \kappa\, m_u - B_u\, \nu - B\, \nu_u) \\
& = (\rho_u\, x + \rho\, x_u) \cdot x_u - \kappa\ (\rho_u\, x + \rho\, x_u) \cdot m_u - B_u\ (\rho_u\, x + \rho\, x_u) \cdot \nu \\
& \qquad - B\ (\rho_u\, x + \rho\, x_u) \cdot \nu_u \\
& = \left(\sin^2 v\right)\ \rho - \kappa\ (\rho_u\, x + \rho\, x_u) \cdot m_u - B\ (\rho_u\, x + \rho\, x_u) \cdot \nu_u \\
& = \left(\sin^2 v\right)\ \rho - \kappa\, r_u \cdot m_u - B\, r_u \cdot \nu_u,
\end{aligned}$$

since $x \cdot x_u = 0, x_u \cdot x_u = \sin^2 v$, and $(\rho_u\, x + \rho\, x_u) \cdot \nu = r_u \cdot \nu = 0$. Also $x_v \cdot x_v = 1$ and $x_u \cdot x_v = 0$, so we obtain similarly

$$N_v \cdot P_v = \rho - \kappa\, r_v \cdot m_v - B\, r_v \cdot \nu_v, \qquad N_u \cdot P_v = -\kappa\, r_u \cdot m_v - B\, r_u \cdot \nu_v.$$

Next, differentiating $r_u \cdot \nu = r_v \cdot \nu = 0$ yields

$$r_u \cdot \nu_u = -r_{uu} \cdot \nu, \qquad r_u \cdot \nu_v = -r_{uv} \cdot \nu, \qquad r_v \cdot \nu_v = -r_{vv} \cdot \nu. \tag{2.34}$$

Therefore,

$$\begin{pmatrix} F_{uu} & F_{uv} \\ F_{vu} & F_{vv} \end{pmatrix} = \begin{pmatrix} -\left(\sin^2 v\right)\ \rho + \kappa\, r_u \cdot m_u - B\, r_{uu} \cdot \nu & \kappa\, r_u \cdot m_v - B\, r_{uv} \cdot \nu \\ \kappa\, r_u \cdot m_v - B\, r_{uv} \cdot \nu & -\rho + \kappa\, r_v \cdot m_v - B\, r_{vv} \cdot \nu \end{pmatrix}$$

$$= -\rho \begin{pmatrix} x_u \cdot x_u & x_u \cdot x_v \\ x_v \cdot x_u & x_v \cdot x_v \end{pmatrix} + \kappa \begin{pmatrix} r_u \cdot m_u & r_u \cdot m_v \\ r_v \cdot m_u & r_v \cdot m_v \end{pmatrix} - B \begin{pmatrix} r_{uu} \cdot \nu & r_{uv} \cdot \nu \\ r_{vu} \cdot \nu & r_{vv} \cdot \nu \end{pmatrix},$$

with $B = (x - \kappa\, m) \cdot \nu$. □

Remark 2.3.2 (Case when m is a constant vector)**.** If $m(u, v) = (m_1, m_2, m_3)$ is constant, then (2.19) is clearly satisfied by any Γ and, in condition (2.20), the second matrix on the right hand side is zero.

Remark 2.3.3. To illustrate the determinant condition in Theorem 2.3.1, let us consider the special case when Γ is a sphere centered at the origin, and m is a constant vector. We have $r(u, v) = R\, x(u, v)$, and $\nu = x(u, v)$. So $r_{uu} = Rx_{uu}$ and similarly for r_{vv} and r_{uv}. Also $B = 1 - \kappa\, m \cdot x$, $x_{uu} \cdot x = -\sin^2 v, x_{uv} \cdot x = 0$, and $x_{vv} \cdot x = -1$. Hence $r_{uu} \cdot x = -R \sin^2 v$, $r_{uv} \cdot x = 0$, and $r_{vv} \cdot x = -R$. Therefore, the determinant in (2.20) equals

$$R^2 \sin^2 v\, (1 - B)^2 = R^2 \kappa^2 \left(\sin^2 v\right) (m \cdot x)^2.$$

For example, if $m = (0, 0, 1)$, i.e., all rays are refracted vertically, then the determinant equals

$$R^2 \kappa^2 \, (\sin v \, \cos v)^2 = \frac{R^2 \kappa^2}{4} \sin^2(2v)$$

which is not zero as long as $v \neq \pi/2$ or zero. This shows also that for the sphere, the phase discontinuity ψ exists and can be obtained by solving the system of equations (2.29). Notice that in this case, a phase discontinuity ψ was calculated explicitly in Section 2.2.2 and given by (2.18).

Remark 2.3.4 (Case when Γ is off centered)**.** A case considered in [2, Section 3]: a sphere of radius R is centered at a point $(0,0,a)$ with $a > R$, and the authors claim there that it is not possible to find a phase discontinuity on such a sphere so that all rays from the origin are refracted into the vertical direction. We believe this claim is in error and in fact, with the method above, will show that for each unit $m = (m_1, m_2, m_3)$ with $m_3 > 0$, there is a phase discontinuity ψ defined in a neighborhood of such a sphere so that its gradient is tangential to the sphere and so that radiation from the origin is refracted into a fixed direction m, see Figure 2.3. In particular, when m is vertical, a phase discontinuity exists. By reversibility of

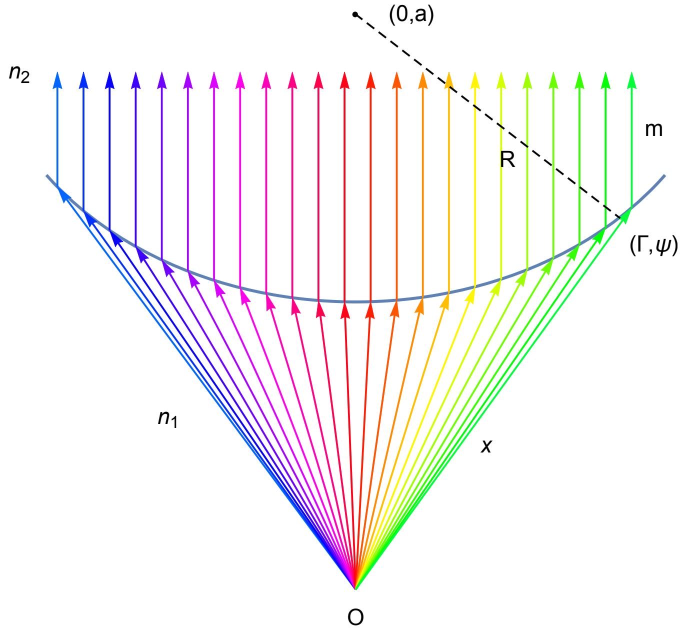

Figure 2.3: Off centered spherical metalens refracting into a fixed direction

optical paths, this shows that the conclusion in [2, Section 3] is incorrect.

First, the lower part of the sphere with center at $(0,0,a)$ and radius R is parametrized by the vector $r(u,v) = \rho(u,v)\, x(u,v)$ with

$$\rho(u,v) = a\,\cos v - \sqrt{R^2 - a^2\,\sin^2 v},$$

where $0 \leq v \leq \arcsin(R/a)$; and the unit normal to the sphere pointing upwards is

$$\nu = \frac{(0,0,a) - \rho(u,v)\, x(u,v)}{R}.$$

To show our claim, we need to verify that the determinant in (2.20) is not zero. From (2.34), we obtain by simple calculations that

$$\begin{aligned} r_{uu}\cdot\nu &= -r_u\cdot\nu_u = \frac{1}{R}\left(\sin^2 v\right)\,\rho^2 \\ r_{uv}\cdot\nu &= -r_u\cdot\nu_v = \frac{1}{R}\,\rho_u\,\rho_v = 0 \\ r_{vv}\cdot\nu &= -r_v\cdot\nu_v = \frac{1}{R}\,\left((\rho_v)^2 + \rho^2\right). \end{aligned}$$

Therefore, the determinant in (2.20) equals

$$\det\begin{pmatrix} F_{uu} & F_{uv} \\ F_{vu} & F_{vv} \end{pmatrix} = \rho\,(\sin^2 v)\,\left(1 + \frac{B}{R}\,\rho\right)\,\left(\rho + \frac{B}{R}\left(\rho^2 + (\rho_v)^2\right)\right), \tag{2.35}$$

with

$$\begin{aligned} B = (x - \kappa\, m)\cdot\nu &= \frac{1}{R}\,(x - \kappa\, m)\cdot((0,0,a) - \rho\, x) \\ &= \frac{1}{R}\left(\sqrt{R^2 - a^2\,\sin^2 v} - \kappa\, a\, m_3 + \kappa\,\rho\,(m\cdot x)\right). \end{aligned}$$

The last determinant is not zero for u, v such that

$$\sin^2 v \neq 0, \quad 1 + \frac{B}{R}\,\rho \neq 0, \quad \text{and } \rho + \frac{B}{R}\left(\rho^2 + (\rho_v)^2\right) \neq 0.$$

Let us take, for example, $m = (0,0,1)$, i.e., rays are refracted vertically, then we get

$$B = \frac{1}{R}\left((1-\kappa\,\cos v)\sqrt{R^2 - a^2\,\sin^2 v} - \kappa\, a\,\sin^2 v\right),$$

so B is independent of u. If $v \approx 0$, then $B \approx 1-\kappa$, $\rho \approx a - R$ and $\rho_v \approx 0$, so

$$1 + \frac{B}{R}\rho \approx 1 + (1-\kappa)\left(\frac{a}{R} - 1\right)$$
$$\rho + \frac{B}{R}\left(\rho^2 + (\rho_v)^2\right) \approx (a-R)\left(1 + (1-\kappa)\left(\frac{a}{R} - 1\right)\right).$$

Recall that $\kappa = n_2/n_1$. If $\kappa < 1$, since $a > R$, we obtain that $1+(1-\kappa)\left(\frac{a}{R} - 1\right) \neq 0$. If $\kappa > 1$, then $1+(1-\kappa)\left(\frac{a}{R} - 1\right) \neq 0$ if and only if $\kappa \neq 1 + \frac{R}{a-R}$. This shows that in these cases, the determinant in (2.35) is not zero for $v \neq 0$ with v close to zero. Therefore, there exists a phase discontinuity ψ, on the sphere centered at $(0,0,a)$ with radius R, defined in a neighborhood of each point of the form $\rho(u,v)\,x(u,v)$ with v close to zero.

2.4 Given a phase discontinuity, find an admissible surface

We now turn to the second question proposed at the beginning of Section 2.2, that is, of finding the surface Γ when the field $V = (V_1, V_2, V_3)$ is given. The unknown surface is given parametrically by

$$r(u,v) = \rho(u,v)\,x(u,v),$$

where $x(u,v)$ are spherical coordinates as before, m is a constant vector, and we seek the polar radius ρ; the value of V along the surface is $V(r(u,v))$. The following theorem gives a sufficient condition for the existence of Γ.

Theorem 2.4.1. *If the field* $V(r(u,v)) = \nabla\psi(r(u,v))$ *for some function* ψ*, and*

$$x \cdot (Ax \times W) = 0, \qquad (2.36)$$

[3]

holds in an open set O *in the variables* (ρ, u, v)*, where* $W(\rho,u,v) = \kappa\, m + V(\rho\, x(u,v))$ *and* $A(\rho,u,v) = \nabla^2\psi(\rho\, x(u,v))$*, then for each* $(\rho_0,u_0,v_0) \in O$ *with* $x(u_0,v_0)\cdot W(\rho_0,u_0,v_0) \neq 1$*, the system* (2.37) *has a unique solution* $\rho(u,v)$ *defined in a neighborhood of* (u_0,v_0) *and satisfying the initial condition* $\rho(u_0,v_0) = \rho_0$*.*

Proof. From the generalized Snell law (2.12), $x(u,v) - \kappa\, m - V(r(u,v))$ is a multiple of the normal ν at $r(u,v)$, so

$$r_u(u,v)\cdot(x(u,v) - \kappa\, m - V(r(u,v))) = 0 \quad \text{and}$$

$$r_v(u,v)\cdot(x(u,v) - \kappa\, m - V(r(u,v))) = 0.$$

We have

$$r_u(u,v) = [\rho(u,v)]_u\, x(u,v) + \rho(u,v)\, x_u(u,v),$$

$$r_v(u,v) = [\rho(u,v)]_v\, x(u,v) + \rho(u,v)\, x_v(u,v),$$

so

$$\begin{aligned} 0 &= r_u(u,v)\cdot(x(u,v) - \kappa\, m - V(r(u,v))) \\ &= ([\rho(u,v)]_u\, x(u,v) + \rho(u,v)\, x_u(u,v))\cdot(x(u,v) - \kappa\, m - V(r(u,v))) \\ &= [\rho(u,v)]_u\,(1 - x(u,v)\cdot[\kappa\, m + V(r(u,v))]) - \rho(u,v)\, x_u(u,v)\cdot[\kappa\, m + V(r(u,v))]\,, \end{aligned}$$

[3]Equivalently $W \cdot (Ax \times x) = Ax \cdot (x \times W) = 0$.

and a similar equation for r_v. That is, $\rho(u,v)$ satisfies the first order nonlinear system of pdes (depending on V)[4]:

$$\begin{cases} \rho_u(u,v) - \dfrac{x_u \cdot [\kappa\, m + V\left(\rho(u,v)\, x(u,v)\right)]}{1 - x(u,v)\cdot[\kappa\, m + V\left(\rho(u,v)\, x(u,v)\right)]}\,\rho(u,v) = 0 \\ \\ \rho_v(u,v) - \dfrac{x_v \cdot [\kappa\, m + V\left(\rho(u,v)\, x(u,v)\right)]}{1 - x(u,v)\cdot[\kappa\, m + V\left(\rho(u,v)\, x(u,v)\right)]}\,\rho(u,v) = 0. \end{cases} \tag{2.37}$$

If $F = (F_1, F_2)$ with

$$F_1(u,v,\rho) = \frac{x_u \cdot [\kappa\, m + V\left(\rho\, x(u,v)\right)]}{1 - x(u,v)\cdot[\kappa\, m + V\left(\rho\, x(u,v)\right)]}\,\rho$$
$$F_2(u,v,\rho) = \frac{x_v \cdot [\kappa\, m + V\left(\rho\, x(u,v)\right)]}{1 - x(u,v)\cdot[\kappa\, m + V\left(\rho\, x(u,v)\right)]}\,\rho,$$

then (2.37) can be written as

$$\nabla\rho = F(u,v,\rho). \tag{2.38}$$

To solve the system (2.38) we need an initial condition, say $\rho(u_0,v_0) = \rho_0$, and use the result from [23, Chapter 6, pp. 117-118] as in the previous section. We will see under what circumstances on the field V, F satisfies condition (2.23), and therefore, the existence of the desired surface $r(u,v)$ will be guaranteed. Set

$$W(u,v,\rho) = \kappa\, m + V\left(\rho\, x(u,v)\right), \tag{2.39}$$

then

$$F_1(u,v,\rho) = \frac{x_u \cdot W(u,v,\rho)}{1 - x(u,v)\cdot[W(u,v,\rho)]}\,\rho,$$
$$F_2(u,v,\rho) = \frac{x_v \cdot W(u,v,\rho)}{1 - x(u,v)\cdot[W(u,v,\rho)]}\,\rho.$$

[4] We are assuming that $1 - x(u,v)\cdot[\kappa\, m + V\left(\rho(u,v)\, x(u,v)\right)] \neq 0$.

We have

$$\frac{\partial F_1}{\partial v} = (x_{uv} \cdot W + x_u \cdot W_v)\,(1 - x \cdot W)^{-1}\rho$$

$$+ (x_v \cdot W + x \cdot W_v)\,(x_u \cdot W)\,(1 - x \cdot W)^{-2}\rho$$

$$\frac{\partial F_2}{\partial u} = (x_{vu} \cdot W + x_v \cdot W_u)\,(1 - x \cdot W)^{-1}\rho$$

$$+ (x_u \cdot W + x \cdot W_u)\,(x_v \cdot W)\,(1 - x \cdot W)^{-2}\rho$$

$$\frac{\partial F_1}{\partial \rho} = (x_u \cdot W)\,(1 - x \cdot W)^{-1}$$

$$+ \left\{(x_u \cdot W_\rho)\,(1 - x \cdot W)^{-1} + (x_u \cdot W)\,(x \cdot W_\rho)\,(1 - x \cdot W)^{-2}\right\}\,\rho$$

$$\frac{\partial F_2}{\partial \rho} = (x_v \cdot W)\,(1 - x \cdot W)^{-1}$$

$$+ \left\{(x_v \cdot W_\rho)\,(1 - x \cdot W)^{-1} + (x_v \cdot W)\,(x \cdot W_\rho)\,(1 - x \cdot W)^{-2}\right\}\,\rho.$$

Hence,

$$\frac{\partial F_1}{\partial v} - \frac{\partial F_2}{\partial u} = \Big\{\,(x_u \cdot W_v - x_v \cdot W_u)\,(1 - x \cdot W)^{-1}$$

$$+ \left[(x \cdot W_v)\,(x_u \cdot W) - (x \cdot W_u)\,(x_v \cdot W)\right]\,(1 - x \cdot W)^{-2}\Big\}\,\rho$$

and

$$\frac{\partial F_1}{\partial \rho}F_2 - \frac{\partial F_2}{\partial \rho}F_1 = \Big[(x_u \cdot W)\,(1 - x \cdot W)^{-1} + \Big\{\,(x_u \cdot W_\rho)\,(1 - x \cdot W)^{-1}$$

$$+ (x_u \cdot W)\,(x \cdot W_\rho)\,(1 - x \cdot W)^{-2}\Big\}\,\rho\Big]\,(x_v \cdot W)\,(1 - x \cdot W)^{-1}$$

$$- \Big[(x_v \cdot W)\,(1 - x \cdot W)^{-1} + \Big\{\,(x_v \cdot W_\rho)\,(1 - x \cdot W)^{-1}$$

$$+ (x_v \cdot W)\,(x \cdot W_\rho)\,(1 - x \cdot W)^{-2}\Big\}\,\rho\Big](x_u \cdot W)\,(1 - x \cdot W)^{-1}$$

$$= \left[(x_u \cdot W_\rho)\,(x_v \cdot W) - (x_v \cdot W_\rho)\,(x_u \cdot W)\right]\,(1 - x \cdot W)^{-2}\,\rho.$$

Therefore, (2.23) holds if

$$
\begin{aligned}
&\frac{\partial F_1}{\partial v} - \frac{\partial F_2}{\partial u} + \frac{\partial F_1}{\partial \rho} F_2 - \frac{\partial F_2}{\partial \rho} F_1 \\
&= \Big\{ (x_u \cdot W_v - x_v \cdot W_u)(1 - x \cdot W)^{-1} + \Big((x \cdot W_v)(x_u \cdot W) \\
&- (x \cdot W_u)(x_v \cdot W) \Big) (1 - x \cdot W)^{-2} \Big\} \rho \\
&+ [(x_u \cdot W_\rho)(x_v \cdot W) - (x_v \cdot W_\rho)(x_u \cdot W)] \; (1 - x \cdot W)^{-2} \rho \\
&= 0.
\end{aligned}
$$

Since we assume $1 - x \cdot W \neq 0$ and $\rho > 0$, this is equivalent to

$$
\begin{aligned}
&(x_u \cdot W_v - x_v \cdot W_u)(1 - x \cdot W) + ((x \cdot W_v)(x_u \cdot W) - (x \cdot W_u)(x_v \cdot W)) \\
&\quad + ((x_u \cdot W_\rho)(x_v \cdot W) - (x_v \cdot W_\rho)(x_u \cdot W)) = 0,
\end{aligned}
$$

that is,

$$
\begin{aligned}
&(x_u \cdot W_v - x_v \cdot W_u)(1 - x \cdot W) \\
&+ \Big\{ [(x \cdot W_v) - (x_v \cdot W_\rho)](x_u \cdot W) - [(x \cdot W_u) - (x_u \cdot W_\rho)] \; (x_v \cdot W) \Big\} = 0.
\end{aligned}
\tag{2.40}
$$

We have

$$
\begin{aligned}
W_u &= \rho \; (\nabla V_1 \cdot x_u, \nabla V_2 \cdot x_u, \nabla V_3 \cdot x_u) \\
W_v &= \rho \; (\nabla V_1 \cdot x_v, \nabla V_2 \cdot x_v, \nabla V_3 \cdot x_v) \\
W_\rho &= (\nabla V_1 \cdot x, \nabla V_2 \cdot x, \nabla V_3 \cdot x) \, .
\end{aligned}
$$

Now

$$
x \cdot W_v = \rho \sum_{k=1}^{3} x_k \, (\nabla V_k \cdot x_v) = \rho \sum_{k=1}^{3} x_k \, \sum_{j=1}^{3} \frac{\partial V_k}{\partial y_j} (x_j)_v = \rho \sum_{k,j=1}^{3} \frac{\partial V_k}{\partial y_j} (x_j)_v \, x_k.
$$

If we let

$$A = \begin{pmatrix} \frac{\partial V_1}{\partial y_1} & \frac{\partial V_1}{\partial y_2} & \frac{\partial V_1}{\partial y_3} \\ \frac{\partial V_2}{\partial y_1} & \frac{\partial V_2}{\partial y_2} & \frac{\partial V_2}{\partial y_3} \\ \frac{\partial V_3}{\partial y_1} & \frac{\partial V_3}{\partial y_2} & \frac{\partial V_3}{\partial y_3} \end{pmatrix},$$

then

$$x \cdot W_v = \rho\, x\, A\, (x_v)^t,$$

where x, x_v are row vectors and t denotes the transpose. Similarly,

$$x \cdot W_u = \rho\, x\, A\, (x_u)^t \quad x_u \cdot W_v = \rho\, x_u\, A\, (x_v)^t$$

$$x_v \cdot W_u = \rho\, x_v\, A\, (x_u)^t \quad x_u \cdot W_\rho = x_u\, A\, (x)^t \quad x_v \cdot W_\rho = x_v\, A\, (x)^t.$$

Since by assumption $V = \nabla\psi$, then $A = \nabla^2\psi$ is a symmetric matrix, so

$$x_u \cdot W_v = x_v \cdot W_u$$

$$(x \cdot W_v) - (x_v \cdot W_\rho) = (\rho - 1)\, x\, A\, (x_v)^t = \frac{\rho - 1}{\rho}\, (x \cdot W_v)$$

$$(x \cdot W_u) - (x_u \cdot W_\rho) = (\rho - 1)\, x\, A\, (x_u)^t = \frac{\rho - 1}{\rho}\, (x \cdot W_u)$$

and (2.40) reads

$$(\rho - 1)\, \left\{ (x\, A\, (x_v)^t)\, (x_u \cdot W) - (x\, A\, (x_u)^t)\, (x_v \cdot W) \right\} = 0, \qquad (2.41)$$[5]

which can be written as

$$\det \begin{pmatrix} x\, A\, (x_u)^t & x\, A\, (x_v)^t \\ x_u \cdot W & x_v \cdot W \end{pmatrix} = \det \begin{pmatrix} x_u \cdot Ax & x_v \cdot Ax \\ x_u \cdot W & x_v \cdot W \end{pmatrix} = 0.$$

[5]Since $(x \cdot W)_u = x_u \cdot W + \rho\, x\, A\, (x_u)^t$ and similarly for $(x \cdot W)_v$, this condition can be re-written as $(\rho - 1) \left\{ x\, A\, (x_v)^t\, (x \cdot W)_u - x\, A\, (x_u)^t\, (x \cdot W)_v \right\} = 0$.

From the Cauchy-Binet formula for cross products[6], this means that

$$(x_u \times x_v) \cdot (Ax \times W) = 0,$$

and since $x_u \times x_v \parallel x$, (2.41) is equivalent to 2.36:

$$x \cdot (Ax \times W) = 0.$$

□

Notice that if $V = V_0$ is a constant field, then $A = 0$, and so (2.41) obviously holds. In this case, (2.37) can be easily integrated, and the solution is

$$\rho(u, v) = \frac{C_1}{1 - x(u, v) \cdot (\kappa m + V_0)} + C_2$$

with C_i constants.

Notice also that with the choice V, as in (2.15), with $h \neq 0$ so $1 - x \cdot W \neq 0$, the system of equations (2.37) becomes

$$\begin{cases} \rho_u(u, v) - \dfrac{\sin u}{\cos u}\,\rho(u, v) = 0 \\ \\ \rho_v(u, v) + \dfrac{\cos v}{\sin v}\,\rho(u, v) = 0, \end{cases}$$

whose solution is $\rho(u, v) = \dfrac{C}{\cos u \, \sin v}$, where the constant C is determined by the point where the solution passes through. This is in agreement with (2.13).

[6] $(a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c)$.

2.5 Near field refracting metasurfaces

The near field case can be regarded as a special case from Section 2.3, when the vector field $m(u, v)$ points towards a fixed point Q, and therefore, the method from that section can be used to derive conditions for the existence of the desired metasurface. In fact, if the surface Γ is parametrized by $r(u, v)$ and $m(u, v) = \dfrac{Q - r(u, v)}{|Q - r(u, v)|}$, then it is easy to see that the compatibility condition (2.19) holds. The existence of the phase discontinuity then follows when the determinant in (2.20) is not zero.

However, the phase discontinuities in the planar and spherical cases can be obtained explicitly as follows; see Figure 2.4 and Figure 2.5.

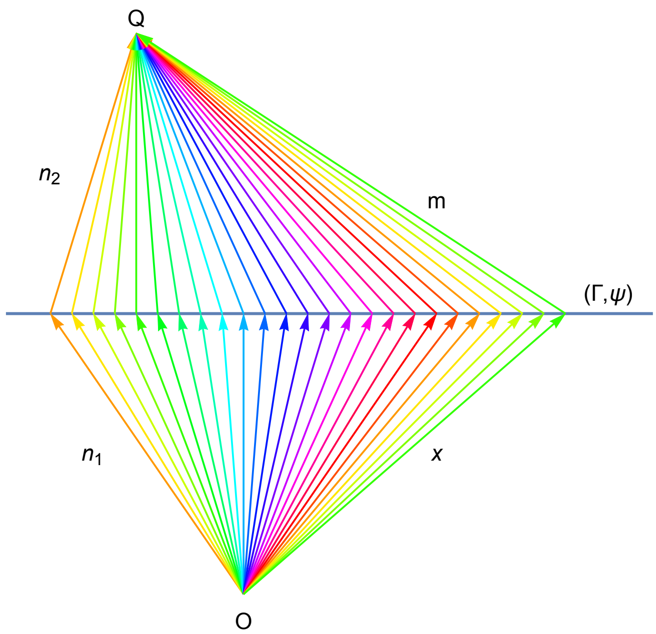

Figure 2.4: Planar metalens in the near field

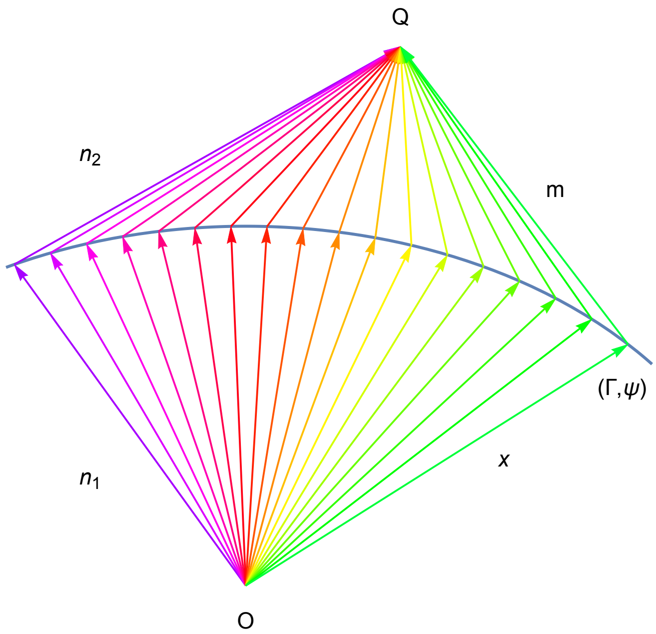

Figure 2.5: Spherical metalens in the near field

2.5.1 Case of a plane interface

Let O be the origin in medium I with index n_1, and let $Q = (q_1, q_2, q_3)$ be a point in medium II with index n_2. Denote by Γ the plane with equation $x_1 = a$ so that it separates the points O and Q. We find the field V so that rays from O are refracted into Q. We know from Section 2.2.1 that Γ is given parametrically by (2.13); the normal $\nu = (1, 0, 0)$. So we seek V such that (2.12) holds. Since the refracted vector from each point $r(u, v)$ on the plane interface to the point Q has unit direction $\dfrac{Q - r(u,v)}{|Q - r(u,v)|}$, V must satisfy

$$\begin{aligned}
\cos u \, \sin v - \kappa \, \frac{q_1 - a}{|Q - r(u,v)|} &= \lambda + V_1 \\
\sin u \, \sin v - \kappa \, \frac{q_2 - a \, \tan u}{|Q - r(u,v)|} &= V_2 \\
\cos v - \kappa \, \frac{q_3 - a/\cos u \tan v}{|Q - r(u,v)|} &= V_3.
\end{aligned}$$

Rewriting these equations in rectangular coordinates yields

$$\begin{aligned}
\frac{a}{\sqrt{a^2 + x_2^2 + x_3^2}} - \kappa \, \left. \frac{q_1 - a}{|Q - (x_1, x_2, x_3)|} \right|_{x_1 = a} &= \lambda + V_1 \\
\frac{x_2}{\sqrt{a^2 + x_2^2 + x_3^2}} - \kappa \, \left. \frac{q_2 - x_2}{|Q - (x_1, x_2, x_3)|} \right|_{x_1 = a} &= V_2 \\
\frac{x_3}{\sqrt{a^2 + x_2^2 + x_3^2}} - \kappa \, \left. \frac{q_3 - x_3}{|Q - (x_1, x_2, x_3)|} \right|_{x_1 = a} &= V_3.
\end{aligned}$$

Therefore, V_i, $i = 1, 2, 3$, are determined:

$$\begin{aligned}
V_1(a, x_2, x_3) &= \partial_{x_1} \left(\sqrt{x_1^2 + x_2^2 + x_3^2} \right) \Big|_{x_1 = a} + \kappa \, \frac{\partial}{\partial x_1} |Q - (x_1, x_2, x_3)| \Big|_{x_1 = a} - \lambda \\
V_2(a, x_2, x_3) &= \partial_{x_2} \left(\sqrt{x_1^2 + x_2^2 + x_3^2} \right) \Big|_{x_1 = a} + \kappa \, \frac{\partial}{\partial x_2} |Q - (x_1, x_2, x_3)| \Big|_{x_1 = a} \\
V_3(a, x_2, x_3) &= \partial_{x_3} \left(\sqrt{x_1^2 + x_2^2 + x_3^2} \right) \Big|_{x_1 = a} + \kappa \, \frac{\partial}{\partial x_3} |Q - (x_1, x_2, x_3)| \Big|_{x_1 = a},
\end{aligned}$$

where λ is chosen arbitrarily. Notice that if we let

$$\psi(x_1, x_2, x_3) = \sqrt{x_1^2 + x_2^2 + x_3^2} + \kappa\,|Q - (x_1, x_2, x_3)|$$

and choose $\lambda = 0$, then $V = \nabla\psi$, and so the plane with the phase discontinuity function ψ does the desired refraction job.

2.5.2 Case of a spherical interface

If Γ is the sphere of radius R centered at the origin, that is, $r(u,v) = R\,x(u,v)$, then the normal $\nu = x$, and from (2.12), we get

$$\left(x - \kappa\,\frac{Q - r(u,v)}{|Q - r(u,v)|} - V\right) \times x = 0.$$

As before, taking cross product with x yields

$$V + \kappa\,\frac{Q - r(u,v)}{|Q - r(u,v)|} - \left(\kappa\left(\frac{Q - r(u,v)}{|Q - r(u,v)|} \cdot x\right) + V \cdot x\right) x = 0.$$

Assuming V is tangential to the sphere,

$$V = -\kappa\,\frac{Q - r(u,v)}{|Q - r(u,v)|} + \kappa\left(\frac{Q - r(u,v)}{|Q - r(u,v)|} \cdot x\right) x.$$

If $V(Rx(u,v)) = (\nabla\psi)(Rx(u,v))$, then

$$\psi_{x_j}(Rx(u,v)) = -\kappa\,\frac{q_j - Rx_j(u,v)}{|Q - Rx(u,v)|} + \kappa\left(\frac{Q - Rx(u,v)}{|Q - Rx(u,v)|} \cdot x\right) x_j, \qquad j = 1, 2, 3. \tag{2.42}$$

Hence,

$$\frac{\partial}{\partial u}\left(\psi(Rx(u,v))\right) = (\nabla\psi)(Rx(u,v)) \cdot Rx_u = -\kappa\,R\,\frac{Q - Rx(u,v)}{|Q - Rx(u,v)|} \cdot x_u, \tag{2.43}$$

and similarly,

$$\frac{\partial}{\partial v}\left(\psi(Rx(u,v))\right) = -\kappa\, R\, \frac{Q - Rx(u,v)}{|Q - Rx(u,v)|} \cdot x_v, \tag{2.44}$$

since $x \cdot x_u = x \cdot x_v = 0$. Since ψ is assumed C^2, we get

$$\left(\frac{Q - Rx(u,v)}{|Q - Rx(u,v)|}\right)_u \cdot x_v = \left(\frac{Q - Rx(u,v)}{|Q - Rx(u,v)|}\right)_v \cdot x_u. \tag{2.45}$$

Integrating (2.43) in u yields

$$\psi(Rx(u,v)) = -\kappa\, R \int \frac{Q - Rx(u',v)}{|Q - Rx(u',v)|} \cdot x_u(u',v)\, du' + h(v),$$

for some function h. To calculate h, we differentiate the integral with respect to v and use (2.45):

$$\begin{aligned}
\frac{\partial}{\partial v}\left(\psi(Rx(u,v))\right) &= -\kappa\, R \int \frac{\partial}{\partial v}\left(\frac{Q - Rx(u',v)}{|Q - Rx(u',v)|} \cdot x_u(u',v)\right) du' + h'(v) \\
&= -\kappa\, R \int \left\{\frac{\partial}{\partial v}\left(\frac{Q - Rx(u',v)}{|Q - Rx(u',v)|}\right) \cdot x_u(u',v)\right. \\
&\quad \left. + \frac{Q - Rx(u',v)}{|Q - Rx(u',v)|} \cdot x_{uv}(u',v)\right\} du' + h'(v) \\
&= -\kappa\, R \int \left\{\frac{\partial}{\partial u}\left(\frac{Q - Rx(u',v)}{|Q - Rx(u',v)|}\right) \cdot x_v(u',v)\right. \\
&\quad \left. + \frac{Q - Rx(u',v)}{|Q - Rx(u',v)|} \cdot x_{vu}(u',v)\right\} du' + h'(v) \\
&= -\kappa\, R \int \frac{\partial}{\partial u}\left(\frac{Q - Rx(u',v)}{|Q - Rx(u',v)|} \cdot x_v(u',v)\right) du' + h'(v) \\
&= -\kappa\, R \left(\frac{Q - Rx(u,v)}{|Q - Rx(u,v)|} \cdot x_v(u,v)\right) + h'(v)
\end{aligned}$$

which implies $h'(v) = 0$ from (2.44). Therefore, the phase discontinuity ψ on the sphere satisfies

$$\psi(Rx(u,v)) = -\kappa\, R \int \frac{Q - Rx(u',v)}{|Q - Rx(u',v)|} \cdot x_u(u',v)\, du' + C$$

$$= \kappa \int \partial_u(|Q - Rx(u',v)|)\, du' + C = \kappa\, |Q - Rx(u,v)| + C,$$

with C a constant. Writing this in rectangular coordinates yields

$$\psi(R(z_1,z_2,z_3)) = \kappa\, |Q - R(z_1,z_2,z_3)| + C, \qquad \text{for } |(z_1,z_2,z_3)| = 1.$$

We now define ψ on a neighborhood of $|z| = R$ so that (2.42) holds. Let

$$\psi(z) = \kappa \left| Q - R\frac{z}{|z|} \right| + C, \text{ for } R - \epsilon < |z| < R + \epsilon. \tag{2.46}$$

We have

$$\nabla\psi(z) = -\kappa R \frac{Q - R\frac{z}{|z|}}{\left|Q - R\frac{z}{|z|}\right|} \frac{1}{|z|} + \kappa R \left(\frac{Q - R\frac{z}{|z|}}{\left|Q - R\frac{z}{|z|}\right|} \cdot \frac{z}{|z|} \right) \frac{z}{|z|^2},$$

so for $z = Rx$, with $|x| = 1$, we obtain

$$\nabla\psi(Rx) = -\kappa \frac{Q - Rx}{|Q - Rx|} + \kappa \left(\frac{Q - Rx}{|Q - Rx|} \cdot x \right) x$$

as desired. Therefore, the phase discontinuity ψ in (2.46) has gradient tangential to the sphere and can be placed on the spherical interface $|z| = R$ so that all rays from the origin are refracted into the point Q.

CHAPTER 3

DERIVATION OF PARTIAL DIFFERENTIAL EQUATIONS FOR METASURFACES

In this chapter we solve the following. Let $\Omega_1 \subset \mathbb{R}^2$, $\Omega_2 \subset S^2$, and Γ a surface given by the graph of the function $u : \Omega_1 \to \mathbb{R}^+$. We are given two intensities, i.e, two non negative functions, f defined in Ω_1 and g defined in Ω_2 satisfying the energy conservation condition

$$\int_{\Omega_1} f(x,y)\,dxdy = \int_{\Omega_2} g(z)\,d\sigma(z),$$

where $d\sigma(z)$ denotes as usual the element of area in the unit sphere of $\mathbb{R}^3$. A collimated beam is emanating from Ω_1. That is, for each $(x,y) \in \Omega_1$ a ray is emitted in the vertical direction $e_3 = (0,0,1)$ with intensity $f(x,y)$ and strikes the surface Γ at the point $(x,y,u(x,y)) = P$. According to the generalized law of reflection in Remark 2.1.2, this ray is reflected by the metasurface (Γ, ψ), into a ray having direction $T(x,y) = e_3 - \lambda\,\nu(P) - \nabla\psi(P)$, where $\nu(P)$ is the normal to Γ at P. The question is then to find a function ψ defined in Γ, called a phase discontinuity, such that the metalens, i.e., the pair (Γ, ψ) reflects all rays from Ω_1 into Ω_2, that is, $T(\Omega_1) = \Omega_2$, and the energy conservation balance

$$\int_E f(x,y)\,dxdy = \int_{T(E)} g(z)\,d\sigma(z) \tag{3.1}$$

holds for each subset E of Ω_1, see Figure 3.1. This problem is solved in Section 3.1.1 when Γ is an horizontal plane above the $x-y$ plane. We also solve similar problems when the incident rays emanate from a point source into a set of unit directions Ω_1, see Figure 3.2. Such a problem is solved in Section 3.1.2 when Γ is an horizontal plane.

In addition, we consider and solve similar problems for refraction using the generalized law (2.6), both in the collimated and point source cases when Γ is an horizontal plane above the $x-y$ plane, see Figures 3.3 and 3.4, Sections 3.2.1 and 3.2.2.

To do this we derive the partial differential equation, for each problem, satisfied by the phase discontinuity ψ and show it is a Monge-Ampère equation. Next we show that the resulting equations have solutions by application of a result by Urbas [40]. The equations corresponding to the four problems considered are (3.5), (3.16), (3.17) and (3.19), and they can be regarded as particular cases of (3.20). A summary of these equations can in found in Section 3.3.

Monge-Ampère equations appear naturally in optics for freeform lens design that have been the subject of recent research, see for example [41]-[21]. Therefore, it is natural that these type of equations appear also for metasurfaces. Monge-Ampère equations have been recently the subject of important mathematical research due to their connections with various topics such us optimal mass transport. We refer the reader to [22] and [10] for details and references therein.

We mention that using the ideas from [17], recent work for reflection is done in [4] to design graphene-based metasurfaces that can be actively tuned between different regimes of operation, such as anomalous beam steering and focusing, cloaking, and illusion optics, by applying electrostatic gating without modifying the geometry of the metasurface.

Finally, if the surface Γ is not necessarily a plane, then is possible to derive the corresponding partial differential equation that the phase discontinuity ψ satisfies, in both the reflection and refraction cases. These are equations of Monge-Ampère type that require a more complicated derivation carried out in Section 3.4 and Section 3.5. Existence of solutions to these equations will be considered in future works.

3.1 Reflection when Γ is a plane

3.1.1 Collimated case

Here, we solve the first problem stated in the introduction. From Remark 2.1.2 with $n_1 = 1$, the vertical ray emanating from the point $(x, y) \in \Omega_1$ is reflected by the metasurface (Γ, ψ) into the unit direction

$$T(x, y) = i(x, y) - \lambda\nu(x, y) - \nabla\psi(x, y), \tag{3.2}$$

where $i(x, y) = (0, 0, 1)$, $\nu(x, y)$ is the normal to $\Gamma = \{z = 1\}$, and

$$\lambda = (i - \nabla\psi) \cdot \nu + \sqrt{1 - (|i - \nabla\psi|^2 - [(i - \nabla\psi) \cdot \nu]^2)} = 1 + \sqrt{1 - \psi_x^2 - \psi_y^2}.$$

We remark that, in the last identity we have used, $\psi_z = 0$ because we seek a phase discontinuity ψ tangential to the surface Γ.

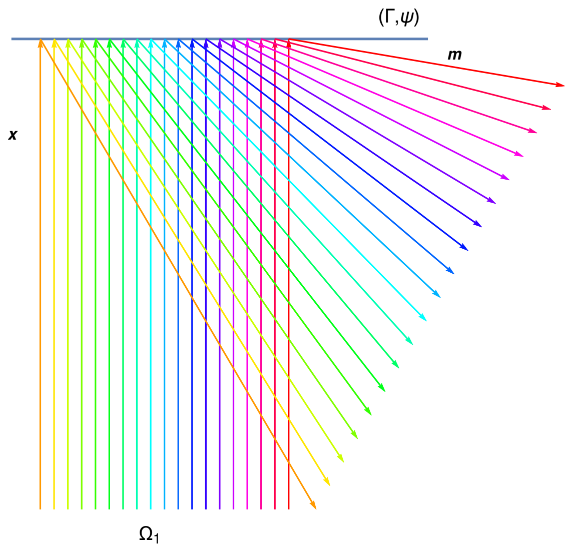

Figure 3.1: Reflection from an extended source (rays are monochromatic; colors are used only for visual purposes).

Therefore,

$$T(x, y) = (T_1, T_2, T_3) = -\left(\psi_x(x, y), \psi_y(x, y), \sqrt{1 - \psi_x^2(x, y) - \psi_y^2(x, y)}\right).$$

From the conservation of energy condition (3.1) and the formula of change of variables for surface integrals

$$\int_E f(x)\, dx = \int_{T(E)} g(y)\, d\sigma(y) = \int_E g(T(z))|J_T(z)|\, dz, \tag{3.3}$$

for each open set $E \subset \Omega_1$, and where $|J_T|= |T_x(x,y) \times T_y(x,y)|$. From (3.3), we obtain

$$f(x) = g(T(x,y))\,|J_T(x,y)| \qquad \text{for } (x,y) \in \Omega_1. \tag{3.4}$$

To calculate $|J_T(x,y)|$, because $|T(x,y)|= 1$, differentiating with respect to x and y yields the equations $T \cdot T_x = T \cdot T_y = 0$. Hence, assuming $T_3(x,y) \neq 0$ and solving these equations in $(T_3)_x$ and $(T_3)_y$, we obtain

$$(T_3)_x = -\frac{T_1(T_1)_x + T_2(T_2)_x}{T_3} \quad \text{and} \quad (T_3)_y = -\frac{T_1(T_1)_y + T_2(T_2)_y}{T_3}.$$

Using these two equations in the determinant defining the cross product $T_x \times T_y$, from an elementary calculation, we obtain

$$T_x \times T_y = \frac{1}{T_3} \det \begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} T.$$

Hence,

$$|J_T|= \frac{1}{|T_3(x,y)|}|\det(D^2\psi)|,$$

where $D^2\psi$ is the matrix of the second derivatives in x and y. Therefore, from (3.4) the phase discontinuity ψ satisfies the following Monge-Ampère equation

$$\frac{1}{\sqrt{1-\psi_x^2(x,y)-\psi_y^2(x,y)}}|\det(D^2\psi)|= \frac{f(x,y)}{g\,(T(x,y))}. \tag{3.5}$$

To show that (3.5) has solutions, we invoke [40, Theorem 2], which says the following:

Theorem 3.1.1. *Let Ω_1 and Ω^* be uniformly convex domains in $\mathbb{R}^n$ with $\partial\Omega_1$, $\partial\Omega^* \in C^{2,1}$ and let $f_1 \in C^{1,1}(\bar{\Omega}_1)$, $f_2 \in C^{1,1}(\bar{\Omega}^*)$ be positive functions satisfying*

$$\int_{\Omega_1} f_1(x)\,dx = \int_{\Omega^*} f_2(p)\,dp. \tag{3.6}$$

Then the boundary value problem

$$\det(D^2u) = \frac{f_1(x)}{f_2(\nabla u)} \quad in \quad \Omega_1, \quad \nabla u(\Omega_1) = \Omega^*,$$

has a convex solution u belonging to $C^{3,\alpha}(\Omega_1) \cap C^{2,\alpha}(\bar{\Omega}_1)$ for any $\alpha \in (0,1)$. Any two such solutions differ by a constant.

In fact, to apply this result to show existence of solutions to (3.5), set $n = 2$, let

$$f_1(x,y) = f(x,y) \text{ for } (x,y) \in D_1 = \Omega_1 \ ,$$
$$f_2(p_1,p_2) = \frac{g\left(-\left(p_1,p_2,\sqrt{1-p_1^2-p_2^2}\right)\right)}{\sqrt{1-p_1^2-p_2^2}},$$

for $(p_1,p_2) \in D_2 = -\Pi(\Omega_2)$, where Π is the orthogonal projection from a set on the unit sphere onto the x,y-plane. In particular, Ω_2 is a subset of the lower unit hemisphere $z \leq 0$. We need to verify (3.6). From the conservation of energy assumption

$$\int_{\Omega_1} f(x)\,dx = \int_{\Omega_2} g(y)\,d\sigma(y),$$

and using the parametrization $q = (q_1,q_2) \to \left(q, -\sqrt{1-|q|^2}\right)$, we can write

$$\int_{\Omega_2} g(y)\,d\sigma(y) = \int_{\Pi(\Omega_2)} \frac{g\left(q, -\sqrt{1-|q|^2}\right)}{\sqrt{1-|q|^2}}\,dq$$
$$= \int_{-\Pi(\Omega_2)} \frac{g\left(-p, -\sqrt{1-|p|^2}\right)}{\sqrt{1-|p|^2}}\,dp = \int_{D_2} f_2(p_1,p_2)\,dp.$$

Therefore, (3.6) holds; hence, the existence of solutions to (3.5) follows.

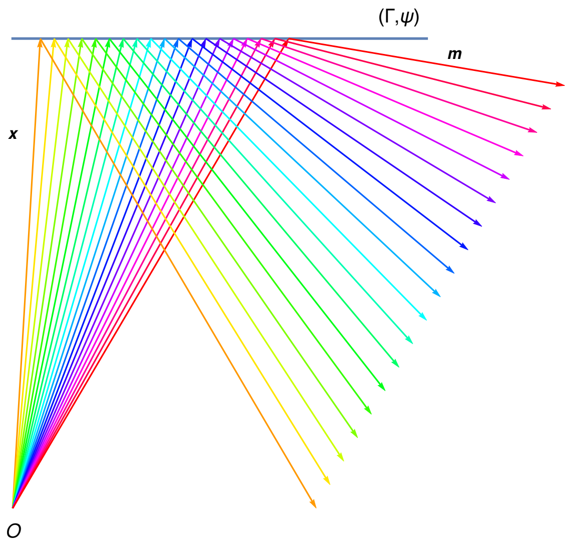

Figure 3.2: Reflection from a point source

3.1.2 Point Source Reflection

We now have a domain Ω_1 of the unit sphere in $\mathbb{R}^3$, and rays emanate from the origin with intensity $f(x) \geq 0$ for each $x \in \Omega_1$. Let Ω_2 be as in the previous section, i.e., a domain of the unit sphere, and let $g > 0$ be a function in Ω_2 such that the following energy conservation condition holds:

$$\int_{\Omega_1} f(x)\, d\sigma(x) = \int_{\Omega_2} g(y)\, d\sigma(y). \tag{3.7}$$

Again, Γ is the plane $z = 1$. Of course, we assume that rays from the origin with unit direction in Ω_1 reach the plane Γ. The question is then to find a phase discontinuity ψ on Γ such that all rays emitted from the origin with direction $x \in \Omega_1$ and intensity $f(x)$ are reflected by the metasurface (Γ, ψ) into Ω_2 such that

$$\int_{E} f(x)\, d\sigma(x) = \int_{T(E)} g(y)\, d\sigma(y) \tag{3.8}$$

for each subset E of Ω_1 and $T(\Omega_1) = \Omega_2$, where T is the reflection map. In order to find the equation ψ satisfies, we parametrize the domains in the sphere using spherical coordinates: $s(u,v) = (\cos u \sin v, \sin u \sin v, \cos v)$, $0 \le u \le 2\pi$, $0 \le v \le \pi/2$. Parametrizing Ω_1 in these coordinates, we obtain $\Omega_1 = s(O)$, for some domain $O \subset [0, 2\pi] \times [0, \pi/2]$. Rewriting the integrals in (3.8) in spherical coordinates, and letting $s(U) = E$, we have

$$\int_U f(s(u,v))|s_u \times s_v| dudv = \int_E f(x)\, d\sigma(x)$$
$$= \int_{T(E)} g(y)\, d\sigma(y) = \int_U g(T(s(u,v)))|(T \circ s)_u \times (T \circ s)_v| dudv.$$

Because this equation must hold for all open sets $U \subset O$, it follows that T satisfies the equation

$$\frac{|(T \circ s)_u \times (T \circ s)_v|}{|s_u \times s_v|} = \frac{f(s(u,v))}{g(T(s(u,v)))}. \tag{3.9}$$

The plane Γ is described in spherical coordinates by the polar radius

$$r(u,v) = \frac{1}{\cos v} s(u,v) = (\cos u \tan v, \sin u \tan v, 1). \tag{3.10}$$

From Remark 2.1.2 with $n_1 = 1$, if the incident ray has direction $i = s(u,v)$, then the reflected ray that has unit direction

$$T(s(u,v)) = s(u,v) - \lambda\nu - \nabla\psi(r(u,v)),$$

where $\nu = (0,0,1)$ is the normal to Γ at the incident point. Because we seek, as before, for a phase ψ tangential to Γ, we have $\nabla\psi(x,y,1) = (\psi_x(x,y,1), \psi_y(x,y,1), 0)$. In addition, from Remark 2.1.2,

$$\lambda = i \cdot \nu + \sqrt{1 - (|i - \nabla\psi|^2 - (i \cdot \nu)^2)}$$

$$= \cos v + \sqrt{1 - (\cos u \sin v - \psi_x(r(u,v)))^2 - (\sin u \sin v - \psi_y(r(u,v)))^2}$$

$$= \cos v + \sqrt{\Delta},$$

where $\Delta = 1 - (\cos u \sin v - \psi_x(r(u,v)))^2 - (\sin u \sin v - \psi_y(r(u,v)))^2$. Therefore, writing T in components

$$T(s(u,v)) = (T_1(s(u,v)), T_2(s(u,v)), T_3(s(u,v)))$$
$$= \left(\cos u \sin v - \psi_x(r(u,v)), \sin u \sin v - \psi_y(r(u,v)), -\sqrt{\Delta}\right). \tag{3.11}$$

Because $|T(s(u,v))|= 1$, it follows as in Section 3.1.1 that

$$|(T \circ s)_u \times (T \circ s)_v|= \frac{1}{|T_3 \circ s|}\left|\det \begin{pmatrix} (T_1 \circ s)_u & (T_1 \circ s)_v \\ (T_2 \circ s)_u & (T_2 \circ s)_v \end{pmatrix}\right|. \tag{3.12}$$

From (3.11),

$$(T_1 \circ s)_u = -\sin u \sin v - \psi_{xx}(r(u,v))(-\sin u \tan v) - \psi_{xy}(r(u,v))(\cos u \tan v),$$

$$(T_1 \circ s)_v = \cos u \cos v - \psi_{xx}(r(u,v)) \left(\frac{\cos u}{\cos^2 v}\right) - \psi_{xy}(r(u,v)) \left(\frac{\sin u}{\cos^2 v}\right),$$

$$(T_2 \circ s)_u = \cos u \sin v - \psi_{xy}(r(u,v))(-\sin u \tan v) - \psi_{yy}(r(u,v))(\cos u \tan v),$$

$$(T_2 \circ s)_v = \sin u \cos v - \psi_{xy}(r(u,v)) \left(\frac{\cos u}{\cos^2 v}\right) - \psi_{yy}(r(u,v)) \left(\frac{\sin u}{\cos^2 v}\right).$$

Inserting these in (3.12) yields

$$|(T \circ s)_u \times (T \circ s)_v|= \frac{1}{|T_3 \circ s|}\left|\det \left(A(u,v) - D^2_{(x,y)}\psi(r(u,v))B(u,v)\right)\right|, \tag{3.13}$$

where

$$A(u,v) = \begin{pmatrix} -\sin u \sin v & \cos u \cos v \\ \cos u \sin v & \sin u \cos v \end{pmatrix},$$

$$B(u,v) = \begin{pmatrix} -\sin u \tan v & \frac{\cos u}{\cos^2 v} \\ \cos u \tan v & \frac{\sin u}{\cos^2 v} \end{pmatrix}.$$

We can rewrite the above quantities in rectangular coordinates, noticing that $x = \cos u \tan v$, $y = \sin u \tan v$, $r(u,v) = (x, y, 1)$, $\sqrt{x^2+y^2+1} = \dfrac{1}{\cos v}$ and $\sqrt{x^2+y^2} = \tan v$. We obtain

$$\begin{aligned}
T_1 &= \frac{x}{\sqrt{x^2+y^2+1}} - \psi_x(x,y,1) = \left(\sqrt{x^2+y^2+1} - \psi(x,y,1)\right)_x, \\
T_2 &= \frac{y}{\sqrt{x^2+y^2+1}} - \psi_y(x,y,1) = \left(\sqrt{x^2+y^2+1} - \psi(x,y,1)\right)_y, \\
|T_3| &= \sqrt{1 - \left(\frac{x}{\sqrt{x^2+y^2+1}} - \psi_x(x,y,1)\right)^2 - \left(\frac{y}{\sqrt{x^2+y^2+1}} - \psi_y(x,y,1)\right)^2} \\
&= \sqrt{1 - \left(\left(\sqrt{x^2+y^2+1} - \psi(x,y,1)\right)_x\right)^2 - \left(\left(\sqrt{x^2+y^2+1} - \psi(x,y,1)\right)_y\right)^2}, \\
A &= \begin{pmatrix} \dfrac{-y}{\sqrt{x^2+y^2+1}} & \dfrac{x}{\sqrt{x^2+y^2+1}\sqrt{x^2+y^2}} \\ \dfrac{x}{\sqrt{x^2+y^2+1}} & \dfrac{y}{\sqrt{x^2+y^2+1}\sqrt{x^2+y^2}} \end{pmatrix}, \\
B &= \begin{pmatrix} -y & \dfrac{x(1+x^2+y^2)}{\sqrt{x^2+y^2}} \\ x & \dfrac{y(1+x^2+y^2)}{\sqrt{x^2+y^2}} \end{pmatrix}.
\end{aligned}$$

Also,

$$\begin{aligned}
|s_u \times s_v| &= \frac{1}{|\cos v|}\left| -\sin^2 u \cos v \sin v - \cos^2 u \cos v \sin v \right| \\
&= \sin v = \frac{\sqrt{x^2+y^2}}{\sqrt{x^2+y^2+1}}.
\end{aligned} \tag{3.14}$$

Now note that

$$\det(A - D^2\psi B) = \det(B)\det(AB^{-1} - D^2\psi),$$

with

$$B^{-1} = -\frac{1}{(x^2+y^2+1)\sqrt{x^2+y^2}} \begin{pmatrix} \frac{y(1+x^2+y^2)}{\sqrt{x^2+y^2}} & \frac{-x(1+x^2+y^2)}{\sqrt{x^2+y^2}} \\ -x & -y \end{pmatrix} \tag{3.15}$$

and

$$AB^{-1} = \begin{pmatrix} \frac{y^2}{b(x)} + \frac{x^2}{c(x)} & \frac{-xy}{b(x)} + \frac{xy}{c(x)} \\ \frac{-xy}{b(x)} + \frac{xy}{c(x)} & \frac{x^2}{b(x)} + \frac{y^2}{c(x)} \end{pmatrix} = D^2\Big(\sqrt{x^2+y^2+1}\Big),$$

where $b(x) = (x^2+y^2)(x^2+y^2+1)^{1/2}$ and $c(x) = (x^2+y^2)(x^2+y^2+1)^{3/2}$. Therefore,

$$\det(A - D^2\psi B) = \det(B)\det\left(D^2\left(\sqrt{x^2+y^2+1} - \psi\right)\right).$$

Letting $\phi(x,y) = \sqrt{x^2+y^2+1} - \psi(x,y)$, using the last equation in (3.13), and using (3.14), we obtain from (3.9) that ϕ satisfies the following equation:

$$\begin{aligned} &\frac{(x^2+y^2+1)^{3/2}}{\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}}|\det(D^2\phi(x,y))| \\ &= \frac{f\left(\frac{1}{\sqrt{x^2+y^2+1}}(x,y,1)\right)}{g\left(\phi_x(x,y),\phi_y(x,y),-\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}\right)}. \end{aligned} \tag{3.16}$$

The above equation holds for $(x,y) \in D$, where D is obtained as follows: for each direction $e \in \Omega_1$, the ray with this direction intersect the plane $z = 1$ at a unique point (x,y), this collection of x and y is D.

We now proceed as in the previous section to show existence of solutions to (3.16). To this end, we need to identity the functions f_1, f_2 in (3.6). Parametrizing Ω_1 by $q : D \to \Omega_1$ with $q(x,y) = \frac{1}{\sqrt{x^2+y^2+1}}(x,y,1)$, we let $f_1(x,y) = \frac{f\,(q(x,y))}{(x^2+y^2+1)^{3/2}}$ for $(x,y) \in D_1 = D$. Also let $f_2(p_1,p_2) = \frac{g\left(p_1,p_2,-\sqrt{1-p_1^2-p_2^2}\right)}{\sqrt{1-p_1^2-p_2^2}}$, for $(p_1, p_2) \in D_2 = \Pi(\Omega_2)$. With these choices and observing that

$$\int_{\Omega_1} f(z)\, d\sigma(z) = \int_D \frac{f\left(\frac{1}{\sqrt{x^2+y^2+1}}(x,y,1)\right)}{(x^2+y^2+1)^{3/2}}\, dxdy,$$

which is a similar calculation as at the end of last section, we obtain that (3.7) is equivalent to (3.6); therefore, the existence of solutions to (3.16) follows as before, invoking Theorem 3.1.1.

3.2 Refraction when Γ is a plane

Here, we solve two problems similar to the ones considered in the previous sections but for refraction.

3.2.1 Collimated case

Incident rays are emitted from an open set Ω_1 of the x-y plane with direction $i(x,y) = e_3 = (0,0,1)$, and Γ is the plane $z = 1$.

From the generalized law of refraction (2.6) and (2.8), the metasurface (Γ, ψ) refracts the incident ray $i(x,y)$ into a ray $r(x,y)$ with direction satisfying

$$n_1 i(x,y) - n_2 r(x,y) = \lambda \nu(x,y) + \nabla\psi(x,y),$$

where n_1 and n_2 are the refractive indices of the two homogeneous and isotropic media separated by the plane Γ, ν is the unit normal to the plane Γ. Also

$$\lambda = (n_1 i - \nabla\psi)\cdot\nu - \sqrt{n_2^2 - |n_1 i - \nabla\psi|^2 + [(n_1 i - \nabla\psi)\cdot\nu]^2}$$
$$= n_1 - \sqrt{n_2^2 - (\psi_x^2 + \psi_y^2)},$$

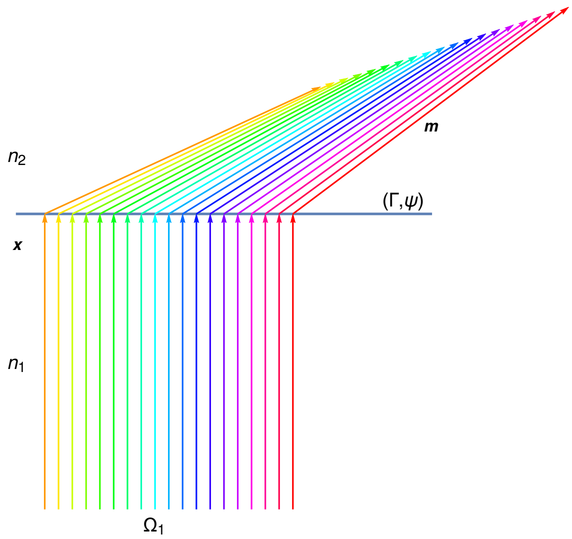

Figure 3.3: Refraction from an extended source

because we seek ψ tangential to Γ; i.e., $\psi_z = 0$. We then let $T : \Omega_1 \to \Omega_2$ to be

$$T(x,y) := r(x,y) = \left(-\frac{1}{n_2}\psi_x(x,y), -\frac{1}{n_2}\psi_y(x,y), \sqrt{1 - \frac{1}{n_2^2}\left(\psi_x^2(x,y) + \psi_y^2(x,y)\right)} \right).$$

We seek ψ defined on Γ with $T(\Omega_1) = \Omega_2$ and satisfying the conservation of energy balance

$$\int_E f(x)\,dx = \int_{T(E)} g(y)\,d\sigma(y) = \int_E g(T(z))|J_T|\,dz \quad \text{for each} \quad E \subset \Omega_1,$$

where $|J_T| = |T_x(x,y) \times T_y(x,y)|$. Because $|T(x,y)| = 1$; similarly, as for reflection, we have that

$$|J_T| = \frac{1}{|T_3(x,y)|}\left|\det\left(D^2\frac{1}{n_2}\psi\right)\right|.$$

Therefore, proceeding as in the reflection case, the phase discontinuity ψ must satisfy the following Monge-Ampére equation

$$\frac{1}{\sqrt{1-\frac{1}{n_2^2}\left(\psi_x^2(x,y)+\psi_y^2(x,y)\right)}}\left|\det\left(D^2\frac{1}{n_2}\psi\right)\right| \tag{3.17}$$
$$= \frac{f(x,y)}{g\left(-\frac{1}{n_2}\psi_x, -\frac{1}{n_2}\psi_y, \sqrt{1-\frac{1}{n_2^2}\left(\psi_x^2+\psi_y^2\right)}\right)};$$

notice that this equation is independent of the value of n_1. Similar to the reflection case, $T(\Omega_1) = \Omega_2$ implies that $\frac{1}{n_2}(\psi_x, \psi_y) \in -\Pi(\Omega_2)$, where Π is once again the orthogonal projection onto the x-y plane. We claim, also in this case, that [40, Theorem 2] can be applied to obtain a solutions ψ to (3.17). Indeed, letting

$$f_1(x,y) = f(x,y) \text{ for } (x,y) \in D_1 = \Omega_1 \,,$$
$$f_2(p_1,p_2) = \frac{g\left(-\frac{1}{n_2}p_1, -\frac{1}{n_2}p_2, \sqrt{1-\frac{1}{n_2^2}\left(p_1^2+p_2^2\right)}\right)}{\sqrt{1-\frac{1}{n_2^2}\left(p_1^2+p_2^2\right)}},$$

for $(p_1, p_2) \in D_2 = -n_2\Pi(\Omega_2)$, and proceeding as before, we obtain that

$$\int_{\Omega_1} f(x)\,dx = \int_{\Omega_2} g(y)\,d\sigma(y)$$

is equivalent to (3.6), and so existence of solutions follows as before.

3.2.2 Point Source Refraction

We now analyze a problem similar to the one in Section 3.1.2 for refraction. That is, rays emanate for a point source and we seek a phase discontinuity ψ defined on the plane $\Gamma = \{z = 1\}$, so that the refraction map T (to be calculated in a moment) satisfies the conservation of energy condition (3.8). As in Section 3.1.2, this implies (3.9), i.e.,

$$\frac{|(T\circ s)_u \times (T\circ s)_v|}{|s_u \times s_v|} = \frac{f(s(u,v))}{g(T(s(u,v)))},$$

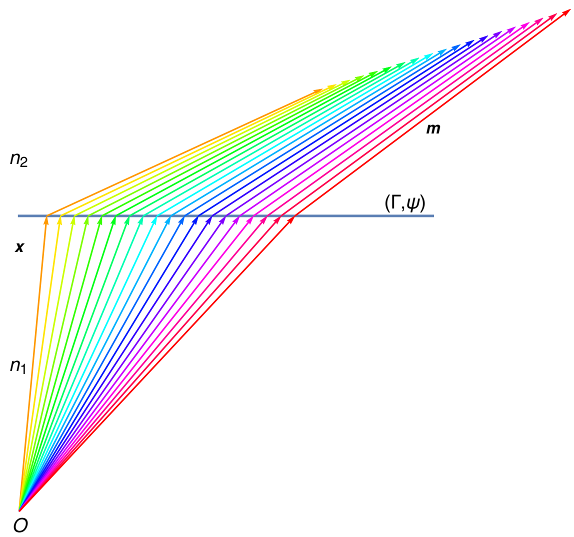

Figure 3.4: Refraction from a point source

and $T(\Omega_1) = \Omega_2$. Let us calculate the refraction map T. As in Section 3.1.2, the plane Γ is described by the polar radius (3.10). Then from (2.6), the refracted ray has unit direction

$$T(s(u,v)) = \frac{n_1}{n_2} s(u,v) - \frac{1}{n_2}\lambda\nu - \frac{1}{n_2}\nabla\psi(r(u,v)),$$

where $\nu = (0,0,1)$ is the normal to Γ at the incident point, $s(u,v)$ are spherical coordinates, $\nabla\psi(x,y,1) = (\psi_x(x,y,1), \psi_y(x,y,1), 0)$ (because we seek a phase discontinuity ψ tangential to Γ), and

$$\begin{aligned}\lambda &= n_1 i\cdot\nu + \sqrt{n_2^2 - (|n_1 i - \nabla\psi|^2 - (n_1 i\cdot\nu)^2)}\\ &= n_1\cos v + \sqrt{n_2^2 - (n_1\cos u\sin v - \psi_x(r(u,v)))^2 - (n_1\sin u\sin v - \psi_y(r(u,v)))^2}\\ &= n_1\cos v + \sqrt{\Delta},\end{aligned}$$

where $\Delta = n_2^2 - (n_1 \cos u \sin v - \psi_x(r(u,v)))^2 - (n_1 \sin u \sin v - \psi_y(r(u,v)))^2$. Therefore,

$$\begin{aligned} T(s(u,v)) &= (T_1(s(u,v)), T_2(s(u,v)), T_3(s(u,v))) \\ &= \left(\frac{n_1}{n_2} \cos u \sin v - \frac{1}{n_2}\psi_x(r(u,v)), \frac{n_1}{n_2} \sin u \sin v - \frac{1}{n_2}\psi_y(r(u,v)), \frac{1}{n_2}\sqrt{\Delta} \right). \end{aligned}$$

Because $|T(s(u,v))| = 1$, we have as in (3.12) that

$$|(T \circ s)_u \times (T \circ s)_v| = \frac{1}{|T_3 \circ s|} \left| \det \begin{pmatrix} (T_1 \circ s)_u & (T_1 \circ s)_v \\ (T_2 \circ s)_u & (T_2 \circ s)_v \end{pmatrix} \right|. \tag{3.18}$$

On the other hand,

$$\begin{aligned} (T_1 \circ s)_u &= -\frac{n_1}{n_2} \sin u \sin v - \frac{1}{n_2}\psi_{xx}(r(u,v))(-\sin u \tan v) \\ &\quad - \frac{1}{n_2}\psi_{xy}(r(u,v))(\cos u \tan v), \\ (T_1 \circ s)_v &= \frac{n_1}{n_2} \cos u \cos v - \frac{1}{n_2}\psi_{xx}(r(u,v)) \left(\frac{\cos u}{\cos^2 v} \right) - \frac{1}{n_2}\psi_{xy}(r(u,v)) \left(\frac{\sin u}{\cos^2 v} \right), \\ (T_2 \circ s)_u &= \frac{n_1}{n_2} \cos u \sin v - \frac{1}{n_2}\psi_{xy}(r(u,v))(-\sin u \tan v) \\ &\quad - \frac{1}{n_2}\psi_{yy}(r(u,v))(\cos u \tan v), \\ (T_2 \circ s)_v &= \frac{n_1}{n_2} \sin u \cos v - \frac{1}{n_2}\psi_{xy} \left(r(u,v))(\frac{\cos u}{\cos^2 v} \right) - \frac{1}{n_2}\psi_{yy}(r(u,v)) \left(\frac{\sin u}{\cos^2 v} \right). \end{aligned}$$

Inserting these into (3.18) yields

$$|(T \circ s)_u \times (T \circ s)_v| = \frac{1}{|T_3 \circ s|} \left| \det \left(A(u,v) - \frac{1}{n_2} D^2_{(x,y)}\psi(r(u,v)) B(u,v) \right) \right|$$

where

$$A(u,v) = \frac{n_1}{n_2} \begin{pmatrix} -\sin u \sin v & \cos u \cos v \\ \cos u \sin v & \sin u \cos v \end{pmatrix},$$

$$B(u,v) = \begin{pmatrix} -\sin u \tan v & \frac{\cos u}{\cos^2 v} \\ \cos u \tan v & \frac{\sin u}{\cos^2 v} \end{pmatrix}.$$

As in the point source reflection case in Section 3.1.2, we can rewrite the above quantities in rectangular coordinates, noticing that $x = \cos u \tan v$, $y = \sin u \tan v$, $r(u,v) = (x, y, 1)$, $\sqrt{x^2+y^2+1} = \dfrac{1}{\cos v}$ and $\sqrt{x^2+y^2} = \tan v$. We obtain

$$\begin{aligned}
T_1 =& \frac{n_1}{n_2}\frac{x}{\sqrt{x^2+y^2+1}} - \frac{1}{n_2}\psi_x(x,y,1) = \left(\frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi(x,y,1)\right)_x, \\
T_2 =& \frac{n_1}{n_2}\frac{y}{\sqrt{x^2+y^2+1}} - \frac{1}{n_2}\psi_y(x,y,1) = \left(\frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi(x,y,1)\right)_y, \\
|T_3| =& \left(1 - \left(\frac{n_1}{n_2}\frac{x}{\sqrt{x^2+y^2+1}} - \frac{1}{n_2}\psi_x(x,y,1)\right)^2 \right. \\
& \left. - \left(\frac{n_1}{n_2}\frac{y}{\sqrt{x^2+y^2+1}} - \frac{1}{n_2}\psi_y(x,y,1)\right)^2\right)^{\frac{1}{2}} \\
=& \left(1 - \left(\left(\frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi(x,y,1)\right)_x\right)^2 \right. \\
& \left. - \left(\left(\frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi(x,y,1)\right)_y\right)^2\right)^{\frac{1}{2}}, \\
A =& \frac{n_1}{n_2}\begin{pmatrix} \frac{-y}{\sqrt{x^2+y^2+1}} & \frac{x}{\sqrt{x^2+y^2+1}\sqrt{x^2+y^2}} \\ \frac{x}{\sqrt{x^2+y^2+1}} & \frac{y}{\sqrt{x^2+y^2+1}\sqrt{x^2+y^2}} \end{pmatrix}, \\
B =& \begin{pmatrix} -y & \frac{x(1+x^2+y^2)}{\sqrt{x^2+y^2}} \\ x & \frac{y(1+x^2+y^2)}{\sqrt{x^2+y^2}} \end{pmatrix}.
\end{aligned}$$

Also, from (3.14), $|s_u \times s_v| = \frac{\sqrt{x^2+y^2}}{\sqrt{x^2+y^2+1}}$. Now note that

$$\det\left(A - \frac{1}{n_2}D^2\psi B\right) = \det(B)\det\left(AB^{-1} - \frac{1}{n_2}D^2\psi\right),$$

with B^{-1} as in (3.15), and

$$AB^{-1} = \frac{n_1}{n_2}\begin{pmatrix} \frac{y^2}{b(x)} + \frac{x^2}{c(x)} & \frac{-xy}{b(x)} + \frac{xy}{c(x)} \\ \frac{-xy}{b(x)} + \frac{xy}{c(x)} & \frac{x^2}{b(x)} + \frac{y^2}{c(x)} \end{pmatrix} = \frac{n_1}{n_2} D^2\left(\sqrt{x^2+y^2+1}\right),$$

where $b(x) = (x^2+y^2)(x^2+y^2+1)^{1/2}$ and $c(x) = (x^2+y^2)(x^2+y^2+1)^{3/2}$. Therefore,

$$\det\left(A - \frac{1}{n_2}D^2\psi B\right) = \det(B)\det\left(D^2\left(\frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi\right)\right).$$

Letting $\phi(x,y) = \frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi(x,y)$, we obtain that ϕ satisfies the following equation

$$\begin{aligned} &\frac{(x^2+y^2+1)^{3/2}}{\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}}|\det(D^2\phi(x,y))| \\ &= \frac{f\left(\frac{1}{\sqrt{x^2+y^2+1}}(x,y,1)\right)}{g\left(\phi_x(x,y),\phi_y(x,y),\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}\right)}. \end{aligned} \tag{3.19}$$

The above equation holds for $(x,y) \in D$, where D is obtained as at the end of Section 3.1.2. The existence of solutions to this equation follows as before, letting $f_1(x,y) = \dfrac{f\left(\frac{1}{\sqrt{x^2+y^2+1}}(x,y,1)\right)}{(x^2+y^2+1)^{3/2}}$ for $(x,y) \in D_1 = D$, and $f_2(p_1,p_2) = \dfrac{g\left(p_1,p_2,\sqrt{1-p_1^2-p_2^2}\right)}{\sqrt{1-p_1^2-p_2^2}}$ for $(p_1,p_2) \in D_2 = \Pi(\Omega_2)$, where Π is once again the orthogonal projection.

Remark 3.2.1. If a ray is emitted from a point Q and strikes the plane $\Gamma = \{z = 1\}$ at the point $P = (x,y,1)$, let $d_Q(x,y)$ be the distance from Q to P. In the collimated case, because all rays are vertical $d_Q(x,y) = 1$. And when the point source Q is the origin, $d_Q(x,y) = \sqrt{x^2+y^2+1}$. Then writing $\phi(x,y) = \dfrac{n_1}{n_2}d_Q(x,y) - \dfrac{1}{n_2}\psi(x,y)$, and noticing that $n_1 = n_2 = 1$ in the reflection cases, the equations (3.5), (3.16), (3.17) and (3.19) can be written as

$$\frac{d_Q^{3/2}(x,y)}{\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}}|\det(D^2\phi(x,y))|=\frac{\tilde{f}\,(x,y)}{g\,(T(x,y))}, \tag{3.20}$$

where $\tilde{f}(x,y)=f(x,y)$ in the collimated case, and $\tilde{f}(x,y)=f\left(\frac{1}{\sqrt{x^2+y^2+1}}(x,y,1)\right)$ in the point source case.

3.3 Summary of the equations in the planar case

Collimated reflection

$$\frac{1}{\sqrt{1-\psi_x^2(x,y)-\psi_y^2(x,y)}}\left|\det(D^2\psi)\right|=\frac{f(x,y)}{g\left(-\psi_x,-\psi_y,-\sqrt{1-\psi_x^2(x,y)-\psi_y^2(x,y)}\right)}.$$

Collimated refraction

$$\frac{1}{\sqrt{1-\frac{1}{n_2^2}\psi_x^2(x,y)-\frac{1}{n_2^2}\psi_y^2(x,y)}}\left|\det(D^2\frac{1}{n_2}\psi)\right|$$
$$=\frac{f(x,y)}{g\left(-\frac{1}{n_2}\psi_x,-\frac{1}{n_2}\psi_y,\sqrt{1-\frac{1}{n_2^2}\psi_x^2(x,y)-\frac{1}{n_2^2}\psi_y^2(x,y)}\right)}.$$

Point source reflection

$$\frac{(x^2+y^2+1)^{3/2}}{\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}}\left|\det(D^2\phi(x,y))\right|$$
$$=\frac{f\left(\frac{x}{\sqrt{x^2+y^2+1}},\frac{y}{\sqrt{x^2+y^2+1}},\frac{1}{\sqrt{x^2+y^2+1}}\right)}{g\left(\phi_x(x,y),\phi_y(x,y),-\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}\right)},$$

where $\phi(x,y)=\sqrt{x^2+y^2+1}-\psi(x,y)$.

Point source refraction

$$\frac{(x^2+y^2+1)^{3/2}}{\sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}}\left|\det(D^2\phi(x,y))\right|$$

$$= \frac{f\left(\frac{x}{\sqrt{x^2+y^2+1}}, \frac{y}{\sqrt{x^2+y^2+1}}, \frac{1}{\sqrt{x^2+y^2+1}}\right)}{g\left(\phi_x(x,y), \phi_y(x,y), \sqrt{1-\phi_x^2(x,y)-\phi_y^2(x,y)}\right)}.$$

where $\phi(x,y) = \frac{n_1}{n_2}\sqrt{x^2+y^2+1} - \frac{1}{n_2}\psi(x,y)$.

3.4 Reflection when Γ is a general surface

In this section we derive the equations for the reflection problems for a given general surface Γ.

3.4.1 Collimated case

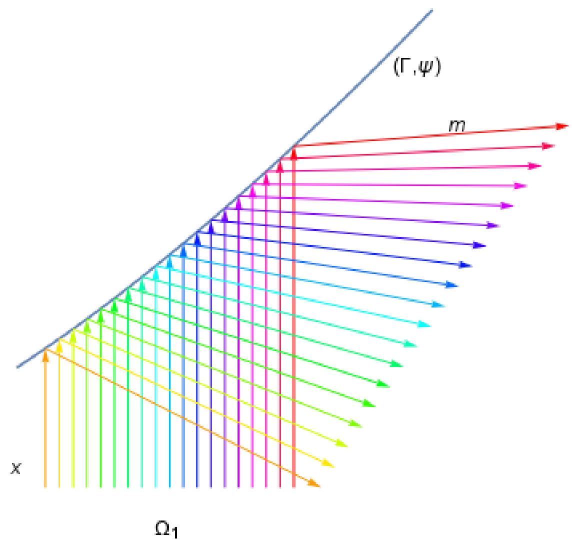

Figure 3.5: Reflection from an extended source

Incident rays are emitted from an open set Ω_1 of the $x-y$ plane with direction $e_3 = (0,0,1)$, and Γ is the surface given by the graph of the function $u : \Omega_1 \to \mathbb{R}^+$. As before, the reflected rays are $r(x,y) = (0,0,1) - \lambda\nu(x,y) - \nabla\psi(x,y)$, where $\nabla\psi = (D\psi, \psi_z)$, $\nu = \dfrac{(-Du, 1)}{\sqrt{1+|Du|^2}}$, and because $\nu \cdot \nabla\psi = 0$, from Remark 2.1.2, we have

$$\lambda = \frac{1}{\sqrt{1+|Du|^2}} + \sqrt{2\psi_z - \psi_x^2 - \psi_y^2 - \psi_z^2 + \frac{1}{1+|Du|^2}}.$$

Because we assume ψ is tangential to Γ, $\psi_z = Du \cdot (\psi_x, \psi_y) = Du \cdot D\psi$. Therefore,

$$\begin{aligned} \lambda &= \frac{1}{\sqrt{1+|Du|^2}} + \sqrt{2Du \cdot D\psi - |D\psi|^2 - (Du \cdot D\psi)^2 + \frac{1}{1+|Du|^2}} \\ &= \frac{1}{\sqrt{1+|Du|^2}} + \sqrt{\Delta}, \end{aligned}$$

where $\Delta = 2Du \cdot D\psi - |D\psi|^2 - (Du \cdot D\psi)^2 + \dfrac{1}{1+|Du|^2}$, and we set

$$\begin{aligned} T(x,y) &= r(x,y) \\ &= \left(\frac{\lambda u_x}{\sqrt{1+|Du|^2}} - \psi_x, \frac{\lambda u_y}{\sqrt{1+|Du|^2}} - \psi_y, 1 - \frac{\lambda}{\sqrt{1+|Du|^2}} - Du \cdot D\psi \right). \end{aligned}$$

As before, we now have to calculate

$$\frac{1}{|T_3|} \det \begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix}.$$

Notice that

$$\begin{aligned} \begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} &= -D^2\psi + \begin{pmatrix} \lambda_x \frac{u_x}{\sqrt{1+|Du|^2}} & \lambda_y \frac{u_x}{\sqrt{1+|Du|^2}} \\ \lambda_x \frac{u_y}{\sqrt{1+|Du|^2}} & \lambda_y \frac{u_y}{\sqrt{1+|Du|^2}} \end{pmatrix} \\ &\quad + \lambda \begin{pmatrix} (\frac{u_x}{\sqrt{1+|Du|^2}})_x & (\frac{u_x}{\sqrt{1+|Du|^2}})_y \\ (\frac{u_y}{\sqrt{1+|Du|^2}})_x & (\frac{u_y}{\sqrt{1+|Du|^2}})_y \end{pmatrix} \\ &= -\nabla^2\psi + \frac{1}{\sqrt{1+|Du|^2}} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \otimes \begin{pmatrix} \lambda_x \\ \lambda_y \end{pmatrix} + \lambda B(Du, D^2u), \end{aligned}$$

where $B(Du, D^2u) = \begin{pmatrix} (\frac{u_x}{\sqrt{1+|Du|^2}})_x & (\frac{u_x}{\sqrt{1+|Du|^2}})_y \\ (\frac{u_y}{\sqrt{1+|Du|^2}})_x & (\frac{u_y}{\sqrt{1+|Du|^2}})_y \end{pmatrix}$. We can calculate the derivatives of λ obtaining,

$$\lambda_x = \left(\frac{1}{\sqrt{1+|Du|^2}}\right)_x + \left(\sqrt{2Du\cdot D\psi - |D\psi|^2 - (Du\cdot D\psi)^2 + \frac{1}{1+|Du|^2}}\right)_x$$

$$= \left(\frac{1}{\sqrt{1+|Du|^2}}\right)_x + \frac{1}{2}\Delta^{-1/2}\Big[2(Du\cdot D\psi)_x - 2(Du\cdot D\psi)(Du\cdot D\psi)_x +$$

$$\left(\frac{1}{1+|Du|^2}\right)_x - (|D\psi|^2)_x\Big]$$

$$= \left(\frac{1}{\sqrt{1+|Du|^2}}\right)_x + \frac{1}{2}\Delta^{-1/2}\Bigg[2((Du)_x\cdot D\psi) - 2(Du\cdot D\psi)((Du)_x\cdot D\psi) +$$

$$\left(\frac{1}{1+|Du|^2}\right)_x\Bigg] + \frac{1}{2}\Delta^{-1/2}\left[2(Du\cdot(D\psi)_x) - 2(Du\cdot D\psi)(Du\cdot(D\psi)_x) - (|D\psi|^2)_x\right]$$

$$= h_1(Du, D^2u, D\psi) + \frac{1}{2}\Delta^{-1/2}\Big[2(Du\cdot(D\psi)_x) - 2(Du\cdot D\psi)(Du\cdot(D\psi)_x)$$

$$- (|D\psi|^2)_x\Big]$$

$$= h_1(Du, D^2u, D\psi) + \frac{1}{2}\Delta^{-1/2}\left[2(1 - Du\cdot D\psi)(Du\cdot(D\psi)_x) - (|D\psi|^2)_x\right],$$

where $\Delta = 2Du\cdot D\psi - |D\psi|^2 - (Du\cdot D\psi)^2 + \frac{1}{1+|Du|^2}$ and $h_1(Du, D^2u, D\psi) = \left(\frac{1}{\sqrt{1+|Du|^2}}\right)_x + \frac{1}{2}\Delta^{-1/2}\left[2((Du)_x\cdot D\psi) - 2(Du\cdot D\psi)((Du)_x\cdot D\psi) + \left(\frac{1}{1+|Du|^2}\right)_x\right]$.
Similarly,

$$\lambda_y = h_2(Du, D^2u, D\psi) + \frac{1}{2}\Delta^{-1/2}\left[2(1 - Du\cdot D\psi)(Du\cdot(D\psi)_y) - (|D\psi|^2)_y\right].$$

Therefore,

$$\frac{1}{\sqrt{1+|Du|^2}}\begin{pmatrix} u_x \\ u_y \end{pmatrix} \otimes \begin{pmatrix} \lambda_x \\ \lambda_y \end{pmatrix} =$$

$$\tilde{H}(Du, D^2u, D\psi) + \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}[(1 - Du\cdot D\psi)(Du\otimes(D^2\psi)(Du)) -$$

$(Du \otimes (D^2\psi)(D\psi))],$

where $\tilde{H} = Du \otimes \dfrac{(h_1, h_2)}{\sqrt{1+|Du|^2}}$.

All in all,

$$\begin{aligned}
\begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} &= -\nabla^2\psi + \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}[(1 - Du\cdot D\psi)(Du \otimes (D^2\psi)(Du)) \\
&\quad - (Du \otimes (D^2\psi)(D\psi))] + \lambda B(Du, D^2u,) + \tilde{H}(Du, D^2u, D\psi) \\
&= [-\mathrm{Id} + \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}((1 - Du\cdot D\psi)Du \otimes Du - Du \otimes D\psi)]D^2\psi \\
&\quad + F(Du, D^2u, D\psi) \\
&= -[\mathrm{Id} + Du \otimes \frac{-\Delta^{-1/2}((1 - Du\cdot D\psi)Du - D\psi)}{\sqrt{1+|Du|^2}}]D^2\psi \\
&\quad + F(Du, D^2u, D\psi) \\
&= -[\mathrm{Id} + Du \otimes A(Du, D\psi)]D^2\psi + F(Du, D^2u, D\psi) \\
&= -[\mathrm{Id} + Du \otimes A)](D^2\psi - [\mathrm{Id} + Du \otimes A]^{-1}F),
\end{aligned}$$

where

$$A(Du, D\psi) = \frac{-\Delta^{-1/2}((1 - Du\cdot D\psi)Du - D\psi)}{\sqrt{1+|Du|^2}},$$

and

$$F(Du, D^2u, D\psi) = \lambda B(Du, D^2u) + \tilde{H}(Du, D^2u, D\psi).$$

Therefore,

$$\begin{aligned}
\det\begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} &= \det\left([\mathrm{Id} + Du \otimes A)]\left(D^2\psi - [\mathrm{Id} + Du \otimes A]^{-1}F\right)\right) \\
&= \det\left(\mathrm{Id} + Du \otimes A\right)\det\left(D^2\psi - [\mathrm{Id} + Du \otimes A]^{-1}F\right).
\end{aligned}$$

Using Sherman-Morrison's formula [1], we obtain

$$\det\begin{pmatrix}(T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y\end{pmatrix} = (1 + Du\cdot A)\,\det\left(D^2\psi - \left[\mathrm{Id} - \frac{Du\otimes A}{1+Du\cdot A}\right]F\right).$$

Therefore, the phase discontinuity ψ satisfies the equation,

$$\frac{(1+Du\cdot A)}{\left|1 - \dfrac{\lambda}{\sqrt{1+|Du|^2}} - Du\cdot D\psi\right|}\left|\det\left(D^2\psi - \left[\mathrm{Id} - \frac{Du\otimes A}{1+Du\cdot A}\right]F\right)\right| = \frac{f(x,y)}{g(T(x,y))}.$$

3.4.2 Point source Reflection

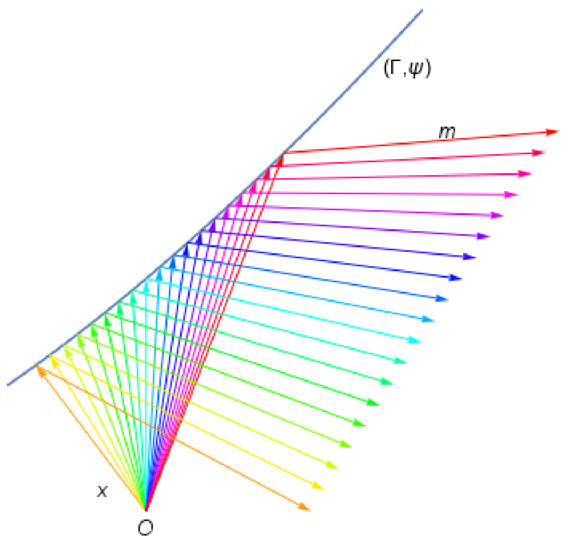

Figure 3.6: Reflection from a point source

Here we derive the equation when the incident rays are emitted from the origin. We use the following parametrization of Ω_1:

$$s : D \to \Omega_1,$$

[1] $\det(A + u\otimes v) = (1 + u^T A^{-1} v)\det(A)$, where A is an invertible matrix and u, v are vectors. https://en.wikipedia.org/wiki/Matrix_determinant_lemma

$$s(x,y) = \left(\frac{x}{\sqrt{x^2+y^2+u^2(x,y)}}, \frac{y}{\sqrt{x^2+y^2+u^2(x,y)}}, \frac{u(x,y)}{\sqrt{x^2+y^2+u^2(x,y)}} \right).$$

In this case $\nu = \dfrac{(-Du,1)}{\sqrt{1+|Du|^2}}$, and $\nabla\psi = (D\psi, Du\cdot D\psi)$, because we assume $\nabla\psi$ is tangential to Γ. Therefore,

$$\begin{aligned}
\lambda &= i\cdot\nu + \sqrt{1-(|i-\nabla\psi|^2-(i\cdot\nu)^2)} \\
&= \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} + \left[1 - \left(\frac{x}{\sqrt{x^2+y^2+u^2}} - \psi_x\right)^2\right. \\
&\qquad - \left(\frac{y}{\sqrt{x^2+y^2+u^2}} - \psi_y\right)^2 - \left(\frac{u}{\sqrt{x^2+y^2+u^2}} - D\psi\cdot Du\right)^2 \\
&\qquad \left. - \left(\frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)^2\right]^{\frac{1}{2}} \\
&= \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} + \left[\frac{2D\psi\cdot((x,y)+uDu)}{\sqrt{x^2+y^2+u^2}} - |D\psi|^2 - (D\psi\cdot Du)^2\right. \\
&\qquad \left. + \left(\frac{u-(x,y)\cdot Du}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)^2\right]^{\frac{1}{2}} \\
&= \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} + \sqrt{\Delta(x,y,u,Du,D\psi)},
\end{aligned}$$

where

$$\begin{aligned}
&\Delta(x,y,u,Du,D\psi) \\
&= \frac{2D\psi\cdot((x,y)+uDu)}{\sqrt{x^2+y^2+u^2}} - |D\psi|^2 - (D\psi\cdot Du)^2 + \left(\frac{u-(x,y)\cdot Du}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)^2.
\end{aligned}$$

We set

$$\begin{aligned}
&T\left(s(x,y)\right) \\
&= \left(\frac{x}{\sqrt{x^2+y^2+u^2}} + \frac{\lambda u_x}{\sqrt{1+|Du|^2}} - \psi_x, \frac{y}{\sqrt{x^2+y^2+u^2}} + \frac{\lambda u_y}{\sqrt{1+|Du|^2}} - \psi_y,\right. \\
&\qquad \left. \frac{u}{\sqrt{x^2+y^2+u^2}} - \frac{\lambda}{\sqrt{1+|Du|^2}} - Du\cdot D\psi\right).
\end{aligned}$$

Because $|T(s(x,y))|= 1$, we have that

$$|(T\circ s)_x \times (T\circ s)_y|= \frac{1}{|T_3\circ s|}\left|\det\begin{pmatrix}(T_1\circ s)_x & (T_1\circ s)_y\\ (T_2\circ s)_x & (T_2\circ s)_y\end{pmatrix}\right|.$$

On the other hand,

$$\begin{pmatrix}(T_1\circ s)_x & (T_1\circ s)_y\\ (T_2\circ s)_x & (T_2\circ s)_y\end{pmatrix} = -D^2\psi + \begin{pmatrix}\lambda_x\frac{u_x}{\sqrt{1+|Du|^2}} & \lambda_y\frac{u_x}{\sqrt{1+|Du|^2}}\\ \lambda_x\frac{u_y}{\sqrt{1+|Du|^2}} & \lambda_y\frac{u_y}{\sqrt{1+|Du|^2}}\end{pmatrix}$$

$$+\lambda\begin{pmatrix}\left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_y\\ \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_y\end{pmatrix}$$

$$+\begin{pmatrix}\left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_y\\ \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_y\end{pmatrix}$$

$$= -D^2\psi + \frac{1}{\sqrt{1+|Du|^2}}\begin{pmatrix}u_x\\ u_y\end{pmatrix}\otimes\begin{pmatrix}\lambda_x\\ \lambda_y\end{pmatrix}$$

$$+\lambda B(Du,D^2u) + C(x,y,u,Du),$$

where $B(Du,D^2u) = \begin{pmatrix}\left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_y\\ \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_y\end{pmatrix}$, and

$C(x,y,u,Du) = \begin{pmatrix}\left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_y\\ \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_y\end{pmatrix}$. We can calculate the derivatives of λ obtaining,

$$\lambda_x = \left(\frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)_x$$

$$+\frac{1}{2}\Delta^{-1/2}\left[2(D\psi)_x\cdot\left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}}\right) + 2(D\psi).\left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}}\right)_x\right.$$

$$- 2(Du \cdot D\psi)(Du \cdot D\psi)_x + \left(\left(\frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} \right)^2 \right)_x - (|D\psi|^2)_x \Bigg]$$

$$= h_1(x,y,u,Du,D^2u,D\psi)$$

$$+ \frac{1}{2}\Delta^{-1/2}\left[2(D\psi)_x \cdot \left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\,Du \right) - \left(|D\psi|^2\right)_x \right],$$

similarly,

$$\lambda_y = h_2(x,y,u,Du,D^2u,D\psi)$$

$$+ \frac{1}{2}\Delta^{-1/2}\left[2(D\psi)_y \cdot \left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\,Du \right) - \left(|D\psi|^2\right)_y \right],$$

where $h_i(x,y,u,Du,D^2u,D\psi)$, with $i+1,2$, are the collections of all the terms that do not contain second derivative of ψ. Therefore,

$$\frac{1}{\sqrt{1+|Du|^2}} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \otimes \begin{pmatrix} \lambda_x \\ \lambda_y \end{pmatrix} = \tilde{H}(x,y,u,Du,D^2u,D\psi)$$

$$+ \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}\Bigg[Du \otimes (D^2\psi)\Bigg(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}}$$

$$- (Du\cdot D\psi)\,Du \Bigg) - (Du \otimes (D^2\psi)(D\psi))\Bigg],$$

where $\tilde{H} = Du \otimes \dfrac{(h_1,h_2)}{\sqrt{1+|Du|^2}}$.

All in all,

$$\begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix}$$

$$= -D^2\psi + \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}[Du \otimes (D^2\psi)\left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\,Du \right)$$

$$- (Du \otimes (D^2\psi)(D\psi))] + \lambda B(Du,D^2u) + C(x,y,u,Du)$$

$$
\begin{aligned}
&+ \tilde{H}(x,y,u,Du,D^2u,D\psi)\\
&= \left[-\operatorname{Id} + \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}} Du \otimes \left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\, Du - D\psi \right) \right] D^2\psi\\
&+ F(x,y,u,Du,D^2u,D\psi)\\
&= -[\operatorname{Id} + Du \otimes A(x,y,u,Du,D\psi)]D^2\psi + F(x,y,u,Du,D^2u,D\psi)\\
&= -[\operatorname{Id} + Du \otimes A)](D^2\psi - [\operatorname{Id} + Du\otimes A]^{-1}F),
\end{aligned}
$$

where

$$
A(x,y,u,Du,D\psi) = \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}} \left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\, Du - D\psi \right), \tag{3.21}
$$

and

$$
\begin{aligned}
&F(x,y,u,Du,D^2u,D\psi) \qquad (3.22)\\
&\qquad = \lambda B(Du,D^2u) + C(x,y,u,Du) + \tilde{H}(x,y,u,Du,D^2u,D\psi).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\det \begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} &= \det \left([\operatorname{Id} + Du\otimes A)] \left(D^2\psi - [\operatorname{Id} + Du\otimes A]^{-1}F \right) \right)\\
&= \det\left(\operatorname{Id} + Du\otimes A\right) \det\left(D^2\psi - [\operatorname{Id} + Du\otimes A]^{-1}F \right).
\end{aligned}
$$

Using Sherman-Morrison's formula [2], we obtain

$$
\det \begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} = (1 + Du\cdot A)\ \det\left(D^2\psi - [\operatorname{Id} - \frac{Du\otimes A}{1+Du\cdot A}]F \right).
$$

[2] $\det(A + u\otimes v) = (1 + u^T A^{-1} v)\det(A)$, where A is an invertible matrix and u, v are vectors. https://en.wikipedia.org/wiki/Matrix_determinant_lemma

Therefore, the phase discontinuity ψ satisfies the equation,

$$\frac{(1+Du\cdot A)}{\tilde{s}(x,y)\left|\frac{u}{\sqrt{x^2+y^2+u}}-\frac{\lambda}{\sqrt{1+|Du|^2}}-Du\cdot D\psi\right|}\left|\det\left(D^2\psi-\left[\mathrm{Id}-\frac{Du\otimes A}{1+Du\cdot A}\right]F\right)\right| = \frac{f(s(x,y))}{g(T(s(x,y)))}$$

where $\tilde{s}(x,y)=|s_x\times s_y|$, A as in (3.21), and F as in (3.22).

3.5 Refraction when Γ is a general surface

We consider the problems in the case of refraction for a given general surface. Even if the calculations are similar to the case of reflection (Section 3.4) , we include them for convenience of the reader.

3.5.1 Collimated case

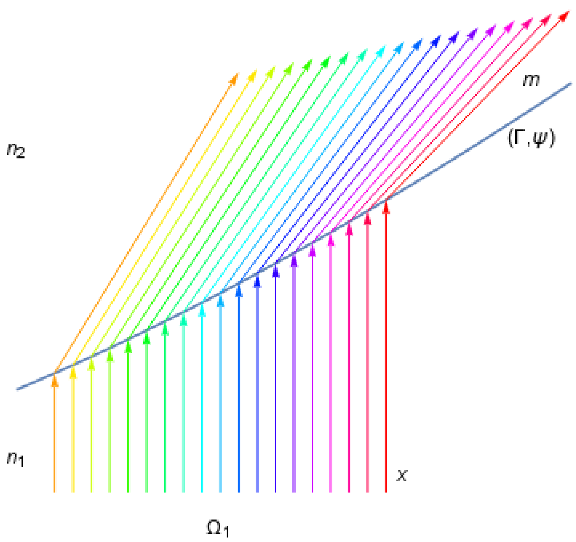

Figure 3.7: Refraction from an extended source

Consider the case in which rays are emitted from an open set Ω_1 of the $x-y$ plane, Γ is the surface given by the graph of the function $u(x,y):\Omega_1\to\mathbb{R}^+$, and the incident rays have direction $e_3=(0,0,1)$. As before $r(x,y)=\dfrac{n_1}{n_2}(0,0,1)-\dfrac{1}{n_2}\lambda\nu(x,y)-\dfrac{1}{n_2}\nabla\psi(x,y)$, where $r(x,y)$ are the refracted rays, $\nabla\psi=(D\psi,\psi_z)$, $\nu=\dfrac{(-Du,1)}{\sqrt{1+|Du|^2}}$ and, because $\nu\cdot\nabla\psi=0$, from [17, Remark 2] we have

$$\lambda=\frac{n_1}{\sqrt{1+|Du|^2}}-\sqrt{n_2^2-n_1^2+2n_1\psi_z-\psi_x^2-\psi_y^2-\psi_z^2+\frac{n_1^2}{1+|Du|^2}}.$$

Using that ψ is tangential to Γ, we have that $\psi_z=Du\cdot(\psi_x,\psi_y)=Du\cdot D\psi$. Therefore,

$$\begin{aligned}\lambda&=\frac{n_1}{\sqrt{1+|Du|^2}}-\sqrt{n_2^2-n_1^2+2n_1Du\cdot D\psi-|D\psi|^2-(Du\cdot D\psi)^2+\frac{n_1^2}{1+|Du|^2}}\\&=\frac{n_1}{\sqrt{1+|Du|^2}}-\sqrt{\Delta},\end{aligned}$$

where $\Delta=n_2^2-n_1^2+2n_1Du\cdot D\psi-|D\psi|^2-(Du\cdot D\psi)^2+\dfrac{n_1^2}{1+|Du|^2}$, and

$$\begin{aligned}&T(x,y)=r(x,y)\\&=\Big(\frac{\lambda u_x}{n_2\sqrt{1+|Du|^2}}-\frac{\psi_x}{n_2},\frac{\lambda u_y}{n_2\sqrt{1+|Du|^2}}-\frac{\psi_y}{n_2},\frac{n_1}{n_2}-\frac{\lambda}{n_2\sqrt{1+|Du|^2}}-\frac{Du\cdot D\psi}{n_2}\Big).\end{aligned}$$

As before, we now have to calculate

$$\frac{1}{|T_3|}\det\begin{pmatrix}(T_1)_x&(T_1)_y\\(T_2)_x&(T_2)_y\end{pmatrix}.$$

Notice that

$$\begin{pmatrix}(T_1)_x&(T_1)_y\\(T_2)_x&(T_2)_y\end{pmatrix}=-\frac{D^2\psi}{n_2}+\frac{1}{n_2}\begin{pmatrix}\lambda_x\frac{u_x}{\sqrt{1+|Du|^2}}&\lambda_y\frac{u_x}{\sqrt{1+|Du|^2}}\\\lambda_x\frac{u_y}{\sqrt{1+|Du|^2}}&\lambda_y\frac{u_y}{\sqrt{1+|Du|^2}}\end{pmatrix}$$

$$+\frac{\lambda}{n_2}\begin{pmatrix}(\frac{u_x}{\sqrt{1+|Du|^2}})_x & (\frac{u_x}{\sqrt{1+|Du|^2}})_y\\ (\frac{u_y}{\sqrt{1+|Du|^2}})_x & (\frac{u_y}{\sqrt{1+|Du|^2}})_y\end{pmatrix}$$

$$= -\frac{D^2\psi}{n_2}+\frac{1}{n_2\sqrt{1+|Du|^2}}\begin{pmatrix}u_x\\ u_y\end{pmatrix}\otimes\begin{pmatrix}\lambda_x\\ \lambda_y\end{pmatrix}+\lambda B(Du,D^2u).$$

where $B(Du,D^2u) = +\frac{1}{n_2}\begin{pmatrix}(\frac{u_x}{\sqrt{1+|Du|^2}})_x & (\frac{u_x}{\sqrt{1+|Du|^2}})_y\\ (\frac{u_y}{\sqrt{1+|Du|^2}})_x & (\frac{u_y}{\sqrt{1+|Du|^2}})_y\end{pmatrix}$. We can calculate the derivatives of λ obtaining,

$$\begin{aligned}\lambda_x &= \left(\frac{n_1}{\sqrt{1+|Du|^2}}\right)_x\\ &\quad -\left(\sqrt{n_2^2-n_1^2+2n_1Du\cdot D\psi-|D\psi|^2-(Du\cdot D\psi)^2+\frac{n_1^2}{1+|Du|^2}}\right)_x\\ &= \left(\frac{n_1}{\sqrt{1+|Du|^2}}\right)_x-\frac{1}{2}\Delta^{-1/2}\Big[2n_1(Du\cdot D\psi)_x-2(Du\cdot D\psi)(Du\cdot D\psi)_x\\ &\quad +\left(\frac{n_1^2}{1+|Du|^2}\right)_x-(|D\psi|^2)_x\Big]\\ &= h_1(Du,D^2u,D\psi)\\ &\quad -\frac{1}{2}\Delta^{-1/2}\Big[2n_1(Du\cdot(D\psi)_x)-2(Du\cdot D\psi)(Du\cdot(D\psi)_x)-(|D\psi|^2)_x\Big]\\ &= h_1(Du,D^2u,D\psi)\\ &\quad -\frac{1}{2}\Delta^{-1/2}\Big[2(n_1-Du\cdot D\psi)(Du\cdot(D\psi)_x)-(|D\psi|^2)_x\Big].\end{aligned}$$

Similarly,

$$\lambda_y = h_2(Du,D^2u,D\psi)-\frac{1}{2}\Delta^{-1/2}\Big[2(n_1-Du\cdot D\psi)(Du\cdot(D\psi)_y)-(|D\psi|^2)_y\Big],$$

where $\Delta = n_2^2-n_1^2+2n_1Du\cdot D\psi-|D\psi|^2-(Du\cdot D\psi)^2+\frac{n_1^2}{1+|Du|^2}$, and $h_i(Du,D^2u,D\psi)$, with $i = 1, 2$, are the collection of all the terms that do not contain second derivative of ψ. Therefore,

$$\frac{1}{n_2\sqrt{1+|Du|^2}} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \otimes \begin{pmatrix} \lambda_x \\ \lambda_y \end{pmatrix} =$$

$$\tilde{H}(Du, D^2u, D\psi) - \frac{\Delta^{-1/2}}{n_2\sqrt{1+|Du|^2}}[(n_1 - Du \cdot D\psi)(Du \otimes (D^2\psi)(Du))$$

$$- (Du \otimes (D^2\psi)(D\psi))],$$

where $\tilde{H} = Du \otimes \dfrac{(h_1, h_2)}{n_2\sqrt{1+|Du|^2}}$.

All in all,

$$\begin{pmatrix} (T_1)_x & (T_1)_y \\ (T_2)_x & (T_2)_y \end{pmatrix} = -\frac{D^2\psi}{n_2} - \frac{\Delta^{-1/2}}{n_2\sqrt{1+|Du|^2}}[(n_1 - Du \cdot D\psi)(Du \otimes (D^2\psi)(Du))$$

$$- (Du \otimes (D^2\psi)(D\psi))] + \lambda B(Du, D^2u) + \tilde{H}(Du, D^2u, D\psi)$$

$$= [-\mathrm{Id} - \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}((n_1 - Du \cdot D\psi)Du \otimes Du - Du \otimes D\psi)]\frac{D^2\psi}{n_2}$$

$$+ F(Du, D^2u, D\psi)$$

$$= -[\mathrm{Id} + Du \otimes \frac{\Delta^{-1/2}((n_1 - Du \cdot D\psi)Du - D\psi)}{\sqrt{1+|Du|^2}}]\frac{D^2\psi}{n_2} + F(Du, D^2u, D\psi)$$

$$= -[\mathrm{Id} + Du \otimes A(Du, D\psi)]\frac{D^2\psi}{n_2} + F(Du, D^2u, D\psi)$$

$$= -[\mathrm{Id} + Du \otimes A](\frac{D^2\psi}{n_2} - [\mathrm{Id} + Du \otimes A]^{-1}F),$$

where

$$A(Du, D\psi) = \frac{\Delta^{-1/2}((n_1 - Du \cdot D\psi)Du - D\psi)}{\sqrt{1+|Du|^2}},$$

and

$$F(Du, D^2u, D\psi) = \lambda B(Du, D^2u) + \tilde{H}(Du, D^2u, D\psi).$$

Therefore,

$$\det\begin{pmatrix}(T_1)_x & (T_1)_y\\ (T_2)_x & (T_2)_y\end{pmatrix} = \det\left([\mathrm{Id} + Du\otimes A)]\left(\frac{D^2\psi}{n_2} - [\mathrm{Id} + Du\otimes A]^{-1}F\right)\right) =$$

$$\det\left(\mathrm{Id} + Du\otimes A\right)\det\left(\frac{D^2\psi}{n_2} - [\mathrm{Id} + Du\otimes A]^{-1}F\right).$$

Using Sherman-Morrison's formula, [3] we obtain

$$\det\begin{pmatrix}(T_1)_x & (T_1)_y\\ (T_2)_x & (T_2)_y\end{pmatrix} = (1 + Du\cdot A)\det\left(\frac{D^2\psi}{n_2} - [\mathrm{Id} - \frac{Du\otimes A}{1 + Du\cdot A}]F\right).$$

Therefore, we obtain that the phase discontinuity ψ satisfies the following equation,

$$\frac{(1 + Du\cdot A)}{\left\|\frac{n_1}{n_2} - \frac{\lambda}{n_2\sqrt{1+|Du|^2}} - \frac{Du\cdot D\psi}{n_2}\right\|}\left|\det\left(\frac{D^2\psi}{n_2} - [\mathrm{Id} - \frac{Du\otimes A}{1 + Du\cdot A}]F\right)\right| = \frac{f(x,y)}{g(T(x,y))}.$$

3.5.2 Point source Refraction

Incident rays are emitted from the origin. We use the following parametrization of Ω_1:

$$s : D \to \Omega_1,$$

$$s(x,y) = \left(\frac{x}{\sqrt{x^2+y^2+u^2(x,y)}}, \frac{y}{\sqrt{x^2+y^2+u^2(x,y)}}, \frac{u(x,y)}{\sqrt{x^2+y^2+u^2(x,y)}}\right).$$

In this case $\nu = \dfrac{(-Du, 1)}{\sqrt{1+|Du|^2}}$, and $\nabla\psi = (D\psi, Du\cdot D\psi)$, because we assume $\nabla\psi$ is tangential to Γ. Therefore,

$$\lambda = n_1 i\cdot\nu + \sqrt{n_2^2 - (|n_1 i - \nabla\psi|^2 - (n_1 i\cdot\nu)^2)}$$

[3] $\det(A + u\otimes v) = (1 + u^T A^{-1} v)\det(A)$, where A is an invertible matrix and u, v are vectors. https://en.wikipedia.org/wiki/Matrix_determinant_lemma

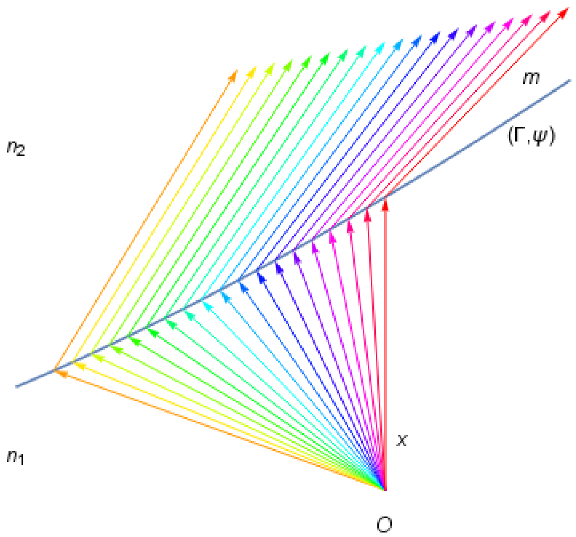

Figure 3.8: Refraction from a point source

$$
\begin{aligned}
&= n_1 \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} + \Bigg[n_2^2 - \left(\frac{n_1 x}{\sqrt{x^2+y^2+u^2}} - \psi_x \right)^2 \\
&\quad - \left(\frac{n_1 y}{\sqrt{x^2+y^2+u^2}} - \psi_y \right)^2 - \left(\frac{n_1 u}{\sqrt{x^2+y^2+u^2}} - D\psi \cdot Du \right)^2 \\
&\quad - \left(n_1 \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} \right)^2 \Bigg]^{\frac{1}{2}} \\
&= n_1 \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} + \Bigg[n_2^2 - n_1^2 \frac{2D\psi \cdot ((x,y) + uDu)}{\sqrt{x^2+y^2+u^2}} \\
&\quad - |D\psi|^2 - (D\psi \cdot Du)^2 + \left(n_1 \frac{u - (x,y)\cdot Du}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} \right)^2 \Bigg]^{\frac{1}{2}} \\
&= n_1 \frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}} + \sqrt{\Delta(x,y,u,Du,D\psi)},
\end{aligned}
$$

where

$$
\Delta(x,y,u,Du,D\psi) = n_2^2 - n_1^2 + n_1 \frac{2D\psi \cdot ((x,y) + uDu)}{\sqrt{x^2+y^2+u^2}} - |D\psi|^2 - (D\psi \cdot Du)^2
$$

$$+\left(n_1\frac{u-(x,y)\cdot Du}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)^2.$$

We set

$$\begin{aligned}
&T\left(s(x,y)\right)\\
&=\left(\frac{n_1x}{n_2\sqrt{x^2+y^2+u^2}}+\frac{\lambda u_x}{n_2\sqrt{1+|Du|^2}}-\frac{1}{n_2}\psi_x,\frac{n_1y}{n_2\sqrt{x^2+y^2+u^2}}\right.\\
&\left.+\frac{\lambda u_y}{n_2\sqrt{1+|Du|^2}}-\frac{1}{n_2}\psi_y,\frac{n_1u}{n_2\sqrt{x^2+y^2+u^2}}-\frac{\lambda}{n_2\sqrt{1+|Du|^2}}-\frac{1}{n_2}Du\cdot D\psi\right).
\end{aligned}$$

Because $|T(s(x,y))|=1$, we have that

$$|(T\circ s)_x\times(T\circ s)_y|=\frac{1}{|T_3\circ s|}\left|\det\begin{pmatrix}(T_1\circ s)_x & (T_1\circ s)_y\\ (T_2\circ s)_x & (T_2\circ s)_y\end{pmatrix}\right|.$$

On the other hand,

$$\begin{aligned}
\begin{pmatrix}(T_1\circ s)_x & (T_1\circ s)_y\\ (T_2\circ s)_x & (T_2\circ s)_y\end{pmatrix}&=\frac{-D^2\psi}{n_2}+\frac{1}{n_2}\begin{pmatrix}\lambda_x\frac{u_x}{\sqrt{1+|Du|^2}} & \lambda_y\frac{u_x}{\sqrt{1+|Du|^2}}\\ \lambda_x\frac{u_y}{\sqrt{1+|Du|^2}} & \lambda_y\frac{u_y}{\sqrt{1+|Du|^2}}\end{pmatrix}\\
&+\frac{\lambda}{n_2}\begin{pmatrix}\left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_y\\ \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_y\end{pmatrix}\\
&+\frac{n_1}{n_2}\begin{pmatrix}\left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_y\\ \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_y\end{pmatrix}\\
&=\frac{-D^2\psi}{n_2}+\frac{1}{n_2\sqrt{1+|Du|^2}}\begin{pmatrix}u_x\\ u_y\end{pmatrix}\otimes\begin{pmatrix}\lambda_x\\ \lambda_y\end{pmatrix}\\
&+\lambda B(Du,D^2u)+C(x,y,u,Du),
\end{aligned}$$

where $B(Du, D^2u) = \frac{1}{n_2}\begin{pmatrix} \left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_x}{\sqrt{1+|Du|^2}}\right)_y \\ \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_x & \left(\frac{u_y}{\sqrt{1+|Du|^2}}\right)_y \end{pmatrix}$ and $C(x, y, u, Du) = \frac{n_1}{n_2}\begin{pmatrix} \left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{x}{\sqrt{x^2+y^2+u^2}}\right)_y \\ \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_x & \left(\frac{y}{\sqrt{x^2+y^2+u^2}}\right)_y \end{pmatrix}$. We can calculate the derivatives of λ obtaining,

$$\begin{aligned}
\lambda_x = n_1 & \left(\frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)_x + \frac{1}{2}\Delta^{-1/2}\Bigg[2n_1(D\psi)_x \cdot \left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}}\right) \\
& + 2n_1(D\psi).\left(\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}}\right)_x - 2(Du\cdot D\psi)(Du\cdot D\psi)_x \\
& + \left(\left(n_1\frac{-(x,y)\cdot Du + u}{\sqrt{x^2+y^2+u^2}\sqrt{1+|Du|^2}}\right)^2\right)_x - (|D\psi|^2)_x\Bigg] \\
= h_1&(x, y, u, Du, D^2u, D\psi) \\
& + \frac{1}{2}\Delta^{-1/2}\left[2(D\psi)_x\cdot\left(n_1\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\,Du\right) - \left(|D\psi|^2\right)_x\right],
\end{aligned}$$

similarly,

$$\begin{aligned}
\lambda_y = h_2&(x, y, u, Du, D^2u, D\psi) \\
& + \frac{1}{2}\Delta^{-1/2}\left[2(D\psi)_y\cdot\left(n_1\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\,Du\right) - \left(|D\psi|^2\right)_y\right],
\end{aligned}$$

where $h_i(x, y, u, Du, D^2u, D\psi)$, with $i + 1, 2$, are the collection of all the terms that do not contain second derivative of ψ. Therefore,

$$\begin{aligned}
\frac{1}{\sqrt{1+|Du|^2}} & \begin{pmatrix} u_x \\ u_y \end{pmatrix} \otimes \begin{pmatrix} \lambda_x \\ \lambda_y \end{pmatrix} = \tilde{H}(x, y, u, Du, D^2u, D\psi) \\
& + \frac{\Delta^{-1/2}}{n_2\sqrt{1+|Du|^2}}\Bigg[Du \otimes (D^2\psi)\left(n_1\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du\cdot D\psi)\,Du\right) -
\end{aligned}$$

$$(Du \otimes (D^2\psi)(D\psi))\Bigg],$$

where $\tilde{H} = Du \otimes \dfrac{(h_1, h_2)}{n_2\sqrt{1+|Du|^2}}$.

All in all,

$$\begin{pmatrix} (T_1)_x & (T_1)_y \\ \\ (T_2)_x & (T_2)_y \end{pmatrix}$$

$$= \frac{-D^2\psi}{n_2} + \frac{\Delta^{-1/2}}{n_2\sqrt{1+|Du|^2}}[Du \otimes (D^2\psi)\left(n_1\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du \cdot D\psi)\, Du\right)$$

$$- (Du \otimes (D^2\psi)(D\psi))] + \lambda B(Du, D^2u) + C(x,y,u,Du) + \tilde{H}(x,y,u,Du,D^2u,D\psi)$$

$$= \left[-\mathrm{Id} + \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}} Du \otimes \left(n_1\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du \cdot D\psi)\, Du - D\psi\right)\right]\frac{D^2\psi}{n_2}$$

$$+ F(x,y,u,Du,D^2u,D\psi)$$

$$= -[\mathrm{Id} + Du \otimes A(x,y,u,Du,D\psi)]\frac{D^2\psi}{n_2} + F(x,y,u,Du,D^2u,D\psi)$$

$$= -[\mathrm{Id} + Du \otimes A)](\frac{D^2\psi}{n_2} - [\mathrm{Id} + Du \otimes A]^{-1}F),$$

where

$$A(x,y,u,Du,D\psi) = \frac{\Delta^{-1/2}}{\sqrt{1+|Du|^2}}\left(n_1\frac{(x,y)+uDu}{\sqrt{x^2+y^2+u^2}} - (Du \cdot D\psi)\, Du - D\psi\right), \tag{3.23}$$

and

$$F(x,y,u,Du,D^2u,D\psi) \tag{3.24}$$

$$= \lambda B(Du, D^2u) + C(x,y,u,Du) + \tilde{H}(x,y,u,Du,D^2u,D\psi).$$

Therefore,

$$\det\begin{pmatrix}(T_1)_x & (T_1)_y\\ (T_2)_x & (T_2)_y\end{pmatrix} = \det\left([\mathrm{Id}+Du\otimes A)]\left(\frac{D^2\psi}{n_2} - [\mathrm{Id}+Du\otimes A]^{-1}F\right)\right)$$
$$= \det\left(\mathrm{Id}+Du\otimes A\right)\det\left(\frac{D^2\psi}{n_2} - [\mathrm{Id}+Du\otimes A]^{-1}F\right).$$

Using Sherman-Morrison's formula [4], we obtain

$$\det\begin{pmatrix}(T_1)_x & (T_1)_y\\ (T_2)_x & (T_2)_y\end{pmatrix} = (1+Du\cdot A)\,\det\left(\frac{D^2\psi}{n_2} - [\mathrm{Id} - \frac{Du\otimes A}{1+Du\cdot A}]F\right).$$

Therefore, the phase discontinuity ψ satisfies the equation,

$$\frac{(1+Du\cdot A)}{\tilde{s}(x,y)E(x,y,u,Du,D\psi)}\left|\det\left(\frac{D^2\psi}{n_2} - \left[\mathrm{Id} - \frac{Du\otimes A}{1+Du\cdot A}\right]F\right)\right| = \frac{f(s(x,y))}{g(T(s(x,y)))}$$

where $\tilde{s}(x,y) = |s_x \times s_y|$,

$E(x,y,u,Du,D\psi) = \left|\dfrac{n_1 u}{n_2\sqrt{x^2+y^2+u}} - \dfrac{\lambda}{n_2\sqrt{1+|Du|^2}} - Du\cdot\dfrac{D\psi}{n_2}\right|$, A as in (3.23), and F as in (3.24).

[4] $\det(A+u\otimes v) = (1+u^TA^{-1}v)\det(A)$, where A is an invertible matrix and u,v are vectors. https://en.wikipedia.org/wiki/Matrix_determinant_lemma

CHAPTER 4

WAVEGUIDES

As we said in the introduction, the problem considered in this chapter is that of modeling energy losses in waveguides. We investigate this problem within the regime of geometric optics in a dielectric into two cases: a straight guide and a circularly curved guide. To model this we use the Fresnel formulas. Let us first explain the set up. Suppose we have two homogeneous media I and II with refractive indices n_1, n_2, respectively, with $n_1 > n_2$, and set $\kappa = \dfrac{n_2}{n_1}$. Suppose media I and II are separated by a smooth surface S. If an incident wave with unit direction x is traveling within medium I and strikes S at a point P, then the wave splits into two waves: one transmitted into medium II and another internally reflected into medium I. The unit directions of these waves are m_t and m_r, respectively, which are determined by the Snell law. If ν is the unit outer normal to the surface S at the point P, then $x - \kappa\, m_t = \lambda\, \nu$ and $m_r = x - 2\,(x \cdot \nu)\nu$, where $\lambda = \Phi(x \cdot \nu)$ and $\Phi(t) = t - \kappa \sqrt{1 - \kappa^{-2}(1 - t^2)}$, see [13].

Therefore, the incident energy E_i carried by the incident wave with direction x splits into two: the transmitted energy E_t carried by the wave having direction m_t and the internally reflected energy E_r carried by the wave having direction m_r, with $E_i = E_t + E_r$, assuming no losses. The percentages of energy carried by the transmitted and internally reflected waves depends on the incident direction x via the Fresnel formulas, a consequence of Maxwell's equations, [5, Section1.5.3]. It is convenient to write these formulas in terms of the vectors x, m_r and m_t as follows, see [13]. Indeed, the percentage of internally reflected energy can be conveniently written for our purposes as

$$r(x) = \frac{1}{(1-\kappa^2)^2}\left(\left[\frac{2\kappa}{x\cdot m_t} - (1+\kappa^2)\right]^2 \frac{I_{\|}^2}{I_{\|}^2 + I_{\perp}^2} + \left[1 - 2\kappa\, x\cdot m_t + \kappa^2\right]^2 \frac{I_{\perp}^2}{I_{\|}^2 + I_{\perp}^2}\right), \tag{4.1}$$

see [13]. Therefore, the percentage of energy transmitted is $t(x) = 1 - r(x)$. Here $I_{\perp}$ and $I_{\|}$ are the coefficients of the amplitude of the incident wave, which might depend of x in a continuous way. Notice that from the Snell law, the function $r(x)$ is a function only depending on the dot product between x and the normal ν. And notice also that the critical angle is when $x \cdot m_t = \kappa$ and for such value of $x \cdot m_t$ we have $r(x) = 1$, that is, all the incident energy is internally reflected.

Given this set up, we want to model the losses of energy within a waveguide confined between two parallel surfaces S_1 and S_2; we assume the dielectric within the two surfaces has refractive index n_1, and the cladding, i.e., the material outside, has refractive index n_2 with $n_1 > n_2$. An incident polarized wave will zig-zag inside between the two surfaces. Depending on the normal to the surfaces at the striking points, one can calculate the energy transmitted and internally reflected by using the formulas above. In other words, the idea is to follow the path of the ray and tally the energy at each striking point on the boundary of the waveguide.

We will first work out this analysis for a straight guide between two parallel planes in Section 4.1. Second, we used this technique to carry out a similar but more difficult analysis when the guide is circular, see Section 4.2. This section contains several subsections analyzing in detail all the geometric possibilities that may arise. This is then used in Section 4.3 to get estimates for the energy internally reflected in the circular guide. The chapter ends showing asymptotics for periodic circular guides, Section 4.4, and final remarks on future research.

4.1 Straight waveguide

We consider an infinite waveguide with the form

$$S = \{(x_1, x_2, x_3) : -a < x_3 < a\}$$

so that, the material inside S has refractive index n_1 and the material outside S has refractive index n_2, where we assume $n_1 > n_2$. We analyze the energy losses for a zig-zagging ray that is internally reflected inside S, and confined to the x_1x_3-plane.

Depending on the normal at the surfaces at the striking points, one can calculate the energy transmitted and internally reflected by using the equation (4.1). In other words, the idea is to follow the path of the ray and tally the energy at each striking point on the boundary of the waveguide. Let $\mathbf{s}^i$ be the incident unit vector impinging the boundary of S at the point $A = (0, 0, -a)$ from inside the guide, and assume $\mathbf{s}^i$ lies on the x_1x_3-plane as in Figure 4.1. The outer unit normal at A is $\nu_A = (0, 0, -1)$, and let θ_i be the angle between ν_A and the incident unit direction $\mathbf{s}^i$. We then have

$$\mathbf{s}^i = -\sin\theta_i\,\mathbf{i} - \cos\theta_i\,\mathbf{k}, \tag{4.2}$$

as usual $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the unit coordinate vectors. We assume the magnetic and electric fields impinging the wave guide at a point A on the boundary $x_3 = -a$ of the guide are plane waves having direction $\mathbf{s}^i$ and traveling in the material n_1. The incident electric field at A is then

$$\begin{aligned} &\mathbf{E}^i_A(\mathbf{r}, t) \\ &= \left(-I^A_{\|}\cos\theta_i, I^A_{\perp}, I^A_{\|}\sin\theta_i\right)\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right) = \mathbf{E}^i_A\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right), \end{aligned} \tag{4.3}$$

where the values $I^A_{\|}$ and $I^A_{\perp}$ are given, since $\mathbf{E}$ is perpendicular to the direction of propagation $\mathbf{s}^i$. Being $\mathbf{E}$ and $\mathbf{H}$ plane waves we also have that

$$\mathbf{E} = -v_1\,\mathbf{s^i} \times \mathbf{H}, \qquad \mathbf{H} = \frac{1}{v_1}\,\mathbf{s^i} \times \mathbf{E}, \tag{4.4}$$

where $\mathbf{H}$ is the magnetic field, and therefore, the Poynting vector is given by

$$\mathbf{S} = \frac{c}{4\pi}\mathbf{E} \times \mathbf{H} = \frac{c}{4\pi v_1}\mathbf{E} \times (\mathbf{s^i} \times \mathbf{E}) = \frac{n_1}{4\pi}|\mathbf{E}|^2\,\mathbf{s^i}.$$

From (4.4)

$$\begin{aligned} &\mathbf{H}^i_A(\mathbf{r}, t) \\ &= \frac{1}{v_1}\left(I^A_{\perp}\cos\theta_i, I^A_{\|}, -I^A_{\perp}\sin\theta_i\right)\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right) = \mathbf{H}^i_A\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right). \end{aligned} \tag{4.5}$$

Let $\mathbf{s}^t$ be the direction of propagation of the transmitted wave, and let θ_t be the angle between the normal ν_A and $\mathbf{s}^t$, so $\mathbf{s}^t = -\sin\theta_t\,\mathbf{i} - \cos\theta_t\,\mathbf{k}$. Similarly, let $\mathbf{s}^r$ be the direction of propagation of the reflected wave and θ_r is the angle between the normal ν_A and $\mathbf{s}^r$. From the Snell law $\theta_r = \pi - \theta_i$ and so

$$\mathbf{s}^r = -\sin\theta_r\,\mathbf{i} - \cos\theta_r\,\mathbf{k} = -\sin\theta_i\,\mathbf{i} + \cos\theta_i\,\mathbf{k}. \tag{4.6}$$

Then the electromagnetic field corresponding to transmission is

$$\begin{aligned}
\mathbf{E}^t_A(\mathbf{r},t) &= \left(-T^A_{\parallel}\cos\theta_t, T^A_{\perp}, T^A_{\parallel}\sin\theta_t\right)\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^t}{v_2}\right)\right) \\
&= \mathbf{E}^t_A\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^t}{v_2}\right)\right) \\
\mathbf{H}^t_A(\mathbf{r},t) &= \frac{1}{v_2}\left(T^A_{\perp}\cos\theta_t, T^A_{\parallel}, -T^A_{\perp}\sin\theta_t\right)\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^t}{v_2}\right)\right) \\
&= \mathbf{H}^t_A\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^t}{v_2}\right)\right);
\end{aligned} \tag{4.7}$$

and similarly the fields corresponding to reflection are

$$\begin{aligned}
\mathbf{E}^r_A(\mathbf{r},t) &= \left(-R^A_{\parallel}\cos\theta_r, R^A_{\perp}, R^A_{\parallel}\sin\theta_r\right)\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right) \\
&= \mathbf{E}^r_A\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right) \\
\mathbf{H}^r_A(\mathbf{r},t) &= \frac{1}{v_1}\left(R^A_{\perp}\cos\theta_r, R^A_{\parallel}, -R^A_{\perp}\sin\theta_r\right)\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right) \\
&= \mathbf{H}^r_A\,\cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right).
\end{aligned} \tag{4.8}$$

Here the values $R^A_{\parallel}, R^A_{\perp}, T^A_{\parallel}$ and $T^A_{\perp}$ are determined from the Fresnel formulas below and depend only on $I^A_{\parallel}$, $I^A_{\perp}$ and the incident direction $\mathbf{s}^i$. In fact, let us replace $\mathbf{s^i}$ by x and $\mathbf{s^t}$ by m, so $\cos\theta_i = x\cdot\nu_A$ and $\cos\theta_t = m\cdot\nu_A$. In addition, setting $\kappa = n_2/n_1 < 1$, from the Snell law $x - \kappa m = \lambda\nu_A$, so the Fresnel equations

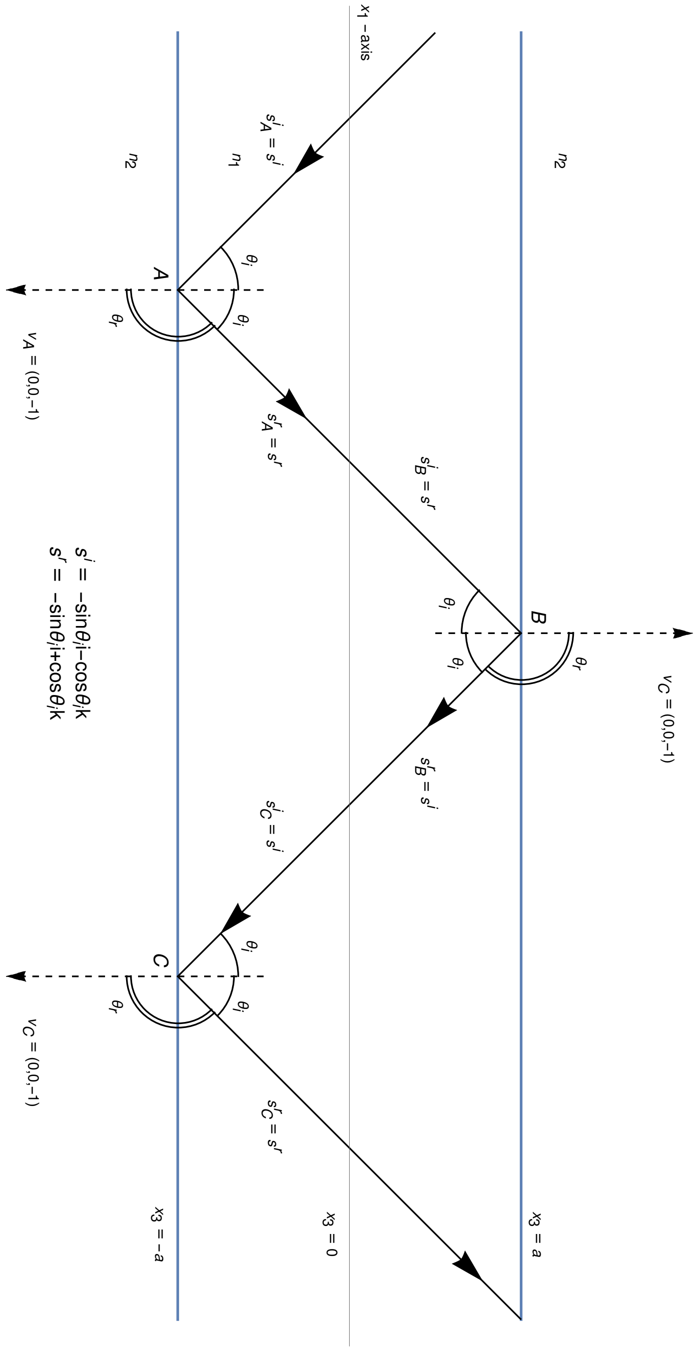

Figure 4.1: Wave guide paths configuration

have the form [1]

$$
\begin{aligned}
T_{\parallel}^{A} &= \frac{2\,x\cdot\nu_A}{\kappa\,x\cdot\nu_A + m\cdot\nu_A}\,I_{\parallel}^{A} = \frac{2\,x\cdot\nu_A}{(\kappa\,x+m)\cdot\nu_A}\,I_{\parallel} = \frac{2\,x\cdot(x-\kappa\,m)}{(\kappa\,x+m)\cdot(x-\kappa\,m)}\,I_{\parallel}^{A}\\
T_{\perp}^{A} &= \frac{2\,x\cdot\nu_A}{x\cdot\nu_A + \kappa\,m\cdot\nu_A}\,I_{\perp}^{A} = \frac{2\,x\cdot\nu_A}{(x+\kappa\,m)\cdot\nu_A}\,I_{\perp}^{A} = \frac{2\,x\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)}\,I_{\perp}^{A}\\
R_{\parallel}^{A} &= \frac{\kappa\,x\cdot\nu_A - m\cdot\nu_A}{\kappa\,x\cdot\nu_A + m\cdot\nu_A}\,I_{\parallel}^{A} = \frac{(\kappa\,x-m)\cdot\nu_A}{(\kappa\,x+m)\cdot\nu_A}\,I_{\parallel}^{A} = \frac{(\kappa\,x-m)\cdot(x-\kappa\,m)}{(\kappa\,x+m)\cdot(x-\kappa\,m)}\,I_{\parallel}^{A}\\
R_{\perp}^{A} &= \frac{x\cdot\nu_A - \kappa\,m\cdot\nu_A}{x\cdot\nu_A + \kappa\,m\cdot\nu_A}\,I_{\perp}^{A} = \frac{(x-\kappa\,m)\cdot\nu_A}{(x+\kappa\,m)\cdot\nu_A}\,I_{\perp}^{A} = \frac{(x-\kappa\,m)\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)}\,I_{\perp}^{A},
\end{aligned}
$$

for $x\cdot m \geq \kappa$, see [13, Section 4] for a derivation of these formulas. The reflection coefficient at A is

$$
\mathcal{R}_A = \left(\frac{|\mathbf{E}_A^r|}{|\mathbf{E}_A^i|}\right)^2,
$$

[5, Eq. (27), Section 1.5.3], representing the percentage of energy internally reflected at A, which can be calculated as follows. We have

$$
|\mathbf{E}_A^i|^2 = (I_{\parallel}^{A})^2 + (I_{\perp}^{A})^2,
$$

and from the Fresnel equations above

$$
\begin{aligned}
|\mathbf{E}_A^r|^2 &= (R_{\parallel}^{A})^2 + (R_{\perp}^{A})^2\\
&= \left[\frac{(\kappa\,x-m)\cdot(x-\kappa\,m)}{(\kappa\,x+m)\cdot(x-\kappa\,m)}\right]^2 (I_{\parallel}^{A})^2 + \left[\frac{(x-\kappa\,m)\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)}\right]^2 (I_{\perp}^{A})^2.
\end{aligned}
$$

Now let

$$
\Delta_1 = \frac{(\kappa\,x-m)\cdot(x-\kappa\,m)}{(\kappa\,x+m)\cdot(x-\kappa\,m)}, \qquad \Delta_2 = \frac{(x-\kappa\,m)\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)},
$$

and since x, m are unit vectors we get by calculation that

$$
\Delta_1 = \frac{1}{1-\kappa^2}\left(\frac{2\,\kappa}{x\cdot m} - (1+\kappa^2)\right), \qquad \Delta_2 = \frac{1}{1-\kappa^2}\left(1 - 2\kappa\,x\cdot m + \kappa^2\right),
$$

[1] Fresnel equations are generally written in terms of angles of incidence. The formulation here is more convenient for our purposes.

i.e., Δ_i depend only on $x \cdot m$. So

$$\mathcal{R}_A = \left(\frac{|\mathbf{E}_A^r|}{|\mathbf{E}_A^i|}\right)^2 = \frac{(R_\parallel^A)^2 + (R_\perp^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2}$$

$$= \frac{1}{(1-\kappa^2)^2}\left(\left[\frac{2\kappa}{x \cdot m} - (1+\kappa^2)\right]^2 \frac{(I_\parallel^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2}\right.$$

$$\left. + \left[1 - 2\kappa\, x \cdot m + \kappa^2\right]^2 \frac{(I_\perp^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2}\right),$$

a function of $x \cdot m$ only. We then have that if the energy of the wave striking at A is $\mathcal{E}$, then the amount of energy internally reflected at A inside the guide is $\mathcal{R}_A\, \mathcal{E}$.

We now follow the path of the reflected ray inside the guide until it strikes the boundary of the wave-guide at a point B on $x_3 = a$ as in Figure 4.1. Notice that the normal at B is $\nu_B = (0,0,1) = -\nu_A$, the new incident unit direction at B is $\mathbf{s}_B^i = \mathbf{s}^r$ given by (4.6) -the direction after reflection at A- and the angle θ_B^i between $\mathbf{s}_B^i$ and ν_B equals θ_i. Since the incident electromagnetic field striking at B are the fields coming after reflection from A we have

$$\mathbf{E}_B^i(\mathbf{r},t) = \mathbf{E}_A^r(\mathbf{r},t) \qquad \mathbf{H}_B^i(\mathbf{r},t) = \mathbf{H}_A^r(\mathbf{r},t). \tag{4.9}$$

Now the incident field at B (perpendicular to $\mathbf{s}_B^i$) is

$$\mathbf{E}_B^i(\mathbf{r},t) = \left(I_\parallel^B \cos\theta_i, I_\perp^B, I_\parallel^B \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}_B^i}{v_1}\right)\right),$$

and

$$\mathbf{H}_B^i(\mathbf{r},t) = \frac{1}{v_1}\mathbf{s}_B^i \times \mathbf{E}_B^i = \frac{1}{v_1}\left(-I_\perp^B \cos\theta_i, I_\parallel^B, -I_\perp^B \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}_B^i}{v_1}\right)\right).$$

So from (4.9) and (4.8) we get

$$\mathbf{E}_B^i(\mathbf{r},t) = \left(I_\parallel^B \cos\theta_i, I_\perp^B, I_\parallel^B \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}_B^i}{v_1}\right)\right)$$

$$= \left(-R_{\parallel}^{A}\cos\theta_r, R_{\perp}^{A}, R_{\parallel}^{A}\sin\theta_r\right)\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right),$$

$$\begin{aligned}
\mathbf{H}_B^i(\mathbf{r},t) &= \frac{1}{v_1}\left(-I_{\perp}^{B}\cos\theta_i, I_{\parallel}^{B}, -I_{\perp}^{B}\sin\theta_i\right)\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}_B^i}{v_1}\right)\right)\\
&= \frac{1}{v_1}\left(R_{\perp}^{A}\cos\theta_r, R_{\parallel}^{A}, -R_{\perp}^{A}\sin\theta_r\right)\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right).
\end{aligned}$$

The internally reflected wave at B has direction $\mathbf{s}_B^r$ with $\mathbf{s}_B^r = \mathbf{s}^i = -\sin\theta_i\,\mathbf{i} - \cos\theta_i\,\mathbf{k}$ and is then given by

$$\begin{aligned}
\mathbf{E}_B^r(\mathbf{r},t) &= \left(-R_{\parallel}^{B}\cos\theta_i, R_{\perp}^{B}, R_{\parallel}^{B}\sin\theta_i\right)\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}_B^r}{v_1}\right)\right) \qquad (4.10)\\
&= \mathbf{E}_B^r\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right),\\
\mathbf{H}_B^r(\mathbf{r},t) &= \frac{1}{v_1}\,\mathbf{s}^i\times\mathbf{E}_B^r(\mathbf{r},t)\\
&= \frac{1}{v_1}\left(R_{\perp}^{B}\cos\theta_i, R_{\parallel}^{B}, -R_{\perp}^{B}\sin\theta_i\right)\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}_B^r}{v_1}\right)\right)\\
&= \mathbf{H}_B^r\,\cos\left(\omega\left(t-\frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right),
\end{aligned}$$

which from the Fresnel formulas at B

$$\begin{aligned}
R_{\parallel}^{B} &= \frac{(\kappa\,x-m)\cdot(x-\kappa\,m)}{(\kappa\,x+m)\cdot(x-\kappa\,m)}\,I_{\parallel}^{B}\\
R_{\perp}^{B} &= \frac{(x-\kappa\,m)\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)}\,I_{\perp}^{B},
\end{aligned}$$

where $x = \mathbf{s}_B^i = \mathbf{s}^r$ and $m = \mathbf{s}_B^r = \mathbf{s}^i$. We have $I_{\parallel}^{B} = R_{\parallel}^{A}$ and $I_{\perp}^{B} = R_{\perp}^{A}$. Also from Fresnel formulas at A

$$\begin{aligned}
R_{\parallel}^{A} &= \frac{(\kappa\,x-m)\cdot(x-\kappa\,m)}{(\kappa\,x+m)\cdot(x-\kappa\,m)}\,I_{\parallel}^{A}\\
R_{\perp}^{A} &= \frac{(x-\kappa\,m)\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)}\,I_{\perp}^{A},
\end{aligned}$$

with $x = \mathbf{s}^i$ and $m = \mathbf{s}^r$.

Since $\mathbf{s}_B^i \cdot \mathbf{s}_B^r = \mathbf{s}^r \cdot \mathbf{s}^i$, we therefore, obtain

$$R_\parallel^B = \left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 I_\parallel^A$$
$$R_\perp^B = \left[\frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\right]^2 I_\perp^A$$

with $x = \mathbf{s}^i$ and $m = \mathbf{s}^r$. So the percentage of energy internally reflected at B is

$$\begin{aligned}
\mathcal{R}_B = \left(\frac{|\mathbf{E}_B^r|}{|\mathbf{E}_B^i|}\right)^2 &= \frac{(R_\parallel^B)^2 + (R_\perp^B)^2}{(I_\parallel^B)^2 + (I_\perp^B)^2} \\
&= \left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 \frac{(I_\parallel^B)^2}{(I_\parallel^B)^2 + (I_\perp^B)^2} \\
&\quad + \left[\frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\right]^2 \frac{(I_\perp^B)^2}{(I_\parallel^B)^2 + (I_\perp^B)^2} \\
&= \frac{1}{(1-\kappa^2)^2}\left(\left[\frac{2\kappa}{x\cdot m} - (1+\kappa^2)\right]^2 \frac{(R_\parallel^A)^2}{(R_\parallel^A)^2 + (R_\perp^A)^2}\right. \\
&\quad \left. + \left[1 - 2\kappa\, x\cdot m + \kappa^2\right]^2 \frac{(R_\perp^A)^2}{(R_\parallel^A)^2 + (R_\perp^A)^2}\right),
\end{aligned}$$

and

$$\begin{aligned}
&\frac{(R_\parallel^A)^2}{(R_\parallel^A)^2 + (R_\perp^A)^2} \\
&= \frac{\left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 (I_\parallel^A)^2}{\left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 (I_\parallel^A)^2 + \left[\frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\right]^2 (I_\perp^A)^2} \\
&= \frac{\left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 (I_\parallel^A)^2}{\mathcal{R}_A\left((I_\parallel^A)^2 + (I_\perp^A)^2\right)}
\end{aligned}$$

$$\begin{aligned}
&\frac{(R_\perp^A)^2}{(R_\parallel^A)^2 + (R_\perp^A)^2} \\
&= \frac{\left[\frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\right]^2 (I_\perp^A)^2}{\left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 (I_\parallel^A)^2 + \left[\frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\right]^2 (I_\perp^A)^2}
\end{aligned}$$

$$= \frac{\left[\dfrac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^2 \; (I_\perp^A)^2}{\mathcal{R}_A\, \left((I_\parallel^A)^2 + (I_\perp^A)^2\right)}.$$

Then

$$\mathcal{R}_B\, \mathcal{R}_A = \left[\frac{(\kappa\, x-m)\cdot(x-\kappa\, m)}{(\kappa\, x+m)\cdot(x-\kappa\, m)}\right]^4 \; \frac{(I_\parallel^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2} + \left[\frac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^4 \; \frac{(I_\perp^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2},$$

where $x = \mathbf{s}^i$ and $m = \mathbf{s}^r$. Hence if the energy of the wave striking A is $\mathcal{E}$, then the energy internally reflected by the guide at the point B equals $\mathcal{R}_B\, \mathcal{R}_A\, \mathcal{E}$.

We now continue in this way and follow the reflected ray inside the guide until it strikes the boundary of the wave-guide at a point C on $x_3 = -a$. The incident direction at C is $\mathbf{s}_C^i = \mathbf{s}^i$ in (4.2), and the reflected direction at C is $\mathbf{s}_C^r = \mathbf{s}^r$ in (4.6); see Figure 4.1. The incident fields striking at C are the fields coming after reflection from B, that is,

$$\mathbf{E}_C^i(\mathbf{r}, t) = \mathbf{E}_B^r(\mathbf{r}, t) \qquad \mathbf{H}_C^i(\mathbf{r}, t) = \mathbf{H}_B^r(\mathbf{r}, t).$$

We then have from (4.10)

$$\begin{aligned}\mathbf{E}_C^i(\mathbf{r}, t) &= \left(-I_\parallel^C \cos\theta_i, I_\perp^C, I_\parallel^C \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}_C^i}{v_1}\right)\right) \\ &= \left(-R_\parallel^B \cos\theta_i, R_\perp^B, R_\parallel^B \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right),\end{aligned}$$

$$\begin{aligned}\mathbf{H}_C^i(\mathbf{r}, t) &= \frac{1}{v_1}\mathbf{s}^i \times \mathbf{E}_C^i(\mathbf{r}, t) = \frac{1}{v_1}\left(I_\perp^C \cos\theta_i, I_\parallel^C, -I_\perp^C \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}_C^i}{v_1}\right)\right) \\ &= \frac{1}{v_1}\left(R_\perp^B \cos\theta_i, R_\parallel^B, -R_\perp^B \sin\theta_i\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^i}{v_1}\right)\right).\end{aligned}$$

The internally reflected wave at C has direction $\mathbf{s}_C^r = \mathbf{s}^r$ and is given by

$$\mathbf{E}_C^r(\mathbf{r},t) = \left(-R_\parallel^C \cos\theta_r, R_\perp^C, R_\parallel^C \sin\theta_r\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}_C^r}{v_1}\right)\right) \tag{4.11}$$

$$= \mathbf{E}_C^r \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right)$$

$$\mathbf{H}_C^r(\mathbf{r},t) = \frac{1}{v_1}\,\mathbf{s}^r \times \mathbf{E}_C^r(\mathbf{r},t) \tag{4.12}$$

$$= \frac{1}{v_1}\left(R_\perp^C \cos\theta_r, R_\parallel^C, -R_\perp^C \sin\theta_r\right) \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right)$$

$$= \mathbf{H}_C^r \cos\left(\omega\left(t - \frac{\mathbf{r}\cdot\mathbf{s}^r}{v_1}\right)\right),$$

with from the Fresnel formulas at C

$$R_\parallel^C = \frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\, I_\parallel^C$$

$$R_\perp^C = \frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\, I_\perp^C,$$

where $x = \mathbf{s}_C^i = \mathbf{s}^i$ and $m = \mathbf{s}_C^r = \mathbf{s}^r$. We have $I_\parallel^C = R_\parallel^B$ and $I_\perp^C = R_\perp^B$ and from Fresnel formulas at B

$$R_\parallel^B = \frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\, I_\parallel^B$$

$$R_\perp^B = \frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\, I_\perp^B,$$

where $x = \mathbf{s}_B^i = \mathbf{s}^r$ and $m = \mathbf{s}_B^r = \mathbf{s}^i$. As before and since Δ_1 and Δ_2 depend only on $x \cdot m$ and $\mathbf{s}_C^i \cdot \mathbf{s}_C^r = \mathbf{s}_B^r \cdot \mathbf{s}_B^i = \mathbf{s}^i \cdot \mathbf{s}^r$, we obtain

$$R_\parallel^C = \left[\frac{(\kappa\, x - m)\cdot(x - \kappa\, m)}{(\kappa\, x + m)\cdot(x - \kappa\, m)}\right]^2 I_\parallel^B$$

$$R_\perp^C = \left[\frac{(x - \kappa\, m)\cdot(x - \kappa\, m)}{(x + \kappa\, m)\cdot(x - \kappa\, m)}\right]^2 I_\perp^B.$$

So

$$\mathcal{R}_C = \left(\frac{|\mathbf{E}_C^r|}{|\mathbf{E}_C^i|}\right)^2 = \frac{(R_\parallel^C)^2 + (R_\perp^C)^2}{(I_\parallel^C)^2 + (I_\perp^C)^2} = \frac{(R_\parallel^C)^2 + (R_\perp^C)^2}{(R_\parallel^B)^2 + (R_\perp^B)^2}$$

$$
\begin{aligned}
&= \left[\frac{(\kappa\, x - m)\cdot(x-\kappa\, m)}{(\kappa\, x + m)\cdot(x-\kappa\, m)}\right]^4 \frac{(I_{\parallel}^B)^2}{(R_{\parallel}^B)^2+(R_{\perp}^B)^2} \\
&\qquad + \left[\frac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^4 \frac{(I_{\perp}^B)^2}{(R_{\parallel}^B)^2+(R_{\perp}^B)^2} \\
&= \frac{\Delta_1^4\,(I_{\parallel}^B)^2+\Delta_2^4\,(I_{\perp}^B)^2}{(R_{\parallel}^B)^2+(R_{\perp}^B)^2} = \frac{1}{\mathcal{R}_B}\,\frac{\Delta_1^4\,(I_{\parallel}^B)^2+\Delta_2^4\,(I_{\perp}^B)^2}{(I_{\parallel}^B)^2+(I_{\perp}^B)^2} \\
&= \frac{1}{\mathcal{R}_B}\Bigg(\left[\frac{(\kappa\, x - m)\cdot(x-\kappa\, m)}{(\kappa\, x + m)\cdot(x-\kappa\, m)}\right]^4 \frac{(R_{\parallel}^A)^2}{(R_{\parallel}^A)^2+(R_{\perp}^A)^2} \\
&\qquad + \left[\frac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^4 \frac{(R_{\perp}^A)^2}{(R_{\parallel}^A)^2+(R_{\perp}^A)^2}\Bigg) \\
&= \frac{1}{\mathcal{R}_B\,\mathcal{R}_A}\Bigg(\left[\frac{(\kappa\, x - m)\cdot(x-\kappa\, m)}{(\kappa\, x + m)\cdot(x-\kappa\, m)}\right]^6 \frac{(I_{\parallel}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2} \\
&\qquad + \left[\frac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^6 \frac{(I_{\perp}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2}\Bigg)
\end{aligned}
$$

and therefore, we get

$$
\begin{aligned}
\mathcal{R}_C\,\mathcal{R}_B\,\mathcal{R}_A = \Bigg(&\left[\frac{(\kappa\, x - m)\cdot(x-\kappa\, m)}{(\kappa\, x + m)\cdot(x-\kappa\, m)}\right]^6 \frac{(I_{\parallel}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2} + \\
&\left[\frac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^6 \frac{(I_{\perp}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2}\Bigg),
\end{aligned}
$$

where $x = \mathbf{s}^i$ and $m = \mathbf{s}^r$.

In general and continuing with this process, if we have a sequence of points $A_1, \cdots, A_N$ along which the wave zig-zags, then

$$
\begin{aligned}
\prod_{j=1}^{N} \mathcal{R}_{A_j} = \Bigg(&\left[\frac{(\kappa\, x - m)\cdot(x-\kappa\, m)}{(\kappa\, x + m)\cdot(x-\kappa\, m)}\right]^{2N} \frac{(I_{\parallel}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2} \\
&+ \left[\frac{(x-\kappa\, m)\cdot(x-\kappa\, m)}{(x+\kappa\, m)\cdot(x-\kappa\, m)}\right]^{2N} \frac{(I_{\perp}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2}\Bigg),
\end{aligned} \tag{4.13}
$$

with $\mathcal{R}_{A_j} = \left(\dfrac{|\mathbf{E}^r_{A_j}|}{|\mathbf{E}^i_{A_j}|}\right)^2$, $A_1 = A$, and $x = \mathbf{s}^i$ and $m = \mathbf{s}^r$. Therefore, if the wave has energy $\mathcal{E}$ when it strikes A_1, then the amount of energy internally reflected at the last point A_N of the zig-zag equals $\prod_{j=1}^{N} \mathcal{R}_{A_j}\,\mathcal{E}$.

Let us rewrite this percentage in terms of the incident direction $\mathbf{s}^i$ at A_1 and the normal to the guide at A_1. We have

$$\frac{(\kappa\,x - m)\cdot(x-\kappa\,m)}{(\kappa\,x + m)\cdot(x-\kappa\,m)} = \frac{1}{1-\kappa^2}\,\left(\frac{2\kappa}{x\cdot m} - (1+\kappa^2)\right),$$

and

$$\frac{(x-\kappa\,m)\cdot(x-\kappa\,m)}{(x+\kappa\,m)\cdot(x-\kappa\,m)} = \frac{1}{1-\kappa^2}\,\left(1-2\kappa\,x\cdot m+\kappa^2\right).$$

From the Snell law $x-\kappa\,m=\lambda\,\nu$, where ν is the outer unit normal and

$$\lambda = \phi(x\cdot\nu),$$

with

$$\phi(t) = t-\kappa\,\sqrt{1-\kappa^{-2}(1-t^2)}.$$

Here we assume $x\cdot\nu \geq \sqrt{1-\kappa^2}$, that is equivalent to say that the angle of incidence is smaller than the critical angle $\theta_c = \arcsin\kappa$. Therefore,

$$x\cdot m = \frac{1}{\kappa}\,(1-\phi(x\cdot\nu)\,(x\cdot\nu)),$$

and we have

$$\frac{1}{1-\kappa^2}\,\left(\frac{2\kappa}{x\cdot m}-(1+\kappa^2)\right) = \frac{1}{1-\kappa^2}\,\left(\frac{2\kappa}{\frac{1}{\kappa}\,(1-\phi(x\cdot\nu)\,(x\cdot\nu))} - (1+\kappa^2)\right)$$

$$:= \Phi_1(x\cdot\nu) \tag{4.14}$$

and

$$\frac{1}{1-\kappa^2}\,\left(1-2\kappa\,x\cdot m+\kappa^2\right) = \frac{1}{1-\kappa^2}\,\left(1-2\kappa\,\left(\frac{1}{\kappa}\,(1-\phi(x\cdot\nu)\,(x\cdot\nu))\right)+\kappa^2\right)$$

$$:= \Phi_2(x \cdot \nu). \tag{4.15}$$

Then (4.13) can be written as

$$\prod_{j=1}^{N} \mathcal{R}_{A_j} = [\Phi_1(x\cdot\nu)]^{2N} \; \frac{(I_{\parallel}^A)^2}{(I_{\parallel}^A)^2 + (I_{\perp}^A)^2} + [\Phi_2(x\cdot\nu)]^{2N} \; \frac{(I_{\perp}^A)^2}{(I_{\parallel}^A)^2 + (I_{\perp}^A)^2}, \qquad A_1 = A. \tag{4.16}$$

Since

$$\mathbf{E}^i_{A_{j+1}} = \mathbf{E}^r_{A_j}$$

for $1 \leq j \leq N-1$, we get

$$\prod_{j=1}^{N} \mathcal{R}_{A_j} \tag{4.17}$$

$$= \left(\frac{|\mathbf{E}^r_{A_N}|}{|\mathbf{E}^i_{A_1}|} \right)^2 = \frac{1}{(I_{\parallel}^A)^2 + (I_{\perp}^A)^2} \; \Big([\Phi_1(x\cdot\nu)]^{2N} \; (I_{\parallel}^A)^2 + [\Phi_2(x\cdot\nu)]^{2N} \; (I_{\perp}^A)^2 \Big). \tag{4.18}$$

Hence

$$\mathcal{R}_{A_N} = \frac{\prod_{j=1}^{N} \mathcal{R}_{A_j}}{\prod_{j=1}^{N-1} \mathcal{R}_{A_j}} = \frac{[\Phi_1(x\cdot\nu)]^{2N} \; (I_{\parallel}^A)^2 + [\Phi_2(x\cdot\nu)]^{2N} \; (I_{\perp}^A)^2}{[\Phi_1(x\cdot\nu)]^{2(N-1)} \; (I_{\parallel}^A)^2 + [\Phi_2(x\cdot\nu)]^{2(N-1)} \; (I_{\perp}^A)^2}.$$

4.2 Circular Waveguide

Let $0 < R_1 < R_2$. We consider a wave guide having a circular section:

$$W = \{(x,y,z) : -\infty < x \leq 0, \; R_1 \leq y \leq R_2, \; -\infty < z < +\infty\}$$

$$\cup \{(x,y,z) : R_1^2 \leq x^2 + y^2 \leq R_2^2, \; x \geq 0, \; y \geq 0, \; -\infty < z < +\infty\}$$

$$\cup \{(x,y,z) : R_1 \leq x \leq R_2, \; -\infty \leq y \leq 0, \; -\infty < z < +\infty\}.$$

The interior of W has refractive index n_1, and the outside refractive index n_2. We assume that $n_1 > n_2$, and we let $\kappa = \dfrac{n_2}{n_1}$. We recall that if a surface S separates two media with refractive indices n_1 and n_2, with $n_1 > n_2$, and if a ray with unit direction x traveling in medium n_1 impinges S at a point P, where the unit normal at P going from medium n_1 to medium n_2 is denoted by ν, then the ray is refracted into medium n_2 into a ray having unit direction m with $x - \kappa m = \lambda \nu$, where $\lambda = x \cdot \nu - \kappa \sqrt{1 - \kappa^{-2} (1 - (x \cdot \nu)^2)}$. Since $\kappa < 1$, then total internal reflection occurs if and only if

$$x \cdot \nu \leq \sqrt{1 - \kappa^2}, \tag{4.19}$$

or in other words, the angle ψ between the unit vectors x and ν satisfies

$$\cos \psi \leq \sqrt{1 - \kappa^2}. \tag{4.20}$$

We also recall that the critical angle of refraction is $\theta_c = \arcsin \kappa$. In case of total internal reflection, the ray is internally reflected within medium n_1 and into the direction $m = x - 2\,(x \cdot \nu)\,\nu$.

We assume that we are having a plane wave traveling upwards in the straight lower part of the waveguide, and want to analyze the energy transmitted on the circular part. We are assuming the vector direction of propagation of the plane wave lies on the xy-plane. This wave is realized as a ray impinging the sides of the lower straight guide that is internally reflected inside the circular section of the guide. In other words, we assume that the ray forms an angle θ with either normal to the walls of the straight guide. We shall then analyze ray tracing and the energy internally reflected in four cases separately:

Case A (to be done in Section 4.2.1): the incoming ray is going towards the circle of radius of radius R_2 and it forms an angle $\theta \geq \theta_c$ with x-axis directed to $+\infty$;

Case B (to be done in Section 4.2.2): the incoming ray is going towards the circle of radius of radius R_1 and it forms an angle $\theta \geq \theta_c$ with x-axis directed to $-\infty$.

Case C (to be done in Section 4.2.3): the incoming ray is going towards the circle of radius of radius R_2 and it forms an angle $\theta < \theta_c$ with x-axis directed to $+\infty$;

Case D (to be done in Section 4.2.4): the incoming ray is going towards the circle of radius of radius R_1 and it forms an angle $\theta < \theta_c$ with x-axis directed to $-\infty$

The analysis of these four cases is not trivial and it requires a careful and somehow long geometrical argument to understand and quantify all possible configurations. The most complicated is Case A, the others somehow follow from this one. For convenience of the reader we present all details.

4.2.1 Analysis of Case A

Let $R_1 \leq x \leq R_2$, $\theta \geq \theta_c$ and let us assume

$$\kappa \leq \frac{R_1}{R_2}. \tag{4.21}$$

Since $\kappa < 1$ is fixed, this condition always holds when R_1 and R_2 are sufficiently close.

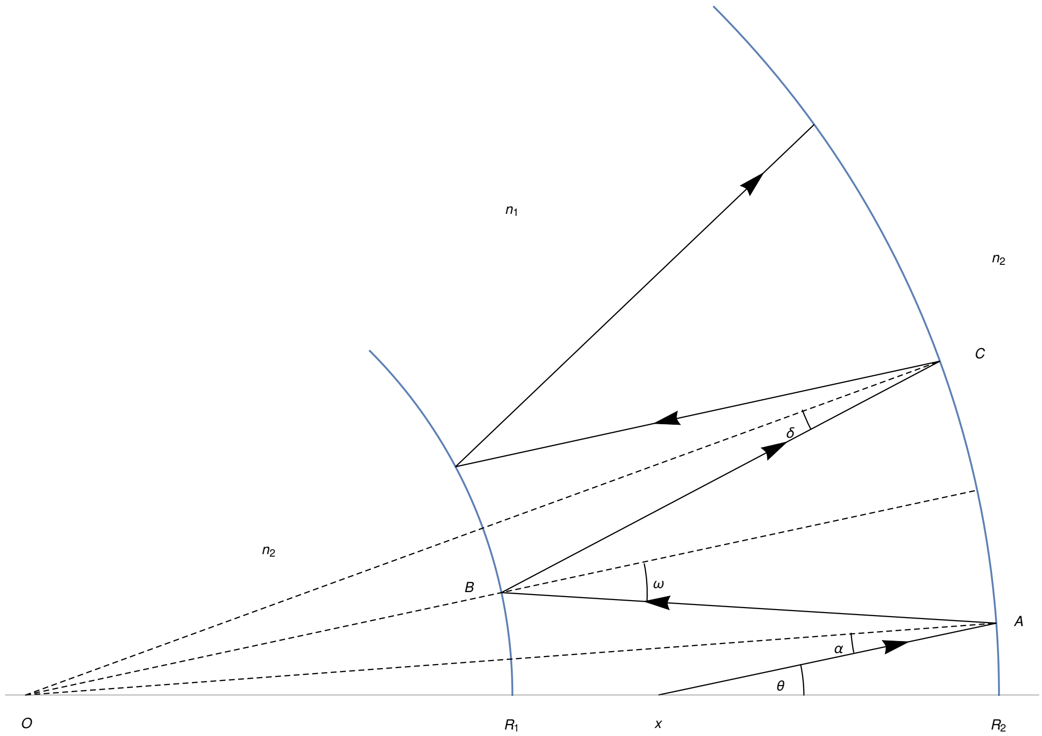

Figure 4.2: Case A and C

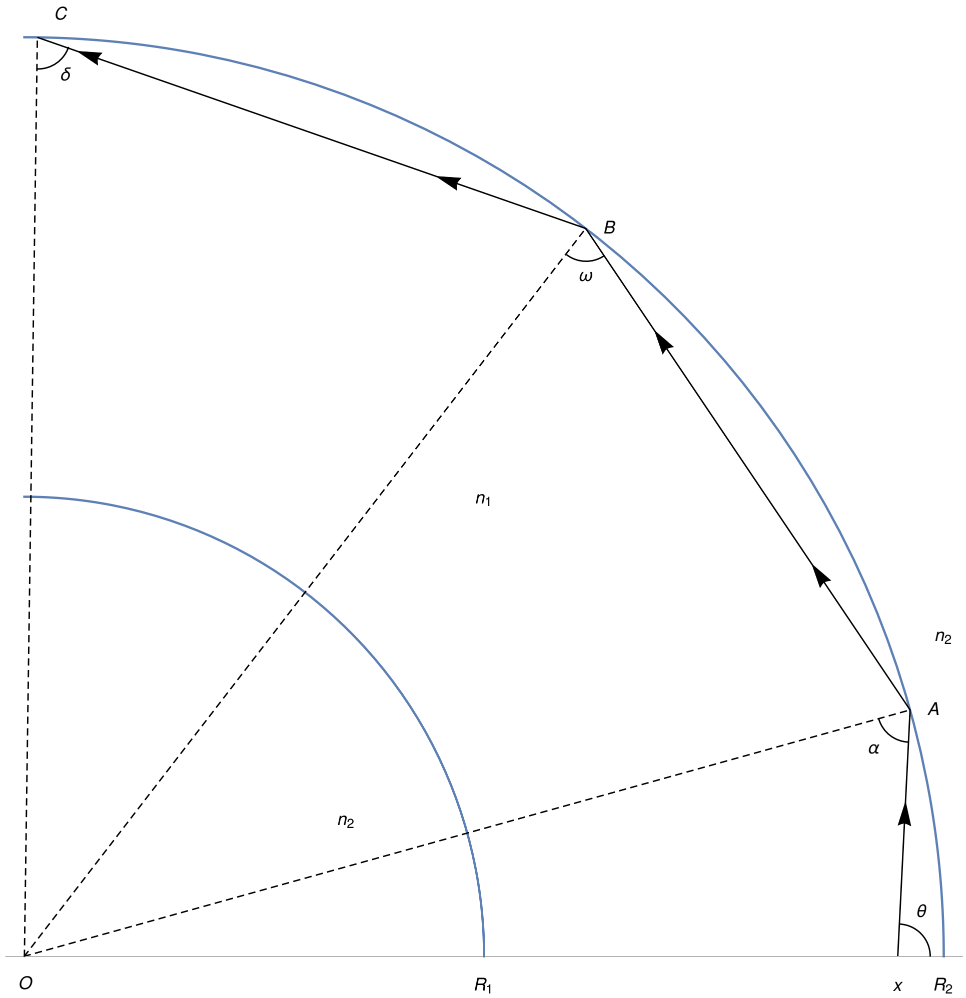

Figure 4.3: Whispering gallery

Ray tracing analysis at the first striking point

Consider a ray $\ell = \ell(x, \theta)$ that passes through $(x, 0)$, making an angle θ with the positive x-axis, and therefore, striking the circle with radius R_2 for the first time at a point A, see Figure 4.2. We denote by θ the angle AxR_2, and let α be the angle $0Ax$, that is, α is the angle of incidence at A. By the sine theorem we have

$$\frac{x}{\sin\alpha} = \frac{R_2}{\sin(\pi - \theta)}$$

which implies

$$\alpha = \arcsin\left(\frac{x}{R_2}\sin\theta\right). \tag{4.22}$$

Energy losses analysis at the first striking point

We first analyze when is there total internal reflection at A. From (4.20), total internal reflection at A occurs if and only if $\cos\alpha \leq \sqrt{1-\kappa^2}$. This means

$$\cos\alpha = \cos\left(\arcsin\left(\frac{x}{R_2}\sin\theta\right)\right) = \sqrt{1 - \left(\frac{x}{R_2}\right)^2 \sin^2\theta} \leq \sqrt{1-\kappa^2},$$

which is equivalent to

$$x \geq \frac{\kappa\, R_2}{\sin\theta}. \tag{4.23}$$

We see where is the point $\frac{\kappa\, R_2}{\sin\theta}$ located relative to the interval $[R_1, R_2]$. First, since $\theta \geq \theta_c = \arcsin\kappa$, we have that $\sin\theta \geq \kappa$, and so $\frac{\kappa\, R_2}{\sin\theta} \leq R_2$. Second, $R_1 \leq \frac{\kappa\, R_2}{\sin\theta}$ holds iff $\sin\theta \leq \kappa\frac{R_2}{R_1}$. Therefore,

$$\text{if } \sin\theta \leq \kappa\frac{R_2}{R_1}, \text{ then total internal reflection at } A \text{ occurs iff } x \in \left[\kappa\frac{R_2}{\sin\theta}, R_2\right]. \tag{4.24}$$

On the other hand,

$$\text{if } \sin\theta > \kappa\frac{R_2}{R_1}, \text{ then there is total internal reflection at } A \text{ for all } x \in [R_1, R_2]. \tag{4.25}$$

Summarizing, (see also Figure 4.4)

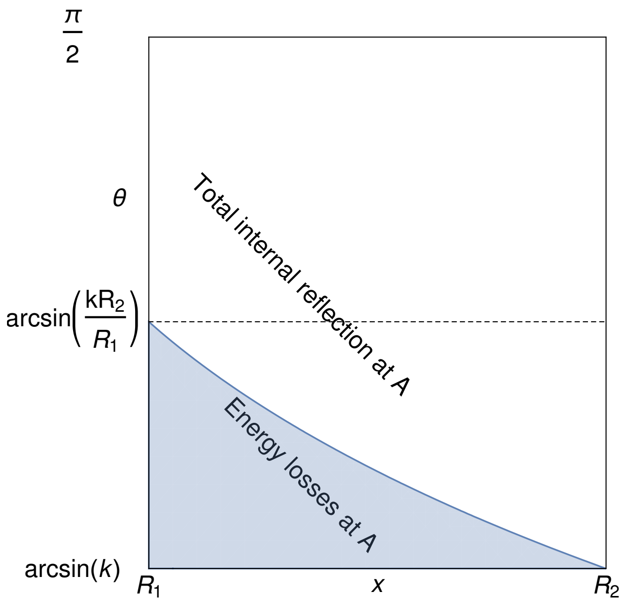

Figure 4.4: Total internal reflection region

(C1) If $\theta \in [\arcsin\kappa, \arcsin(\kappa R_2/R_1)]$, then we have total internal reflection at A if and only if $x \in \left[\kappa\frac{R_2}{\sin\theta}, R_2\right]$.

(C2) If $\theta \in (\arcsin(\kappa R_2/R_1), \pi/2]$, then we have total internal reflection at A for all $x \in [R_1, R_2]$.

Ray tracing analysis at the second striking point

We next analyze the trajectory of the ray ℓ to see where strikes next. After being reflected at the point A, the ray ℓ continues traveling and either strikes the inner circle C_1 (see Figure 4.2) or the outer circle C_2 (see Figure 4.3) at a point denoted by B. From the point A the angle α for which the ray ℓ is possibly tangential to the smaller circle is α_T with $\sin\alpha_T = \dfrac{R_1}{R_2}$. So ℓ strikes the smaller circle iff $\alpha \le \alpha_T$ and ℓ strikes the larger circle iff $\alpha > \alpha_T$. Suppose $\alpha \le \alpha_T$. So $\sin\alpha \le \dfrac{R_1}{R_2}$, and from (4.22) we get $\sin\theta \le \dfrac{R_1}{x}$, that is, the ray ℓ strikes the circle C_1 if and only if $x \le \dfrac{R_1}{\sin\theta}$. Suppose that $R_2 \le \dfrac{R_1}{\sin\theta}$, that is, $\sin\theta \le \dfrac{R_1}{R_2}$. Recall that $x \le R_2$. So if $\sin\theta \le \dfrac{R_1}{R_2}$, then the ray ℓ strikes C_1 for all $x \in [R_1, R_2]$. On the other hand, if $\dfrac{R_1}{\sin\theta} < R_2$, that is, when $\dfrac{R_1}{R_2} < \sin\theta$, then ℓ strikes C_1 for $x \in \left[R_1, \dfrac{R_1}{\sin\theta}\right]$ and ℓ strikes C_2, when $x \in \left(\dfrac{R_1}{\sin\theta}, R_2\right]$. Summarizing, (see also Figure 4.5)

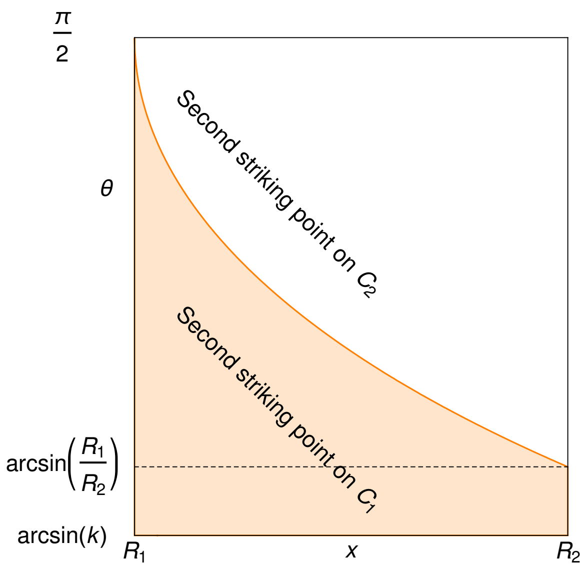

Figure 4.5: Second striking point region

(C3) If $\theta \in [\arcsin \kappa, \arcsin(R_1/R_2)]$, then ℓ strikes C_1 for all $x \in [R_1, R_2]$.

(C4) If $\theta \in (\arcsin(R_1/R_2), \pi/2]$, then ℓ strikes C_1 for $x \in \left[R_1, \dfrac{R_1}{\sin\theta}\right]$, and ℓ strikes C_2 when $x \in \left(\dfrac{R_1}{\sin\theta}, R_2\right]$.

Combining (C1), (C2), (C3) and (C4) we get the following Figures (4.6, 4.7) and summary:

Case 1. See Figure 4.6. Suppose $\kappa \leq \left(\dfrac{R_1}{R_2}\right)^2$. This means $\kappa\dfrac{R_2}{R_1} \leq \dfrac{R_1}{R_2}$. We analyze what happens when θ belongs to various intervals.

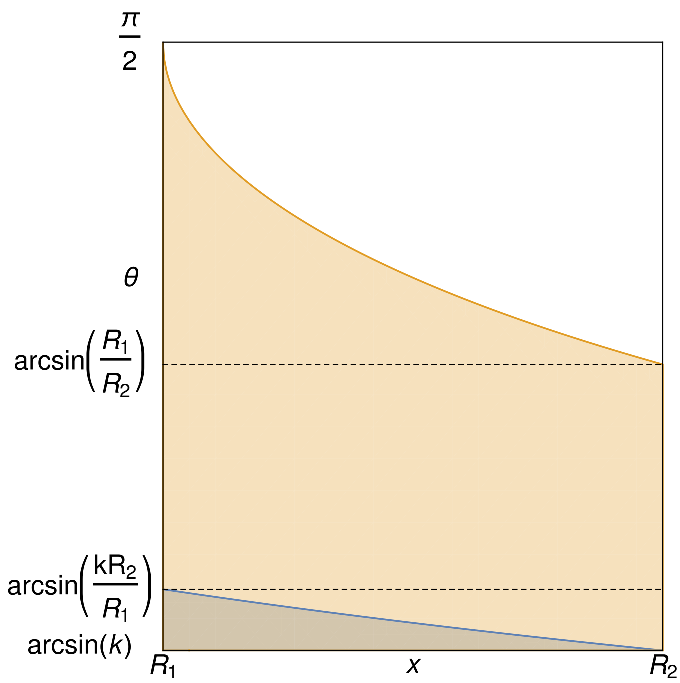

Figure 4.6: Case 1

Suppose $\kappa \leq \sin\theta \leq \kappa\dfrac{R_2}{R_1}$. From (C3) ℓ strikes C_1 for all $x \in [R_1, R_2]$, and from (4.24) there is total internal reflection at A for all $x \in \left[\kappa\dfrac{R_2}{\sin\theta}, R_2\right]$.

Suppose $\kappa\dfrac{R_2}{R_1} < \sin\theta \leq \dfrac{R_1}{R_2}$. From (C3) all rays strike C_1 and from (4.25) there is total internal reflection at A for all $x \in [R_1, R_2]$.

Suppose $\dfrac{R_1}{R_2} < \sin\theta$. From (C4) ℓ strikes C_1 for $x \in \left[R_1, \dfrac{R_1}{\sin\theta}\right]$ and strikes C_2 for $x \in \left(\dfrac{R_1}{\sin\theta}, R_2\right]$, and from (4.25) there is total internal reflection at A for all $x \in [R_1, R_2]$.

Case 2. See Figure 4.7. Suppose $\left(\dfrac{R_1}{R_2}\right)^2 < \kappa \leq \dfrac{R_1}{R_2}$. This implies $\kappa\dfrac{R_2}{R_1} > \dfrac{R_1}{R_2}$.

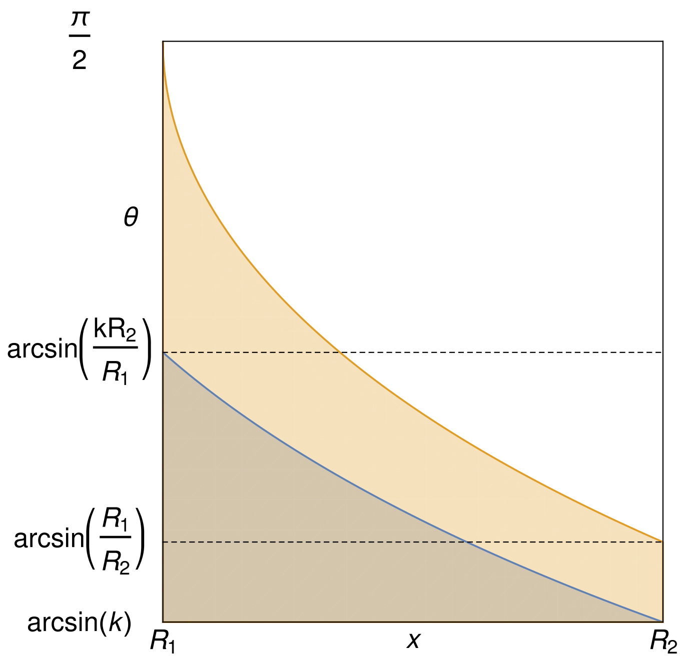

Figure 4.7: Case 2

Suppose $\kappa \leq \sin\theta \leq \dfrac{R_1}{R_2}$. From (C3) ℓ strikes C_1 for all $x \in [R_1, R_2]$, and from (4.24) there is total internal reflection at A for all $x \in \left[\kappa\dfrac{R_2}{\sin\theta}, R_2\right]$.

Suppose $\dfrac{R_1}{R_2} < \sin\theta \leq \kappa\dfrac{R_2}{R_1}$. By (C4), ℓ strikes C_1 for $x \in \left[R_1, \dfrac{R_1}{\sin\theta}\right]$, and ℓ strikes C_2 when $x \in \left(\dfrac{R_1}{\sin\theta}, R_2\right]$. From (4.21), we have $\kappa\dfrac{R_2}{\sin\theta} \leq \dfrac{R_1}{\sin\theta}$. From (4.24), total internal reflection at A occurs iff $x \in \left[\kappa \frac{R_2}{\sin\theta}, R_2\right]$. Also if $x \in \left[\kappa \frac{R_2}{\sin\theta}, \frac{R_1}{\sin\theta}\right]$, then ℓ strikes C_1; and if $x \in \left(\frac{R_1}{\sin\theta}, R_2\right]$, then ℓ strikes C_2.

Suppose $\kappa \frac{R_2}{R_1} < \sin\theta \leq 1$. From (C4), ℓ strikes C_1 when $x \in \left[R_1, \frac{R_1}{\sin\theta}\right]$, and ℓ strikes C_2 when $x \in \left(\frac{R_1}{\sin\theta}, R_2\right]$, and from (4.25) there is total internal reflection at A for all $x \in [R_1, R_2]$.

Energy losses analysis at the second striking point

Let us analyze what happens at B with internal reflection.

Suppose $B \in C_1$. From (C3) and (C4) this happens only when $\sin\theta \leq R_1/R_2$ and for each $x \in [R_1, R_2]$ or when $\sin\theta > R_1/R_2$ and for each $x \in [R_1, R_1/\sin\theta]$. Let ω be the angle of incidence at B, then the angle OBA is $\pi - \omega$ and the angle $BA0$ is α. By the sine theorem

$$\frac{R_2}{\sin(\pi - \omega)} = \frac{R_2}{\sin\omega} = \frac{R_1}{\sin\alpha},$$

and so from (4.22)

$$\omega = \arcsin\left(\frac{R_2}{R_1}\sin\alpha\right) = \arcsin\left(\frac{R_2}{R_1}\frac{x}{R_2}\sin\theta\right) = \arcsin\left(\frac{x}{R_1}\sin\theta\right). \tag{4.26}$$

Notice that to have the function arcsin well defined we need $\frac{x}{R_1}\sin\theta \leq 1$, which holds because of the assumption that $B \in C_1$. Then

$$\cos\omega = \sqrt{1 - \left(\frac{x}{R_1}\right)^2 \sin^2\theta}.$$

We have that $\sqrt{1 - \left(\frac{x}{R_1}\right)^2 \sin^2\theta} \leq \sqrt{1-\kappa^2}$ is equivalent to $x \geq \frac{\kappa R_1}{\sin\theta}$. Since $x \geq R_1$ and $\kappa/\sin\theta \leq 1$, such inequality always holds and therefore, there is total internal reflection at B for any x for which B is on C_1 (regardless of weather or not there is total internal reflection at A).

Next suppose $B \in C_2$. From (C4), this happens when $\sin\theta > R_1/R_2$ and $x \in (R_1/\sin\theta, R_2]$. Let ω be the angle $0BA$ and let α be the angle $0AB$. By the sine theorem $\dfrac{R_2}{\sin\alpha} = \dfrac{R_2}{\sin\omega}$, and therefore, $\omega = \alpha$. From (4.21) $\dfrac{R_1}{\sin\theta} \geq \kappa\dfrac{R_2}{\sin\theta}$. Since $x > R_1/\sin\theta$, it follows that x satisfies (4.23) and therefore, there is total internal reflection at B. Summarizing,

(C5) Regardless of weather $B \in C_1$ or $B \in C_2$, there is total internal reflection at B for each $x \in [R_1, R_2]$.

Ray tracing and energy losses analysis at the third and successive striking points

We now analyze the next striking point. If $B \in C_1$, then the ray ℓ will continue traveling and will strike C_2 at some point C, as in Figure 4.2. If $B \in C_2$, then the ray ℓ will continue traveling, it will never strike C_1 again and it will strike C_2 at some point C, as in Figure 4.3.

Let us first analyze the case when $B \in C_1$. The ray ℓ then continues traveling and strikes C_2 at some point C. We show that there is total internal reflection at C only if (4.23) holds. In fact, let δ be the angle $0CB$, we have that $0BC$ equals $\pi - \omega$. Again by the sine theorem

$$\frac{R_1}{\sin\delta} = \frac{R_2}{\sin\omega}$$

so from (4.26)

$$\delta = \arcsin\left(\frac{R_1}{R_2}\sin\omega\right) = \arcsin\left(\frac{x}{R_2}\sin\theta\right).$$

Therefore,

$$\cos\delta = \sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} \leq \sqrt{1-\kappa^2}$$

if and only if (4.23) holds.

Suppose $B \in C_2$. Let δ be the angle $0CB$ and ω be the angle CBO. By the sine theorem $R_2/\sin\delta = R_2/\sin\omega$ and so $\delta = \omega$. Therefore, the incident angle at C equals the incident angle at A since $\omega = \alpha$, and there is total internal reflection at C if and only if there is total internal reflection at A. Now the analysis of the striking points follows the same pattern, in other words, the angles of incidence repeat themselves. Summarizing, *under* (4.23) *there are no losses at both all the striking points on the inner circle and all the striking points on the outer circle.* We remark that the points B, C, etc. can all be on the outer circle C_2. From (C4) this is the case when $\dfrac{R_1}{R_2} < \sin\theta$, and $x \in \left(\dfrac{R_1}{\sin\theta}, R_2\right]$. In fact, in this case the striking points are all on the outer circle and have the form

$$R_2\left(\cos\left(\theta-(2N+1)\alpha+N\pi\right), \sin\left(\theta-(2N+1)\alpha+N\pi\right)\right), \quad N = 0, 1, \cdots.$$

This sequence of points resembles the points on a whispering gallery, see [35, Section 287, p. 115].

Next we consider the complementary case to (4.23), that is, when

$$R_1 \leq x < \frac{\kappa R_2}{\sin\theta}. \tag{4.27}$$

Let us first analyze internal reflection at the first striking point A. Since the angle of incidence at A is α, then from (4.22) we have

$$\cos\alpha = \sqrt{1 - \left(\frac{x}{R_2}\right)^2 \sin^2\theta}$$

which from (4.27) is greater than or equal to $\sqrt{1-\kappa^2}$ and therefore, there is no total internal reflection at A and *so there are losses at A.* At the next striking point B it was shown before that if $B \in C_1$, then there is total reflection at B. On the other hand, if $B \in C_2$, then from (C4) and (C3) this happens only when $x > R_1/\sin\theta$. Then from (4.27) we must have $\dfrac{\kappa R_2}{\sin\theta} > x > \dfrac{R_1}{\sin\theta}$, which is impossible from the assumption (4.21), and therefore, B cannot be on C_2. Then at the next point C we must have $C \in C_2$. We will show that there is no total internal reflection at C. In fact, let δ be the angle $0CB$, we have that $0BC$ equals $\pi - \omega$. Again by the sine theorem

$$\frac{R_1}{\sin\delta} = \frac{R_2}{\sin\omega}$$

so from (4.26)

$$\delta = \arcsin\left(\frac{R_1}{R_2}\sin\omega\right) = \arcsin\left(\frac{x}{R_2}\sin\theta\right) = \alpha.$$

Therefore, from (4.27)

$$\cos\delta = \sqrt{1 - \left(\frac{x}{R_2}\right)^2 \sin^2\theta} \geq \sqrt{1-\kappa^2}$$

implying that *there are losses at C.* In summary, *under* (4.27) *there are no losses at all the striking points on the inner circle and there are losses at all the striking points on the outer circle.* We will calculate now a formula for the points where the ray ℓ strike the outer circle under the assumption (4.27). The first point on C_2 is A and we have $A = R_2\,(\cos(\theta-\alpha), \sin(\theta-\alpha))$. The next striking point B is on the circle C_1 and the radius $0B$ forms with the positive x-axis an angle $\delta' + \theta - \alpha$, where $\delta' = \omega - \alpha$ with ω being the incident angle at B given by (4.26). The next point C is on C_2 and has the form

$$C = R_2\,\left(\cos(\theta + 2\omega - 3\alpha), \sin(\theta + 2\omega - 3\alpha)\right).$$

The next striking point is on C_1 and after that the following striking point is on C_2 and has the form

$$\begin{aligned} &R_2\,\big(\cos(4\delta' + \theta - \alpha), \sin(4\delta' + \theta - \alpha)\big) \\ &= R_2\,\left(\cos(\theta + 4\omega - 5\alpha), \sin(\theta + 4\omega - 5\alpha)\right), \end{aligned}$$

since the δ' repeats itself. Continuing in this way we get that the Nth-striking point on the outer circle is given by

$$P_N = R_2\,\left(\cos(\theta + 2N\omega - (2N+1)\alpha), \sin(\theta + 2N\omega - (2N+1)\alpha)\right), \quad N = 0, 1, \cdots \tag{4.28}$$

where α is given by (4.22) and ω given by (4.26). The angle of incidence at P_N is α for $N \geq 0$.

Summarizing,

(L1) If $x \in \left[\dfrac{\kappa\,R_2}{\sin\theta}, R_2\right]$, there are no losses at both all the striking points on the inner circle and all the striking points on the outer circle.

(L2) If $x \in \left[R_1, \dfrac{\kappa\,R_2}{\sin\theta}\right)$ there are no losses at all the striking points on the inner circle and there are losses at all the striking points on the outer circle.

Notice that if $\theta > \arcsin\left(\kappa\frac{R_2}{R_1}\right)$, then $\left[R_1, \dfrac{\kappa R_2}{\sin\theta}\right)$ is empty.

4.2.2 Analysis of Case B

The analysis is similar to Case A (Section 4.2.1) and we obtain the same results (L1) and (L2).

Ray tracing and energy losses analysis at the first striking point. Consider a ray ℓ that passes through $(x, 0)$, $R_1 \leq x \leq R_2$, making an angle $\theta \geq \theta_c$ with the x-axis going to $-\infty$. There are two possibilities: the ray first strikes C_1 or it first strikes C_2. Notice that the ray can be tangential to C_1, in fact, this happens at $\theta_T(x)$ with $\sin\theta_T(x) = \dfrac{R_1}{x}$. This is possible for $\theta_T \geq \theta_c$ because from (4.21), $\dfrac{R_1}{\kappa} \geq R_2 \geq x$. If the ray passing through $(x, 0)$ has an angle $\theta > \theta_T(x)$, then the ray does not strike C_1 and strikes C_2. On the other hand, if the angle $\kappa \leq \theta \leq \theta_T(x)$ then the ray strikes C_1 at a point A. This means that $x \leq \dfrac{R_1}{\sin\theta}$, and in this case, the ray is reflected at A with an angle of incidence α such that $x/\sin(\pi - \alpha) = R_1/\sin\theta$. That is, $\alpha = \arcsin\left(\dfrac{x}{R_1}\sin\theta\right)$. Total internal reflection at A occurs if and only if $\cos\alpha \leq \sqrt{1-\kappa^2}$. That is, $\sqrt{1 - \left(\dfrac{x}{R_1}\right)^2 \sin^2\theta} \leq \sqrt{1-\kappa^2}$, which holds if and only if $x \geq \kappa\dfrac{R_1}{\sin\theta}$. This holds for all $x \geq R_1$ since $\kappa\dfrac{R_1}{\sin\theta} \leq R_1$ because $\theta \geq \theta_c$. If $\sin\theta > \dfrac{R_1}{R_2}(\geq \kappa)$, then $R_2 > \dfrac{R_1}{\sin\theta}$. On the other hand, if $\sin\theta \leq \dfrac{R_1}{R_2}$, then $R_2 \leq \dfrac{R_1}{\sin\theta}$. Summarizing, if $\sin\theta > \dfrac{R_1}{R_2}$ and $x \in \left[R_1, \dfrac{R_1}{\sin\theta}\right]$, then the ray strikes C_1 at a point A where there is total internal reflection. If $\sin\theta \leq \dfrac{R_1}{R_2}$, then the ray strikes C_1 at a point A where there is total internal reflection for all $x \in [R_1, R_2]$. Let us then assume now that $\theta > \theta_T(x)$. Then the ray strikes C_2 into a point A. If

ω is the angle $0Ax$, then by the sine theorem we have

$$\frac{x}{\sin\omega} = \frac{R_2}{\sin\theta},$$

which gives

$$\omega = \arcsin\left(\frac{x}{R_2}\sin\theta\right).$$

Total internal reflection at A happens if and only if $\cos\omega \leq \sqrt{1-\kappa^2}$, that is, if and only if $x \geq \kappa\frac{R_2}{\sin\theta}$. Since $\theta > \theta_T(x)$, we have $\sin\theta > \frac{R_1}{x}$, so $x > \frac{R_1}{\sin\theta} \geq \kappa\frac{R_2}{\sin\theta}$, where the last inequality follows from (4.21). If $\sin\theta \leq R_1/R_2$, then $R_1/\sin\theta \geq R_2$ and so x cannot be bigger than $R_1/\sin\theta$. Summarizing, if $\sin\theta > \frac{R_1}{R_2}$ and $x \in \left(\frac{R_1}{\sin\theta}, R_2\right]$, then the ray strikes the circle C_2 at which there is total internal reflection.

Ray tracing and energy losses analysis at the second and successive striking points. We next analyze what happens at the next striking point. Suppose the first strike is on a point A in C_1, which means $x \leq \frac{R_1}{\sin\theta}$. Then the ray strikes next a point B in C_2. Let ω be the angle $0BA$ and $\pi-\alpha$ be the angle $0Ax$. By the sine theorem $R_1/\sin\omega = R_2/\sin\alpha$. So $\omega = \arcsin\left(\frac{R_1}{R_2}\sin\alpha\right) = \arcsin\left(\frac{x}{R_2}\sin\theta\right)$. Total internal reflection at B happens iff $x \geq \kappa\frac{R_2}{\sin\theta}$. We then have total internal reflection when $x \in \left[\kappa\frac{R_2}{\sin\theta}, \frac{R_1}{\sin\theta}\right]$. Notice that $\frac{R_1}{\sin\theta} < R_2$ iff $\sin\theta > R_1/R_2$. If $\sin\theta \leq R_1/R_2$, then total internal reflection happens for $x \in [\kappa\frac{R_2}{\sin\theta}, R_2]$. When the first striking point is on C_2, that is when $x \in \left(\frac{R_1}{\sin\theta}, R_2\right]$, we have total internal reflection at A for all $x \in \left(\frac{R_1}{\sin\theta}, R_2\right]$. We also have total internal reflection at any successive point as in the Analysis of case A (Section 4.2.1). For the third and successive striking points the analysis is the same as in case A (Section 4.2.1).

Therefore, we have obtained,

(L1) If $x \in \left[\frac{\kappa R_2}{\sin\theta}, R_2\right]$, there are no losses at both all the striking points on the inner circle and all the striking points on the outer circle.

(L2) If $x \in \left[R_1, \frac{\kappa R_2}{\sin\theta}\right)$ there are no losses at all the striking points on the inner circle and there are losses at all the striking points on the outer circle.

Notice that if $\theta > \arcsin\left(\kappa\frac{R_2}{R_1}\right)$, then $\left[R_1, \frac{\kappa R_2}{\sin\theta}\right)$ is empty.

4.2.3 Analysis of Case C

This is the analogue of Case A (Section 4.2.1) for $\theta < \theta_c$ but we will indicate the differences in the analysis. Let $R_1 \leq x \leq R_2$ and let us assume $\kappa \leq \frac{R_1}{R_2}$. Since $\kappa < 1$ is fixed, this condition always holds when R_1 and R_2 are sufficiently close.

Consider a ray $\ell = \ell(x, \theta)$ that passes through $(x, 0)$, making an angle θ with the positive x-axis, and therefore, striking the circle with radius R_2 for the first time at a point A. We denote by θ the angle AxR_2, and let α be the angle $0Ax$, that is, α is the angle of incidence at A. By the sine theorem we have

$$\frac{x}{\sin\alpha} = \frac{R_2}{\sin(\pi - \theta)}$$

which implies

$$\alpha = \arcsin\left(\frac{x}{R_2}\sin\theta\right).$$

We first analyze when is there total internal reflection at A. From (4.20), total internal reflection at A occurs if and only if $\cos\alpha \leq \sqrt{1-\kappa^2}$. This means

$$\cos\alpha = \cos\left(\arcsin\left(\frac{x}{R_2}\sin\theta\right)\right) = \sqrt{1 - \left(\frac{x}{R_2}\right)^2 \sin^2\theta} \leq \sqrt{1-\kappa^2},$$

which is equivalent to

$$x \geq \frac{\kappa\, R_2}{\sin\theta}.$$

Since $\theta < \theta_c = \arcsin\kappa$, we have $\dfrac{\kappa\, R_2}{\sin\theta} > R_2$. Therefore, we have losses at A for every $\theta < \theta_c$ and for each $x \in [R_1, R_2]$.

We next analyze the trajectory of the ray ℓ to see where strikes next. After being reflected at the point A, the ray ℓ continues traveling and strikes the inner circle C_1 at a point denoted by B. Differently from Case A (Section 4.2.1) the ray ℓ cannot strike C_2. From the point A the angle α for which the ray ℓ is possibly tangential to the smaller circle is α_T with $\sin\alpha_T = \dfrac{R_1}{R_2}$. So ℓ strikes the smaller circle iff $\alpha \leq \alpha_T$ and ℓ strikes the larger circle iff $\alpha > \alpha_T$. Notice that in this case α cannot bigger than α_T

$$\alpha > \alpha_T \iff \frac{x\sin\theta}{R_2} > \frac{R_1}{R_2} \iff \theta > \arcsin\frac{R_1}{x},$$

this is impossible since we assume $\theta < \arcsin\kappa < \arcsin\frac{R_1}{R_2}$. Therefore, ℓ strikes C_1 for every $\theta < \theta_c$ and for each $x \in [R_1, R_2]$.

Let us analyze what happens at B with internal reflection. Let ω be the angle of incidence at B, as for Case A (4.2.1) we have

$$\omega = \arcsin\left(\frac{x}{R_1}\sin\theta\right).$$

We have total internal reflection if $\cos\omega \leq \sqrt{1-\kappa^2}$ that is $\sqrt{1-\left(\dfrac{x}{R_1}\right)^2\sin^2\theta} \leq \sqrt{1-\kappa^2}$ which is equivalent to $x \geq \dfrac{\kappa R_1}{\sin\theta}$. Notice that $\dfrac{\kappa R_1}{\sin\theta} > R_2$ if $\sin\theta < \kappa\dfrac{R_1}{R_2}$. Therefore, If $\theta < \arcsin\left(\kappa\dfrac{R_1}{R_2}\right)$ we have losses for each $x \in [R_1, R_2]$.

If $\arcsin\left(\kappa\frac{R_1}{R_2}\right) \leq \theta < \arcsin\kappa$ we have losses for each $x \in \left[R_1, \kappa\frac{R_1}{\sin\theta}\right)$ and total internal reflection for $x \in \left[\kappa\frac{R_1}{\sin\theta}, R_2\right]$.

As in Case A (Section 4.2.1), after these two striking points the ray will keep to zig-zag between C_2 and C_1 and the angles of incidence will repeat, that is if the striking point is on C_2 the angle of incidence will be α, if the striking point is on C_1 the angle of incidence will be ω. Summarizing,

(L3) If $x \in \left[\frac{\kappa R_1}{\sin\theta}, R_2\right]$, there are no losses at all the striking points on the inner circle and there are losses at all the striking points on the outer circle.

(L4) If $x \in \left[R_1, \frac{\kappa R_1}{\sin\theta}\right)$ there are losses at both all the striking points on the inner circle and all the striking points on the outer circle.

Notice that if $\theta < \arcsin\left(\frac{\kappa R_1}{R_2}\right)$, then $\left[\frac{\kappa R_1}{\sin\theta}, R_2\right]$ is empty.

4.2.4 Analysis of Case D

The analysis is similar to Case C (Section 4.2.3) and we obtain the same results (L3) and (L4). Consider a ray ℓ that passes through $(x, 0)$, $R_1 \leq x \leq R_2$, making an angle $\theta < \theta_c$ with the x-axis going to $-\infty$. While in Case B (Section 4.2.2) there are two possibilities: the ray first strikes C_1 or it first strikes C_2, in this case the ray must strikes C_1. Indeed, notice that in order for the ray to strikes C_2 we need $\sin\theta \geq \sin\theta_T(x) = \frac{R_1}{x}$. This is impossible since $\sin\theta < \kappa$ by assumption and $\kappa < \frac{R_1}{R_2} < \frac{R_1}{x}$. Call the first striking A. As in case Case B (Section 4.2.2), the angle of incidence at A is $\alpha = \arcsin\left(\frac{x}{R_1}\sin\theta\right)$. As in the section above, we have total internal reflection at A if $\cos\alpha \le \sqrt{1-\kappa^2}$, that is $\sqrt{1-\left(\frac{x}{R_1}\right)^2 \sin^2\theta} \le \sqrt{1-\kappa^2}$, which is equivalent to $x \ge \frac{\kappa R_1}{\sin\theta}$. Notice that $\frac{\kappa R_1}{\sin\theta} > R_2$ if $\sin\theta < \kappa\frac{R_1}{R_2}$. Therefore, if $\theta < \arcsin\left(\kappa\frac{R_1}{R_2}\right)$ we have losses at A for every $x \in [R_1, R_2]$. If $\arcsin\left(\kappa\frac{R_1}{R_2}\right) \le \theta < \arcsin\kappa$ we have losses at A for every $x \in \left[R_1, \kappa\frac{R_1}{\sin\theta}\right)$ and total internal reflection for $x \in \left[\kappa\frac{R_1}{\sin\theta}, R_2\right]$.

The second striking point B has to be in C_2, and as in Case B (Section 4.2.2) the angle of incidence is $\omega = \arcsin\left(\frac{x}{R_2}\sin\theta\right)$. From (4.20), total internal reflection at A occurs if and only if $\cos\omega \le \sqrt{1-\kappa^2}$. This means

$$\cos\omega = \cos\left(\arcsin\left(\frac{x}{R_2}\sin\theta\right)\right) = \sqrt{1-\left(\frac{x}{R_2}\right)^2 \sin^2\theta} \le \sqrt{1-\kappa^2},$$

which is equivalent to

$$x \ge \frac{\kappa\, R_2}{\sin\theta}.$$

Since $\theta < \theta_c = \arcsin\kappa$, we have $\frac{\kappa\, R_2}{\sin\theta} > R_2$. Therefore, we have losses at B for every $\theta < \theta_c$ and for each $x \in [R_1, R_2]$. As in all the cases above, after these two striking points the ray will keep to zig-zag between C_1 and C_2 and the angles of incidence will repeat, that is if the striking point is on C_2 the angle of incidence will be ω, if the striking point is on C_1 the angle of incidence will be α. Summarizing,

(L3) If $x \in \left[\frac{\kappa\, R_1}{\sin\theta}, R_2\right]$, there are no losses at all the striking points on the inner circle and there are losses at all the striking points on the outer circle.

(L4) If $x \in \left[R_1, \frac{\kappa\, R_1}{\sin\theta}\right)$ there are losses at both all the striking points on the inner circle and all the striking points on the outer circle.

Notice that if $\theta < \arcsin\left(\frac{\kappa\, R_1}{R_2}\right)$, then $\left[\frac{\kappa\, R_1}{\sin\theta}, R_2\right]$ is empty.

4.3 Quantitative estimates of the internally reflected energy, circular guide

From (L1),(L2),(L3)and (L4) we conclude there are three different regions in which we have energy losses (pictured in Figure 4.8);

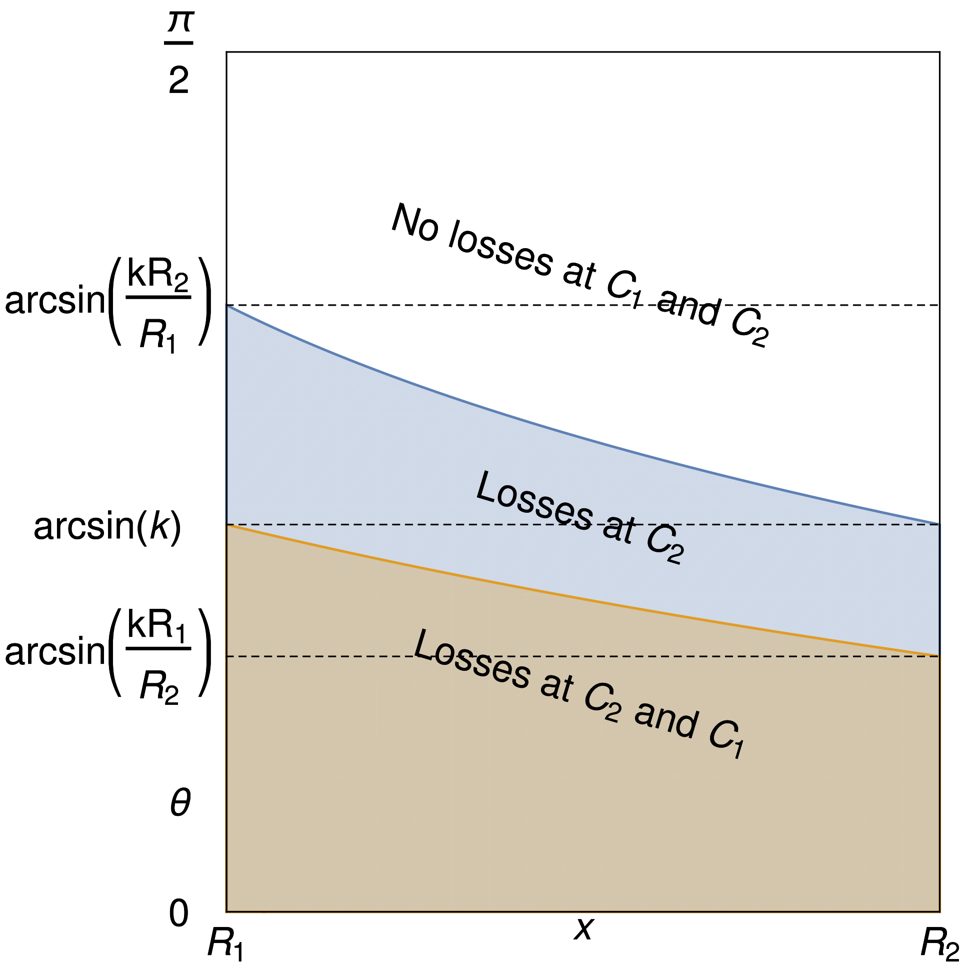

Figure 4.8: Losses regions

(R1) If $\arcsin \kappa \leq \theta < \arcsin\left(\kappa \frac{R_2}{R_1}\right)$, then we have losses at C_2 for each $x \in \left[R_1, \kappa \frac{R_2}{\sin\theta}\right)$, and we have no losses at C_1.

(R2) If $\arcsin\left(\kappa\frac{R_1}{R_2}\right) \leq \theta < \arcsin\kappa$, then we have losses at C_2 for each $x \in [R_1, R_2]$ and we have losses at C_1 for each $x \in \left[R_1, \kappa\frac{R_1}{\sin\theta}\right)$.

(R3) If $\theta < \arcsin\left(\kappa\frac{R_1}{R_2}\right)$, then we have losses at C_1 and C_2 for each $x \in [R_1, R_2]$.

We now give quantitative estimates for the internally reflected energy in each of the regions above. Without loss of generality we assume the ray ℓ is going towards the circle C_2. The analysis of region (R2) follows from the analysis of regions (R1) and (R3), for this reason it will be carry out as last.

Internally reflected energy for region (R1): Fix θ. We denote by P_N the Nth-striking point on the outer circle. At each of these striking points, the ray ℓ internally reflects energy according to Fresnel's formula. Let E be the energy that the ray ℓ is carrying from the straight waveguide into the circular waveguide. This ray is incident for the first time on the boundary of the circular waveguide at the point $A = P_0 \in C_2$. From Fresnel formulas there is an amount E_0 that is internally reflected at P_0. The amount E_0 will be the incident energy at $C = P_1 \in C_2$ (since we have total internal reflection at $B \in C_1$) and so a percentage of this energy, denoted by E_1, will be internally reflected at P_1. Next E_1 will be the incident energy at $P_2 \in C_2$ and then a percentage E_2, will be the internally reflected energy at P_2. We can now apply formula (4.17) relating the incident energy at the first point with the energy at the last point. With the notation of Section 4.1, we note that in the circular case we have that the angle between the incident unit direction x at each P_j and the normal at P_j equals α (the angle of incidence at $P_0 = A$ and so at P_j).

We then have from (4.16) or (4.17)

$$\begin{aligned}\frac{E_N}{E} &= \left(\frac{|\mathbf{E}^r_{P_N}|}{|\mathbf{E}^i_{P_0}|}\right)^2\\ &= [\Phi_1(\cos\alpha)]^{2(N+1)}\ \frac{(I^A_\parallel)^2}{(I^A_\parallel)^2+(I^A_\perp)^2} + [\Phi_2(\cos\alpha)]^{2(N+1)}\ \frac{(I^A_\perp)^2}{(I^A_\parallel)^2+(I^A_\perp)^2},\end{aligned}$$

where Φ_1 and Φ_2 are given in (4.14) and (4.15) respectively. This means that the energy internally reflected at the point P_N equals

$$E_N = \left([\Phi_1(\cos\alpha)]^{2(N+1)}\ \frac{(I^A_\parallel)^2}{(I^A_\parallel)^2+(I^A_\perp)^2} + [\Phi_2(\cos\alpha)]^{2(N+1)}\ \frac{(I^A_\perp)^2}{(I^A_\parallel)^2+(I^A_\perp)^2}\right) E.$$

From (4.22)

$$\cos\alpha = \sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta}.$$

Recall that

$$\phi(t) = t-\kappa\sqrt{1-\kappa^{-2}(1-t^2)},$$

and

$$\frac{1}{1-\kappa^2}\left(\frac{2\kappa}{\frac{1}{\kappa}\,(1-\phi(\cos\alpha)\,\cos\alpha)} - (1+\kappa^2)\right) = \Phi_1(\cos\alpha),$$

$$\frac{1}{1-\kappa^2}\left(1-2\kappa\left(\frac{1}{\kappa}\,(1-\phi(\cos\alpha)\,\cos\alpha)\right)+\kappa^2\right) = \Phi_2(\cos\alpha).$$

We have

$$\begin{aligned}\phi(\cos\alpha) &= \sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} - \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}\\ &= \frac{1-\kappa^2}{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta}+\sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}},\end{aligned}$$

$$1-\phi(\cos\alpha)\,\cos\alpha$$

$$= 1 - \frac{1-\kappa^2}{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta}$$

$$= \frac{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta} - (1-\kappa^2)\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}$$

$$= \frac{\sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \kappa^2\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}.$$

Then

$$\Phi_1(\cos\alpha) = \frac{\kappa^2\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} - \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}{\kappa^2\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}} := G_1\left(\frac{x}{R_2}\sin\theta\right),$$

and

$$\Phi_2(\cos\alpha) = \frac{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} - \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}}{\sqrt{1-\left(\frac{x}{R_2}\right)^2\sin^2\theta} + \sqrt{\kappa^2-\left(\frac{x}{R_2}\right)^2\sin^2\theta}} := G_2\left(\frac{x}{R_2}\sin\theta\right).$$

Let

$$F(x,\theta,N) = [\Phi_1(\cos\alpha)]^{2(N+1)}\ \frac{(I_{\parallel}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2} + [\Phi_2(\cos\alpha)]^{2(N+1)}\ \frac{(I_{\perp}^A)^2}{(I_{\parallel}^A)^2+(I_{\perp}^A)^2}, \tag{4.29}$$

which represents the fraction of energy internally reflected at the point P_N.

We propose the following integral as an average measure of the fraction of energy internally reflected by all rays ℓ having direction θ and passing through each x satisfying (4.27):

$$\frac{1}{\kappa\frac{R_2}{\sin\theta} - R_1}\int_{R_1}^{\kappa\frac{R_2}{\sin\theta}} F(x,\theta,N)\,dx.$$

Making the change of variables $u = \dfrac{x}{R_2}\sin\theta$, the last integral equals

$$\frac{1}{\frac{R_2}{\sin\theta} - R_1}\frac{R_2}{\sin\theta}\int_{\frac{R_1}{R_2}\sin\theta}^{\kappa} G(u,N)\,du \tag{4.30}$$

where

$$G(u,N) = G_1(u)^{2(N+1)}\frac{(I_{\parallel}^A)^2}{(I_{\parallel}^A)^2 + (I_{\perp}^A)^2} + G_2(u)^{2(N+1)}\frac{(I_{\perp}^A)^2}{(I_{\parallel}^A)^2 + (I_{\perp}^A)^2},$$

$$G_1(u) = \frac{\kappa^2\sqrt{1-u^2} - \sqrt{\kappa^2-u^2}}{\kappa^2\sqrt{1-u^2} + \sqrt{\kappa^2-u^2}},$$

$$G_2(u) = \frac{\sqrt{1-u^2} - \sqrt{\kappa^2-u^2}}{\sqrt{1-u^2} + \sqrt{\kappa^2-u^2}}.$$

Remark 4.3.1. *Consider the case* $R_1 = R_2 - \epsilon$ *for some* $\epsilon > 0$. *We showed that if*

$$\theta \in [\arcsin\kappa, \arcsin\kappa\frac{R_2}{R_1}] \quad \textit{and} \quad R_1 \le x < \frac{\kappa R_2}{\sin\theta},$$

then there are no losses at all the striking points on the inner circle and there are losses at all the striking points on the outer circle. This implies that we have losses if

$$R_2 - \epsilon \le x < \frac{\kappa R_2}{\sin\theta}.$$

Notice that for $R_2 > \dfrac{\epsilon}{1 - \dfrac{\kappa}{\sin\theta}}$, *the interval above is empty, therefore, taking* R_2 *big enough we will have no losses for every* x *and* θ.

Remark 4.3.2. *Consider the case* $R_1 = aR_2$ *for some* $0 < a < 1$. *If* $a > \dfrac{\kappa}{\sin\theta}$, *then we have no losses. If* $a \le \dfrac{\kappa}{\sin\theta}$ *we will have losses for* $aR_2 \le x < \dfrac{\kappa}{\sin\theta}R_2$. *Recall that from* (4.28) *the point* P_N *is at the end of the arc of the circle* C_2 *starting from* $(R_2, 0)$ *and so this arc has length*

$$L = R_2\left(\theta + 2N\omega - (2N+1)\alpha\right).$$

Consider $c > 0$ such that $L = c\pi R_2$. We obtain

$$c\pi = \theta + 2N\omega - (2N+1)\alpha.$$

We can solve the above equation for N obtaining

$$N = \frac{c\pi - \theta + \alpha}{2(\omega - \alpha)}.$$

Recall that $\alpha = \arcsin\left(\frac{x\sin(\theta)}{R_2}\right)$ and $\omega = \arcsin\left(\frac{x\sin(\theta)}{R_1}\right)$. Making the change of variable $u = \frac{x\sin(\theta)}{R_2}$ and since R_1 is aR_2 we obtain

$$N = \frac{c\pi - \theta + \arcsin(u)}{2(\arcsin(u + \frac{1-a}{a}u) - \arcsin(u))}.$$

Using the Mean value theorem we have

$$N = \frac{c\pi - \theta + \arcsin(u)}{2\frac{1-a}{a}u}\sqrt{1-\xi^2}$$

for some $\xi \in (u, u + \frac{1-a}{a}u)$. Since we have losses for $aR_2 \leq x < \dfrac{\kappa}{\sin\theta}R_2$, equivalently we have losses for $a\sin\theta \leq u < \kappa$. Therefore, we obtain the following estimates:

$$\frac{c\pi - \theta + \arcsin(a\sin\theta)}{2\dfrac{1-a}{a}\kappa}\sqrt{1-\left(\frac{\kappa}{a}\right)^2} \leq N \leq \frac{c\pi - \theta + \arcsin(\kappa)}{2(1-a)\sin\theta}.$$

Note that these estimates are independent from x and R_2.

Internally reflected energy for region (R3): In this region we have losses at all the striking point on C_1 and C_2. Let E be the energy that the ray ℓ is carrying from the straight waveguide into the circular waveguide. This ray is incident for the first time on the boundary of the circular waveguide at the point $P_0 \in C_2$. From Fresnel formulas there is an amount E_0 that is internally reflected at P_0. The amount E_0 will be the incident energy at $P_1 \in C_1$ and so a percentage of this energy, denoted by E_1, will be internally reflected at P_1. Next E_1 will be the incident energy at $P_2 \in C_2$ and then a percentage E_2, will be the internally reflected energy at P_2. The angle of incidence at P_j is equals to α if $P_j \in C_2$ and it is equals to ω if $P_j \in C_1$. Therefore,

$$\begin{aligned} E_{2n} &= \left([\Phi_1(\cos\alpha)]^{2(n+1)} \frac{(I_{\|}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} + [\Phi_2(\cos\alpha)]^{2(n+1)} \frac{(I_{\perp}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} \right) \\ &\times \left([\Phi_1(\cos\omega)]^{2n} \frac{(I_{\|}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} + [\Phi_2(\cos\omega)]^{2n} \frac{(I_{\perp}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} \right) E \\ &= H_{n+1}(\cos\alpha) H_n(\cos\omega) E \\ E_{2n+1} &= \left([\Phi_1(\cos\alpha)]^{2(n+1)} \frac{(I_{\|}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} + [\Phi_2(\cos\alpha)]^{2(n+1)} \frac{(I_{\perp}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} \right) \\ &\times \left([\Phi_1(\cos\omega)]^{2(n+1)} \frac{(I_{\|}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} + [\Phi_2(\cos\omega)]^{2(n+1)} \frac{(I_{\perp}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} \right) E \\ &= H_{n+1}(\cos\alpha) H_{n+1}(\cos\omega) E, \end{aligned}$$

where Φ_1 as in (4.14), Φ_2 as in (4.15) ,and

$$H_k(t) = [\Phi_1(t)]^{2k} \frac{(I_{\|}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2} + [\Phi_2(t)]^{2k} \frac{(I_{\perp}^A)^2}{(I_{\|}^A)^2 + (I_{\perp}^A)^2}.$$

Let

$$H(x,\theta,N) = \begin{cases} H_{n+1}(\cos\alpha) H_n(\cos\omega) & \text{if } N = 2n \\ H_{n+1}(\cos\alpha) H_{n+1}(\cos\omega) & \text{if } N = 2n+1, \end{cases} \tag{4.31}$$

which represents the fraction of energy internally reflected at the point P_N. We propose the following integral as an average measure of the fraction of energy internally reflected by all rays ℓ having direction θ and passing through each $x \in [R_1, R_2]$:

$$\frac{1}{R_2 - R_1} \int_{R_1}^{R_2} H(x, \theta, N)\, dx. \tag{4.32}$$

Internally reflected energy for region (R2): In this region we have energy losses at all the striking points on C_1 and C_2 if $x \in \left[R_1, \kappa \frac{R_1}{\sin\theta}\right)$ but energy losses only at the striking point on C_2 if $x \in \left[\kappa \frac{R_1}{\sin\theta}, R_2\right]$. Therefore, for the rays passing through $x \in \left[R_1, \kappa \frac{R_1}{\sin\theta}\right)$ the analysis is the same as for the ray in (R3), and for the rays passing through $x \in \left[\kappa \frac{R_1}{\sin\theta}, R_2\right]$ the analysis is the same as for the ray in (R1). We propose the following integral as an average measure of the fraction of energy internally reflected by all rays ℓ having direction θ and passing through each $x \in [R_1, R_2]$:

$$\frac{1}{R_2 - R_1} \left(\int_{R_1}^{\kappa \frac{R_1}{\sin\theta}} H(x, \theta, N)\, dx + \int_{\kappa \frac{R_1}{\sin\theta}}^{R_2} F(x, \theta, N)\, dx \right), \tag{4.33}$$

where $H(x, \theta, N)$ as in (4.31), and $F(x, \theta, N)$ as in (4.29).

4.4 Asymptotics

Suppose to have a periodic circular waveguide as in Figure 4.9. The ray passing through $(x, 0)$ with an angle θ, with $\theta \in \left[\arcsin\kappa, \arcsin\left(\kappa \frac{R_2}{R_1}\right)\right]$ and $x \in \left[R_1, \kappa \frac{R_2}{\sin\theta}\right)$, will keep to zig-zag between C_1 and C_2 until it will lose all of its energy (see Section (R1)). The problem we solve on this section is the following: fix $\theta \geq \theta_C$ and assume the ray ℓ is going towards the circle C_2, what is the asymptotics of integral in (4.30) for $N \to \infty$?

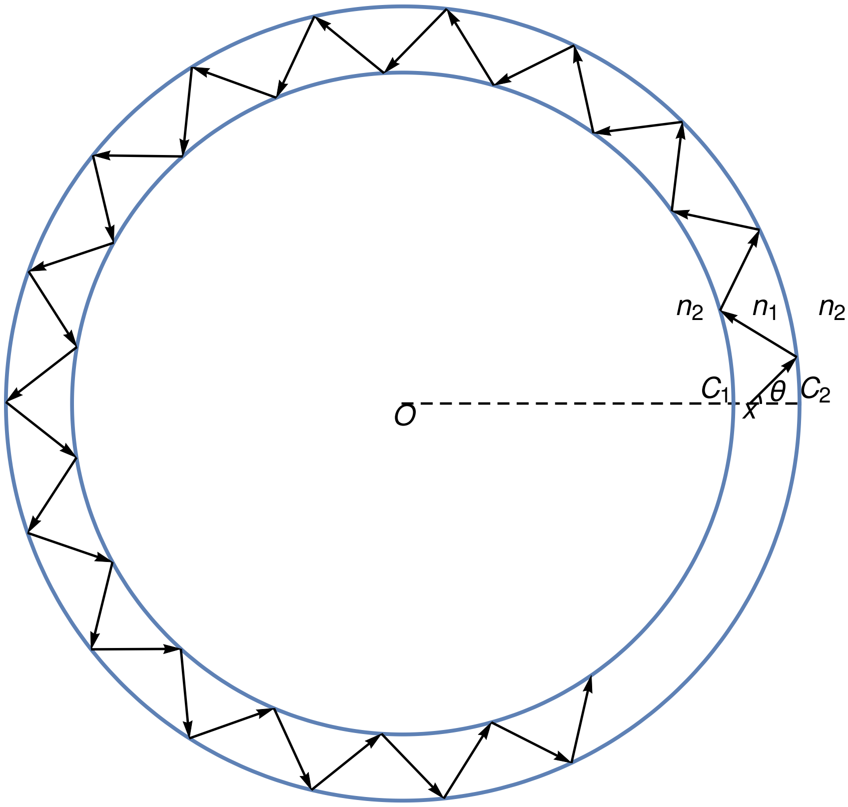

Figure 4.9: Circular waveguide

From the two theorems below (Theorem 4.4.1 and Theorem 4.4.3) will follow that

$$\int_{R_1\,\sin\theta/R_2}^{\kappa} G(u,N)\,du \sim \frac{C}{N^2} \quad \text{as} \quad N \to \infty,$$

for some constant C depending only on κ, where we recall that

$$G(u,N) = G_1(u)^{2(N+1)}\,\frac{(I_\parallel^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2} + G_2(u)^{2(N+1)}\,\frac{(I_\perp^A)^2}{(I_\parallel^A)^2 + (I_\perp^A)^2},$$

$$G_1(u) = \frac{\kappa^2\sqrt{1-u^2} - \sqrt{\kappa^2-u^2}}{\kappa^2\sqrt{1-u^2} + \sqrt{\kappa^2-u^2}},$$

$$G_2(u) = \frac{\sqrt{1-u^2} - \sqrt{\kappa^2-u^2}}{\sqrt{1-u^2} + \sqrt{\kappa^2-u^2}}.$$

Theorem 4.4.1. *Given $0 < \kappa < 1$ and $0 < b < \kappa$, we have that*

$$\lim_{x\to\infty} x^2 \int_b^\kappa G_2(u)^x du = \frac{1-\kappa^2}{4\kappa}.$$

Proof. Write by integration by parts:

$$\begin{aligned}
&\int_b^\kappa G_2(u)^x du = \int_b^\kappa e^{x\ln G_2(u)} du = \frac{1}{x}\int_b^\kappa \frac{G_2(u)}{G_2'(u)} d(e^{x\ln G_2(u)}) \\
=&\frac{1}{x}\frac{G_2(u)}{G_2'(u)} e^{x\ln G_2(u)}\Big|_{u=b}^{u=\kappa} - \frac{1}{x}\int_b^\kappa e^{x\ln G_2(u)} \frac{d}{du}\left(\frac{G_2(u)}{G_2'(u)}\right) du \\
=&\frac{1}{x}\frac{G_2(\kappa)}{G_2'(\kappa)} e^{x\ln G_2(\kappa)} - \frac{1}{x}\frac{G_2(b)}{G_2'(b)} e^{x\ln G_2(b)} - \frac{1}{x^2}\int_b^\kappa \frac{G_2(u)}{G_2'(u)}\frac{d}{du}\left(\frac{G_2(u)}{G_2'(u)}\right) d(e^{x\ln G_2(u)}) \\
=&(*).
\end{aligned}$$

Now $G_2(\kappa) = 1$, and $G_2'(u) \to +\infty$ as $u \to \kappa$ because

$$G_2'(u) = \frac{2u}{1-\kappa^2}\left(\frac{\left(\sqrt{1-u^2} - \sqrt{\kappa^2-u^2}\right)^2}{\sqrt{1-u^2}\sqrt{\kappa^2-u^2}}\right),$$

so

$$\begin{aligned}
(*) = -\frac{1}{x}\frac{G_2(b)}{G_2'(b)} G_2(b)^x - \frac{1}{x^2}\frac{G_2(u)}{G_2'(u)}\frac{d}{du}\left(\frac{G_2(u)}{G_2'(u)}\right) e^{x\ln G_2(u)}\Big|_{u=b}^{u=\kappa} \\
+ \frac{1}{x^2}\int_b^\kappa e^{x\ln G_2(u)} \frac{d}{du}\left(\frac{G_2}{G_2'}\frac{d}{du}\left(\frac{G_2}{G_2'}\right)\right) du = (**).
\end{aligned}$$

Now

$$\begin{aligned}
\frac{G_2(u)}{G_2'(u)} &= \frac{\sqrt{1-u^2}\sqrt{\kappa^2-u^2}}{2u} \\
\frac{d}{du}\left(\frac{G_2(u)}{G_2'(u)}\right) &= \frac{-\kappa^2+u^4}{2u^2\sqrt{1-u^2}\sqrt{\kappa^2-u^2}}
\end{aligned}$$

so

$$\frac{G_2(u)}{G_2'(u)}\frac{d}{du}\left(\frac{G_2(u)}{G_2'(u)}\right) = \frac{-\kappa^2+u^4}{4u^3}.$$

If $u = \kappa$ then

$$\frac{G_2}{G_2'}\frac{d}{du}\left(\frac{G_2}{G_2'}\right) = \frac{\kappa^2 - 1}{4\kappa},$$

so

$$\begin{aligned}
(**) &= -\frac{1}{x}\frac{G_2(b)}{G_2'(b)}G_2(b)^x - \frac{1}{x^2}\frac{\kappa^2-1}{4\kappa} + \frac{1}{x^2}\left(\frac{-\kappa^2+b^4}{4b^3}\right)G_2(b)^x \\
&\quad + \frac{1}{x^2}\int_b^\kappa e^{x\ln G_2(u)}\frac{d}{du}\left(\frac{-k^2+u^4}{4u^3}\right)du \\
&= -\frac{1}{x}\frac{G_2(b)}{G_2'(b)}G_2(b)^x - \frac{1}{x^2}\frac{\kappa^2-1}{4\kappa} + \frac{1}{x^2}\left(\frac{-\kappa^2+b^4}{4b^3}\right)G_2(b)^x \\
&\quad + \frac{1}{x^2}\int_b^\kappa e^{x\ln G_2(u)}\left(\frac{3k^2+u^4}{4u^4}\right)du.
\end{aligned}$$

Multiplying by x^2, we get

$$\begin{aligned}
&x^2\int_b^\kappa G_2(u)^x du = \\
&-\frac{xG_2(b)}{G_2'(b)}G_2(b)^x + \frac{1-\kappa^2}{4\kappa} + \left(\frac{-\kappa^2+b^4}{4b^3}\right)G_2(b)^x + \int_b^\kappa G_2(u)^x\left(\frac{3k^2+u^4}{4u^4}\right)du.
\end{aligned}$$

By Lebesgue dominated convergence, the last integral tends to zero as $x \to \infty$, and since $G_2(b) < 1$ we are done. □

Remark 4.4.2. *The integration by parts are justified because*

$$\frac{d}{du}\left(\frac{G_2}{G_2'}\right) \quad \text{and} \quad \frac{d}{du}\left(\frac{G_2}{G_2'}\frac{d}{du}\left(\frac{G_2}{G_2'}\right)\right)$$

are both integrable functions in $[b, \kappa]$.

Similarly,

Theorem 4.4.3. *Given* $0 < \kappa < 1$ *and* $0 < b < \kappa$, *we have that*

$$\lim_{x\to\infty} x^2\int_b^\kappa G_1(u)^x du = \frac{1}{4}\kappa^3(1-\kappa^2).$$

Proof. As before,

$$\begin{aligned}
&\int_b^\kappa G_1(u)^x du \\
=&\frac{1}{x}\frac{G_1(\kappa)}{G_1'(\kappa)}e^{x\ln G_1(\kappa)} - \frac{1}{x}\frac{G_1(b)}{G_2'(b)}e^{x\ln G_1(b)} - \frac{1}{x^2}\int_b^\kappa \frac{G_1(u)}{G_1'(u)}\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right) d(e^{x\ln G_1(u)}) \\
=&(*).
\end{aligned}$$

Now, $G_1(\kappa) = 1$ and $G_1'(u) \to +\infty$ as $u \to \kappa$ because

$$\frac{d}{du}\left(G_1(u)\right) = -\frac{2k^2\left(k^2-1\right)u}{\sqrt{1-u^2}\sqrt{k^2-u^2}\left(k^2\sqrt{1-u^2}+\sqrt{k^2-u^2}\right)^2},$$

so

$$\begin{aligned}
(*) = &-\frac{1}{x}\frac{G_1(b)}{G_1'(b)}G_1(b)^x - \frac{1}{x^2}\frac{G_1(u)}{G_1'(u)}\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right)e^{x\ln G_1(u)}\Bigg|_{u=b}^{u=\kappa} \\
&+ \frac{1}{x^2}\int_b^\kappa e^{x\ln G_1(u)}\frac{d}{du}\left(\frac{G_1}{G_1'}\frac{d}{du}\left(\frac{G_1}{G_1'}\right)\right)du = (**).
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{G_1(u)}{G_1'(u)} &= \frac{\sqrt{1-u^2}\sqrt{k^2-u^2}\left(k^2\left(u^2-1\right)+u^2\right)}{2k^2u} \\
\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right) &= \frac{k^4+3\left(k^2+1\right)u^6+k^2\left(k^2+1\right)u^2-\left(2k^4+5k^2+2\right)u^4}{2k^2u^2\sqrt{1-u^2}\sqrt{(k-u)(k+u)}},
\end{aligned}$$

so

$$\begin{aligned}
&\frac{G_1(u)}{G_1'(u)}\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right) \\
=&\frac{-k^6+3\left(k^2+1\right)^2u^8+k^2\left(3k^4+7k^2+3\right)u^4-2\left(k^6+5k^4+5k^2+1\right)u^6}{4k^4u^3}.
\end{aligned}$$

If $u = \kappa$, then

$$\frac{G_1}{G_1'}\frac{d}{du}\left(\frac{G_1}{G_1'}\right) = \frac{1}{4}k^3\left(k^2-1\right),$$

so

$$(**)$$

$$= -\frac{1}{x}\frac{G_1(b)}{G_1'(b)}G_1(b)^x - \frac{1}{x^2}\left(\frac{1}{4}k^3\left(k^2-1\right)\right) + \frac{1}{x^2}\left(\frac{G_1(u)}{G_1'(u)}\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right)\Big|_{u=b}\right)G_1(b)^x$$

$$+ \frac{1}{x^2}\int_b^\kappa e^{x\ln G_1(u)}p(u)du$$

where,

$$p(u) = \frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right)\right)$$

$$= \frac{3k^6 + 15\left(k^2+1\right)^2u^8 + k^2\left(3k^4+7k^2+3\right)u^4 - 6\left(k^6+5k^4+5k^2+1\right)u^6}{4k^4u^4}.$$

Multiplying by x^2, we get

$$x^2\int_b^\kappa G_1(u)^x du =$$

$$-\frac{xG_1(b)}{G_1'(b)}G_1(b)^x + \frac{1}{4}k^3\left(k^2-1\right) + \left(\frac{G_1(u)}{G_1'(u)}\frac{d}{du}\left(\frac{G_1(u)}{G_1'(u)}\right)\Big|_{u=b}\right)G_1(b)^x$$

$$\int_b^\kappa e^{x\ln G_1(u)}p(u)du.$$

By Lebesgue dominated convergence, the last integral tends to zero as $x \to +\infty$, and since $G(b) < 1$ we are done. □

Remark 4.4.4. *The integration by parts are justified because*

$$\frac{d}{du}\left(\frac{G_1}{G_1'}\right) \quad \textit{and} \quad \frac{d}{du}\left(\frac{G_1}{G_1'}\frac{d}{du}\left(\frac{G_1}{G_1'}\right)\right)$$

are both integrable functions in $[b, \kappa]$.

4.5 Finals remarks

The analysis carried out in this chapter is motivated by the work done by C. Koos and collaborators in [28]. Their goal is to optimize the shape of a curved waveguide connecting two straight guides in a photonic circuit so that the energy losses are minimal. Guides with kinks are not a good option because they lead to trapped modes and therefore, high losses [31]. Based on numerical experimentation only, Koos et al. propose that the power loss or attenuation coefficient is approximately χ^q, where χ is the curvature of the outer contour of the waveguide, and q is a material parameter with $q \in (2,3)$. Therefore, they reduce the problem to the minimization of the variational integral $\int_\gamma \chi^q \, ds$ over all curves γ joining two fix points and with prescribed initial and final tangent lines, where s is the arc length. However, even if they propose $q \in (2,3)$, they use $q = 2$ in their paper and for the calculation of the optimal shape of the waveguide. Using this optimal shape, they numerically show that for 180°-bends, the energy loss remains below 1% of the initial energy. While, with a 180° circular bend, they calculate an energy loss equals to 4.5% of the initial energy. A theoretical reason that would explain their numerical results is not provided and is an interesting open question. Our analysis in this thesis is based on theoretical foundations using Maxwell's equations. However our application is made only to the linear and circular geometries. To apply these ideas to other more complicated geometries seems an interesting but challenging question. The minimization of the variational integral $\int_\gamma \chi^q \, ds$, when $q = 2$, is the well-known classical variational problem proposed by Daniel Bernoulli and Leonhard Euler in

1744, and the curve solution is the so called *elastica* (see, e.g., [38]). Further very interesting research in this direction is done by Horn [25], and Melhum [30] to apply the notion of elastica to computer graphics and computer aided design. In addition, the elastica has applications to stress in the design of aircraft fuselages [9]. Therefore, the mathematical understanding of the variational problem for $q \neq 2$ is of interest, and we plan to return to it in the future.

REFERENCES

[1] F. Aieta, P. Genevet, N. Yu, M.A. Kats, Z. Gaburro, and F. Capasso. Out-of-plane reflection and refraction of light by anisotropic optical antenna metasurfaces with phase discontinuities. *Nano letters*, 12(3):1702–1706, 2012.

[2] F. Aieta, A. Kabiri, P. Genevet, N. Yu, M.A. Kats, Z. Gaburro, and F. Capasso. Reflection and refraction of light from metasurfaces with phase discontinuities. *Journal of Nanophotonics*, 6(1):063532–063532, 2012.

[3] H. Ammari, M. Ruiz, W. Wu, S. Yu, and H. Zhang. Mathematical and numerical framework for metasurfaces using thin layers of periodically distributed plasmonic nanoparticles. In *Proc. R. Soc. A*, volume 472. The Royal Society, 2016.

[4] S. R. Biswas, C.E. Gutiérrez, A. Nemilentsau, In-Ho Lee, Sang-Hyun Oh, P. Avouris, and Tony Low. Tunable Graphene Metasurface Reflec-

tarray for Cloaking, Illusion, and Focusing. *Physical Review Applied*, 034021, volume 9, 2018.

[5] M. Born and E. Wolf, *Principles of optics, electromagnetic theory, propagation, interference and diffraction of light*, seventh (expanded), 2006 ed., Cambridge University Press, 1959.

[6] H.T. Chen, A.J. Taylor, and N. Yu. A review of metasurfaces: physics and applications. *Reports on Progress in Physics*, 79(7):076401, 2016.

[7] M.P. do Carmo. *Differential Geometry of Curves and Surfaces.* Prentice-Hall, 1976.

[8] R. De Leo, C.E. Gutiérrez and Henok Mawi. On the Numerical Solution of the Far Field Refractor Problem, *Nonlinear Analysis*, Vol. 157, July 2017, pp. 123-145.

[9] I.D. Faux and M.J.Pratt. Computational Geometry for Design and Manufacture. *Ellis Horwood Ltd.*, Chichester, England, 1979.

[10] A. Figalli. *The Monge-Ampère Equation and Its Applications.* EMS Zurich Lectures in Advanced Mathematics, European Mathematical Society, vol. 22, 2017.

[11] P. Genevet, F. Capasso, F. Aieta, M. Rezakhorasaninejad, and R. Devlin. Recent advances in planar optics: from plasmonic to dielectric metasurfaces. *Optica*, 4(1):2334–2536, January 2017.

[12] C.E. Gutiérrez and Q. Huang. The refractor problem in reshaping light beams, *Arch. Rational Mech. Anal.* **193** (2009), no. 2, 423–443.

[13] C.E. Gutiérrez and H. Mawi. The far field refractor with loss of energy, *Nonlinear Analysis: Theory, Methods & Applications* **82** (2013), 12–46.

[14] C.E. Gutiérrez and A. Sabra. The reflector problem and the inverse square law, *Nonlinear Analysis: Theory, Methods & Applications* **96** (2014), 109–133.

[15] C.E. Gutiérrez and F. Tournier. The parallel refractor, *Development in Mathematics* **28** (2013), 325–334.

[16] C.E. Gutiérrez. Refraction problems in geometric optics. In *Lecture Notes in Mathematics*, volume 2087, pages 95–150. Springer, 2014.

[17] C.E. Gutiérrez, L. Pallucchini, and E. Stachura. General refraction problems with phase discontinuities on nonflat metasurfaces, *JOSA A* 34(7), 1160-1172 (2017).

[18] C.E. Gutiérrez, L. Pallucchini. Reflection and refraction problems for metasurfaces related to Monge–Ampŕe equations, *JOSA A* 35(9), 1523-1531 (2018).

[19] C.E. Gutiérrez. Refraction problems in geometric optics. *Lecture Notes in Mathematics, vol. 2087*, pages 95–150. Springer-Verlag, 2014.

[20] C.E. Gutiérrez and A. Sabra. Aspherical lens design and imaging. *SIAM J. Imaging Sci.*, 9(1):386–411, 2016.

[21] C.E. Gutiérrez and A. Sabra. Freeform lens design for scattering data with general radiant fields. *Archive for Rational Mechanics and Analysis*, pp. 341-399, May 2018, Volume 228(2).

[22] C.E. Gutiérrez. *The Monge-Ampère equation.* Birhäuser, Boston, 2016, 2nd Edition.

[23] P. Hartman. *Ordinary Differential Equations*, volume 38 of Classics in Applied Mathematics. SIAM, 2002.

[24] D. Hondros, P. Debye. Elektromagnetische Wellen an dielektrischen Drähten, *Ann. d. Phys* 32, 465, 1910.

[25] B.K.P Horn. The curve of least energy. *A. I. Memo 612*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., Jan. 1981.

[26] M. Khorasaninejad, W. T. Chen, R.C. Devlin, J. Oh, A.Y. Zhu, and F. Capasso. Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging. *Science*, 352(6290):1190–1194, June 2016.

[27] M. Khorasaninejad, A. Y. Zhu, C. Roques-Carmes, W. T. Chen, J. Oh, I. Mishra, R.C. Devlin, and F. Capasso. Polarization-insensitive metalenses at visible wavelengths. *Nano Lett.*, 16(11):7229–7234, 2016.

[28] C. Koos, C. G. Poulton, L. Zimmermann, L. Jacome, J. Leuthold, and W. Freude. Ideal bend contour trajectories for single-mode operation of low-loss overmoded waveguides,*IEEE Photonics Technology Letters* 19 (2007), no. 11, 819–821.

[29] R.K. Luneburg. *Mathematical theory of optics.* Univ of California Press, 1964.

[30] Mehlum, E., Curve and Surface Fitting Based on Variational Criteriae for Smoothness. *Central Institute for Industrial Research*, Oslo, Norway, 1969.

[31] S.A. Nazarov and A.V. Shanin. Trapped modes in angular joints of 2D waveguides. *Applicable Analysis*, Vol. 93, No. 3, 572–582.

[32] K.S. Packard. The origin of waveguides: A case of multiple rediscovery. *IEEE Transaction on microwave theory and techniques*, Vol. MTT-32, No. 9, 1984.

[33] A.V. Pogorelov. *Differential Geometry.* Noordhoff, 1959.

[34] Lord Rayleigh. On the passage of electric waves through tubes, or the vibrations of dielectric cylinders, *Phil. Mag.*, vol. XLIII, pp.125-132, 1897.

[35] Lord Rayleigh. The theory of sound `https://gallica.bnf.fr/ark:/12148/bpt6k95131k/f125.item`, vol. 2, Macmillan (London), 1877-1878.

[36] Lord Rayleigh. The theory of sound, vol. 2,*Macmillan and Co.*, 1896.

[37] The runners-up. *Science*, 354(6319):1518–1523, 2016.

[38] D. A. Singer. Lectures on Elastic Curves and Rods. *AIP Conference Proceedings* 1002, 3 (2008).

[39] J. J. Thomson. Notes on Recent Researches in Electricity and Magnetism. *Oxford: Clarendon, 1893, (reprinted London: Dawsons,1968)*, pp. 344-347.

[40] J. Urbas. On the second boundary value problem for equations of Monge-Ampére type, *J. Reine Angew. Math.* 487 (1997), 115-124.

[41] X.J. Wang. On the design of a reflector antenna, *Inverse Problems* 12 (1996), 351-375.

[42] N. Yu and F. Capasso. Flat optics with designer metasurfaces. *Nature Materials*, 13(2):139–150, 2014.

[43] N. Yu, P. Genevet, M.A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro. Light propagation with phase discontinuities: generalized laws of reflection and refraction. *Science*, 334(6054):333–337, 2011.