UC Merced

UC Merced Electronic Theses and Dissertations

Title

The Optics for Solving Prescribed Illumination Problems

Permalink

Author

Ricketts, Melissa Nicole

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

The Optics for Solving Prescribed Illumination Problems

A dissertation submitted in partial satisfaction of the requirements
for the degree of Doctor of Philosophy

in

Physics

by

Melissa N Ricketts

Committee in charge:

Professor Linda S Hirst, Chair
Professor Arnold Kim
Professor Kevin Mitchell
Professor Roland Winston

Fall 2017

The Optics for Solving Prescribed Illumination Problems

The dissertation of Melissa N Ricketts, titled The Optics for Solving Prescribed Illumination Problems, is approved:

Chair _____   Date _____
                        Linda S. Hirst

_____   Date _____
                        Arnold D. Kim

_____   Date _____
                        Kevin Mitchell

_____   Date _____
                        Roland Winston

University of California, MERCED

Abstract

The Optics for Solving Prescribed Illumination Problems

by

Melissa N Ricketts

Doctor of Philosophy in Physics

University of California, MERCED

Professor Linda S Hirst, Chair

The need for controlled illumination arises from emerging efficiency standards and increasing light pollution. When the illumination sector diverged from imaging optics finding solutions instead in nonimaging optics, the field of illumination engineering greatly evolved. Light optics can now minimize light waste, improve light quality, and enhance light aesthetics. And because illumination optics is concerned with the transferring of light, fundamental concepts in nonimaging optics lead to solutions without imposing the constraints found in imaging optics.

This dissertation is largely concerned with nonimaging optics. An overview of this field will be given, addressing topics such as edge-ray theory, strings method, étendue, phase space, angular space, thermodynamics, and flow lines. New advances will be discussed, specifically the theoretical advances pertaining to the asymmetric compound parabolic concentrator (ACPC). Although similar to the compound parabolic concentrator, the ACPC has differing acceptance angles, making it versatile for both the fields of solar concentration and illumination. For solar concentration, its asymmetry can be utilized for areas of the world far from the equator, where more extreme seasons are experienced. Also, in regards to illumination, the ACPC offers more specialized control in non-symmetric instances.

Here, a method to determine the acceptance angles based on the design angles for the ACPC is provided. The étendue, phase space, and angular acceptance for the ACPC is then shown. Two cases for each of these results, and a way to predict these cases will be discussed. Flow lines for this asymmetric design are also discussed, pushing the boundaries of this relatively new nonimaging optics topic.

The ACPC could potentially help in reducing light pollution once further analysis has been completed. Light pollution is a growing problem worldwide. The valley in Yosemite National Park is one example of a place in need of lighting reform. Nonimaging optics offers ways to improve the light quality there. Using a wedge design as a primary optic to transform phase space for a compound parabolic concentrator (CPC), illumination for an equipment yard was controlled to reduce stray light. This nonimaging optics solution was both quick

and inexpensive to produce. Furthermore, its small size allowed for retrofitting, which is an ideal way to fix the lights in Yosemite.

Another optic that will be discussed utilizes total internal reflection (TIR) to control illumination. Nicknamed "The Jellyfish" for its shape, this novel aplanatic lens is one of a kind. Impressively, the Jellyfish can be used as either an illuminator or a solar concentrator because its optics work in both forward and reverse scenarios. When designed on a small scale, this optic becomes useful for micro-optic scale concentrating photovoltaic (CPV) solutions. As a light source, its adjustable size, acceptance angle, and thickness can be increased to meet various lighting standards. When designed for ideal cases, emerging rays exit the surface nearly parallel to one another. In fact, high efficiencies are seen for rays to within two degrees of the optical axis. This is due in large part to the design method, which is carried out using the concepts first developed by Ernest Abbe. The Abbe Sphere offers a starting point, after which, ideas of reflection and refraction can be utilized at front and back surfaces to guid light via TIR to its exit points.

Work documented here takes the Jellyfish and optimizes it for illumination solutions. It is adjusted to work with an extended source (LED) and meet MR-16 standards. Design and simulation processes are given, along with prototyping results.

Finally, design methods in freeform optics offers solutions that can be tailored for even the most complicated illumination distributions. One method, the Supporting Quadrics Method (SQM), takes light rays and directs them to designated locations on a target. The quadrics used for these designs can be ellipsoids, hyperbolids, or parabolids. Numbers of them can used in conjunction with one another to create a desired distribution, after which an envelope is taken to generate a final surface. When the number of these quadrics increase, they must become smaller to accommodate the overall size of the lens. This leads to the question of diffraction effects. Because each quadric is its own aperture, does diffraction play a role in disrupting what should be a precise distribution? Preliminary analysis is done to address this question.

All the work completed within this dissertation falls into nonimaging optics for illumination. With the growing prevalence of energy standards, optical design is important for controlling the light emitted from LEDs. This relatively new field provides the fundamental concepts necessary to design solutions for preventing light pollution, creating prescribed distributions, and achieving high efficiencies.

To my husband, Matthew Mitchell, for your support throughout the pursuit of my Ph.D.

Contents

List of Figures

List of Tables

Acknowledgments

My pursuit of a career in physics was influenced by many who left their mark upon my life through support, encouragement, motivation, and inspiration. But specifically, I would like to acknowledge some of my wonderful educators, starting first with my high school AP physics teacher, Mr. Michael Boykin. He showed me that physics was something to be celebrated, not feared. My love of physics was further stoked by Mrs. Lana Jordan, my undergraduate general physics instructor, who challenged and supported me during those intimidating initial years of higher education.

There are many at UC Merced who have played a key role in my education. From Dr. Carrie Menke, who offered me memorable advice and encouraged me to stick with physics when I was struggling, to Dr. Linda Hirst, who welcomed me into her lab to conduct my senior thesis research project. Dr. Hirst still remains a wonderful role model for me. But most of all, I would like to acknowledge Dr. Roland Winston, who recognized my love of research, and encouraged me to give graduate school a chance. His guidance and wisdom these past six years, along with the mentoring of my committee members: Dr. Linda Hirst, Dr. Kevin Mitchell, Dr. Arnold Kim, and Dr. Roland Winston, has been a valuable asset to both my research, and my career.

The research and learning environment that Dr. Winston cultivates for his research group and colleagues is one that I will be sad to say goodbye to. I am lucky to have been a part of such a great group and all the unique opportunities encountered as part of his team at UC Solar. I wish all my group members the best, and I am eternally grateful for both their friendship and collaboration during my Ph.D.

Chapter 1

Introduction

With movements toward more efficient lighting, our society has become increasingly proactive regarding the elimination of energy waste. As such, optimally designed optics for lighting are attractive. A light emitting diode (LED) is only as good as the optics it is housed within. Via illumination engineering, one can generate optics for lighting that minimize light waste, improve light quality, and enhance light aesthetics.

In the past, optical engineering has been largely interested in imaging and/or lens design. Thus, the literature available for the field of illumination dealt mostly with the application of lighting, limiting the available breadth of knowledge to design methods and suggestions, whilst avoiding the necessary theoretical development of design principles. Fortunately, with the introduction of energy efficient LEDs and the need to house them in efficient optics, the last two decades have seen significant advances in illumination design.

This subfield is concerned with transferring light (visible wavelengths) or more broadly speaking, electromagnetic radiation, from a source to a target. Furthermore, this transfer is to be carried out efficiently, limiting losses. Of course, light transfer has also been a necessity in imaging systems as well, yet imaging systems are not optimal for illumination applications because they do not maximize concentration. Also, imaging systems are constrained by imaging requirements. No such requirements exist in the field of nonimaging optics.

This introduction chapter will offer an overview of illumination design while including the following:

- Types of optics to be seen in the chapters to follow
- Unit convention used in illumination
- Light sources used in the research covered here
- Structure of this dissertation

1.1   Illumination Design

Illumination design is chiefly concerned with the transfer of light from a source to a target. There are two important parameters involved here: transfer efficiency and target distribution. Often, there is a direct trade off between the two. Transfer efficiency is important because of the rising energy efficiency demands, as well as concerns over environmental impact. In addition, achieving a prescribed distribution of illumination at the target is desirable. A designer must keep both of these parameters in mind, along with lesser criteria. These include color, volume requirements, fabrication costs, etc. Thus, all of these parameters must be dealt with to produce an end product that is quick and cheap to fabricate, whist maintaining a high level of transfer efficiency, desired illumination distribution, etc. In the end, these various criteria act as the platform by which optical illumination designs and systems are compared.

Transfer and Optical Efficiency

Throughout illumination optics, both transfer and optical efficiency are used to gauge the success of an optic. Transfer efficiency, η_t, is a ratio of flux at the target, Φ_{target}, to flux at the input, which is usually the flux emitted by a light source, Φ_{source}:

$$\eta_t = \frac{\Phi_{target}}{\Phi_{source}} \tag{1.1}$$

Here flux is measured in either watts (W), when dealing with radiometry, or lumens (lm) when dealing with photometry. A section will be devoted to the difference between these two units of flux.

Assuming that an optic is lossless, meaning, all the light from the source reaches the optic's output aperture, we can treat the output aperture of the optic as the source. Then, when it comes to measuring the transfer efficiency, the measurement is made based on what is seen coming from the optic arriving to the target.

Optical efficiency concerns itself with the optic. It represents how well the light inside of the optic travels from the source to the output aperture. In other words, it is a measure of losses within the system.

$$\eta_o = \frac{\Phi_{aperture}}{\Phi_{source}} \tag{1.2}$$

Given that most systems suffer from various losses due to reflection and absorption, this is a useful measurement.

Illumination Distribution Uniformity

Illumination distribution analysis offers a way to show how well the measured distribution of an optic agrees with the desired or prescribed distribution during the design procedure.

This is known as uniformity. An analysis can be determined by a comparison between sampled measurement of a model's distribution and goal distribution. There are a number of comparative methods currently used. The one used here will be the variance of the model distribution with the design goal:

$$\sigma^2 = \frac{1}{mn-1} \sum_{i=1}^{n} \sum_{j=1}^{m} [f_{model}(i,j) - f_{goal}(i,j)]^2 \tag{1.3}$$

where f_{model} and f_{goal} define the values obtained (measured in flux / unit area) for distributions of the fabricated model and the goal, respectively. The terms i and j are the counters for the m by n samples, respectively, over the two orthogonal axis. A value of 0 for σ^2 would indicate a perfect agreement between the desired and experimental distributions. In other words, complete uniformity between the values of each position on the two targets.

1.2 Optical Devices

Predominantly, there are five categories of optical systems in illumination optics: refractive optics such as lenses, reflective optics such as mirrors, total internal reflection (TIR) optics such as light pipes, scattering optics such as diffusers, and hybrid optics such as LED pseudo collimators. The work in this dissertation fits into the first three categories, and a brief overview of each will be given here.

Reflective Optics

Reflective optics provide the most "power" to spread light because of their potential for high concentration, but this is at the expense of tolerance demands as well as higher absorption losses. However, in illumination optics, the tolerance demands are lower than those in imaging optics. Reflective optics also tend to be bulkier, taking up larger amounts of space. Thus, although they are not necessarily popular devices for imaging, they are a great fit for the illumination industry and are highly prevalent within this field.

There are many different types of reflective optics used for illumination. These originate most often from either edge-ray design principles [1] or tailored edge-ray designs [1, 2]. Some examples of edge-ray designs would be the Compound Parabolic Concentrator (CPC) and Compound Elliptical Concentrator (CEC). Edge-ray designs work well as two-dimensional troughs, where as tailored edge-ray designs tend to work better for three-dimensional cases. Some examples of reflectors used for illumination design are: luminaires for indoor and outdoor lighting, headlamps, and emergency warning lights.

Fundamental to any reflector is the law of reflection, which states that the incident ray angle, θ, must be equivalent to the reflected ray angle, θ''.

$$\theta = \theta'' \tag{1.4}$$

where single prime notation indicates refraction and double primed notation indicates reflection. Reflection optics will be discussed in more detail in Chapters 2, 3, and 4.

Refractive Optics

The shortcomings of reflective optics are oftentimes mitigated with refractive optics, making them the top choice for most imaging systems. Refractive optics can make for quite compact designs. For example, many new optical systems employ micro-lens arrays, which can be on the order of several hundred μm in size. However, refractive (lens) optics are always more expensive due to alignment issues and post production manufacturing requirements.

Some examples of refractive optics that are used in illumination are: Fresnel lens for displays, projection headlamps, and pillow lens arrays for transportation applications. One of the biggest reasons that imaging refractive optics are not employed for illumination is, they often do not meet the high concentrations that can be achieved with designs originating in nonimaging optics.

Refractive optics rely on differing materials to transport light. Refraction is dictated by Snell's law, also known as the law of refraction. This law states that for two differing materials of indexes n, n', the index of refraction, n, multiplied by sine of the angle of an incoming ray, $\mathbf{r}$, must be equivalent to the index of refraction in the secondary material, n', multiplied by the sine of the refracted angle in that material, θ'.

$$n \sin \theta = n' \sin \theta' \tag{1.5}$$

TIR Optics

In theory, TIR optics can be an optimal choice between refractive and reflective optics, except when issues arise with fulfilling the critical-angle condition at all interfaces of the device. These devices rely on refraction to trap light within a material substrate. The critical angle,θ_c, is the angle needed for a ray to become trapped, and is governed by Snell's law. It is based entirely off the index of refraction of a material and the material's surrounding environment, relying on the fact that the material have a higher index than its surrounding environment. As long as a ray, $\mathbf{r}$, strikes the surface within some material at an angle, $\geq \theta_c$, it will reflect within the device as $\mathbf{r}$".

$$\theta_c = arcsin\left(\frac{n'}{n}\right) \tag{1.6}$$

where n is the higher index material, and n' is the lower index material. Some examples of TIR devices include: lightpipes, optical fibers, light guides for display applications, and brightness enhancing films. Chapter 5 will discuss an example of a TIR aplanatic optic collimator.

1.3   Unit Convention

In illumination optics, there are essentially two types of quantities used: radiometric terms and photometric terms. The terminology, often bewildering, can lead to confusion. Thus, it is important to distinguish between the two. Radiometric terms are deterministic quantities based on the physical nature of light (electromagnetic radiation) and always use the watt for rate of energy transfer. Photometric terms use the lumen, and are quantities based on the human visual system such that only the visible electromagnetic radiation spectrum is considered. Both of these terms are connected through the response of the human eye, which is standardized by the International Commission on Illumination. The terms that will most often be used in this dissertation will be given in two tables, one for radiometric and one for photometric.

Term and description	Symbol	Functional form	SI units
Radiant energy	Q_e		J
Radiant power	Φ_e	$\frac{dQ_e}{dt}$	J/s or W
Irradiance	E_e	$\frac{d\Phi_e}{dA_{target}}$	W/m^2
Radiant intensity	I_e	$\frac{d\Phi_e}{d\Omega}$	W/sr
Radiance	L_e	$\frac{d^2\Phi_e}{dA_{s,proj}d\Omega}$	$W/m^2/sr$

Table 1.1: Radiometric terms and their characteristics

For Tab. 1.1, the subscript "e" is given for electromagnetic quantities.

For photometric terms, which are based on the visual spectrum of electromagnetic radiation, 360 - 830 nm, there is an analogous set of terms. Just as with radiometric terms, the word luminous precedes any term referring to photometric quantities. All terms relating to photometric quantities will be subscripted with "v" to denote visual.

If one wishes to convert between radiometric and photometric quantities, this is accomplished by taking into account the CIE standard observer. The functional form for this is given by:

$$f_v(\lambda) = K(\lambda)f_e(\lambda), \tag{1.7}$$

where $f_v(\lambda)$ is the spectral photometric quantity of interest, $f_e(\lambda)$ is the analogous spectral radiometric term, and $K(\lambda)$ is the luminous efficacy. The luminous efficacy is a function of wavelength, λ and its units are in lm/W; it represents the CIE observer response to visible chromatic radiation. In general, from the definition of the lumen, one watt of radiant energy at the wavelength of maximum visual sensitivity (550 nm) is equivalent to 680 lumens. That pertains to what is called photopic vision. There is also dark-adapted scotopic vision which

Term and description	Symbol	Functional form	SI units
Luminous energy	Q_v		T
Luminous power	Φ_v	$\frac{dQ_v}{dt}$	lm
Luminous exitance	M_v	$\frac{d\Phi_v}{dA_{source}}$	lx
Luminous intensity	I_v	$\frac{d\Phi_v}{d\Omega}$	cd
Illuminance	E_v	$\frac{d\Phi_v}{dA_{target}}$	lx
Luminance	L_v	$\frac{d^2\Phi_v}{dA_{s,proj}d\Omega}$	nt

Table 1.2: Photometric terms and their characteristics

has its peak at approximately 507 nm. Most often though, the conversion only deals with photopic vision.

1.4 Light Sources

All optical tests carried out in this research employ either point sources (unphysical / theoretical), or light emitting diodes (LEDs). A point source is straight forward; it emits equal amounts of electromagnetic radiation in all directions radially. Point sources often simplify optical designs and are much easier to work with. However, point sources do not exist in nature, and so they are only used in the ideal sense. Because an LED is considered an extended source, meaning, it possesses significant dimensions, it cannot be treated like a point source. Any small source in possession of an area will radiate some amount of power per unit of solid angle. Thus, an LED's radiation characteristics are expressed in terms of power per unit solid angle per unit area, which is called radiance if using watts. Or, if using lumens, it would called luminance, and be expressed in lumens per unit solid angle per unit area, which is known as the unit nit (nt),and is shown at the bottom of Tab. 1.2.

LEDs inherently possess what is called a lambertian distribution. To lead into the discussion of lambertian distribution, we first discuss intensity and radiance/luminance.

Intensity

Once familiar with both radiometric and photometric terminology, it is prudent to address some of the terms used in Tab. 1.1 and Tab. 1.2 to understand lambertian distributions. These terms will be used throughout this dissertation, thus, they deserve a short discussion.

Radiant or luminous intensity describes light distribution as a function of solid angle, $d\Omega$. Specifically, it is the flux per unit solid angle. The unit of solid angle is the steradian,

sr, and is expressed as a cone subtending an area dA on the surface of a sphere of radius r, with the cone's vertex at the center of the sphere. The definition of solid angle is given by:

$$d\Omega = \frac{dA}{r^2} = \sin\theta d\theta d\phi. \tag{1.8}$$

where θ is the polar angle, and ϕ is the azimuthal angle.
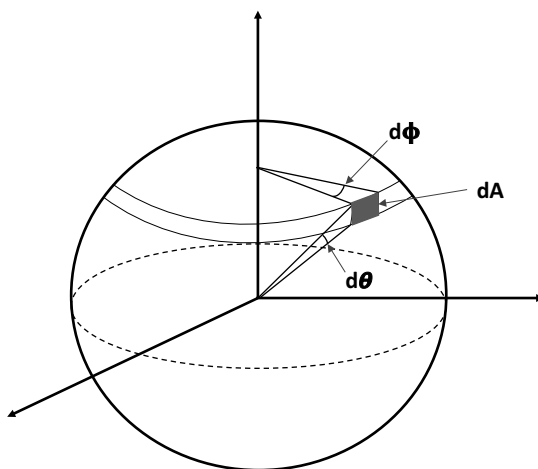
Figure 1.1: Unit sphere for understanding units of steradian. The angle θ is the polar angle, ϕ is the azimuthal angle, and dA is the area on the sphere. The radius of the sphere, r, is equal to 1.

For a cone that subtends the entire sphere, i.e. $\theta = \pi$, the solid angle is 4π. For a cone that subtends a hemisphere, i.e. $\theta = \pi/2$, the solid angle is 2π.

In ray trace software such as Light Tools, intensity can be measured in either watts/sr or lumens/sr when using what is called a "far field receiver" or a "near field receiver". Both receivers are spherical in nature, and allow measurements to be taken as power per some angular subtense. Therefore, when measuring intensity in an experimental set up, one must measure the power, and then take into account the angular subtense of the detector at some distance r from the source.

Illuminance and Irradiance

Both illuminance and irradiance integrate the radiance and luminance over an angular component. Thus, they describe the spacial distribution of power in either watts or lumens. The best way to understand this is by way of an example. Imagine a point source some distance r away from a receiver of area dA. We know from Tab. 1.2 that illuminance, E_v, is,

$$E_v = \frac{d\Phi}{dA}. \tag{1.9}$$

Also known from Eq. 1.8, and taking into account the projection angle, θ, we have,

$$d\Omega = \frac{dA_{proj}}{r^2} = \frac{dA\cos\theta}{r^2}. \tag{1.10}$$

Thus, after substituting for dA, Eq. 1.9 becomes,

$$E_v = \frac{d\Phi}{dA} = \frac{d\Phi\cos\theta}{r^2 d\Omega} = \frac{I\cos\theta}{r^2}. \tag{1.11}$$

The final substitution is made using the definition for irradiance from Tab. 1.2. Notice that this equation shows that the flux density at the target falls off as $1/r^2$, which is known as the inverse-square law. The cosine factor simply denotes the orientation of the target with respect to the source.
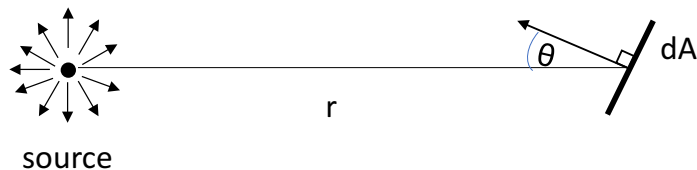
Figure 1.2: Diagram for finding the irradiance of a point source. The source is a distance r from the target which has an area dA with its normal at some angle θ from the distance line r.

Radiance and Luminance

In a lossless system, these two fundamental terms are always conserved. Another word for radiance and luminance is brightness [3]. Radiance distribution of a source, for example, an LED, describes the emission of light from each point on the LED as a function of angle. Thus, it is an important term because knowing the radiance distribution allows one to determine the propagation of radiation through an optical system. As such, radiance and luminance are the most fitting quantities to utilize in the illumination design process. With radiance, there are two types of distributions: Lambertian and isotropic.

Lambertian Distribution

The intensity from an LED source follows the cosine law, meaning that its intensity in any given direction is proportional to the maximum intensity emitted from the LED, multiplied by the cosine of the angle taken with respect to the normal of the LED surface.

$$I_{lam} = I_0 cos(\theta), \tag{1.12}$$

where I_{lam} represents the lambertian intensity and I_0 is the overall intensity of the LED. Figure 1.3 illustrates this concept.
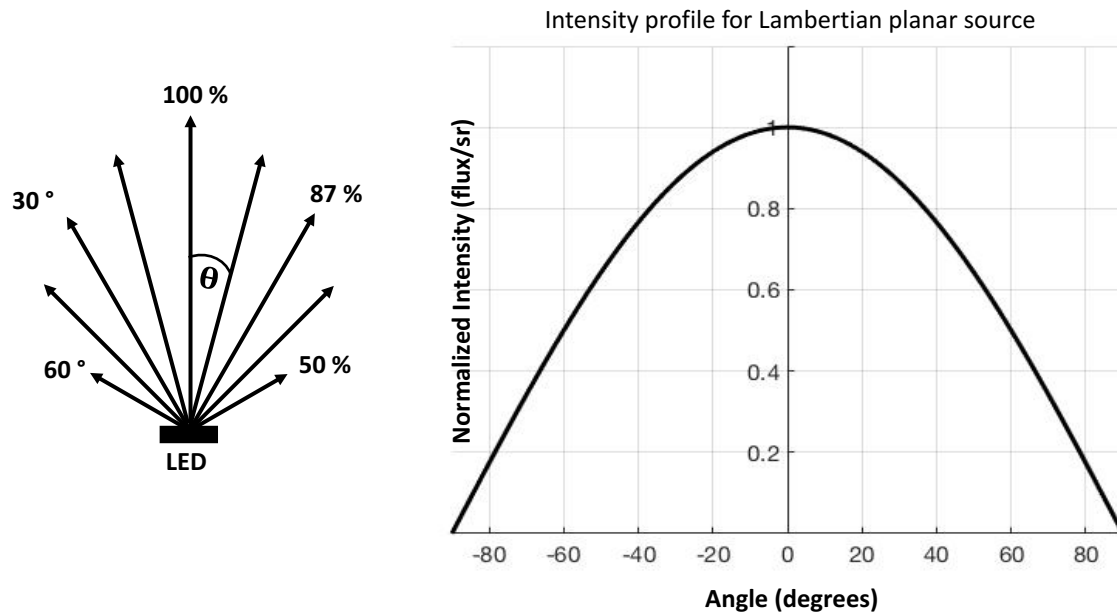
Figure 1.3: Left: Lambertian intensity of an LED. Right: Normalized intensity plotted versus angle taken with respect to the normal of the LED source

When a source is said to be lambertian, it implies that its emission profile is independent of direction. In other words, in terms of radiance, a lambertian surface possesses the same radiance when viewed from any angle. In other words, the emission profile does not depend on direction. So although intensity emitted from the LED surface is reduced by the cosine of the emission angle, the solid angle of the observed area as seen by an observer, increases by a corresponding amount. Thus, radiance, measured in $W/m^2/sr$ remains the same. This is why when viewing an LED from any direction, its brightness will appear isotropic.

1.5   Structure of Dissertation

This dissertation will focus on three topic fields to achieve designs: nonimaging optics, aplanatic optics, and freeform optics. Freeform optics is the newest of these. Nonimaging optics follows, and while it is further developed, there are still many advances to be made, particularly in flow line theory. Aplanatic optics is the oldest of the topics, having been used for the last two hundred of years in imaging optics, yet, new devices employing aplanatic

theory have emerged within the last ten years for concentration and illumination. Each of these will be discussed in more depth throughout the next chapters.

The structure of this dissertation is as follows. Chapter 2 will offer a detailed look at fundamental nonimaging concepts, delving into the basics required to understand recent advances. Nonimaging optics lies at the heart of illumination design. Therefore, the discussion should begin there. This chapter will give the reader a deeper understanding of the ideas used to construct concentrators that can be used in both solar and illumination applications. It will cover topics such as the strings method, étendue, phase space, angular space, thermodynamics, geometrical optics, and flow lines.

Following, Chapter 3 will address the advances made in nonimaging optics. The concepts used in Chapter 2 will be expanded upon. Chapter 4 will discuss controlled illumination, providing an application for reducing light pollution in Yosemite National Park. Here, a concentrator designed using principles from chapters 2 and 3 is used for illumination purposes.

Chapter 5 moves into aplanatic optics. Aplanatic optics have been around the longest, some two-hundred years. However, they have more recently moved into the solar industry because of their fast imaging capabilities, making them great concentrators and illuminators. An example of a design from theoretical development to to the prototyping process will be provided. This aplanatic design, known as the Jellyfish, has been successfully prototyped. Testing results will be offered and analyzed.

Finally, chapter 6 will delve into freeform optics - a new way of designing optics that allows one to further tailor designs. Freeform optics is a fairly new niche field of nonimaging optics. It is fast becoming a popular solution in the illumination industry. A discussion of the Supporting Quadrics Method will be given, along with a study of the diffraction effects for this method.

To tie everything together, a final chapter is provided in conclusion, that addresses future developments of the field of illumination optics.

Chapter 2

Nonimaging Optics

In its early years, nonimaging optics was developed for use in solar applications. At this time, its concepts were not widely known. In the nineties, illumination engineers began applying its principles to illumination design, and the field exploded. There were other reasons for this growth too, such as photovoltaics, and the idea that concentrators could be utilized for concentrating photovoltaic (CPV) technologies. But with the illumination industry's prevalence and wide-spread growth, specifically due to the rapid development of the LED, nonimaging optics has become a fundamental necessity in lighting.

Since its infancy, a variety new applications have emerged in this field, and while new theories have developed, there are still many ideas that remain mysterious and unanswered. In this chapter, the fundamental concepts inherent to nonimaging optics will be addressed. This will pave the way for the chapter to follow, which will address new concepts.

2.1   Background

Nonimaging optics is very much the science of thermodynamics. Unlike its name suggests, rather than possessing fundamental principles based in optics, its theories arise from the transfer of heat (radiation), probabilities, and the mapping of extended sources.

Take, for example, imaging optics. Imaging optics relies on the concept that an object is mapped to an image, via some optic. This is done point by point, with a one-to-one point mapping ratio. It is conducted using geometrical optics where radiation is represented as rays, which emanate from a point and terminate to a point. Yet, neither of these mathematical constructs exist in nature. As such, this can lead to paradoxes, such as the ellipsoid paradox [4]. The assumption that optical objects can be treated as simple points producing rays is non-physical; these points can neither produce nor receive rays in radiation heat transfer. Certainly, their usage can aid in conceptualizing and analyzing an optical system. But the solution is to also consider thermodynamics when treating the non-physical aspects of geometrical optics. Nonimaging optics at its core addresses such a challenge.

Nonimaging optics uses first principles governed by thermodynamics to solve the para-

doxes created by geometrical optics. As such, it does not treat points. Rather, extended bodies are used as sources and sinks. Then, the edges of the source are mapped to the edges of a receiver (sink). This is known as the edge-ray principle and will be discussed in detail.

The main difference between imaging and nonimaging is that no image is created. Why is this? It is because all the information within the boundary of the edge mapping - information that is carried along - is not necessarily organized.

One way to think about this is the alphabet soup example. Consider a bowl of alphabet soup whose letters are either organized or jumbled. The bowl itself is the boundary, or edge rays. The contents of the soup is all the information within. If you move the bowl from one place to another, all the contents go with it. But due to possible disorganization of the letters, there won't necessarily be any image creation at the end of its journey. This concept will become more clear after a discussion on the "string method".

Curiously enough, although the governing principles used to design nonimaging concentrators are rooted in thermodynamics, geometrical optics seems to know this. The two, though very different subjects, in the end will agree. And this can make for quite an intriguing discussion.

A good place to start our explanation of nonimaging optics is with the second law of thermodynamics and the edge-ray principle, or "strings" method. This ruling concept gives the limitations that must be accounted for when designing any nonimaging concentrator for illumination or otherwise.

2.2 Strings Method

The edge-ray principle, also known as the "Strings" method, is governed by thermodynamics. To design a concentrator using this principle, one must understand the thermodynamic origin of nonimaging optics. Nonimaging optics concerns itself with transferring energy from source to target, whilst maintaining the second law of thermodynamics.

Due to the nonimaging properties of a concentrator, radiation transfer can be accomplished in either forward(solar concentration), or reverse (illumination) directions. In the case of an illuminator, radiation is transfered in reversed such that it emanates from a concentrator's end point (absorber), and travels outward, exiting from the concentrator's entrance aperture. Thus, one need only replace the concentrator's absorber with a lambertian source. Both instances are shown in Fig. 2.1.

Probabilities for Obtaining Concentration Ratio

We begin with probabilities. The general setup of a concentrating system, as seen in Fig. 2.1, requires three components: a radiation source, a concentrator aperture, and an absorber. We assume that the source is a perfect lambertian emitter. Thus, the probability of radiation leaving the source and arriving at the receiver is defined as,
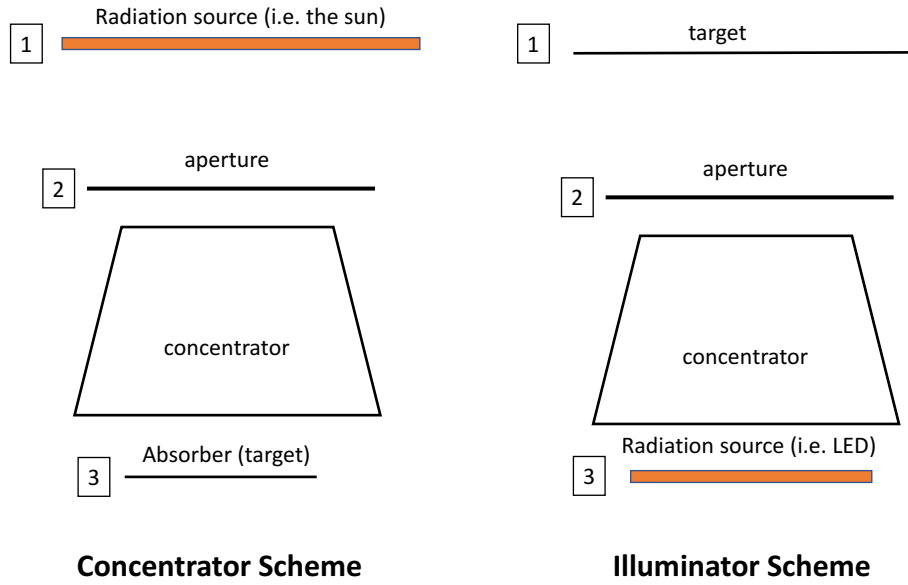
Figure 2.1: Two schemes, one for solar concentration and one for illumination. In this diagram, both employ the same concentrator. The steps labeled 1 - 3 indicate the direction with the left side being the forward direction and the right side being the reverse.

$$P(1,3) = \frac{Radiation\ incident\ on\ 3\ from\ 1}{Total\ radiation\ emitted\ from\ 1} = \frac{\Phi_{e_{1,3}}}{\Phi_{e_1}}, \tag{2.1}$$

where Φ_e is the radiant power in units of watts, as seen in Tab. 1.1. The more general form of Eq. 2.1, which can be applied to any set of surfaces, i, j would be:

$$P(i,j) = \frac{Radiation\ incident\ on\ j\ from\ i}{Total\ radiation\ emitted\ from\ i} = \frac{\Phi_{e_{i,j}}}{\Phi_{e_i}}. \tag{2.2}$$

According to the Stefan-Boltzmann law, the irradiance, E_e, is,

$$E_e = \frac{\Phi_{e_1}}{A_1} = \sigma T_1^4, \tag{2.3}$$

where A_1 is the area of the source, σ is the Stefan-Boltzmann constant, and T_1 is the temperature of the source. Therefore, the radiation from the source to the absorber is:

$$\Phi_{e_{1,3}} = \Phi_{e_1} P(1,3) = A_1 \sigma T_1^4 P(1,3), \tag{2.4}$$

The principle of reciprocity is a direct consequence of the second law of thermodynamics. With it we have,

$$\sigma T_1^4 A_1 P(1,3) = \sigma T_3^4 A_3 P(3,1), \tag{2.5}$$

where governing principles dictate that $T_1 = T_3$. This gives,

$$A_1 P(1,3) = A_3 P(3,1). \tag{2.6}$$

Equation 2.6 represents a fundamental concept: given two blackbodies at the same temperature, the radiative power from one to the other must be equivalent. From this, we can derive the radiative heat flux, q, at the surface of the absorber.

$$q_{1,3} = \frac{\Phi_{e_{1,3}}}{A_3} = \frac{A_1 P(1,3)}{A_3} \sigma T_1^4 = P(3,1) \sigma T_1^4 \tag{2.7}$$

Because $P_{3,1} \leq 1$, then $q_{1,3}$ can reach a maximum radiative flux equal to that of the source when $P_{3,1} = 1$, which is constant with the second law of thermodynamics.

Therefore, we have,

$$q_{1,3} = \sigma T_1^4 \iff P(3,1) = 1 \tag{2.8}$$

For an ideal concentrator, all of the energy passing from a radiation source through a concentrator will arrive to its absorber. Note that this does not account for optical losses due to materials. For a non-ideal system, refer to [5]. Therefore, for an ideal concentrator, we have:

$$\Phi_{e_{1,2}} = \Phi_{e_{1,3}}. \tag{2.9}$$

From this equation, we see that all the radiation passing from the source to the aperture is the same as the radiation passing from the source to the absorber. Expanded, we have,

$$\sigma T_1^4 A_1 P(1,2) = \sigma T_1^4 A_1 P(1,3), \tag{2.10}$$

which gives,

$$A_1 P(1,2) = A_1 P(1,3) \tag{2.11}$$

Finally, to obtain the maximum concentration ratio, which shows the highest amount of concentration one can have with any given source, we once more use the principle of reciprocity:

$$A_1 P(1,2) = A_2 P(2,1). \tag{2.12}$$

Combining Eqs. 2.6 2.11 and 2.12, we have,

$$C = \frac{A_2}{A_3} = \frac{P(1,3)}{P(2,1)} \leq \frac{1}{P(2,1)} \tag{2.13}$$

First note that the left portion of Eq. 2.13 represents what is called the geometrical concentration ratio. One can determine any geometrical concentration ratio if the concentrator

aperture and concentrator absorber area are known. Second, it is interesting that from thermodynamics and probabilities, we arrive at an equation for maximum concentration based solely on geometry. Equation 2.13 represents the design principle for any ideal nonimaging concentrator. By requiring that the optical efficiency of a concentrator be maximized and that concentrated flux be maximized at the absorber, then one obtains a concentration ratio of $1/P(2,1)$. This equation also shows that such a design is achieved when $P_{1,3} = 1$, according to Eq. 2.8. Thus, the thermodynamic approach of maximizing the absorber temperature such that it is equavalent to the radiation source, or, $P(3,1) = 1$, serves as the underlying principle for most nonimaging designs.

Although the thermodynamic understanding of nonimaging optics presented above does not offer a direct design method to generate a nonimaging concentrator, it does provide a theoretical limit and some intuition for the designing process. It also serves as a guideline on how nonimaging optical system can be designed, as shown below.

The String Method

Hoyt Hottel, an MIT engineer working on the theory of furnaces, [6], showed a convenient method for calculating radiation transfer between walls in a furnace using "strings". The theory can be expanded to any surface in a system that exchange radiation, and is explained as follows:

Suppose you want to know the probability of radiation traveling from one surface to another. Shown in Fig. 2.2, by connecting strings c, d to the end points of the surfaces A, B (short strings), and connecting cross strings a, b to the surfaces A, B (long strings), the lengths of these strings can be used to determine probability of radiation reaching one surface to another. The probability of radiation from one surface reaching the other is the sum of the long strings minus the sum of the short strings, divided by two times the area of the surface where radiation is emitted (Equation 4.48 in [7])
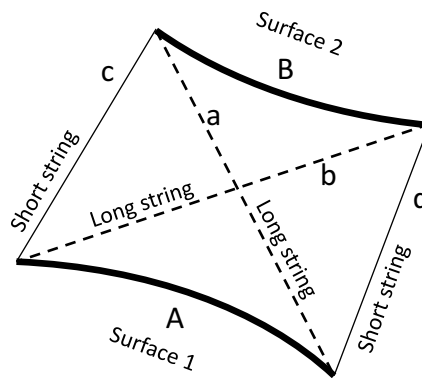
Figure 2.2: Two surfaces connected by strings.

$$P(1,2) = (A_a + A_b) - (A_c + A_d)/2A_1.$$ (2.14)

By connecting the two surfaces A, B by strings, while ensuring that the probability of radiation traveling from one surface to the other maximized to one, we can construct any concentrator that is thermodynamically efficient. The short strings become the walls of the concentrator. The curves of these short strings arise from mapping the rays (representative of electromagnetic radiation) from the edges of a source to the edges of a receiver, and vice versa. This is how the compound parabolic and compound elliptical concentrators are constructed.

As an example of this method, the compound elliptical concentrator (CEC) will be used [8]. Two surfaces, a source and a receiver, are constructed asymmetrically (see Fig.2.3). The cross strings, or long strings, are first constructed joining opposite sides of the two surfaces. Radiation from the edge of surface three at point c' is mapped to the edge of surface one at point a' via an elliptical reflector bc. The reflector is elliptical because that is the shape that sends radiation from the source to the receiver. In other words, the elliptical curve in the figure is constructed segment by segment so as to reflect radiation from one edge point of the source to the other.
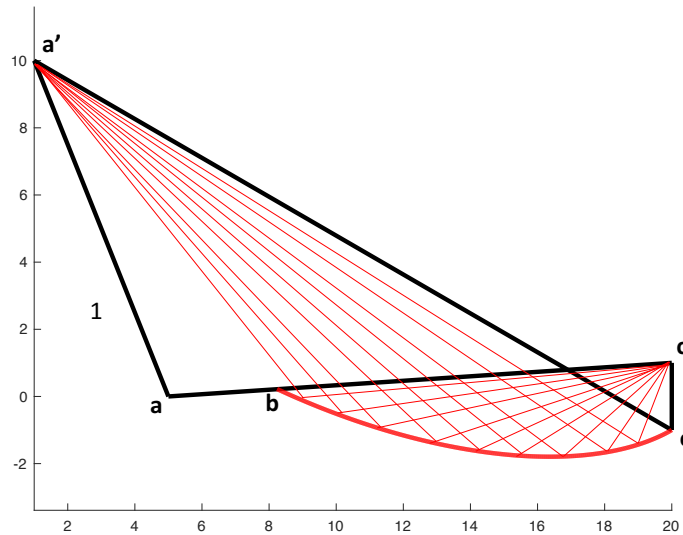
Figure 2.3: Demonstrating the edge-ray principle using strings. Rays are mapped from c' to a' and the resultant string is bc.

The same can be done with the opposite edge points to generate another elliptical reflector $b'c'$. Thus, a concentrator forms, whose aperture is then bb' as shown in Fig. 2.4. The combination of the two elliptical reflectors along with surface three make what is known as an asymmetrical compound elliptical concentrator (ACEC) [9].
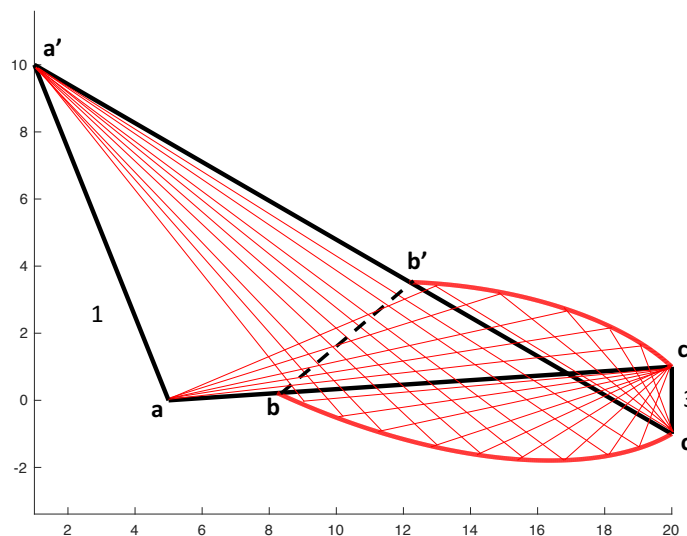
Figure 2.4: Using the same method, rays are mapped to generate the opposite side of the asymmetrical compound elliptical concentrator

The strings method, which uses the edge-ray principle, is the simplest way to construct a concentrator. So simple, that one can use household items such as a ball of twine, some thumb-tacks, and a pencil, to construct concentrator profiles that efficiently transport radiation. This efficient transportation of radiation is called "efficient étendue engineering," and will be discussed in the following section.

2.3   Étendue

Étendue is a measure of power accepted by an optical system concentrated onto an aperture. Although all optical systems conserve étendue (assuming there is no loss of flux), not all systems exhibit efficient étendue engineering: power maximized over a concentrator's aperture. When a concentrator is designed to be ideal, the étendue has been engineered efficiently. In a nonideal system, étendue is wasted. A good example would be a parabolic trough, where the power over the target is not maximized. This creates an unfavorable distribution for concentration purposes. To illustrate this idea, a comparison is given in Fig. 2.5 between a parabolic trough and the first concentrator invented using nonimaging optics: the compound parabolic concentrator (CPC).

Étendue is a geometric quantity. To grasp the concept of it, one must understand that light traveling through any optical system requires space, which is represented by two components: area and angular extent. Together, these components are the étendue. Thus, "étendue" is an apt name, because it is derived from the french term étendue géométrique, meaning "geometrical extent". To provide an example of the geometrical representation of étendue, consider light crossing through an area dA, which is at some angle θ to the normal
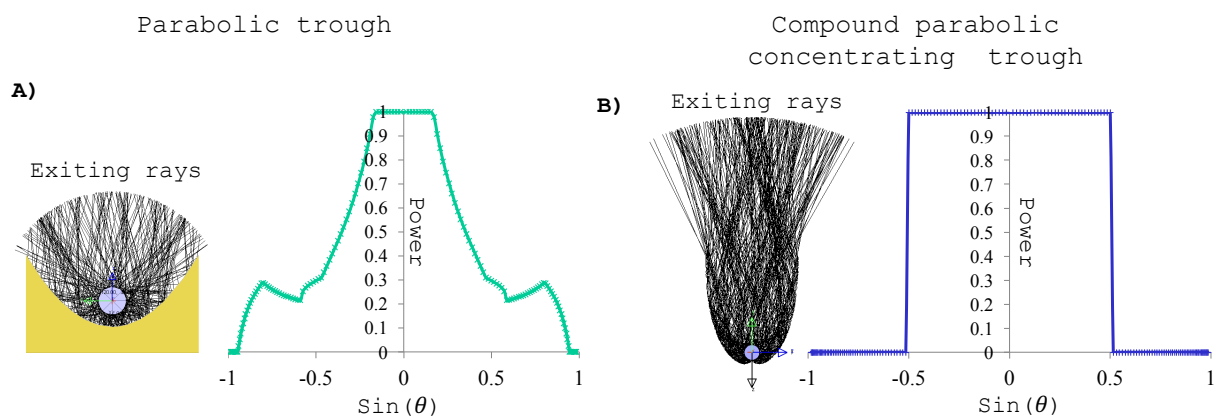
Figure 2.5: A comparison between two concentrators. A) shows the power distribution vs. angle for a parabolic trough and is an example of inefficient étendue engineering. B) shows the power distribution for a compound parabolic concentrator trough and is an example of efficient étendue engineering.

of dA. This light will see the projected area $dAcos\theta$ as the available space to pass through. This is its spatial room. The spatial room multiplied by the angular room, $d\Omega$ which is defined by the solid angle of the beam, will give us the étendue. Here we assume an index of refraction, n, of unity.
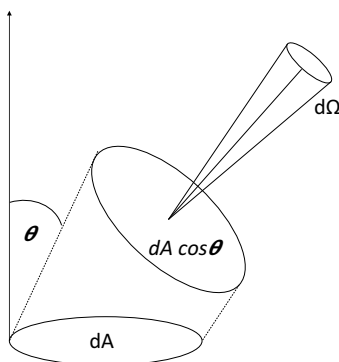
Figure 2.6: The geometrical construct of étendue requires both area and angular extent.

In differential form we have,

$$dU = dAcos\theta d\Omega. \tag{2.15}$$

Because it is invariant, if the area available for the light is increased, the solid angle must be decreased, and vice versa.

To think about this conceptually, consider Fig. 2.7. Light passes from a circular source into a box. The angular extent of the beam is 2θ and the opening at the box is AB. If AB decreases, θ must increase and vice versa. Thus, the étendue remains constant.
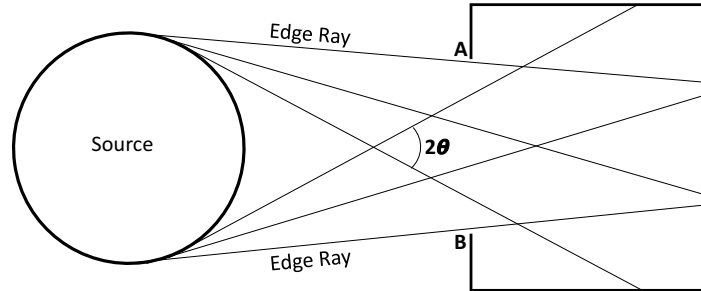
Figure 2.7: An example of the relationship between the angular extent of a light beam and the area through which it passes.

For small angles, étendue is defined as the quantity $A^2\theta^2$ (for a refractive index of unity) where A is the length of an aperture through which light passes, and θ is the semi-angle of the beam accepted by the system[10].

For nonimaging optics purposes, étendue can be defined in terms of area and optical momentum. As shown in Fig. 2.8, when given the path of a ray, the optical momentum is a vector defined at each point along its path [2]. For a ray traveling through a uniform medium, the optical momentum, $\mathbf{p}$, in three dimensional Cartesian space is,

$$\mathbf{p} = (p_x, p_y, p_z) = (n\cos\theta_x, n\cos\theta_y, n\cos\theta_z) = (nL, nM, nN), \qquad (2.16)$$

where n is the index of refraction of the medium, $\|\mathbf{p}\| = n$, the angles $\theta_x, \theta_y, \theta_z$ are the angles the vector makes with each respective axis, and L, M, N are the direction cosines.

The differential element of étendue is then given by[11],

$$dU = dx\,dy\,dp_x\,dp_y. \qquad (2.17)$$

The étendue entering an optical system is conserved when

$$\int_A dx\,dy\,dp_x\,dp_y = \int_{A'} dx\,dy\,dp_x\,dp_y, \qquad (2.18)$$

where A and A' are the entrance and exit apertures of a light collector[12].

Several proofs exist and are documented for Eq. 2.18 [1, 2, 13], however, the most intuitive proof cannot be found in any text book. Rather, it can be derived from Enrico Fermi's own handwritten notes, which were published shortly after his death [14]. Few in the field of nonimaging optics have seen his notes regarding étendue, and fewer still have applied them
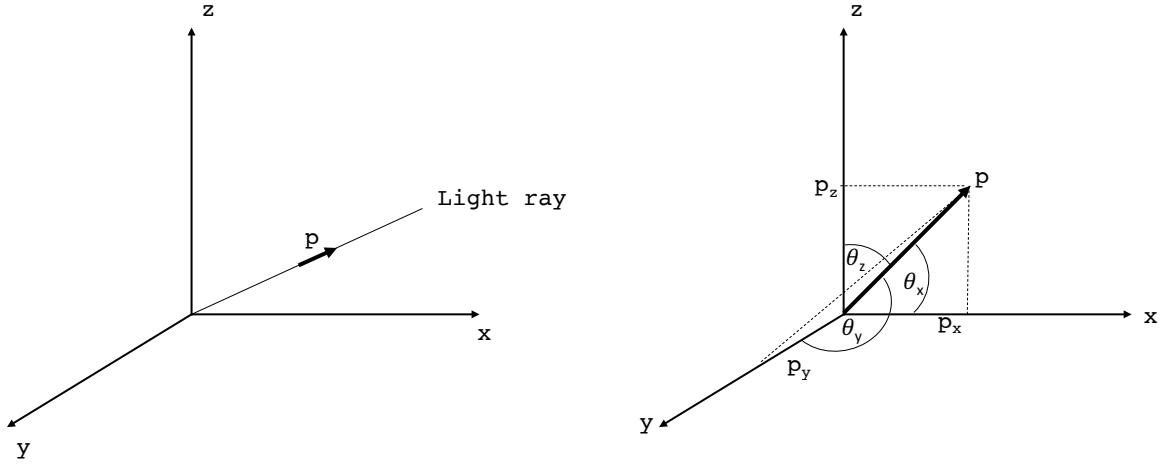
Figure 2.8: The optical momentum is the vector **p**, tangent to the ray path.

to the proof in its entirety. Thus, no formal publications exist to document it. So, it has been derived here in more detail.

To begin, we first consider the Legendre transformation of the Optical Legrangian:

$$p_x\dot{x} + p_y\dot{y} - \mathcal{L}(x, y, \dot{x}, \dot{y}) = \mathcal{H} \tag{2.19}$$

Taking the derivative, we have

$$d\mathcal{H} = p_x d\dot{x} + dp_x\dot{x} + p_y d\dot{y} + dp_y\dot{y} - \left(\frac{\partial\mathcal{L}}{\partial x}dx + \frac{\partial\mathcal{L}}{\partial y}dy + \frac{\partial\mathcal{L}}{\partial\dot{x}}d\dot{x} + \frac{\partial\mathcal{L}}{\partial\dot{y}}d\dot{y}\right) \tag{2.20}$$

$$d\mathcal{H} = -\dot{p}_x dx - \dot{p}_y dy + \dot{x}dp_x + \dot{y}dp_y. \tag{2.21}$$

This brings us to the familiar Hamiltonian equations, $\frac{\partial\mathcal{H}}{\partial x} = -\dot{p}_x$, $\frac{\partial\mathcal{H}}{\partial y} = -\dot{p}_y$, $\frac{\partial\mathcal{H}}{\partial p_x} = \dot{x}$, $\frac{\partial\mathcal{H}}{\partial p_y} = \dot{y}$. Now we can construct a vector, $\vec{W} = (\dot{x}, \dot{y}, \dot{p}_x, \dot{p}_y)$. If the divergence is taken, we get

$$\nabla \cdot \vec{W} = \left(\frac{\partial W}{\partial\dot{x}} + \frac{\partial W}{\partial\dot{y}} + \frac{\partial W}{\partial\dot{p}_x} + \frac{\partial W}{\partial\dot{p}_y}\right) = \frac{\partial\dot{p}_x}{\partial p_x} + \frac{\partial\dot{p}_y}{\partial p_y} + \frac{\partial\dot{x}}{\partial x} + \frac{\partial\dot{y}}{\partial y} = 0, \tag{2.22}$$

because we know that from Hamilton's equations and Eq. 2.21, $-\frac{\partial}{\partial p_x}\left(\frac{\partial\mathcal{H}}{\partial x}\right) = \frac{\partial\dot{p}_x}{\partial p_x} = -\frac{\partial}{\partial x}\left(\frac{\partial\mathcal{H}}{\partial p_x}\right) = \frac{\partial\dot{x}}{\partial x}$, etc. Thus, the field of the four-dimensional vector $\vec{W}$ has an important property, namely, its divergence, $\nabla \cdot \vec{W}$, is zero. In other words, the four-dimensional hyperspace of (x, y, p_x, p_y) conserves volume as light rays evolve in an optical system.
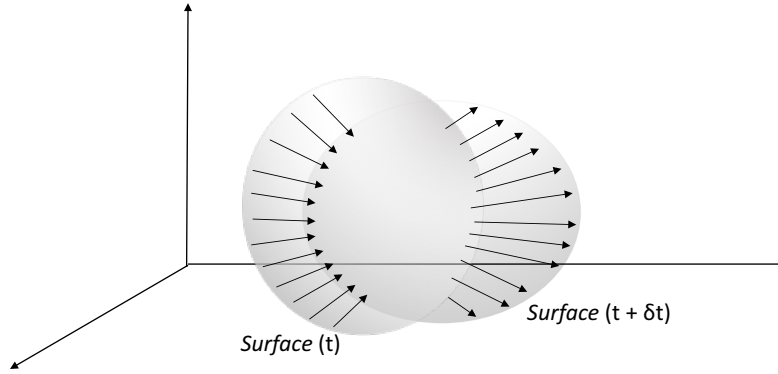
Figure 2.9: Change in the fluid's surface while conserving volume.

An analogy can be offered. Consider an incompressible fluid.

This fluid possesses volume, V, which is surrounded by a closed *surface*. We then have the velocity field, $\vec{v}$ of the small elements within the fluid. As the fluid begins to flow according to the change in t, and the $surface(t)$ begins to evolve into $surface(t + \delta t)$, the enclosed volume of V, will also change as:

$$V(t + \delta t) - V(t) = \oiint\limits_{surface(t)} \vec{v}\delta t \cdot d\vec{s}, \tag{2.23}$$

where $d\vec{s}$ is the surface vector pointing outward along the normal direction. Utilizing Gauss's Divergence Theorem we have,

$$\oiint\limits_{surface(t)} \vec{v}\delta t \cdot d\vec{s} = \iiint \nabla \cdot \vec{v}d\tau. \tag{2.24}$$

where $d\tau$ is the volume element. If $\nabla \cdot \vec{v} = 0$ everywhere, then it is known that $V(t + \delta t) - V(t) = 0$, or, $V(t) = constant$. Applying this idea to étendue conservation, it is found that the four-dimensional volume (x, y, p_x, p_y) remains constant as it evolves with time. The vector field $\vec{W} = (\dot{x}, \dot{y}, \dot{p}_x, \dot{p}_y)$, replaces $\vec{v}$, and $\nabla \cdot \vec{W} = 0$ just as $\nabla \cdot \vec{v} = 0$. This not only implies that the étendue is conserved, but also that there is no source or sink of field $\vec{W}$.

For the design of a nonimaging concentrator, holding étendue constant such that it is not lost, acts as a constraint. By designing an optic such that power over a target is maximized, the étendue is not only conserved, but also engineered efficiently. To further understand étendue and its importance, delving into phase space is necessary.

2.4 Phase Space

To better understand phase space, consider a two-dimensional concentrator, specifically in this case a CPC, as shown in Fig. 2.10. A CPC accepts all rays with projected angles in the $x - z$ plane $\leq \theta_{max}$, the design cut-off angle. All rays outside of the accepted angular range will be rejected. All rays within will arrive at the aperture A' after a single reflection to one of the walls of the concentrator.
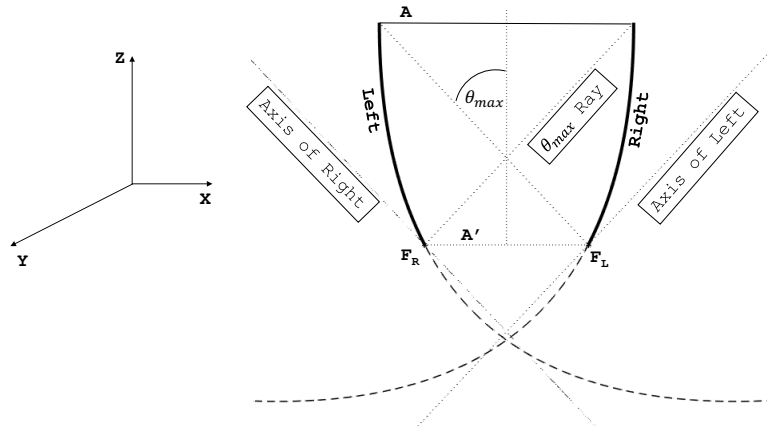
Figure 2.10: Symmetric CPC where A and A' are the entrance and exit apertures, $Left$ and $Right$ are the left and right parabolas, and F_L and F_R are the focus points of the left and right parabolas.

Rays propagating from entrance aperture A to exit aperture A' are preserved (not accounting for absorption). These rays at the entrance and exit apertures occupy areas in phase space [12]. In real space, the rays are lines, but in phase space, they are points. The populated points correspond to each position at the entrance and exit apertures for a specific angular value (direction) of each light ray. The result is a rectangle in phase space possessing axis p_x and x. For any collector possessing sharp cut-off angles, the phase space geometry will be rectangular. The length along the x axis is equivalent to the aperture's length, and the height along the p_x axis is equivalent to $\sin \theta_{max}$ values. The phase space areas of both apertures must be equivalent because they represent the étendue, which is conserved.
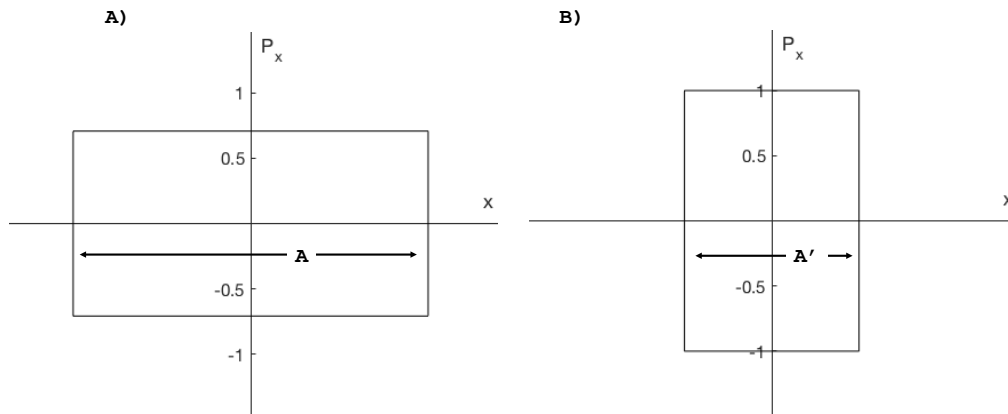
Figure 2.11: Phase space areas of a two dimensional collector. The area in phase space occupied by the left diagram is equivalent to the area occupied by the right, as it must be for conservation of étendue. A and A' are the entrance and exit aperture diameters, respectively.

Given a specific location on an aperture, a range of angles are accepted, equivalent to $\leq \theta_{max}$. In phase space, this will populate as points along a vertical line cutting through the $x - axis$. These vertical lines of identical length will occur along the $x - axis$ for each position on the aperture for a two-dimensional concentrator. When rays enter the entrance aperture at a specific angle, the phase space area is occupied by those rays. This is the the étendue at the entrance aperture. The output aperture is a smaller length, so the angle at the output aperture must increase. The CPC and many other nonimaging concentrators are designed such that the output aperture A' is 90°, because that is the angle of a lambertian emitter. Naturally, this constraint requires that the length of the output aperture A' shrink to accommodate.

Angular Acceptance

Considering the CPC as shown above (Fig. 2.10), we can take the ideas of étendue and phase space a step further to consider angular acceptance. For a two-dimensional light collector in possession of sharp cut-off angles, one can plot the optical momentum L versus M in what is called "angular space". From Fig. 2.12, we see that to determine $\tan\theta_{max}$ of the acceptance angle, we project the light ray onto the x and z axis. This gives the following relationship,

$$\tan(\theta_{max}) \geq \frac{L}{N}. \tag{2.25}$$

This equation can be squared on both sides. And because the three optical momenta (L, M, N) cannot be larger than unity such that collectively the largest they may be is less than or equal to unity, we have, $L^2 + M^2 + N^2 \leq 1$. This equation gives a sphere with radius of unity in three-dimensional L, M, N space, or a circle with radius of unity in two-dimensional

L, M space, when the values are at their maximum. Solving for N^2 and substituting this into Eq. 2.25, we obtain

$$L^2 \leq \tan^2(\theta_{max})(1 - L^2 - M^2). \tag{2.26}$$

With algebra we can rearrange and simplify such that,

$$\frac{L^2}{\tan^2(\theta_{max})} \leq (1 - L^2 - M^2), \tag{2.27}$$

$$\frac{L^2}{\tan^2(\theta_{max})} + L^2 + M^2 \leq 1, \tag{2.28}$$

$$L^2 \left(\frac{1}{\tan^2(\theta_{max})} + 1 \right) + M^2 \leq 1, \tag{2.29}$$

$$L^2 \left(\frac{\cos^2(\theta_{max})}{\sin^2(\theta_{max})} + \frac{\sin^2(\theta_{max})}{\sin^2(\theta_{max})} \right) + M^2 \leq 1. \tag{2.30}$$

Finally, we arrive at the equation for the angular acceptance of the CPC [15],

$$\frac{L^2}{sin^2(\theta_{max})} + M^2 \leq 1 \tag{2.31}$$

where the angular acceptance fills an ellipse of semi-minor axis equal to $sin(\theta_{max})$ and semi-major axis equal to unity. The area of this ellipse is then $\pi sin(\theta_{max})$.
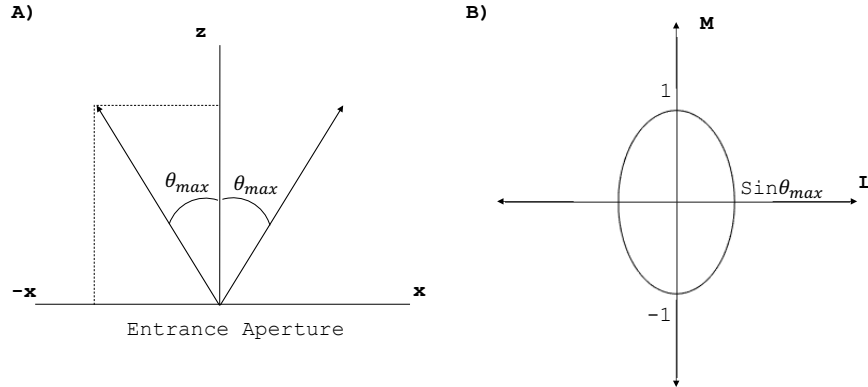
Figure 2.12: A: Condition of angular symmetry at the entrance aperture. B: Angular acceptance in L, M space.

The result shown in Fig. 2.12 is for symmetric cases when the concentrator adheres to symmetry about the optical axis. Cases where symmetry is broken will be discussed in further detail in Chapter 3.

2.5 Flow Lines

In the scheme of geometrical optics, the propagation of the totality of light rays can be pictured as a fluid flow in six-dimensional phase space [1]. The six-dimensional phase space, (p, x), has components of generalized coordinates (x_1, x_2, x_3) and generalized momenta (p_1, p_2, p_3). In this sense, it is considered a geometric vector flux, but for simplicity the term flow lines was coined. Flow lines have been used in illumination as the "photic field" [16], and also in computer graphic rendering as a light field [17, 18]. The concept was introduced as early as 1893 [19] as simply a light vector.

Flow Line Definition

The flow of geometric vector flux is subject to conservation theorems loosely referred to as Liouville's theorem, but more appropriately called the integral of invariants of Poincaré, and is defined as [1],

$$\mathbf{J} = \left(\int \int dp_y dp_z, \int \int dp_x dp_z, \int \int dp_x dp_y \right), \tag{2.32}$$

where $\mathbf{J}$ is invariant such that it has zero divergence.

The geometrical vector flux, $\mathbf{J}$, always points in the direction of the flow line. $\mathbf{J}$ is always the bisector of edge rays, with a direction and magnitude defined by the angle between the two rays. The magnitude of $\mathbf{J}$ is derived in The Photic Field [16] and later in [1, 2] to be,

$$\|\mathbf{J}\| = 2n \sin \theta, \tag{2.33}$$

where n is the index of refraction through the material, and θ is the angle between the two bisected rays shown in Fig. 2.13.
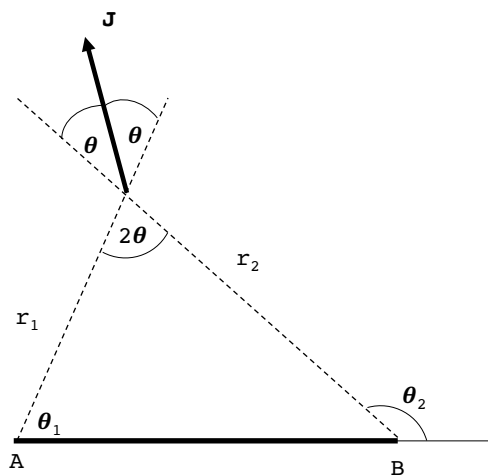
Figure 2.13: The bisector, **J** bisects the two extreme rays r_1 and r.

The flow line never disrupts the flow of rays. Mirrored surfaces can be placed along the flow line without altering the light. To further illustrate this concept, we apply it to an extended lambertian source as in Fig. 2.14. Edge rays emanate from points A and B and make an angle θ between them at point X. As X moves along the flow line towards AB, the angle θ grows larger. At AB, θ becomes π because the source is lambertian with angles of $\pm\frac{\pi}{2}$ to its normal. Therefore, the flow lines from a flat lambertian source are hyperbolas; it is an inherent property of the hyperbola to bisect the orthocenter of a triangle, in this case, triangle AXB. The same can be said of the reverse case where point X approaches Y such that the angle, θ, gets smaller.
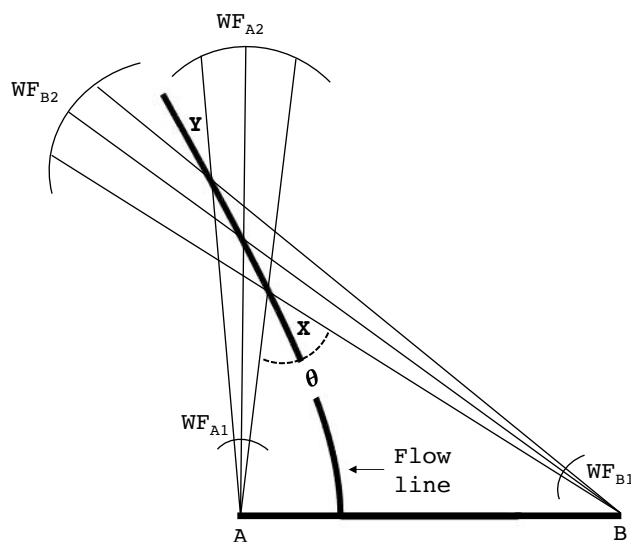
Figure 2.14: Derivation of flow lines from a lambertian source. Wave fronts emanate from edge rays and are used to determine the flow lines.

In Fig. 2.14, all the rays flowing from WF_{A1} arrive at the flow line. If there is no mirrored surface along the flow line, the wave front continues on eventually becoming WF_{A2}. However, if mirrors are placed along the flow line intercepting WF_{A1}, then this wave front will reflect according to the law of reflection, and become WF_{B2}. The same is true for WF_{B1}. With no disruption, it becomes WF_{B2}. But if mirrors are placed along the flow line, it becomes WF_{A2}. In this way, placing mirrors does not disrupt the system, because the system still produces the same secondary wave fronts regardless of the mirror placement.
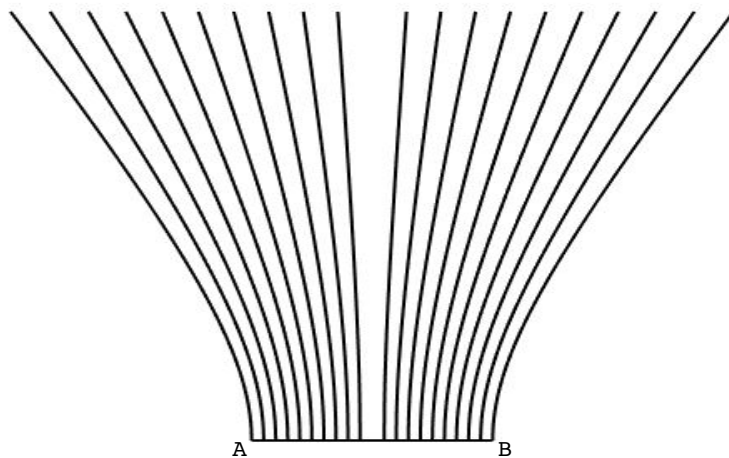
Figure 2.15: Flow lines from a lambertian source.

Figure 2.15 shows a lambertian source with flow lines. These flow lines bisect all of the rays emitted from the lambertian source. For any flat linear or disk-like source, the flow lines will be lambertian.

Finally, we consider the remarkable results when drawing strings and flow lines between two surfaces A_1, A_2 as shown in Fig. 2.16. Note first that A_1 and A_2 are different lengths. The strings are the dotted lines consisting of both the long cross strings and the short strings. As mentioned in Eq. 2.14, the probability of radiation from one of the surfaces reaching the other is the sum of the short string lengths subtracted from the sum of the long string lengths normalized by twice the area/length [20].
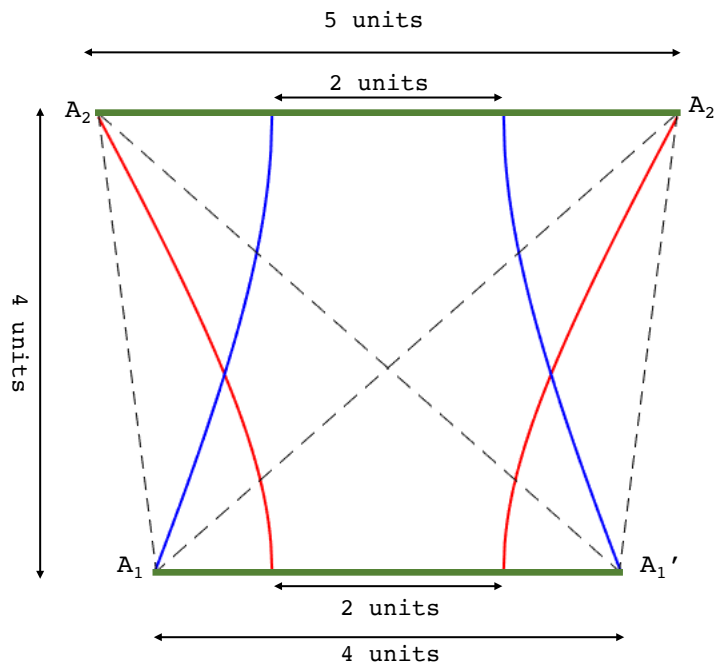
Figure 2.16: Flow lines from two lambertian sources each of a different size. The top source is 5 units while the bottom source is 4 units. The flow lines still begin with the same width of 2 units from each.

The short and long string emerging from A_2 and terminating at both ends of A_1 are bisected by a hyperbola, as was mentioned previously, this is an inherent property of the hyperbola. The same can be said about the short and long string emerging from A_2'. Pairs of strings can be drawn from the edges of each surface to various points along the hyperbolas and their difference will remain constant. In this example, that constant length is 2 units. In other words, even though their lengths are different, the amount of radiation A_2 receives from A_1 must be the same as the amount of radiation A_1 receives from A_2. If it weren't, this example would defy the second law of thermodynamics. And that makes this example all the more intriguing. The flow lines are purely a geometric concept, where as the strings

are derived using thermodynamics. Yet they both agree, as if geometry knows and follows the laws of thermodynamics. The fact that this agreement exists is not trivial, nor is it coincidental.

2.6 Conclusion

The theories inherent to nonimaging optics are rooted in thermodynamics. Yet, geometry can be used as a secondary proof to explain its results. This is a fascinating idea. Examples are given in this chapter using the concentration ratio (Eq. 2.13), as well as the relation of strings and flowlines from lambertian sources.

Concepts such as probabilities of radiation transfer, étendue, phase space, angular space, and flowlines are all fundamental concepts of nonimaging optics, and understanding each is crucial for nonimaging optical design. This chapter discusses each of these topics, developing the reader's understanding for future work in this dissertation. Most important of these topics is the edge-ray design principle, because it can be used to design nonimaging optical designs. This principle utilizes the idea that, by accounting for the edge rays of an emitter, information will be included within. Most impressive is that there is no need for organization of this information. And this concept is what gives nonimaging optics its name.

The most mysterious of nonimaging topics discussed here are flowlines. Flowlines are a way to represent light as if it were a fluid flow in six-dimensional phase space. Flowline theory has huge potential, lending itself to the development of new concentrators, which will be discussed in the following chapter.

Finally, it is important to understand that nonimaging optics derives its fundamental theories from the idea that extended sources are mapped to sinks via edge-ray mapping. This is a thermodynamic process, meaning, radiation is transferred from one surface to another. So despite the suggestion of its name, radiation transfer is at the heart of this field.

Chapter 3

Advances in Nonimaging Optics

The first nonimaging optical concentrator, the compound parabolic concentrator, was developed in the ninteen-sixties for Čerenkov radiation detection [21, 22, 23]. It was later applied to solar collection for producing heat as a two-dimensional design extruded into the third dimension. Today these are known as compound parabolic concentrators (CPC). With the invention of the CPC came the modification of symmetry and the asymmetrical compound parabolic concentrator emerged. The ACPC was evaluated in the seventies as a viable option for solar thermal generation and has since been used for specialized purposes of solar thermal and solar photovoltaic collection. The discussion in this chapter pertains to the ACPC.

3.1   The Asymmetric Compound Parabolic Concentrator

The Asymmetric Compound Parabolic Concentrator (ACPC) is similar to the CPC in that it has two parabolic sides and is constructed in much the same way using the edge ray method. However, the ACPC possesses differing acceptance angles θ_L and θ_R (Fig.3.1), whereas the CPC has identical acceptance angles θ_{max}, which happen to be centered about the optical axis. Unique to the asymmetry of the ACPC is the fact that the design angles φ_L and φ_R are not the same as acceptance angles θ_L and θ_R, which are always taken with respect to the concentrator's entrance aperture, A. Highlighting this distinction is necessary because using the design angles inappropriately as the acceptance angles will give inaccurate values for concentration and étendue. Therefore, it is best to remember that acceptance angles are always taken with respect to the concentrator aperture's normal.
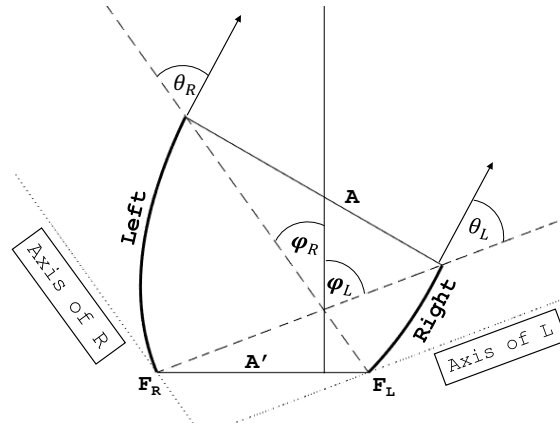
Figure 3.1: Diagram of an ACPC with left and right design angles φ_L and φ_R taken with respect to the normal of the absorber A', and acceptance angles of θ_L and θ_R taken with respect to the normal of the entrance aperture A. The focus points of the left and right parabolas are labeled F_L, F_R respectively.

Due to the differing acceptance angles, the geometric concentration of the ACPC is not the same as Eq. 2.13,

$$C = A'/A = \frac{\sin(\theta_L) + \sin(\theta_R)}{2}. \tag{3.1}$$

The difference in acceptance angles makes the ACPC an ideally tuned concentrator for winter-summer use, when the sun's location differs in the sky [24, 25, 26]. For example, the sun is higher in the sky during summer than in winter. Thus, assuming the ACPC trough is arranged such that the transverse direction is east-west, we can see by looking at Fig. 3.1 that radiation incident in the direction of $\leq \theta_R$ would be ideally suited for summer collection, and radiation incident in the direction of $\leq \theta_L$ would be ideally suited for winter collection.

Various studies have addressed the ACPC as a viable option for solar concentration [27, 28, 29, 30, 31]. For example, the ACPC can be mounted to the side of a building or wall when designed for extreme asymmetry, taking advantage of sunlight as it passes overhead. They can be useful in latitudes that have longer winters for use in heating, or longer summers for use in cooling, while still taking advantage of both seasons. And finally, they can demonstrate higher peak concentrations [32], which is an impressive result.

Use of the ACPC in illumination is not prevalent in the literature. The literature also lacks theoretical backing and development for the ideas inherent to the ACPC. This has been mitigated [33], and the results will be discussed in following sections.

Constructing the ACPC and calculating its acceptance angles

The construction of the ACPC begins with three parameters: the absorber size, and the two design angles. In a following section, it will be shown that the acceptance angles can be calculated using the design angles. If the ACPC is designed using the "string method", the absorber length is selected and drawn first, followed by the selection of the design angles which allows for the construction of the diagonal strings as seen in Fig. 3.2. With these three parameters, the parabolic sides can be constructed point by point from the edges of the absorber to the diagonal strings.
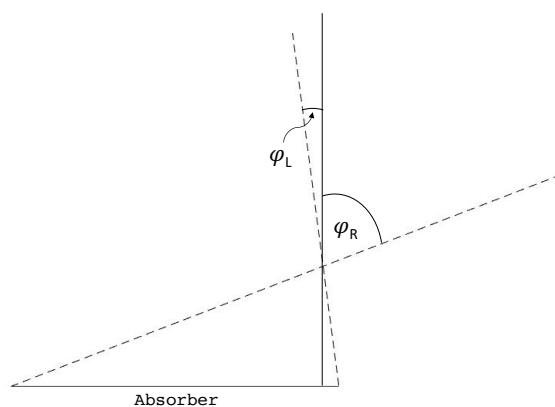
Figure 3.2: The ACPC construction begins with an absorber of some length l, and two extreme rays (dashed lines) defined by the design angles of φ_L and φ_R.

Like the CPC, the ACPC is also an ideal concentrator: étendue is transported efficiently to maximize power over the target. Figure 3.3 demonstrates ideal concentration with the transmission-angle curve for a 2D ACPC traced in forward and reverse directions. Figure 3.4 shows images of the design being traced.
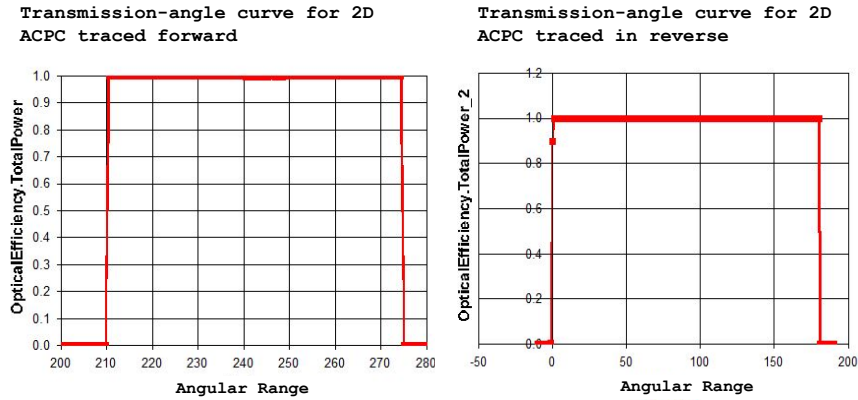
Figure 3.3: Left and Right show ideal concentration using forward and reverse ray trace over a range of angles (with respect to the object's orientation in the ray trace program). The ACPC has design angles of $\varphi_L = 60°$, and $\varphi_R = 5°$, images of which can be seen in Fig. 3.4 below. The left image is the range of angles accepted $= 65°$ and the right is the range of angles received $= 90°$, as expected.
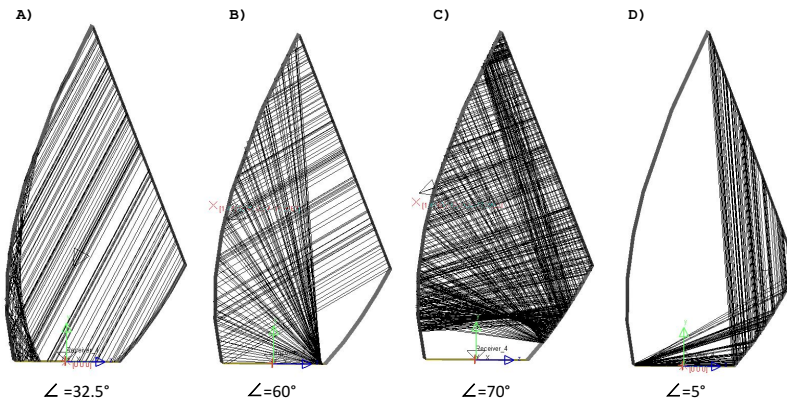
Figure 3.4: Images A) - D) are all the same ACPC with design angles of $\varphi_L = 60°, \varphi_R = 5°$ Four instances of forward ray trace are given with different incident ray angles. The incident rays for c) exceed the maximum acceptance angle, thus, rays never arrive at the absorber.

Determining the Acceptance Angles With Respect to Aperture

Like for the CPC, the ACPC's étendue can be calculated and the angular acceptance derived. It is important to note that the entrance aperture of the ACPC is no longer symmetric about the $z-$axis as it is for a CPC. Only what is seen coming out of the entrance aperture of the

ACPC (for illumination) or entering the aperture (for solar concentration), matters. It is assumed that nothing behind this surface in the $-z'$ direction is known.
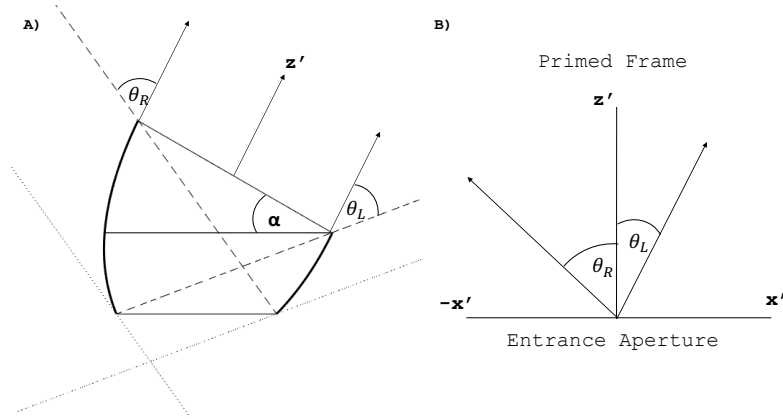
Figure 3.5: Primed Frame for calculating θ_L and θ_R. Figure A) shows the location of the angle α between the entrance aperture and the horizontal. Figure A) shows the relationship between primed angles and the entrance aperture for the ACPC.
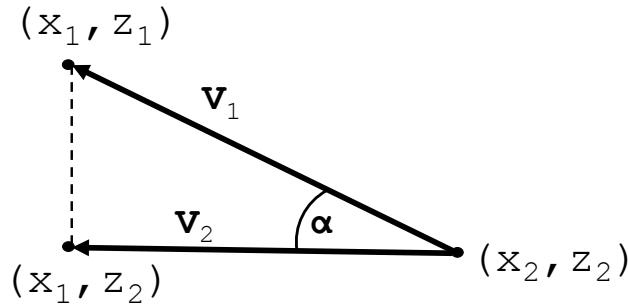
Figure 3.6: Vector diagram for calculating α using the points from curves generated for the sides of the ACPC.

The new z-axis, z', is perpendicular to the ACPC's aperture, and the new x-axis, x', is along the entrance aperture. Thus, we have effectively rotated $z \to z'$ by an angle α. We will call α our primed frame rotation. Acceptance angles θ_L and θ_R are taken with respect to this primed axis orientation, specifically, z'. To determine them, we must find the axis rotation, α. This rotation can then be applied to the design angles φ_L and φ_R to

get the acceptance angles. Alpha can be found once the curve's points for each parabolic side are generated. These curves can be generated using the String Method for asymmetric concentrators [34].The end points of the curves $Left$ and $Right$ are then taken as (x_1, z_1) and (x_2, z_2), and the vectors $\mathbf{v}_1, \mathbf{v}_2$ can be obtained. With these vectors, α is determined.

$$cos(\alpha) = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\|} \tag{3.2}$$

The new relationship between the primed frame α, the design angles, and acceptance angles are as follows:

$$\theta_L = \varphi_L - \alpha \tag{3.3}$$

$$\theta_R = \varphi_R + \alpha \tag{3.4}$$

It should be mentioned that φ_L and φ_R are taken as positive because the subscripts L, R denote on which side of the vertical axis their maximum angle rays fall. After calculating the acceptance angles, if α's value causes one of the acceptance angles to become negative, this means that the acceptance angle "flips" to the other side of the z'-axis. Retaining the negative $(-)$ sign in these instances is important because they result in special cases.

Distinction of Two Cases

We now distinguish between two cases of the ACPC and highlight what happens in each. The difference between φ_L and φ_R plays an important role in the tilt of the entrance aperture, and thus α. As the difference between these two design angles $(\varphi_L - \varphi_R)$ increases, α increases, causing the entrance aperture to become more tilted. Eleven ACPC designs with varying design angles were used to determine this relationship. The results are given in Fig. 3.7. Based on the graph, as the difference between the two design angles increases, a point is reached where z' is at an angle α that surpasses our largest design angle ($\varphi_L = 60°$). When this happens, θ_L flips to the other side of z' such that both θ_L and θ_R lie to the left side of z', resulting in a flipped case. The diagrams from Fig.3.4 are an example of a design where both acceptance angles fall to the left of the $z'-$axis.
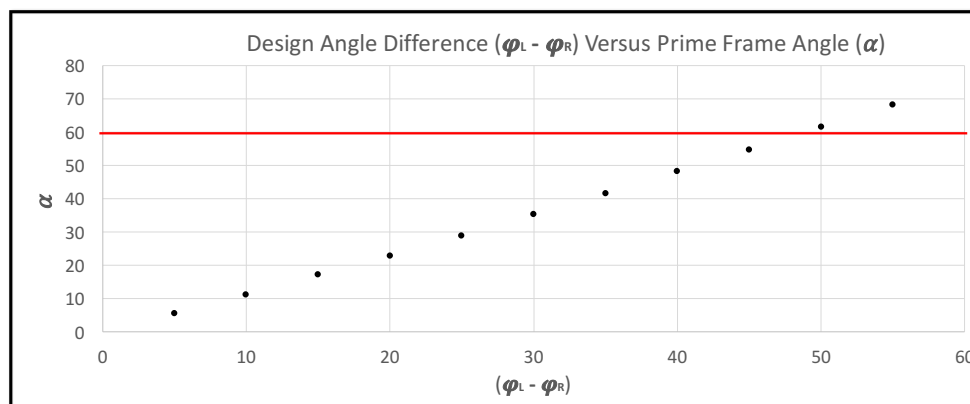
Figure 3.7: Depicted here is the relationship between the difference in design angles, and the tilt of the entrance aperture (primed frame rotation). The horizontal axis is the difference in design angle, and the vertical axis is the primed frame rotation angle. As the difference in φ_L and φ_R increases, the entrance aperture becomes more tilted, thus α increases. The red line represents the shift from one case to the next, when α is $\geq$ the larger of the design angles, in our case $\varphi_L = 60°$. After this point, the aperture is so tilted that the largest acceptance angle makes an acute angle with the aperture rather than an obtuse angle. In other words, it has flipped to the other side of z'.

3.2   Phase Space and Angular Acceptance Regarding Asymmetry

To derive the angular acceptance using direction cosines, we recall that the sum of the squares of the direction cosines is equal to 1,

$$L^2 + M^2 + N^2 = 1. \tag{3.5}$$
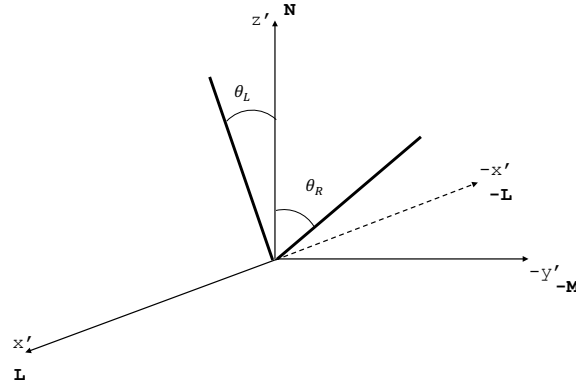
Figure 3.8: Three dimensional view of the acceptance angles for angular space derivation

From Fig. 3.8, we see the following relations for the two differing primed acceptance angles,

$$tan(\theta_L) \geq \frac{L}{N} \qquad \text{and} \qquad tan(\theta_R) \geq \frac{-L}{N}. \tag{3.6}$$

Squaring these we get,

$$tan^2(\theta_L) \geq \frac{L^2}{N^2} \qquad \text{and} \qquad tan^2(\theta_R) \geq \frac{(-L)^2}{N^2}. \tag{3.7}$$

It then follows that,

$$L^2 \leq tan^2(\theta_L)(1 - L^2 - M^2) \qquad \text{and} \qquad L^2 \leq tan^2(\theta_R)(1 - L^2 - M^2). \tag{3.8}$$

Here Eq. 3.5 can be utilized to obtain,

$$L^2 \left(\frac{1}{tan^2(\theta_L)} + 1 \right) + M^2 \leq 1 \qquad \text{and} \qquad L^2 \left(\frac{1}{tan^2(\theta_R)} + 1 \right) + M^2 \leq 1. \tag{3.9}$$

Trigonometric identities are then used, giving,

$$L^2 \left(\frac{cos^2(\theta_L)}{sin^2(\theta_L)} + \frac{sin^2(\theta_L)}{sin^2(\theta_L)} \right) + M^2 \leq 1 \qquad \text{and} \qquad L^2 \left(\frac{cos^2(\theta_R)}{sin^2(\theta_R)} + \frac{sin^2(\theta_R)}{sin^2(\theta_R)} \right) + M^2 \leq 1. \tag{3.10}$$

Finally, we arrive at,

$$L^2 \left(\frac{1}{sin^2(\theta_L)} \right) + M^2 \leq 1 \qquad \text{and} \qquad L^2 \left(\frac{1}{sin^2(\theta_R)} \right) + M^2 \leq 1, \tag{3.11}$$

which is analogous to Eq. 2.31 for a CPC, except with two differing acceptance angles. When θ_L and θ_R are both positive (falling to their respective right and left sides of z') we have case 1:

$$\frac{L^2}{sin^2(\theta_L)} + M^2 \leq 1 \quad \text{for} \quad L \geq 0, \tag{3.12}$$

for θ_L. Likewise, for θ_R,

$$\frac{L^2}{sin^2(\theta_R)} + M^2 \leq 1 \quad \text{for} \quad L \leq 0. \tag{3.13}$$

When θ_L and θ_R are negative and positive, respectively, meaning (in our case) that θ_L flips to the left side of the z'-axis resulting in case 2, we have:

$$\frac{L^2}{sin^2(\theta_L)} + M^2 \leq 1 \quad \text{for} \quad L \leq 0, \tag{3.14}$$

for θ_L. Likewise, for θ_R we have,

$$\frac{L^2}{sin^2(\theta_R)} + M^2 \leq 1 \quad \text{for} \quad L \leq 0. \tag{3.15}$$

We recognize these equations as those pertaining to ellipses. The angular (L,M) space diagrams are given below for both cases (Fig. 3.9). In the first case, the angular acceptance fills an asymmetric ellipse with semi-major axis of unity and semi-minor axis of $sin(\theta_L)$and $sin(\theta_R)$. The area of the asymmetric ellipse is,

$$area = \frac{\pi(sin(\theta_L) + sin(\theta_R))}{2}. \tag{3.16}$$

For second case resulting in the folded ellipse, it too has a semi-major axis of unity and semi-minor axis of $sin(\theta_L)$and $sin(\theta_R)$. The area is also the same as that above, but note that now θ_L is negative, resulting in an overall negative value for $sin(\theta_L)$.
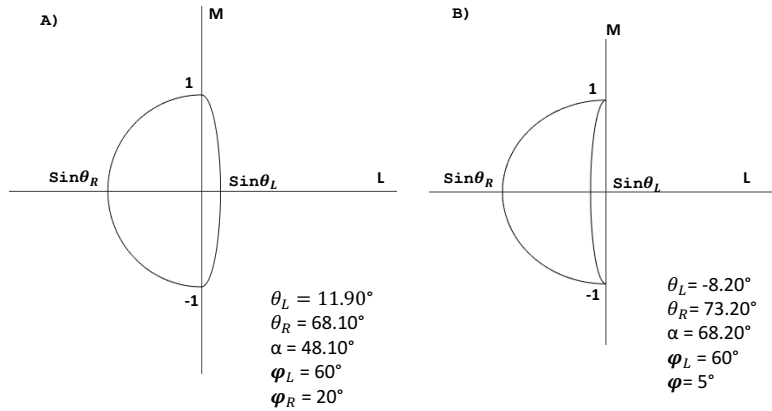
Figure 3.9: Angular acceptance for cases 1 and 2. A) asymmetrical ellipse and B) folded ellipse

It should be mentioned that both cases depicted occur when $\varphi_L \geq \varphi_R$. If the angles are switched, i.e. $\varphi_R \geq \varphi_L$, then case 2 will change such that both θ_L, θ_R will instead fall to the right side of the primed frame axis, rather than the left.

Phase Space Area

As with the angular acceptance ellipses, the difference in acceptance angles means that the phase space area for the entrance aperture of the ACPC will no longer be symmetric about $x-$axis. For the first case, when both θ_L and θ_R fall to their respective locations (to the right and left of the $z'-$ axis), we see the phase space area as antisymmetric about the $x-$ axis, with a larger portion of its area falling below the $x-$axis. However, in the second case, we see that the entire phase space area falls below the $x-$axis, again, showing the antisymmetry about the $x-$axis.

A)

$\theta_L = 11.90°$
$\theta_R = 68.10°$
$\alpha = 48.10°$
$\varphi_L = 60°$
$\varphi_R = 20°$

B)

$\theta_L = -8.20°$
$\theta_R = 73.20°$
$\alpha = 68.20°$
$\varphi_L = 60°$
$\varphi_R = 5°$
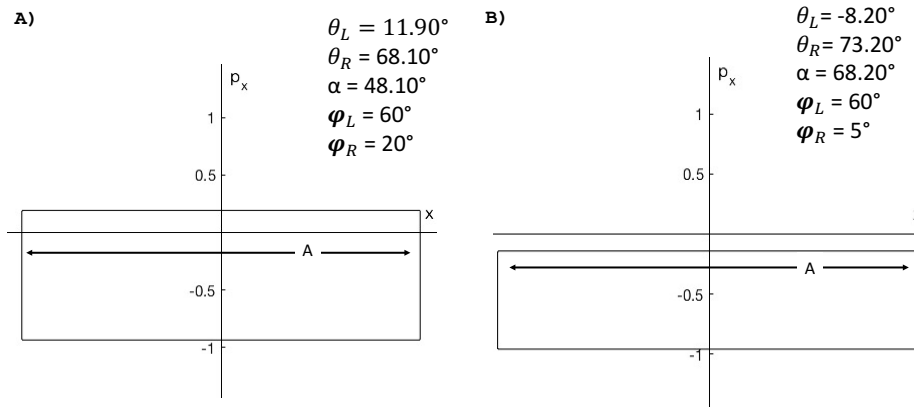
Figure 3.10: Phase space for cases 1 and 2. A) phase space area is asymmetric about $x-$axis and B) phase space area falls entirely below the $x-$axis making it again, antisymmetric about the $x-$axis. It should be noted that these are both for the entrance aperture, A. The exit aperture's phase space diagram will look like that of Fig. 2.11 B), where the p_x- axis is unity and the $x-$axis length is equivalent to A', the length of the exit aperture.

3.3 Phase Space and Angular Acceptance Ray Trace

The purpose for utilizing ray trace is two-fold. Firstly, ray trace offers a different and perhaps more illustrative way to demonstrate the results obtained in Fig.3.9 and Fig.3.10. Second, the intensity distribution [W/Sr] mapped to both phase space and angular space goes beyond the boundaries of what is obtained mathematically. Understanding what these distributions look like is not necessarily intuitive, especially for the angular space regime.

Angular Acceptance

ACPC models are built in Light Tools corresponding to the parameters defined in Fig. 3.9 and Fig. 3.10. Rays are traced in the reverse direction from a lambertian source at A' onto a hemispherical, far-field receiver. A far-field receiver measures the number of rays over solid angle as $dn/d\Omega$. In terms of direction cosines, $d\Omega = dLdM/N$. To plot in LM space, the data should be number of rays (Watts) per $dLdM$. Therefore, the data can be multiplied by $1/N$ to achieve this.

MATLAB is used to map the intensity mesh data [W/Sr] into angular space [W/dLdM]. The light exiting the entrance aperture is "limited lambertian" and thus, as expected, the boundaries of the angular space intensity map are identical to those boundaries obtained in Fig. 3.9. The distribution in angular space is uniform, as it should be for any ideal concentrator. The presence of non-uniformity is a strong indication that a more optimal design may be available.
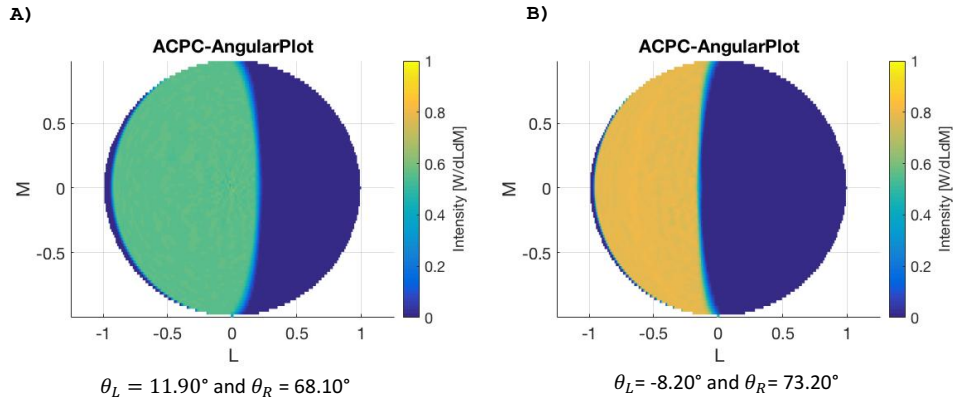
Figure 3.11: Case 1 and 2 angular acceptance for ACPC designs. A) depicts an asymmetric
ellipse and B) depicts a folded ellipse. The color bar to the right gives the W/dLdM for the
two designs.

Étendue Ray Trace

If we think of étendue as an incompressible fluid in phase space, the rays in a ray tracer
possess étendue information, i.e. each ray is a drop of étendue. Thus, for traces of ideal
concentrators, all rays or drops of étendue leaving our light source will reach our receiver
with an efficient distribution, maximizing power for the range of angles accepted.
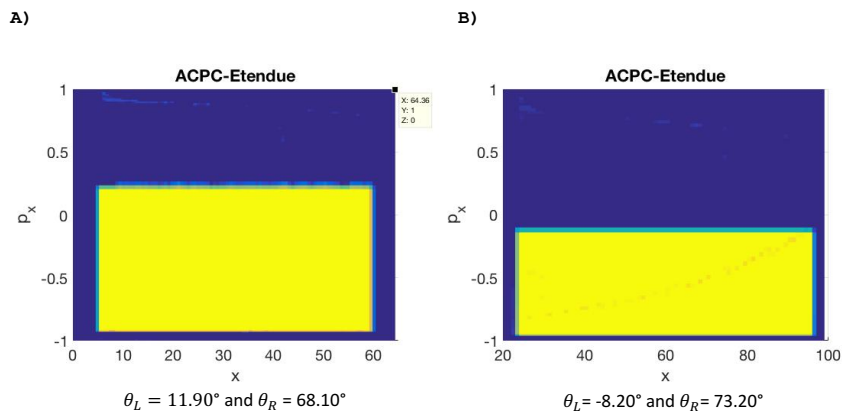
Figure 3.12: Phase space for cases 1 and 2.

The ACPC is set up in Light Tools such that a receiver sits at the output aperture (A')
and a planar light source at the entrance aperture (A). The étendue is mapped onto the
receiver using the parameter sensitivity function. Optimization variables for this function

are: the length along the entrance aperture and the range of acceptance angles. The merit function is then the power received at the receiver. The strip of light is then moved along the length of the aperture. For each movement, a range of accepted angles are traced. The phase space mesh data generated from Light Tools is given in Fig.3.12. The ACPC design parameters correspond to the parameters in Fig. 3.9 and Fig. 3.10. One thing to notice is that the ACPC takes the phase space generated by a CPC and transforms it, moving it either up or down along the p_x- axis. Because the distribution of rays in phase space is uniform for ideal concentrators, just as for the angular acceptance ray trace mapping, this phase space mapping technique can also be utilized to test concentrator performance and determine ideal vs. non-ideal for constructed designs.

3.4   Flow Line Generator

It was first shown in [35, 36] that any pair of two flow lines may be taken and together, form a new ideal concentrator. This exciting new realization leads to the ability to form new ideal nonimaging optics concentrators. However, this method means that the flow lines are limited by the design of the optic they reside within, meaning, one must begin with a concentrator to form a new concentrator. A truly impressive feat would entail using flow lines directly without beginning with an existing optic. In other words, if one could simply use a flow line method to construct new ideal concentrators without limitations from already existing designs, the boundaries of nonimaging optics might expand. This would be revolutionary for the illumination sector, which is constantly seeking new solutions to prescribed illumination problems.

Perhaps the key lies with the flow line generator. We have already seen one flow line generator, the flat lambertion source in Fig. 2.15, which generates confocal hyperbolas. Pairs of these hyperbolas can be selected to make new concentrators that are ideal. Symmetric pairs need not be chosen, any two will do. Several other flow line generators are mentioned in [37, 11, 1]. No flow line generator has yet been derived for the ACPC. It is now derived here to expand the library of flow line generators in hopes of furthering the understanding of how to construct and use them.

Understanding Flow Line Behavior in the Sections of an Ideal Concentrator

We start by building the intuition for the flow line generator. As seen in [1], given a truncated wedge with absorber length QQ', and sides R and R' prime extending to infinity, the flow line directions can be determined for every point within the regions noted in Fig. 3.13. The regions are determined by the dotted lines, which mark the long cross strings taken at the angle for which the concentrator is designed for. In this case, the truncated wedge with sides extending to infinity is for the CPC. The sides extend to infinity because the source(sun) is

at infinity. The wedge angle is $2\theta_{max}$, where θ_{max} is the design angle and in the case of the CPC, the acceptance angle.
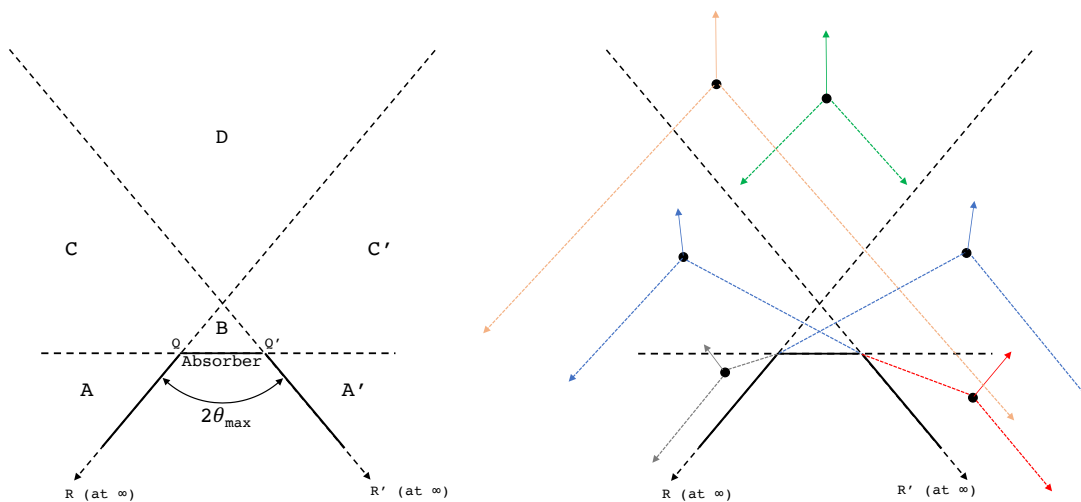
Figure 3.13: Flow line generator for the CPC. Each region sees a different part of the truncated wedge. For each point in each region a flow line vector can be constructed.

Focusing on a single line of flow, we obtain Fig. 3.14. Here one side of the parabola is obtained. There are a few arrows labeled P to indicate the flow line vectors. These are tangent to the parabolic surface, and always bisect the two lines which, in this case, connect from Q and R' at infinity.
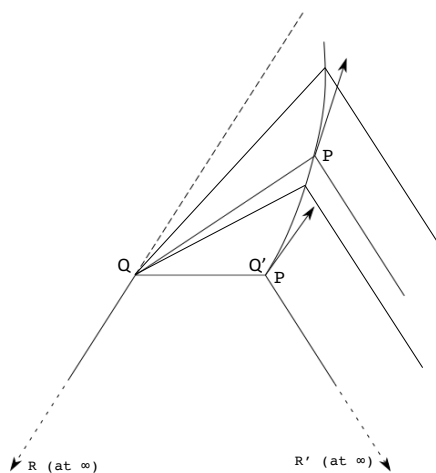
Figure 3.14: Flow lines for a CPC with a 45° acceptance angle.

The final result is a number of flow lines as seen in Fig. 3.15. These can be as dense or as sparse as one chooses.
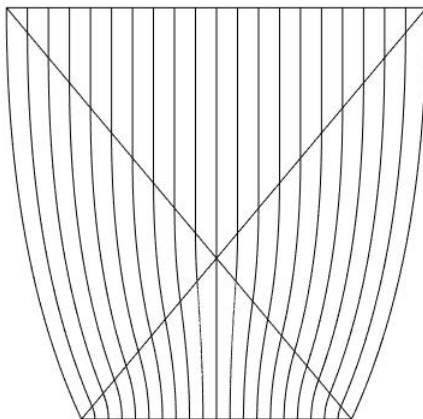
Figure 3.15: Flow lines for a CPC with a 45° acceptance angle.

Next, we apply this same principle to construct an asymmetric CPC generator. Because of the asymmetry, the design angles φ_L and φ_R are different, it would make sense that the diagonal cross strings that are constructed, which continue past the absorber to infinity, will also be asymmetric. Thus, we will have an asymmetric truncated wedge. Again we have different regions for which we can take lines and connect the visible portions of the truncated wedge to obtain the flow line direction. This is done in Fig. 3.16
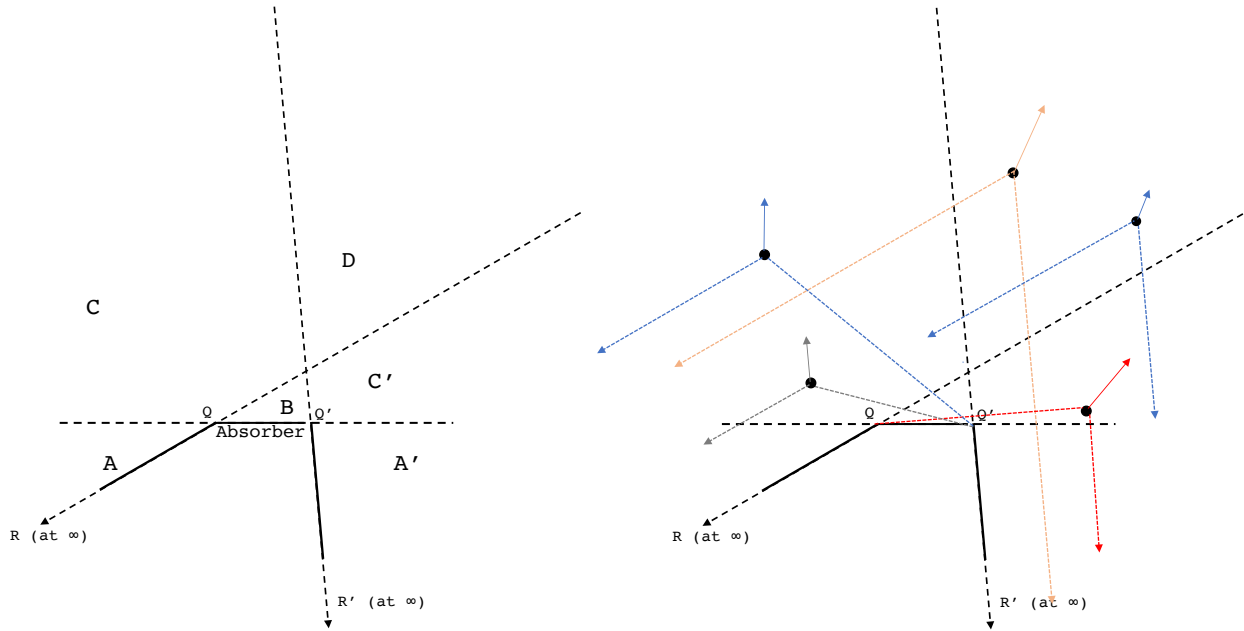
Figure 3.16: Flow lines for a CPC with a 45° acceptance angle.

Using the same procedure as above, we can determine as many flow lines as desired. For each point in each of the regions, flow line vectors can be determine. With the differing design angles, the walls of the truncated wedge merely change in angle. These walls, R and R' still extend out to infinity since the source remains at infinity. In the event that the source moves closer such that it is no longer infinite, the parabolic walls tend to ellipses. Figure 3.17 shows the flow lines for the asymmetric CPC.

Understanding how the asymmetric CPC generator differs from the symmetric case, and further, understanding how these two generators differ from those that have come before help to pave the road to determining new generators that will allow for new concentrator designs.
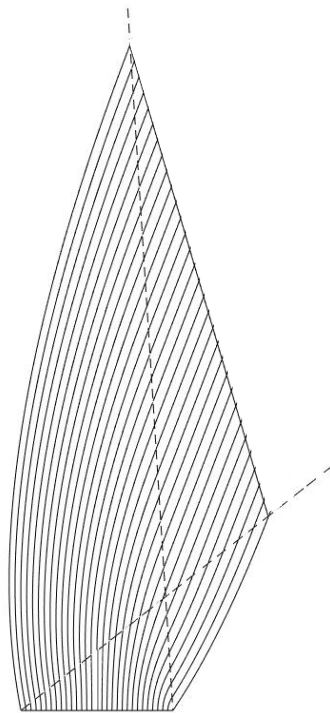
Figure 3.17: Flow lines are constructed for the ACPC. The lines of flow, **J** are smooth and unbroken. They originate as confocal hyperbolas with the focus points at the edges of the absorber, then after crossing the long cross strings (dotted lines), they become confocal parabolas, finally, upon crossing the cross strings again, they become straight segments.

3.5 Conclusion

The asymmetric CPC has been useful in specialized cases of solar thermal and solar concentration applications. For a time, the theoretical background to support the ACPC was sparse. The work done in this chapter has helped to round out the information and theory behind this variation of the CPC. Understanding that the design angles, φ_L, φ_R, for the ACPC differ from the acceptance angles, θ_L, θ_R, which are always taken with respect to the entrance aperture, is important when calculating concentration, étendue, and angular acceptance. Here, a tool has been provided for determining acceptance angles after a design is constructed based on the entrance aperture's orientation. Because the ACPC is asymmetric about the $z-$axis, the differing acceptance angles result in both asymmetric angular acceptance and asymmetric étendue. Derivations for both are carried out. For angular acceptance, two cases occur in angular space. These two cases depend on the difference in acceptance angles θ_L and θ_R. In the first case, an asymmetric ellipse is obtained, and in the second case, a folded ellipse is obtained. Ray tracing is performed to offer a different way of visualizing the results, while also demonstrating the uniformity of distributions for both the

angular acceptance as well as the étendue. The idea that an ideal concentrator will always give a uniform distribution may be intuitive, but beyond that, mapping intensity meshes from ray trace into both phase and angular space can be a powerful tool for analyzing new concentrator designs. Furthermore, determining the flow line generator for the ACPC will help our understanding in developing flow line generators to accommodate new designs that solve prescribed illumination problems.

Chapter 4

Night Sky Preservation and Controlled Illumination using Concentrators

National parks hold some of the last remaining Milky Way starscapes in the United States. With the emergence of excessive light pollution, even these escapes that harbor darkness are under threat. In fact, the Night Skies division of the National Park Service (NPS) states that, the experience of viewing the "natural lightscape" at night in the absence of human caused light is an endangered resource[38].

The night sky with its brilliant starscape and views of the Milky Way can be one of the most awe inspiring sights we will ever experience. Many people visit Yosemite National Park to enjoy this diminishing resource. Unfortunately, with the growing tourism in the Yosemite's valley and the need for safe lighting, the natural lightscape, like many other national park lightscapes, is now under threat.

The best way to combat this problem is to reduce light pollution by controlling stray illumination. As seen in Fig. 4.1, the need for such a measure is present. Strict guidelines exist regarding park lighting. Yet, they do not solve the illumination pollution issue. For example, these guidelines in Yosemite require that all lights must have hoods to shield them from emitting upward. This does not solve the problem of light spreading uncontrollably sideways. Mounting complaints from guests and park staff regarding the views from the valley at night are evidence that despite regulations in place, Yosemite Valley is still suffering from unwanted amounts of light pollution. What has happened is, a significant amount light illuminating more than the necessary areas (i.e. parking structures and pathways) now illuminates surrounding grasses, trees, and granite cliff faces. This effects the flora and fauna in direct proximity to these sources.

However, implementing change is more difficult than simply installing state of the art luminaires. A dominant rule in Yosemite park requires that new technology must remain within the "circa 1860s" guidelines. This ensures that the appearance of the park appears similar to the way it did when it was founded. Thus, it is difficult to implement new ideas.

Illumination pollution in Yosemite Valley

Sunset (partial darkness)   Night time (full darkness)
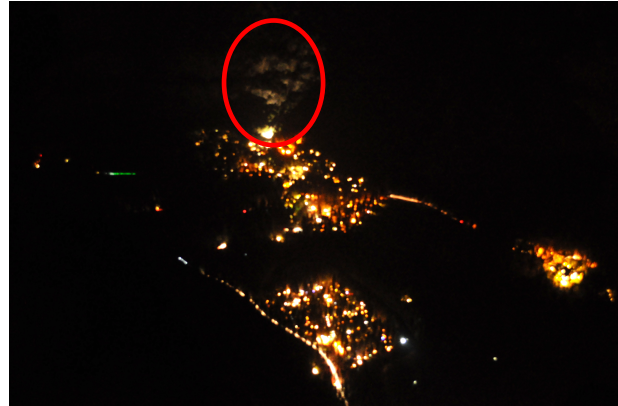
Figure 4.1: Yosemite Valley light pollution as the sun is setting (Left) and after the sun has
set and darkness has fallen (Right). The red circle is light from a single flood lamp in the
park that illuminates a snow plow area. It can be seen shining onto the adjacent cliff face
that is next to Yosemite Falls.

Fortunately, a work-around can be found by implementing ideas that have seemingly little
impact on appearance while offering great improvements otherwise.

In regards to overall lighting, while the Yosemite park has started to make the switch
to LED technology, a majority of its lighting is outdated. The solution offered in this work
would allow for the implementation of retrofitted reflectors that are inexpensive and easy to
fabricate during Yosemite's switch-over to LEDs. LED lamps with advanced optics can be
expensive and that cost is a deterrent for limited budjets. Having a reliable solution that
can be implemented on cheaper LED lighting boosts the ability to implement change.

Furthermore, this work offerers a unique opportunity to bring awareness to light pollution
in National Parks - specifically Yosemite - by improving a pollutant light in Yosemite Valley
which obstructed views of the Milky Way. The offending light along with its impact upon the
cliff face beside Yosemite Falls is circled in red in Fig. 4.1. A solution was determined to make
it possible to implement non-obtrusive changes to this light while eliminating its negative
effect on surrounding wildlife. The hope is that such a solution might be implemented on
other lighting around the park to diminish overall light pollution. Thus the candidate light
was selected to lead as an example. The solution developed was a quick, cost efficient, and
effective way to solve the presented problem using limited resources. In this chapter, we look
at the candidate light as an example of a way to use solar concentration for illumination to
achieve the goal of reducing light pollution.

4.1 Statement of the Problem

Due to the culture in Yosemite, the park is strict on introducing new technology. For lighting, implementation must be done in the least obtrusive manner possible. We were not permitted to install new light posts, nor change the location of existing posts and lights. We worked around these obstacles by retrofitting reflectors into the candidate light chosen for improvement.

For our purposes, we chose a particularly obtrusive light away from the park visitor sites. This flood lamp, is located in an equipment yard (see Fig. 4.2), and is currently used to illuminate the large sand/salt piles used by the snow plows. Anchored to a 25-foot pole, the light is mounted to point directly at the trees and surrounding cliff faces, not downward towards the ground (see Fig. 4.3). This caused a significant amount of stray light in unnecessary areas, including the iconic Yosemite Falls.

Figure 4.2: The equipment yard where the candidate light is located. The light pole is located on the right side of the image and is approximately 25 feet high.
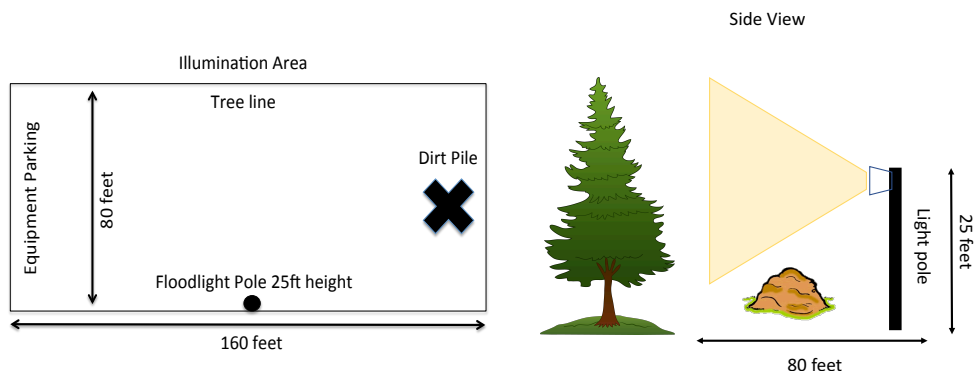
Figure 4.3: Sketch of the target area to be illuminated (Left). The right image shows the orientation of the light pole, and the direction the light travels from the flood lamp that is mounted on the pole.

Granite is highly reflective. As a result, the light from this one flood lamp was observed far from the source location. Using high definition resolution images (HDRI), long exposure photography illustrated the extent of the illumination on the nearby cliff face adjacent to Yosemite Falls (see Fig. 4.4). Furthermore, many in the park complained that the light was nuisance to park visitors coming to enjoy the natural darkness of Yosemite.

Figure 4.4: Illumination from the candidate light falling upon the cliff face adjacent to the iconic Yosemite Falls. HDRI photography was used to take this image with a thirty second exposure rate.

To ensure worker safety in this equipment area, any improvements needed to maintain adequate illuminance levels. The proper levels of illumination were retained using nonimaging optics, while eliminating the wasted light to the trees and cliff face. This was accomplished by guiding the illumination to the places where it is needed. The size of ground area requiring illumination was measured. This became the target area. It is rectangular in shape, and covers 160ft x 80ft. The 25ft pole is located to the side of this area as shown in Fig. 4.3.

Finally, before developing an optical solution, measurements of illuminance were taken over various distances in the target area. The most important location was the edge of the target area; this was the place where illumination needed to reach its lowest levels to avoid spilling onto surrounding areas. That became the purpose of this work: eliminate light at the edge of the target and beyond.

4.2 Design Solution

The optical design solution developed was based on the orientation of the LED flood lamp and target area to be illuminated. The devised solution did not require any changes to the

existing light orientation nor its mounting pole. Changes made were done by retrofitting two reflectors into the flood light housing.

Preliminary Designs

To achieve a desired design outcome, two tasks were undertaken. Firstly, the illumination emerging from the LED chips required an angular transformation to point the light downward at the target plane (the ground). Second, the emitted light required tighter focus to illuminate the target and not the surrounding flora and fauna. Transforming the light angularly can be done with a semicircular wedge or torus shape [2] [39]. Implementation of a semi-circular wedge does not disturb the light. Rather, the shape of the optics transforms the emission plane by some angle. In other words, phase space is transformed. Focusing the light for tight control to within an angle θ can be done with a compound parabolic concentrator. The two optical solutions were combined to generate a final design (see Fig. 4.5) [40]. It should be noted that because of the angular rotation, the light on the target will not be head on, and a predicted cosine effect will take place.

Matlab was utilized to generate the two dimensional curves extruded into a third dimension, which were then imported into ray trace software (LightTools). Iterations of this procedure were carried out based on simulation results during the design process. The ideal design angle of both the wedge and the CPC were determined. The final model possessed a 63° semi-circle wedge, which rotates the optic by 27° from the downward direction, and a 45° angular acceptance on the CPC. These two pieces were then combined. The size of the LED chips were each 30 mm in diameter. The final geometric concentration of the optic was 1.41, where Eq. 2.13 was used to determine this.
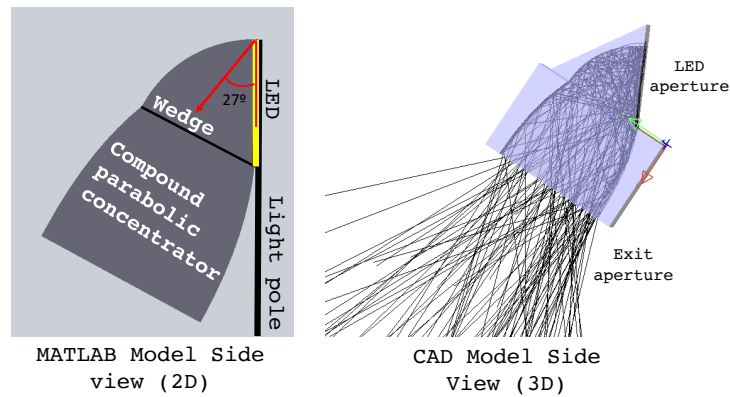
Figure 4.5: An optical solution that consists of a semicircular wedge subtending a 63°, effectively rotating the optic 27° from pointing downward, and a compound parabolic concentrator (CPC) with a 45° acceptance angle. The image on the right shows a version with side depth as a translucent view in LightTools as rays are traced. The side depth was given for ease of 3D printing for the initial prototyping.

flowlines

To confirm that the wedge implementation with the CPC did not disrupt the flow of light, the design was evaluated with flowlines. The flowlines travel from the lambertian light source through the semicircular wedge to the CPC. It is clear that the lines undergo smooth transitions from the wedge to the CPC, and thus, there is no disruption to the light. This demonstrates that the design is not losing any light.
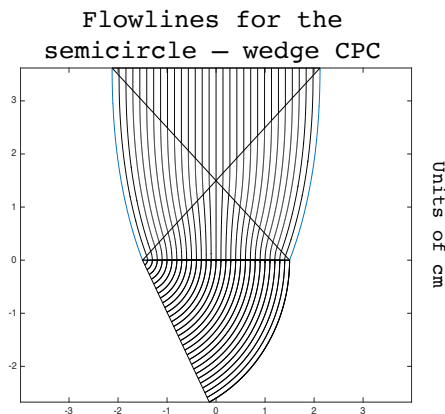
Figure 4.6: Flowlines traveling through the design.

As mentioned previously, any pair of flowlines can be selected and used to make new

ideal concentrators. The same can be done here, and the new concentrator is tested and
confirmed to be ideal (power is maximized over the target).
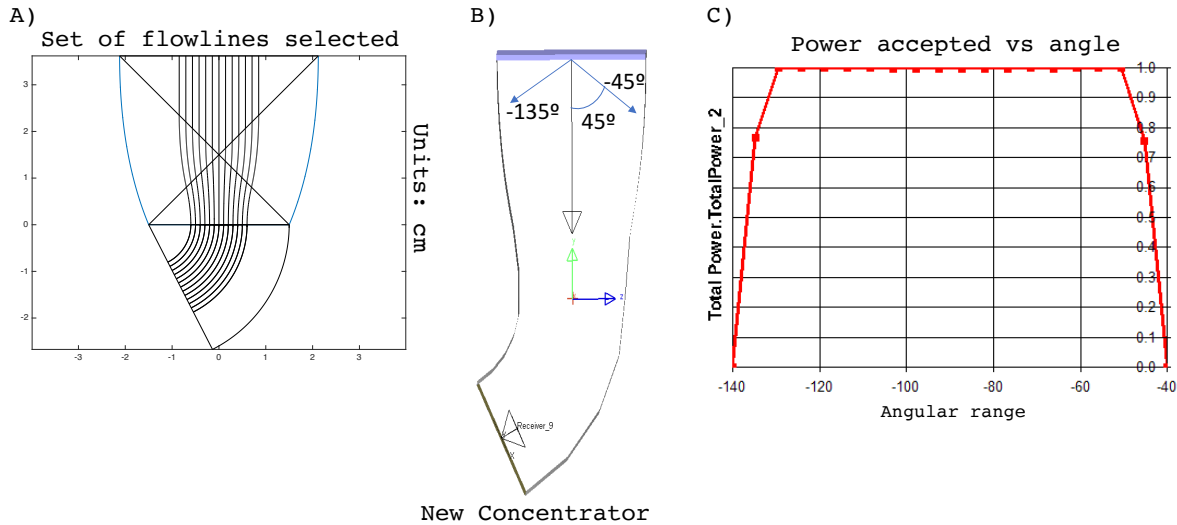
Figure 4.7: Flowlines traveling through the design.

As seen in the power accepted vs. angle image in Fig. 4.7, all the power from the LED
is incident on the blue receiver at the aperture of the concentrator over the range of angles
accepted by this new concentrator design.

4.3 Optical Analysis

This optical analysis confirms two things. First, it confirms that the fabricated model behaves
as expected based on ray trace results. It is a way to ensure that adequate fabrication
has been performed. And second, the analysis confirms that the illumination distribution
improves by implementing optics.

Optical performance is first evaluated and optimized in ray trace software. Then a model
is produced and similar data is collected in a controlled environment. For both the ray
trace and experimental analysis, the illuminance distribution of light on a target plane is
evaluated. In both instances, the illumination distribution is recorded by breaking up the
target into "bins" or bits of area. The size of these bins depends on both the number of bins
and the area of the target plane. Increasing the number of bins over a set target area leads
to greater accuracy of illuminance measurement.

The illuminance is a measure of lx, or $lumens/m^2$. These values give an m x n array of
"mesh data". Exact values in lx for each bin can then be compared. Experimentally, these
readings are taken with a lux meter.

A lux meter is a device that measures the amount of illuminance (lx or $lumens/m^2$))
falling onto a surface at any given location on that surface. Because it is a photometric device,
it takes into account the sensitivity of the human eye. Most meters including the one used in
this work consist of a photodetector and a digital readout display. The material used in the
photodetector, which is often selenium or silicon, determine illuminance photovoltaically: a
current is generated that is proportional to the photons received. The lux meter used in this
work employs a silicon based photodetector, which needs to amplify the voltage generated by
the light exposure to take an accurate reading. Shown in Fig. 4.8, this lux meter is designed
to take readings in the range 0.1 to 19,900 lx.

Figure 4.8: Lux meter used in this work. Meter is made by Konica Minolta Sensing Inc. It
is the TL-1 model.

Initial prototype testing

For all evaluations, the following parameters were used: The optical design was placed at a
height of 700 mm (2.3 ft) over a target area of 600 mm by 600 mm (2 ft by 2 ft), which was
broken up into equally sized bins. This measurement was done with 8 bins by 8 bins for a
total of 64 bins, with a million rays traced. Such a large number of rays ensures minimal
error per bin. For a million rays, the error per bin is the square root of the number of rays
arriving to each bin ($error = \sqrt{n}$) averaged over all 64 bins. The averaged error for the ray
trace measurement was then 1.94%.

In this work, a greater number could have been used for better target resolution, but it
was prudent to choose a number that would be reasonable for taking manual experimental
measurements. The illuminance mesh data was obtained as a mesh of lx values. For the
experimental test, three sets of readings were taken and averaged.

Data for each analysis was imported into MATLAB and color maps were generated for
comparisons as shown in Fig. 4.9. The first set compares the performance of the prototype
model's distribution with the goal distribution. This is used to determine how well the model
adheres to the predicted behavior.
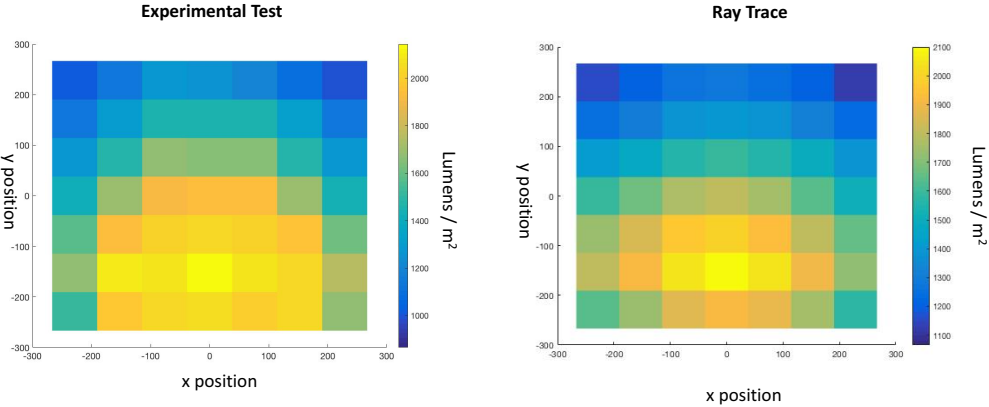
Figure 4.9: Color charts representing the mesh grid. This compares the goal distribution on
the right, with the achieved model distribution on the left.

The concept of uniformity mentioned in Chapter 1 is used here. The modeled distribution
is compared to the goal distribution by way of subtracting one mesh from the other. It is
also represented mathematically by Eq.4.1. The mesh grid should be uniform for a model
that produces a distribution identical to the goal distribution . As shown in Fig. 4.10, the
uniformity is not perfect. Highest values approach 300 lx. The test light used has an output
of approximately 3,200 lumens. And values of lux on the distributions shown in Fig. 4.9
approach 2,100 lx.
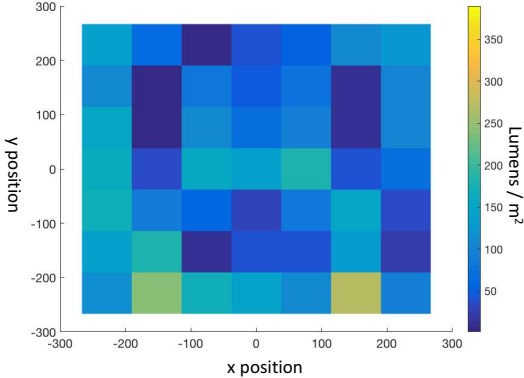
Figure 4.10: Data results for uniformity. A uniformly colored mesh would indicate perfect
match where the data obtained with the fabricated model match identically to the optical
results obtained in ray trace. This would give a uniformity value of $\sigma = 0$.

To determine σ, we use Eq. 4.1.

$$\sigma^2 = \frac{1}{mn-1} \sum_{i=1}^{n} \sum_{j=1}^{m} [f_{model}(i,j) - f_{goal}(i,j)]^2 \tag{4.1}$$

Taking the square root to find standard deviation, a value of $\sigma = 114.6$ lux is obtained. Perfect uniformity would result in a value of 0.

Another way to compare the prototyped model to the goal can be done by using slices of data in either the x or y directions along the mesh grid. Because the mesh grid data is given in x and y with illuminance values at each location, taking a slice of data (either from x or y) and plotting distance versus illuminance can be accomplished. The goal data can be plotted with the model data as shown in Fig. 4.11.
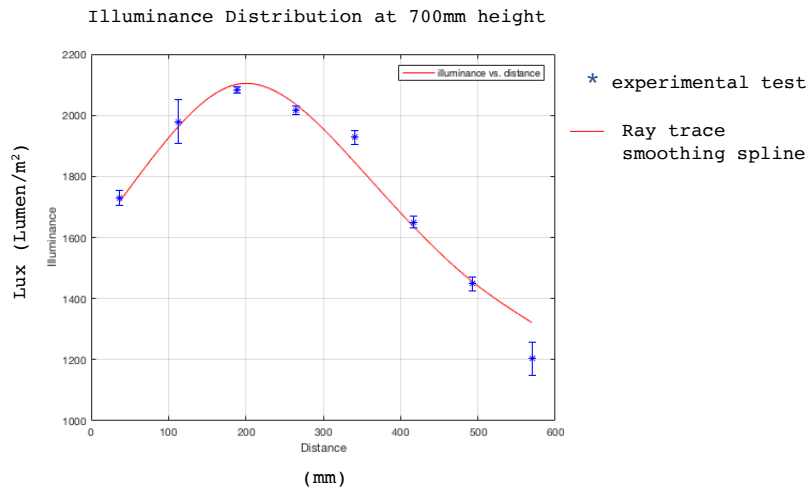
Figure 4.11: Data results for the theoretical and experimental tests done for the optical design with a fixed x value and illuminance values taken along y. The red line corresponds to the goal data generated in ray trace software. The blue stars correspond to the vertical slice of values obtained experimentally.

For the goal data obtained in ray trace, points obtained were used with a curve fitting tool that created a smoothing spline (shown in red in Fig. 4.11). The smoothing spline s is constructed for the specified smoothing parameter p and weights w_i which in this case were set to $w_i = 1$. The value for p ranges from 0 to 1. In this case, MATLAB automatically selected p to give the smoothest curve. The equation for the smoothing spline is,

$$p \sum_{i} w_i (y_i - s(x_i))^2 + (1-p) \int \left(\frac{d^2 s}{dx^2}\right)^2 dx \tag{4.2}$$

where x and y are the values used in this case, for distance, and illuminance on the plot. To evaluate how close the smoothing spline line (shown in red) is to the data used to create

it, a value of $R - Square = 0.997$ was obtained. This value is very close to 1, meaning that the generated smoothing spline fits the individual data points. The points in blue represent the experimental data obtained with the prototyped model.

The error bars on the experimental data are calculated using the RMS values for the measurements taken. Aside from statistical error given by the error bars, there was also systematic error. The systematic error was due to the measurement instrument, and the LED employed. The lux meter possesses two types of error which is not shown in the graph. First, the lux meter has a cosine error of $\pm 3\%$ at 30° and $\pm 10\%$ at 60°. This is because light entering the photodetector at an angle can refract, causing measurement error. Because the light is not facing the target directly, the bins that are further from the light are at a larger angle from it, and suffer a higher cosine effect. These bins will have a higher error associated with them. The second error associated with the lux meter is related to the spectral response. The lux meter as an error of $\pm 4\%$ from the CIE photopic standard observer curve. This means that for different color temperatures, and even different types of lights, the correction may be slightly off, and thus arises the spectral response error.

Next, a color chart comparison is done to demonstrate the improvement achieved by implementing optics to the LED light. The purpose of implementing optics is to create a sharper cutoff between the target and surrounding area. The size of the target area remained the same (600mm by 600 mm). Color charts are provided in Fig. 4.12.
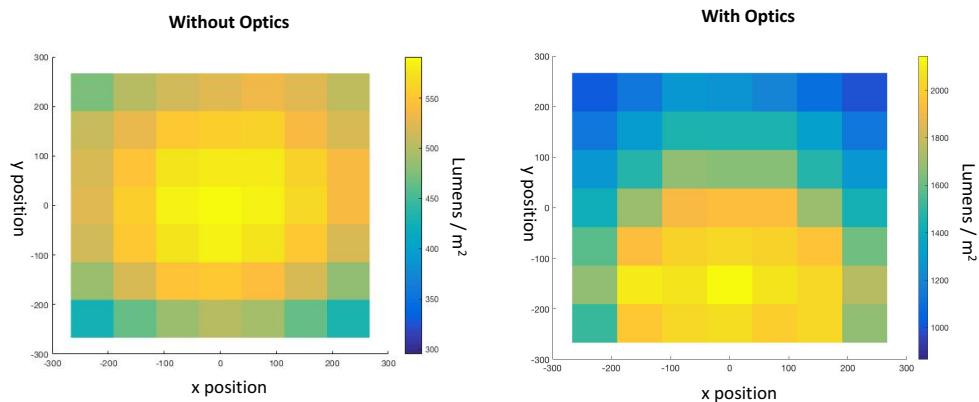
Figure 4.12: Mesh grids showing the results for the bare LED before and after optics are implemented. Note the differing color scales on each.

For the LED without optics, we see that the illumination is fairly evenly distributed with no real fall off over the two foot range. Also, the lux values are fairly uniform. For the color map showing the light with optics, there is a sharper fall off nearer to the edges of the target. Also, higher lux values are achieved because the light is more focused.

Final Prototype Testing

After preliminary testing, the plastic 3D printed prototypes were implemented into the flood lamp to be used in Yosemite. The specs of the flood lamp are shown in Fig. 4.13. Due to the heat of the lamp, these began melting too quickly before any preliminary tests could be conducted in the actual light to be used. However, since preliminary tests from the smaller version light on a smaller scale were promising, the reflectors were fabricated using thermally sound materials.

LED SYSTEM

	160W System	140W System
Calculated L_{70} (TM-21)	74,000 hours	74,000 hours
Delivered Lumens	15,530 lm	13,760 lm
Total Input Watts	160.0 W	137.4 W
Luminaire Efficacy Rating (LER)	97 lm/W	100 lm/W
Correlated Color Temperature (CCT)	4000 K	4000 K
Color Rendering Index (CRI)	> 80	> 80
Max Ambient Temp	110˚ F	110˚ F
Universal Driver	120-277 V	120-277 V

LED System data above based on FDL-160WLED-UNIV-4000K & FDL-140WLED-UNIV-4000K
(1) LED Lumen Maintenance Estimates based on TM-21 projections for the light source at 25°C ambient
(2) Specific Configurations Listed on DLC.
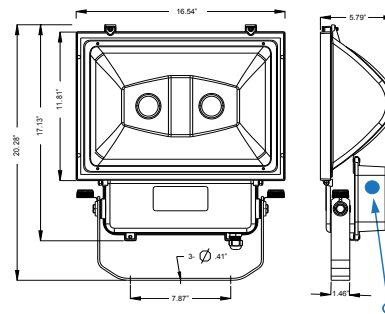
ILP Large Dual Flood
160W/140W LED

Figure 4.13: The left shows the exact specifications of the flood lamp used for the optical design retrofit. The right image shows its physical measurements and design, including the location of the round LED chip arrays.

Figure 4.14: Plastic 3D fabricated models implemented in the flood lamp before they melted.

The final optic created is made from Anolux MIRO-silver sheet metal, the pieces of which were bent and shaped, and then JB welded together. Both the sheet metal and JB weld were thermally sound and easily withstood the $80°C$ inside of the flood lamp. The aluminum-silver sheets have a reflectivity of 98%, making them the highest reflective material available on the market.
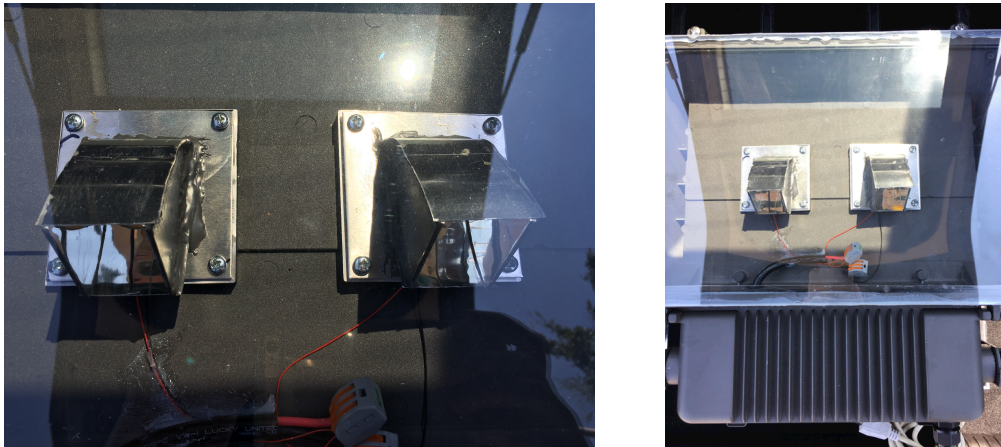
Figure 4.15: Anolux MIRO-silver prototypes permanently attached to the flood lamp.

After the prototypes were fabricated and attached to the flood lamp, the flood lamp
was replaced. Using a lux meter, readings were taken directly outward from the pole to
the edge of the target as a vertical slice from x = 0 along the y range. Due to the large
amount of equipment scattered around the target area, locations for taking readings were
extremely limited. This direction for the slice of data was the only unencumbered path
available for readings. Readings were taken in 5 feet increments then averaged. The same
slice of readings were obtained in the simulation software for generating a theoretical curve.
The experimental values were plotted around these in the same way as done above for the
initial optical testing.

The experimental data very closely matches the curve. The illuminance values near the
edge of the target plane closest to the light do not peak as highly with the experimental
data as the theoretical data. This could be for a number of reasons, but mainly because the
housing around the light was difficult to model in the simulation software and did restrict
some of the light that fell directly adjacent to the pole.
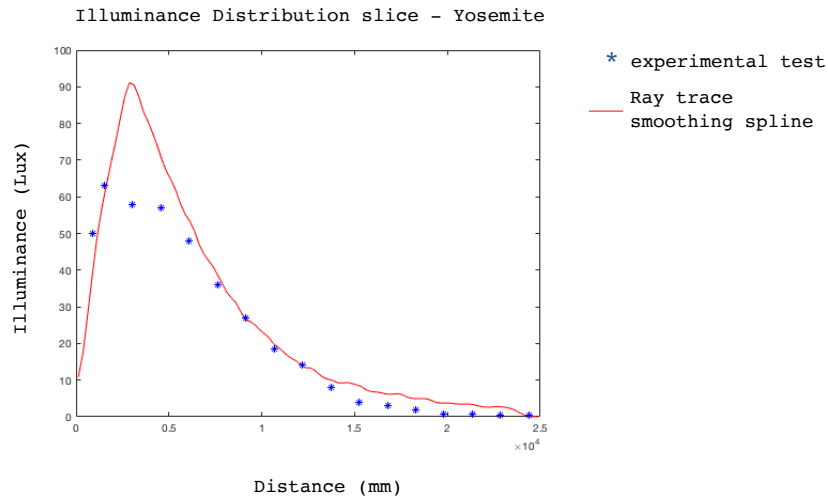
Figure 4.16: The curve in red shows the plot of the theoretical data taken from a slice of target data in Light Tools. The same line of data was collected experimentally in 5 feet increments using a Lux meter. This data was plotted and correlated very closely to the theoretical line.

The main purpose of implementing reflectors on the LED was to create a sharp cut off of light at the target's edge. Lux measurements were taken before and after the implementation of the retrofitted reflectors. At the target's edge (80 feet) the original lux value was 6 lx. The final lux value after the reflectors were installed was 0.5 lx. This is a great improvement.

Finally, an after image was taken to show the difference in light control before and after optics were implemented. These are shown side by side for comparison.

Without Optics   **With Optics**
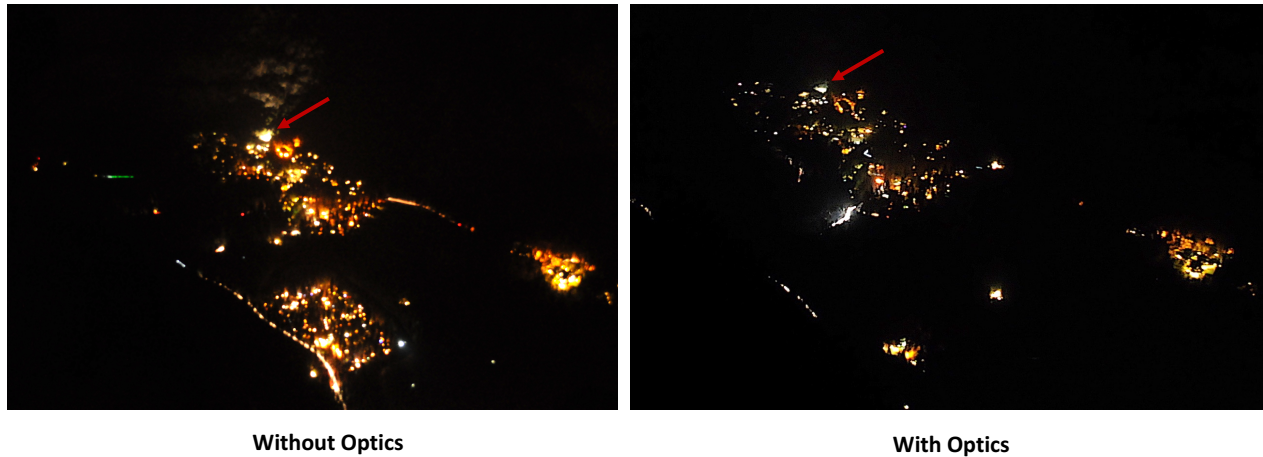
Figure 4.17: Images of Yosemite Valley from Glacier Point showing night time illumination. The flood light can be seen in both, with a great deal of illumination control shown in the after image.

4.4 Conclusion

Light pollution is a growing problem for communities everywhere. Thus, it is important to bring about awareness by implementing solutions. A great place to do this is in the iconic Yosemite National Park. The purpose of this project was to deliver a quick, cost efficient, and effective optical solution for reducing light pollution of a designated candidate light in Yosemite's valley. The fix took less thanfour months from the conception of the idea, to its development and implementation and testing in Yosemite Vally. Resources on hand were utilized, such that there was no need to outsource labor, materials, or designs. The results were positive, and a drastic reduction of light pollution was seen with the new retrofitted reflectors into the candidate light. Also, optical analysis shoed that the optics fabricated behaved very similarly to the theoretical scenarios.

From the outset of this project, a degree of craftiness was necessary. There were restrictions in place for the types of solutions that could be implemented. For example, keeping the original light pole and its set-up was necessary. Another hinderance was that the test area was filled with various machinery and equipment, making it difficult to take illuminance measurements on the ground during all stages of the project. These obstacles required a work-around. Preliminary optical analysis tests were conducted in a more controlled environment before implementing the final prototypes. Measurements on these prototypes were done in a way that could be compared to theoretical values, by taking lux readings of only a vertical slice of the target plane directly outward from the light to the target's edge. Readings showed improvements. And visual tests done at Glacier point confirmed that the illuminated cliff face shown in Fig. 4.4 was no longer illuminated. The light is currently being

used and will be used for the duration of its lifetime, some years to come. This work has inspired the idea that future UC Merced capstone teams will pick a new light in Yosemite each semester and make improvements. Ideally, illumination pollution in Yosemite's valley will be improved one light at a time.

Chapter 5

The Jellyfish - Aplanatic Total Internal Reflection Optic

An aplanat is a reflecting surface free from spherical aberration and linear coma [41]. It was Ernst Abbe who coined the term, however, the first known aplanat was invented by CA Steinheil [42]. Since their invention, aplanats have most commonly been employed as imaging devices, having been developed for use in telescopes [43, 44, 45]. However, more recently, they have found a place in solar concentration as dual mirror systems capable of approaching the thermodynamic limit to flux concentration at high collection efficiencies [46, 47, 48, 49]. This makes them appealing for concentrating photovoltaics (CPV).

As is the general case, good concentrators make good illuminators. Aplanats are not currently used in the illumination industry. The work in this chapter evaluates the feasibility of using an aplanat for illumination.

To carry out this research, a code was written in MATLAB to construct two dimensional aplanatic curves. The code also optimizes these curves, which is important for the design phase. Using this code, optimized curves were generated and then used to build a three dimensional optic. Once constructed, the aplanatic optic was tested as an illuminator using ray trace software. When results were not optimal, the optimization code was revisited. After obtaining reasonable results, the optic was prototyped and tested.

5.1   Background

The Jellyfish optic is a novel aplanatic lens that utilizes total internal reflection (TIR) for a number of applications beyond imaging[50]. In fact, it has been designed to work in both forward and reverse, and thus, is a good fit for both in both the fields of concentration as well as illumination [51]. In general, an aplanat often consists of two mirrored surfaces. This can be a problem when it comes to collection or illumination of light, because the front surface blocks a portion of incident/reflected light. As seen in Fig. 5.1, rays 1 and 2 cross the Abbe sphere and reflect off the back surface, but ray 0 is blocked by the front surface. To work

around this, the front mirrored surface can be made into a "one-way" mirror by filling the
area between the two surfaces with a dielectric medium of a larger refractive index than
air. The device will then admit light(concentration) or release light(illumination). The light
inside the device will be trapped. This leads to TIR.
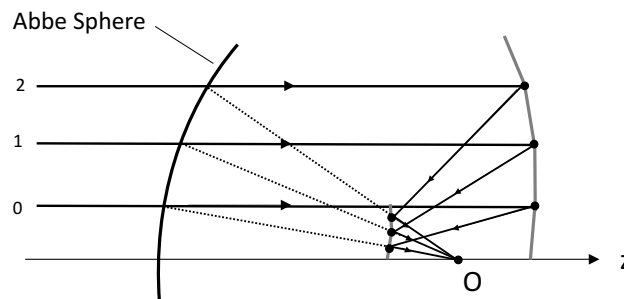
Figure 5.1: Depiction of the Luneburg method for constructing a two-mirrored aplanat.
The two constructed curves are the those with the points. Rays 0, 1, and 2 enter across
the sphere and reflect off the back surface before arriving to the front surface, where they
once more reflect before striking the point O. If they are traced back from the front surface
(dotted lines) then their backwards traces also intersect the Abbe Sphere. Notice that ray
0 is blocked by the front surface. This can be corrected by changing the front surface to a
"one-way" mirror.

TIR occurs within the device when the ray escape angle, also known as the critical
angle, θ_C, is smaller than the angle at which the trapped rays strike the top surface (see
Fig. 5.2). More impressive still, these escaping rays emerge nearly parallel, to within 2° (see
Fig. 5.3. For illumination, this 2° makes the device a good collimator; for concentration, a
good concentrating photovoltaic (CPV) with tolerant tracking. As such, a small portion of
reflective coating is required at the center of the top surface to block the middle-most rays
from emerging at larger angles than 2° (see Fig. 5.4). Rays emerging from a light source in
this area will strike the small reflective coating in the middle of the top surface, and then
be reflected down to the back surface, where they will begin their TIR process that will
eventually lead them out of the device at the correct angle.
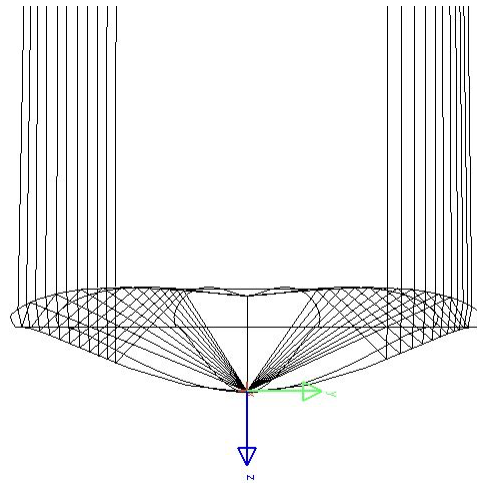
Figure 5.2: Rays undergo TIR. The rays that would strike the center coating will bounce
back and forth between the reflective back surface and the mirrored center. Eventually many
will be absorbed into the LED. The percentage absorbed depends on the efficiency of the
Jellyfish. The rays striking the center coating have been removed for the purpose of this
picture to show the TIR. (Image from Light Tools)
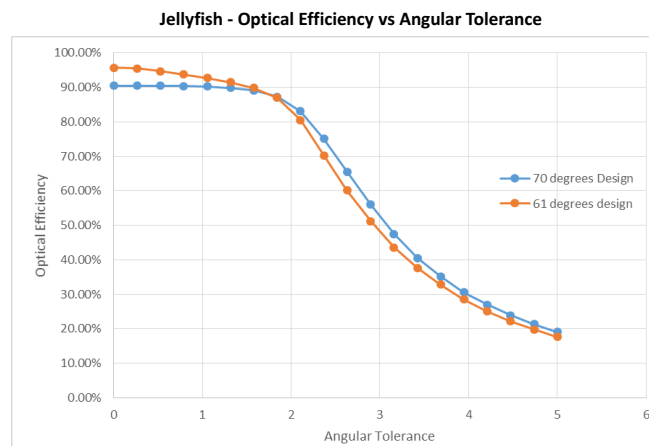
Figure 5.3: The angular tolerance for two different Jellyfish designs. Both data sets were
obtained in Light Tools using parameter sensitivity. As seen, both designs have very high
efficiencies to within two degrees of incoming or outgoing light.)
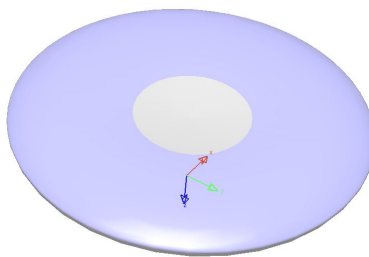
Figure 5.4: Front view of the Jellyfish. The silver circle in the center is the reflective coating.
(Image from Light Tools)

For illumination, a portion of rays will never exit the device; for concentration, a portion
of the rays will never enter the device. This is due to the reflective coating on the front
surface. And this ratio of reflective coating area to the total area of the top surface will give
the overall theoretical optical efficiency η of the device:

$$\eta = 1 - \frac{A_c}{A_a} = 1 - \frac{\eta_{trap}}{P_{tot}}, \tag{5.1}$$

where A_c is the area of the reflective coating on the top surface, A_a is the total area
of the top surface, η_{trap} is the amount of light (in Watts) trapped within the device, and
P_{tot} is the total amount of light emitted (in Watts) from the illumination source. The first
part of the equation pertains best to the concentration problem. Having a portion of the
incoming radiation blocked by the small reflective coating in the middle means that part of
the phase space is lost. Likewise, for the second portion, the illumination representation of
the efficiency, an equal portion of the light must be blocked from escaping. It is easy to see
that in the concentration problem, the loss is simply a rejection of rays. In the illumination
problem, it is not so easy: where do the rejected rays go? One can only assume that they
are absorbed by the LED as heat.

This chapter will look at the physics behind the Jellyfish design, the methods used to
design the optic including optimization processes for fabrication, and finally the results of a
device that was fabricated using selected optimization parameters.

5.2   Physics of the Jellyfish

It should first be mentioned that the Jellyfish satisfies the "Abbe Sine Condition", which
helps to explain why it suffers from no spherical aberrations or coma. As such, this ensures
good off-axis performance. This is done by requiring an on axis object of infinity as well
as fulfillment of the aplanatic condition, thus, the reason why there are two surfaces. The
Luneburg method for designing a two surface aplanat is iterative [52], and will be explained
in detail in the next section. This section is devoted to explaining the "Abbe Sine Condition"
and reasons for imaging free of spherical aberrations/coma.

The Abbe Sine Condition

The Abbe Sine Condition relies on stigmatic imaging, in which a single point source in the object plane is focused to a single point in the image plane [53]. For the purpose of the Jellyfish (concentration), we wish to take incoming light from an image at infinity(such as the sun), and focus it to a single point inside the optic (image plane). Thus, a special condition requiring an infinite image will be employed here. But first, let us deal with the general case.

We begin by setting up our system. O_0 is an axial object point, and P_0 is any point in its neighborhood (not necessarily on the axis). We must assume that this system images these two points stigmatically. Let O_1 and P_1 be the stigmatic images. Then we have (x_0, y_0, z_0) and (x_1, y_1, z_1) as the coordinates of points P_0, P_1, respectively. The $z-$direction is taken along the axis of the system. This system lies in a meridional plane $(y - z)$ and x is out of the page.
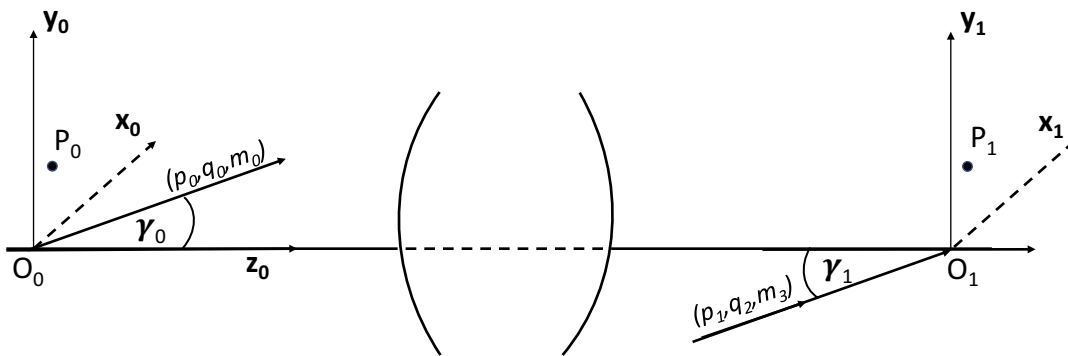
Figure 5.5: Illustrating the sine condition

The principle of equal optical path length (OPL) applies here, thus, the OPLs of all the rays joining P_0 and P_1 are the same. And so, we can construct what is called the point characteristic formula V for the medium. The characteristic function is defined as the OPL $[P_0, P_1]$ of the ray between the two points, considered as a function of their coordinates. This gives,

$$V(x_0, y_0, z_0; x_1, y_1, z_1) = \int_{P_0}^{P_1} nds, \tag{5.2}$$

where n is the index of refraction as a function of the path, s. This leads to,

$$V(x_0, y_0, z_0; x_1, y_1, z_1) - V(0, 0, 0; 0, 0, 0) = F(x_0, y_0, z_0; x_1, y_1, z_1), \tag{5.3}$$

where $V(0, 0, 0; 0, 0, 0)$ represents the starting position for rays from O_0 to O_1 which both lie along the axis at their respective origins. F is a function representative of time difference

between two starting points of a wave front, such as $\overline{O_0 P_0}$. The wavefront itself is the same at O_0 and P_0 because they are in the same neighborhood. Therefore, it should be noted that F does not depend on ray components; in other words, the function is independent of the direction the rays travel along their path to their corresponding point in the image plane. Rays come from O_0 and arrive at O_1 with ray components (p_0, q_0, m_0) and (p_1, q_1, m_1), which are the direction cosines [53]:

$$\frac{\partial V}{\partial x_0} = p_0; \frac{\partial V}{\partial y_0} = q_0; \frac{\partial V}{\partial z_0} = m_0; \frac{\partial V}{\partial x_1} = p_1; \frac{\partial V}{\partial y_1} = q_1; \frac{\partial V}{\partial z_1} = m_1. \tag{5.4}$$

And,

$$p_0 = ncos(\alpha), q_0 = ncos(\beta), m_0 = ncos(\gamma), ... \tag{5.5}$$

where α, β, γ are the angles made with the x, y, z axis. Now, if a ray coming from O_0 is a projection onto the $z - axis$, we can multiply that projection to the distance $\overline{O_0 P_0}$. And that gives us our equation for F:

$$\overline{O_1 P_1}(p_1, q_1, m_1) - \overline{O_0 P_0}(p_0, q_0, m_0) = F(x_0, y_0, z_0; x_1, y_1, z_1). \tag{5.6}$$

This equation is representative of the time difference that will occur when we trace rays from points P_0 and P_1 instead of rays from O_0 and O_1. F is constant, no matter what the ray path is. Thus, our equation can be summarized as follows,

$$\nabla V \cdot dr = constant, \tag{5.7}$$

which is an equation synonymous with the derivation of Hamilton's Equation via Huygen's Principle [54]. We can then rewrite F:

$$(p_1 x_1 + q_1 y_1 + m_1 z_1) - (p_0 x_0 + q_0 y_0 + m_0 z_0) = F(x_0, y_0, z_0; x_1, y_1, z_1). \tag{5.8}$$

Note that this is a taylor expansion that neglects terms above the first power in distance. The reason for this is because, when something is axially stigmatic, there is an implication that all the terms in the expansion of the characteristic function to higher powers than the first will be so small (due to the off axis distance) that they can be neglected. And of course, this implies the absence of spherical aberration of all orders. When combined with the Abbe Sine Condition (if met), all terms in the characteristic function dependent on the first power of off-axis distance must also vanish.

This brings us to our special case of interest: when P_0 and P_1 lie in the planes $z_0 = z_1 = 0$.

Special Case $z_0 = z_1 = 0$

Considering our system above, in which our points lie in a meridional plane $(x_0 = x_1 = 0)$, we look at the points P_0 and P_1 when they lie on $z_0 = 0$ and $z_1 = 0$, respectively. Our equation for F then becomes,

$$(q_1y_1) - (q_0y_0) = F(0, y_0, 0; 0, y_1, 0). \tag{5.9}$$

This equation holds for each pair of conjugate rays. Thus, it must also hold for the axial pairs. This pair here is the pair along the optical axis. Thus, $F(0, y_0, 0; 0, y_1, 0)$ must equal 0 and so we then have,

$$q_1y_1 = q_0y_0, \tag{5.10}$$

or more explicitly,

$$n_1y_1sin(\gamma_1) = n_0y_0sin(\gamma_0), \tag{5.11}$$

where γ_0, γ_1 are the angles which the corresponding rays through O_0, O_1 make with the optical axis. This equation is known as "The Abbe Sine Condition", and it is the required condition under which a small region of the object plane in the neighborhood of the axis is imaged sharply by a pencil of rays with any angular divergence.

Special Case: Object at Infinity

If the object lies at infinity (condition for the Jellyfish), the sine condition takes a different form. With Fig. 5.5 in mind, we now move the object point to infinity. If z_0 is the abscissa of the point with respect to the axis at the first focus, and h_0 is the height above the axis at which the ray from the axial point meets the first surface, then we know the rays are coming in parallel, thus, we want to eliminate the angle γ_0.
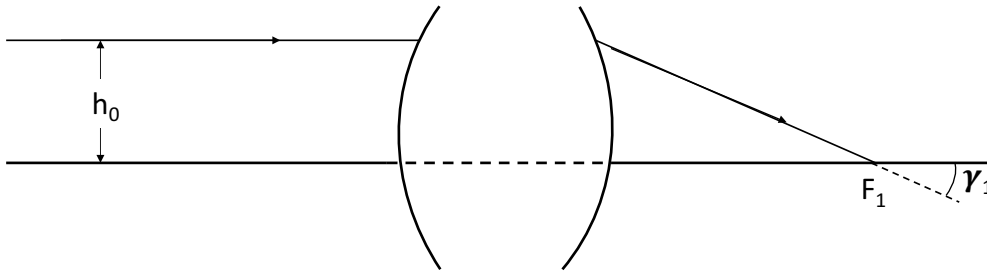
Figure 5.6: Special case of an object at infinity

We can represent $sin(\gamma_0) - h_0/z_0$ or, more precisely, $z_0sin(\gamma_0)/h_0 \to -1$ as $z_0 - \infty$ whilst h_0 is fixed. And so, if z_0 is kept large enough, then Eq. 5.11 becomes:

$$\frac{y_1}{y_0}z_0sin(\gamma_1) = \frac{n_0}{n_1}h_0. \tag{5.12}$$

Aside from eliminating the first angle γ_0, we also want to eliminate our y dependence, and refrain from using any kind of information from the object side other than h_0 (height of parallel light). By eliminating these variables from the object side, we can reduce our previous equation. Our first reduction comes from using Newton's Equation. This equation can be derived using geometrical optics, and found in various texts on optics [53, 55].

$$\frac{y_1}{y_0} = \frac{f_0}{z_0} = \frac{z_1}{f_1} \tag{5.13}$$

Utilizing the first two terms we have:

$$\frac{y_1}{y_0} z_0 = f_0 \tag{5.14}$$

Since we intend to be rid of the term f_0 because it lies in the object space, we introduce another equation, that of the ratio of focal lengths:

$$\frac{f_1}{f_0} = -\frac{n_1}{n_0} \tag{5.15}$$

Next we come to Maxwell's elongation formula, which imply that the longitudinal magnification is equal to the square of the lateral magnification multiplied by either ratio of index of refraction, or focal length. These equations are taken from the text on optics written by Born and Wolf [53].

$$\frac{dz_1}{dz_0} = \frac{n_1}{n_0} \left(\frac{dy_1}{dy_0} \right)^2, \tag{5.16}$$

and

$$\frac{dz_1}{dz_0} = \frac{f_1}{f_0} \left(\frac{dy_1}{dy_0} \right)^2, \tag{5.17}$$

allow us to arrive at our final equation:

$$\frac{y_1}{y_0} z_0 = -\frac{n_0}{n_1} f_1 \tag{5.18}$$

Ignoring the ratio of index of refraction for now, if we take our object out to infinity, i.e. $z_0 \to \infty$ then we see that y_1 for our image point must go to zero. In other words, by taking the object out to infinity, the image moves to the focus point in the image space. By doing this, we bring all parallel incoming rays to a point on the $z - axis$ in the image space at f_1.

$$\frac{y_1}{y_0} = -\frac{f_1}{z_0} \to \frac{0}{y_0} = \frac{f_1}{\infty} \tag{5.19}$$

Finally, substituting 5.18 into 5.12, we arrive at our final equation,

$$\frac{h_0}{sin\gamma_1} = f_1 \tag{5.20}$$

Known as the Sine Condition for an object lying at infinity. It implies that incident rays in the direction parallel to the axis intersect conjugate rays on a sphere of radius f_1, which is concentrated at the focus F_1 as in Fig. 5.6. Axial points which are stigmatic images of each other, and which also have the property that conjugate rays passing through them satisfy the sine condition, form what is called an aplanatic pair.

Finally, we have assumed the condition of axial stigmatism, which means that we have kept higher order terms that do not depend on "off-axis" distance of the object out of the characteristic function. As such, there is an absence of spherical aberration of all orders. When we met the sine condition, we required that all first order terms also became zero, effectively eliminating circular coma. And so, our system is perfectly imaging.

5.3 Constructing the Jellyfish

The Jellyfish is constructed using an iterative process to create mathematical curves which are the two-dimensional side profiles. The curves are first constructed in MATLAB software using a combination of the Abbe Sine Condition, the law of reflection, and Snell's Law. Then the curves can be imported into a three-dimensional CAD program, or in this case a ray trace program called Light Tools. There the curves are turned into a three-dimensional device.

Iterative Curves

Matlab was used to construct iterative segments that made up each Jellyfish curve. Here we will go through the steps used to construct these mathematical curves. Figure 5.7 will be used to illustrate each of the steps.
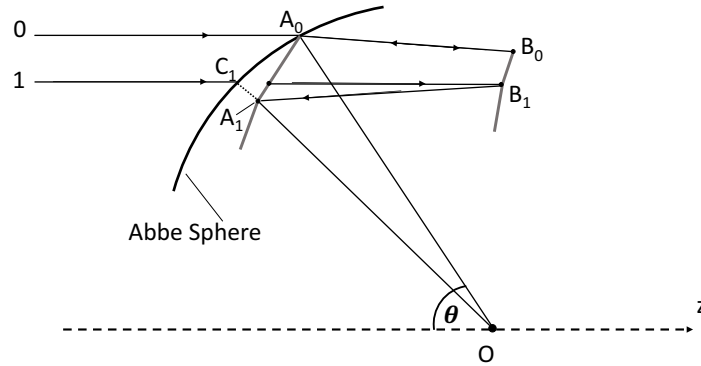
Figure 5.7: Constructing the Jellyfish using the Abbe Sphere via the Luneburg Method where iterative segments are used.

First: Define your Abbe Sphere. The focus O will be the center of the sphere. The variable f_1 is the radius of the sphere.

Second: Define incoming parallel "seed ray", labeled as ray 0 in Fig. 5.7 intersect it with the Abbe Sphere.

Third: Choose a location for the back surface. This is equivalent to setting the "thickness" of the Jellyfish.

Forth: Take ray 0 and refract it towards the back surface based on Snell's Law. Find θ_R doing this.

$$\theta_R = arcsin\left(\frac{n_i}{n_r}sin\theta_i\right) \tag{5.21}$$

Fifth: Determine the y position where the ray strikes the back surface.

Sixth: Determine the slope of the back surface kB_0 of B_0 by finding the surface tangent there. A description of how to find surface tangents (slopes) will follow these steps in the next section.

Seventh: Reflect the ray from the back surface as a "retroreflection", doubling it back on itself such that it arrives at the initial point it crossed the Abbe Sphere at A_0.

Eighth: Send the ray from A_0 to the focus O of the Abbe Sphere. It should reflect from the front surface at the same angle as θ_R because of the reflection law.

Ninth: Determine at A_0 the slope kx_0. This is done using Snell's Law and determining the geometric relationships between angles.

The three relationships are as follows:

$$n_i sin\theta_i = n_r sin\theta_r, \tag{5.22}$$

$$kx_0 = tan(\theta_i), \tag{5.23}$$
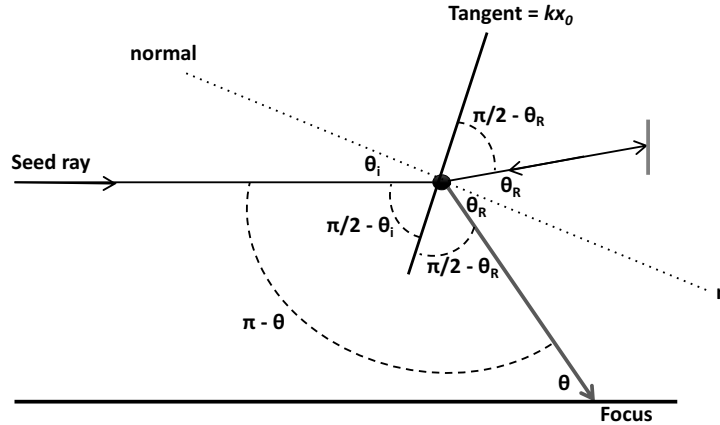
Figure 5.8: Geometric relations for determining the slope of the front surface, kx_0.

$$\pi - \theta = \frac{\pi}{2} - \theta_i + \frac{pi}{2} - \theta_r, \tag{5.24}$$

such that we arrive at the front surface slope:

$$kx_0 = \frac{\frac{n_i}{n_r} + cos\theta}{sin\theta}. \tag{5.25}$$

Tenth: Finally, using A_0, B_0, and the slopes kx_0 and Bx_0, extend the front and back surfaces and repeat the process of steps for new seed rays, with the exception that these no longer retroreflect. Rather, they bounce off the surfaces with a new angle allowing the reflected ray from B_1 to arrive at the end of the first tangent segment at A_1. At this point A_1, a new segment is drawn based on Eq. 5.25.

Steps one through ten iterated allow one to determine two aplanatic surface curves. These curves define only half of the aplanat, however they can be rotated about their $z - axis$ such that a three-dimensional optic is generated. Following this, reflective coating is applied to the back surface as well as the central region of the top surface. The optic is given the correct index of refraction, and it is then ready for simulation testing.

5.4 Optimizing and Prototyping the Jellyfish

The prototyping process for the Jellyfish is done in such a way as to cater to the lighting industry. For example, the size of the Jellyfish can be made to fit what is called MR16 standards. MR stands for multifaceted reflector, which is an optic containing an inside surface that is covered by reflective coating. Generally these reflecting surfaces are faceted
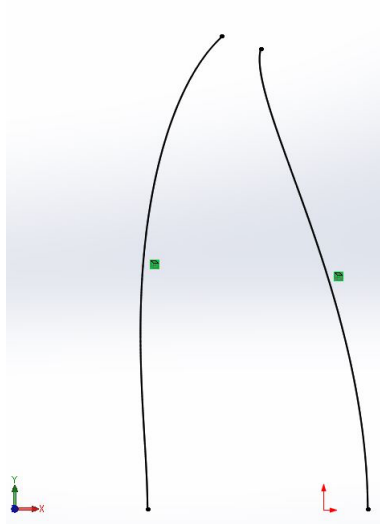
Figure 5.9: Mathematical curves of the Jellyfish optic.

to provide optical control. However, some lamps have smooth surfaces yet are still called
MR by convention. As a guideline, the following prototyping processes have followed these
standards.

MATLAB Optimization

To coincide with MR16 standards, the diameter of the Jellyfish design is restricted to $45mm$.
A material index of $n = 1.49$ is set. The only other free parameters that may be adjusted
during the design process are the edge thickness of the optic, and the half acceptance angle
for the LED, θ. Changing either of these parameters will dictate the size of the reflective
coating on the front surface, and thus, the theoretical efficiency based on Eq 5.1. So the
idea is to get the reflective coating a small as possible to maximize the theoretical efficiency.
The relationship between reflective coating, half acceptance angle, and edge thickness are as
follows:

1.) As the Jellyfish thickness increases, so does the radius of the reflective coating.

2.) As the Jellyfish half LED acceptance angle increases, the reflective coating area
decreases.

Both relationships are given in Fig. 5.10.
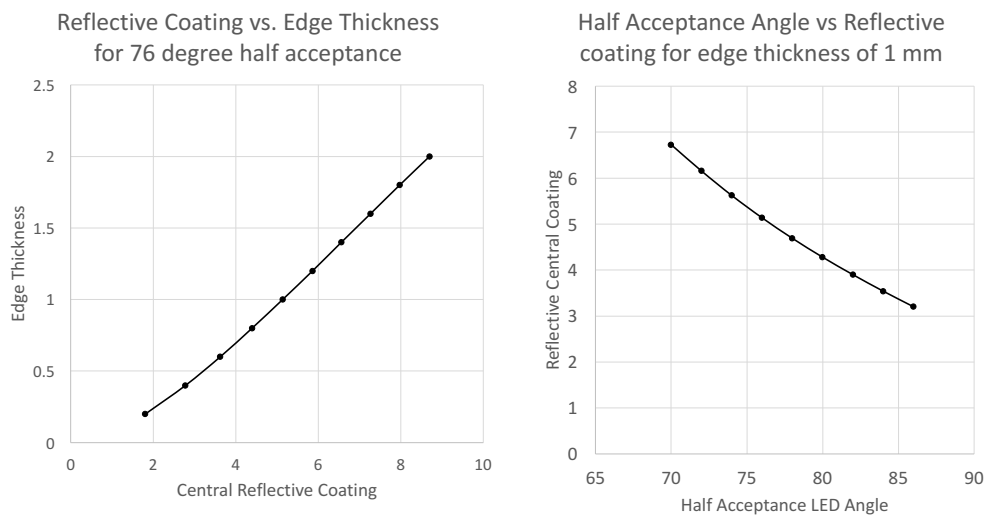
Figure 5.10: Graphs showing the relationship between the central reflective coating on the top surface of the Jellyfish and how it is influenced by the edge thickness (left) and half acceptance angle θ (right). All units here are in mm for the left graph and both mm and degrees for the right graph.
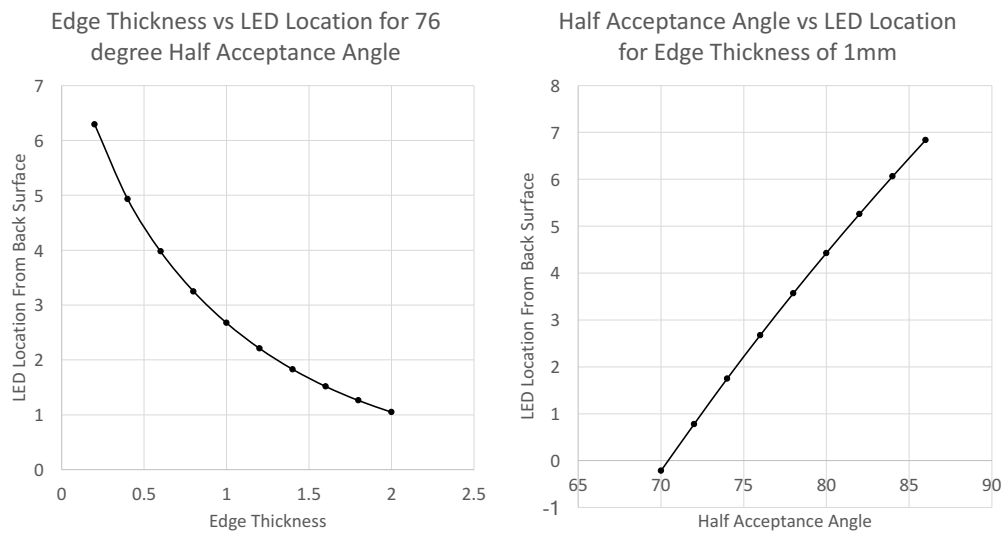
Figure 5.11: Graphs showing the relationship between the LED location with respect to the back surface of the Jellyfish, and how it is influenced by the edge thickness (left) and half acceptance angle θ (right). The right graph shows the LED location going negative. This happens for various parameter settings and is obviously unphysical, as we cannot place the LED outside of the optic.

The final requirement in terms of generating a working prototype, requires that the LED source must be placed directly at the back of the Jellyfish for heat sink purposes. This further limits design parameters. Increasing the edge thickness will inherently move the location of the LED closer to the back surface. Increasing the half acceptance angle of the design will move the LED forward, away from the back surface; we cannot allow this because then the LED will be too far inside the Jellyfish with nowhere to diffuse its heat. With each of these limitations in mind, the optic is designed to accommodate for all of them.

Figure 5.12 shows various locations for the LED with respect to the back surface of the Jellyfish. The LED is always located at the origin $(0,0)$. By adjusting the aforementioned parameters, the location for the origin with respect to the back surface is also adjusted. These images were generated in MATLAB whilst building the curves for front and back surfaces.
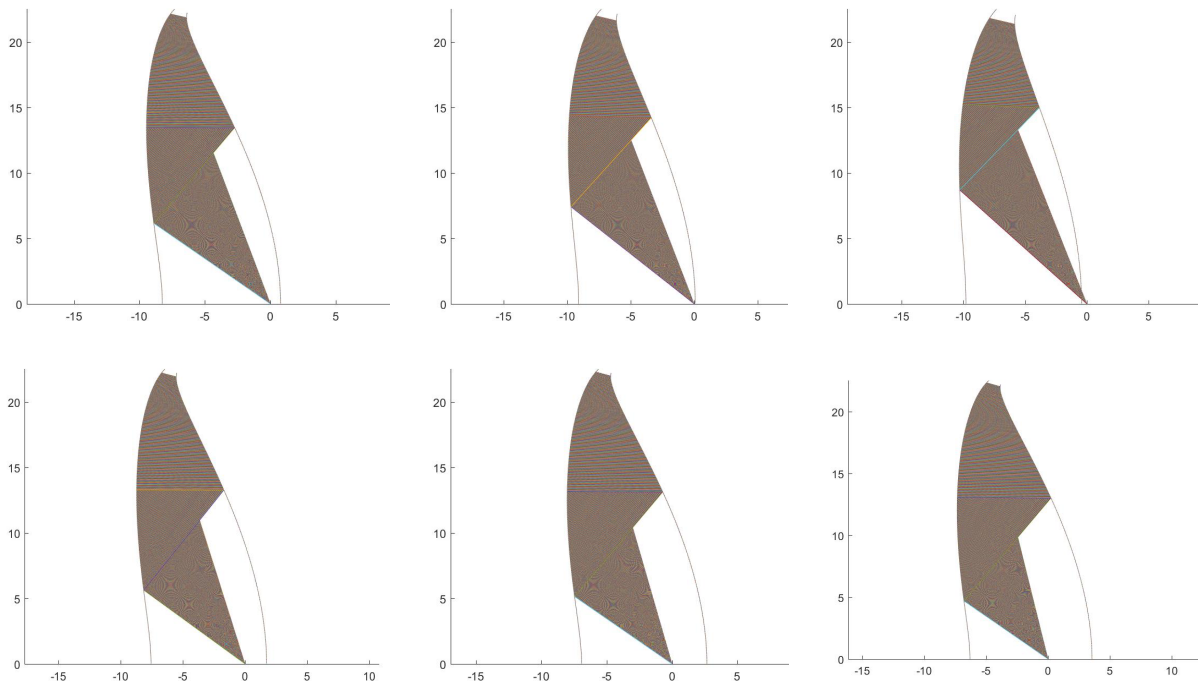
Figure 5.12: Jellyfish optics with various LED locations $(0,0)$. The top row shows increasing edge thickness while holding the half acceptance angle constant; the LED moves closer to the back surface until it is outside of the back surface. The bottom row shows a constant edge thickness while increasing the half acceptance angle; the LED moves further away from the back surface. The dark shading are the ray paths in the optic. (Images from MATLAB)

Once a model is constructed in MATLAB to fit the requirements mentioned above, its curves can be used to generate a three-dimensional optic.

Using Light Tools to generate a three-dimensional optic

The Jellyfish optic is rotationally symmetric. This is favorable for construction in three dimensions. Using ray tracing software (Light Tools), two dimensional y, z curves can be given to build a three-dimensional swept entity, which is an option given for an optic that can be rotated about a third dimension, x. Three curves are used, two for the front surface, and one for the back surface. The front surface curves are split so that one is the reflective coating and the other is the transparent dielectric material. Points at the top of the front and back surfaces are connected to close the optic's curves. The built in feature in Light Tools automatically rotates the optic in the third dimension. The result is a three dimensional Jellyfish optic.
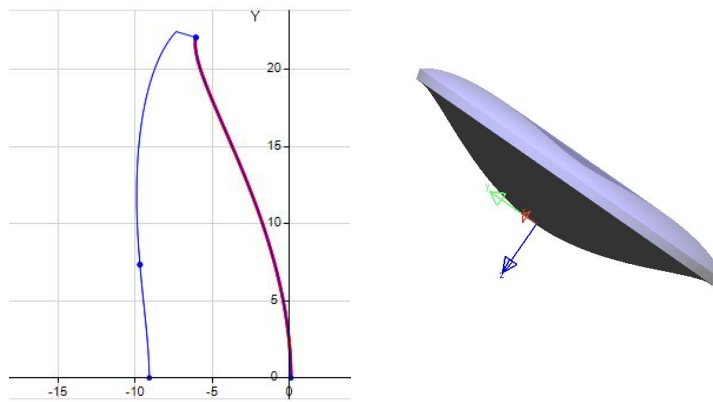
Figure 5.13: The left shows the three segments of curves imported into Light Tools. The right shows the three dimensional solid generated from the curves.

Material of appropriate index is assigned, and optical properties for reflecting and transmitting portions are also designated. To check the integrity of any optical design during the prototyping process, theoretical tests are first carried out. The Jellyfish is designed using a point source rather than an extended source, so for initial tests, a point source of illumination is used. The illumination receiver employed is called a far field receiver, which is a spherical receiver of infinite radius that collects angular variations of radiant intensity in watts/ster. To ensure that the majority of light exiting the device falls within $2°$, a parameter sensitivity test is run, which allows for multiple iterations of simulation (see Fig. 5.14). For each iteration, the far field receiver takes measurements over an incrementally increasing angular extent, and generates a graph that shows the power received over the full range of angles.

A)   Point Source Parameter Sensitivity   **B) Point Source Parameter Sensitivity**

Acceptance Angle (θ) = 72°
Edge Thickness = 1.3mm
Reflective Coating radius = 7.41mm

Acceptance Angle (θ) = 78°
Edge Thickness = 1.9mm
Reflective Coating radius = 7.63mm
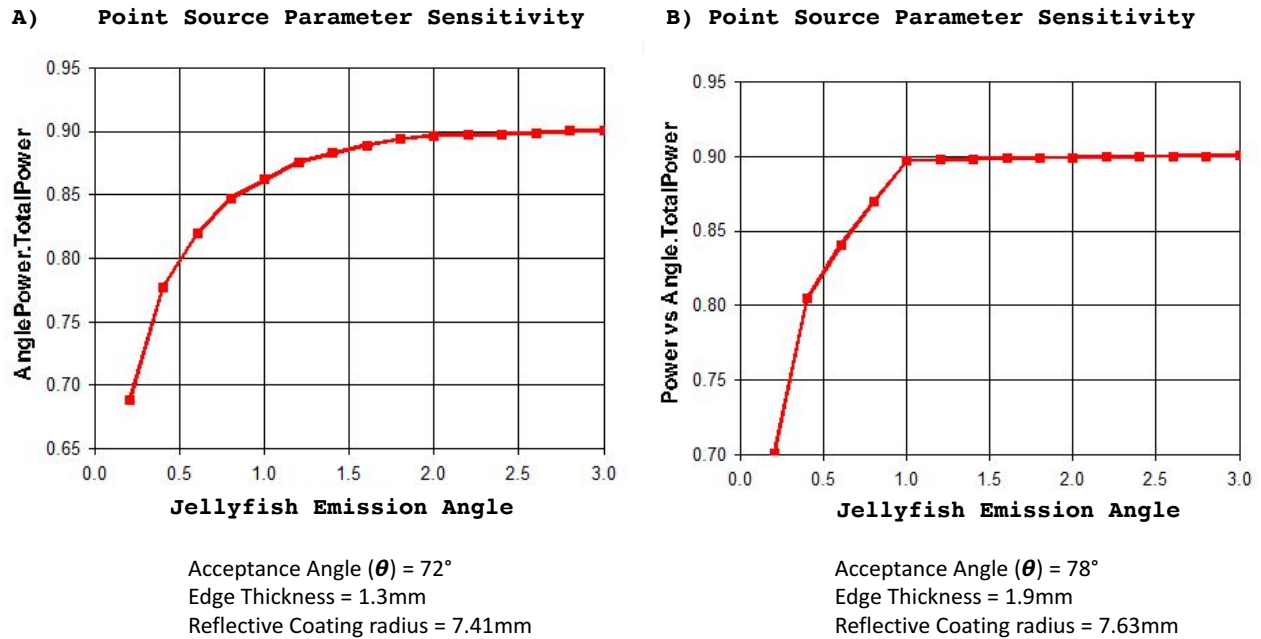
Figure 5.14: Parameter sensitivity test for ideal testing. Jellyfish optic uses a point source for illumination. Increments of increasing angle are measured with the far field receiver and the range of angle increments is plotted vs. power (Watts/ster).

After theoretical testing, material properties are adjusted in Light Tools' built in optical properties manager to reflect realistic manufacturing capabilities. For example, there are no reflective coatings that achieve 100% reflectivity. A more realistic percentage would range from 92% to 96%. A ray data source is then used in place of the point source. A ray data source for an LED is a set of points that emit rays as a lambertian distribution. These are generally provided by a manufacturer of LEDs for the purposes of testing in ray trace softwares. The ray data source was first tested to determine its emission angle and results are shown in Fig. 5.15.

The same parameter sensitivity tests are run for the LED ray data source in the Jellyfish. Results are analyzed and if a poor optical efficiency is obtained, the design process is repeated until a design meets criteria while possessing an optical efficiency that is greater than 80%. This efficiency requirement need not be within the 2° angular extent. For an extended source, there will be a large deviation from the ideal case where rays exit the device to within 2° of the normal to the surface. As long as nearly all the illumination exits $\geq$ than 80% within a 90° extent, illumination manufacturers are happy. For parameter sensitivity tests the angular range is increased to 90°.
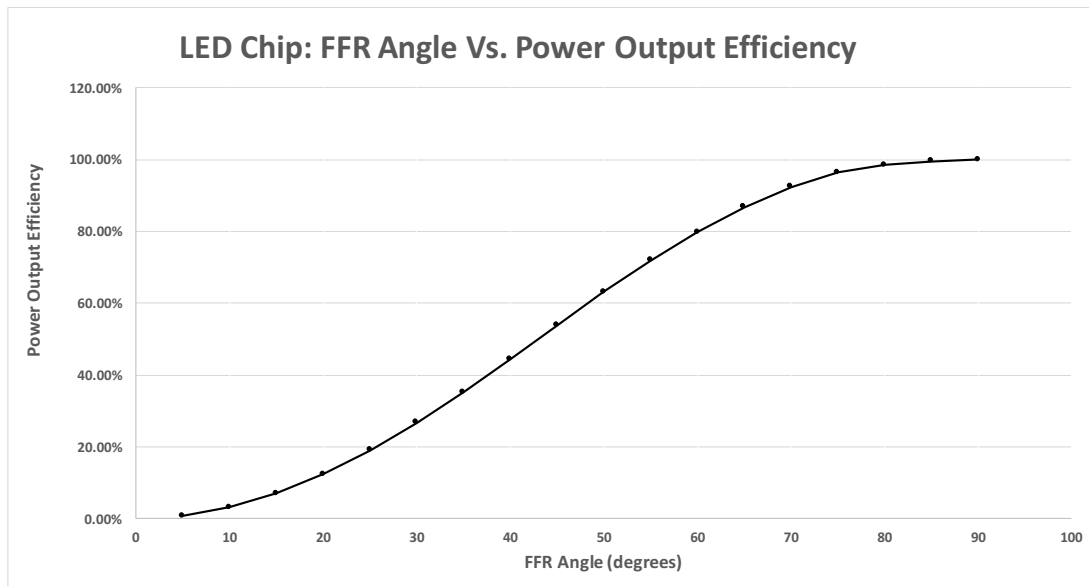
Figure 5.15: Parameter sensitivity test data from Light Tools shows the Far Field Receiver
angle vs. power emission for the LED ray data source that is used. The LED emits .25watts
of power and at 90° the receiver collects 100%.
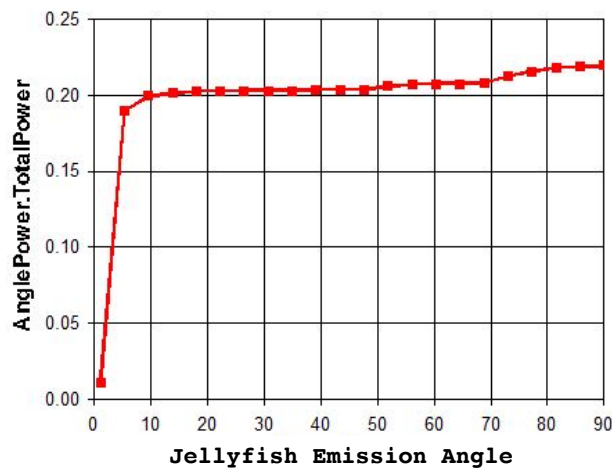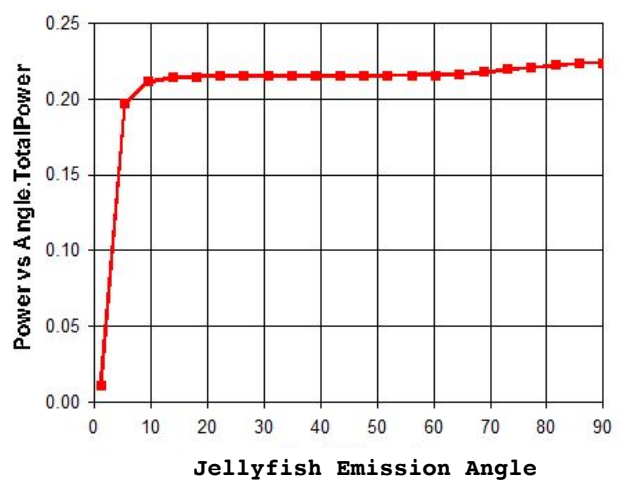
Figure 5.16: Parameter Sensitivity test for LED ray data source using Light Tools. Because
an LED is an extended source, the Jellyfish, which is designed based on point source optics,
will not emit light directly parallel to the surface normal. For manufacturing purposes, as
long as at least 80% of the optic's light is emitted into space, the optic is considered adequate
for manufacturing. Here the LED power is .25W and one can see that at approximately 90°
the nearly 88% of the light escapes the device.

Finally, when looking at the near field illumination distribution of the Jellyfish, it is
seen that it creates a nice "flashlight" collimation beam (see Fig. 5.17 and Fig. 5.18). At
approximately six feet away, the size of the beam is 1foot by 1 foot. Also note that although
the Jellyfish optic is round, the illumination pattern it makes is square. This is because the
Jellyfish is an imaging device and it is imaging the square ray data source LED.
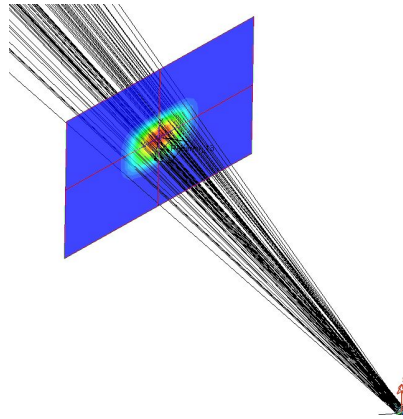
Figure 5.17: Image of the Jellyfish in Light Tools traced in forward with a near field rectangular target that displays false color
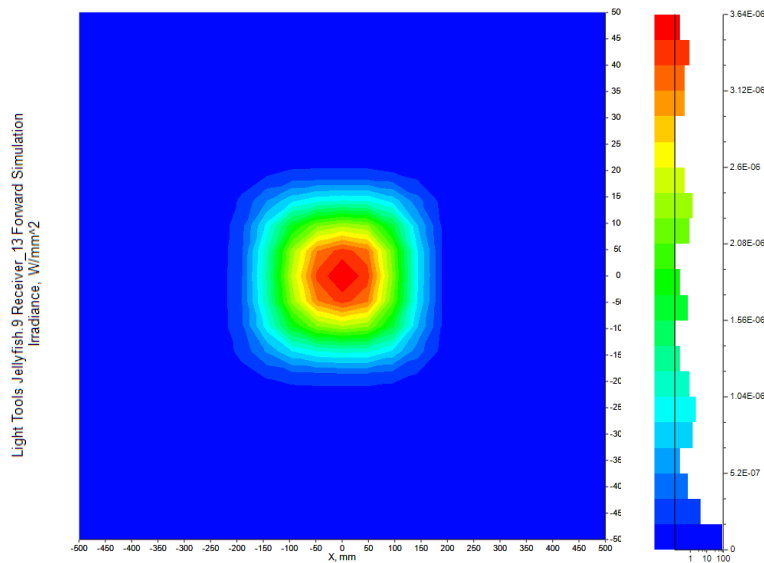
Figure 5.18: From the display, it can be seen that the length and width of the illumination
on the target is approximately 400mm by 400mm.

Prototyping

Once results are adequate, CAD mockups are generated for prototyping. Files are sent to be prototyped, the process takes about six weeks. The process is done by a method called "injection molding," where a mold is made using the mathematical curves of the design and the CAD drawings. The mold is done precisely through machining such as milling and grinding, using what is called a five-axis diamond turning machine. After it is complete, the PMMA plastic is injected into the mold. After it cures, a reflective coating is applied to the back surface and the small middle portion of the front surface. An image of final prototype of the Jellyfish is given in Fig. 5.19. The Jellyfish name came from the fact that upside down, the shape of the back surface semi-resembles the top of a jellyfish.
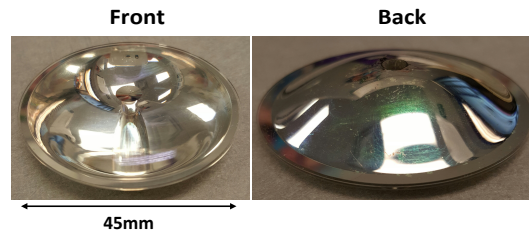
Figure 5.19: Image of prototyped Jellyfish optic. Diameter of this prototype is $45mm$.

5.5   Experimental Results

Testing for the generated prototype was done at Teledine. Two different LEDs were used to complete separate tests using an integrating sphere, which measures the amount of light leaving the optic in an isolated, naturally dark environment. The optical efficiency for each test is given. Optical efficiency is the amount of light that exits a device divided by the amount of light that enters the device. The LEDs are first tested in the sphere to determine their power and lumen output. Then, they are coupled to the Jellyfish using optical gel. The purpose of optical gel is to allow the light to flow from the silicone epxoy dome of the LED, directly into the refractive index material of the Jellyfish, without any air gap losses.

The optical efficiency obtained for the prototyped Jellyfish optic was much lower than anticipated. After receiving the results, the optic was reevaluated. There were no flaws found with the code to generate the curves, nor with the testing. After doing some visual tests on the optic, it appeared to have some shaded spots when viewed head on. We were resigned to conclude that the low efficiency was a result of manufacturing capabilities in the following areas: reflectivity of the coating, lower quality PMMA, and or errors in the diamond turning process.

Integrating Sphere Testing	(*Lumens*)	(*Watts*)
LED 1	167.7	0.49
LED 1 with Jellyfish optic	102.7	0.30
Optical Efficiency	61%	

Table 5.1: Optical efficiency test for the Jellyfish optic (test 1).

Integrating Sphere Testing	(*Lumens*)	(*Watts*)
LED 1	169.2	0.49
LED 1 with Jellyfish optic	111.6	0.32
Optical Efficiency	66%	

Table 5.2: Optical efficiency test for the Jellyfish optic (test 2).

5.6   Conclusion

The Jellyfish optic is an aplanat design that is constructed using the Abbe sine condition and utilizes total internal reflection. It does not suffer from spherical or chromatic aberrations, making it a perfect imaging device. Aside from imaging, it can be used in a wide variety of fields such as solar concentration and illumination. For this work, it is used as an illuminator. Curves for its front and back surface are generated using iterations in MATLAB software. These curves are imported into ray trace software (Light Tools) and revolved in a third dimension to generate a three dimensional revolved optic. Tests are then conducted using illumination sources to determine the optics theoretical optical efficiency. For an ideal design, where a point source is employed, the device achieves its theoretical maximum efficiency for rays exiting to within 2° from the optic's normal. When an LED source, which is an extended source, is substituted, the rays no longer come out nearly parallel, but deviate. The deviation for the majority of the devices power falls to within 10° and at the full angular extent of 90°, approximately 90% of the power is emitted. When it comes to the prototyping process, there are many flaws that require attention for this optic to be a functioning light. For manufacturing purposes, flaws in the prototyping process such as materials used, reflectivity ability, and machining limitations, should all be addressed. The optic has high theoretical efficiencies that are not reflected in the prototypes. With a drop of nearly 20% of efficiency, as it stands now, this device does not make a good illuminator in practice. Once the industry

sector resolves the limitations placed on the design's potential, it should in theory make a great illuminator both esthetically pleasing and free from glare.

Chapter 6

Freeform Optics

The optics manufacturing industry has never before been so technologically dynamic, allowing the field of optics to change rapidly [56]. This has opened the door to a slew of new optics that depart from symmetry. Because of these advancements, an entirely new field known as "freeform optics" has taken shape to accommodate the growing need for higher precision of optical devices.

6.1   Background

A number of differing definitions can be found regarding what qualifies a surface to be "freeform". The simplest and most encompassing of these is given here. A freeform surface is one that leverages a third independent axis during the design process to create an optical surface with nonsymmetric features[57]. Freeform surfaces offer more degrees of freedom than the traditional spherically symmetric options do, and thus, both imaging and nonimaging systems can benefit from the advances; however, nonimaging optics employs far more solutions, utalizing freeform optical systems in streetlights, automotive headlights, secondary optics for LEDs, and laser beam shaping [58]. More impressive still are the number of methods that have taken form to generate these new surfaces. The most prevalent of these will be briefly discussed here.

Methods

The first breakaway from traditional optical designs came in the early 1990s, when Winston suggested that designs(surfaces) could be leveraged and thus tailored for illumination [59, 60]. The field continued to evolve as the desire for solutions to prescribed illumination problems grew. The next major breakthrough method, which ushered in the field now known as freeform optics, happened in the early 2000s. Oliker [61], along with Reis and Muschaweck [62], published on numerical surface solutions to the non-linear second-order partial differential equation of the Monge-Ampere type. Each proposed a different approach.

Oliker's solution utilized iteration, where the optical surface consisted of quadrics (these being ellipses for the near field case, or parabolas for the far field case). Oliker's method, known as the Supporting Quadrics Method (SQM), will be discussed in more detail in the next section. Reis and Muschawek offered a solution based on wave fronts solved numerically using multigrid methods, however, they did not give enough detail to reproduce results, likely because the method is proprietary.

Others have used flux mapping for the design of freeform lenses and reflectors [63, 64, 65]. To map the flux, a source and its target are broken up into matching, equi-flux grids. In this way, flux from each area on the source is mapped to a corresponding area on the target. This creates uniform flux distribution, and allows one to determine where each ray from the source will land on the target.

The Simultaneous Multiple Surface (SMS) method offers another way to generate freeform surfaces that depart from symmetry in three dimensions [66, 67, 68]. The SMS method calculates multiple surfaces simultaneously beginning with an initial pre-defined area [69]. Multiple wave fronts are utilized, and extended sources are taken into account. This method is most commonly used for designing waveguides, lenses, and reflectors.

Finally, more direct methods using optimization have been used when one of the above methods may not apply [70]. Optimization is dependent upon the starting surface. Iterations using design software are used progressively to determine the final surface shape. However, the methods differ based on the desired design.

Less prominent methods also exist; as the field continues to evolve, newer methods will emerge. The increasing usefulness of freeform optics has opened the door to a constant stream of new solutions for optical design. These designs are made possible by the rapidly growing manufacturing industry, which is now capable of producing optics at the level of precision necessary [71].

In this chapter, an overview of the Supporting Quadrics Method will be given. Because of the inherent characteristics of the SQM, diffraction effects become more relevant. The question is: are the diffraction effects influential enough to be taken into consideration? This will be evaluated here.

6.2 Supporting Quadrics Method

The method for supporting quadrics arose from the need for beam shaping [72]. Solutions to the beam shaping problem in the general case resulted in a nonlinear second order partial differential equation of Monge-Ampère type, which remained unsolved for a time. Using geometric optics and ray mapping, Oliker solved this problem without imposing any symmetry assumptions or paraxiallity on the input and output beams and their irradiance [73]. He named his method the Supporting Quadrics Method (SQM).

A quadric is a surface defined by an equation of the second degree, such as a parabola, hyperbola, ellipse, etc. Inherent to an ellipse is the property that rays emerging from one focus will arrive it at an ellipse's second focus. This can be seen in Fig. 6.1. The SQM utilizes

this for the creation of discrete sets of illumination points. Supporting quadrics method can utilize either elliptical surfaces or parabolic surfaces, depending on whether the illumination target is in the near field or the far field. When a target is in the far field, or at infinity, the quadrics are parabolic in nature. As the target moves closer to a finite distance, the quadrics tend towards ellipses.
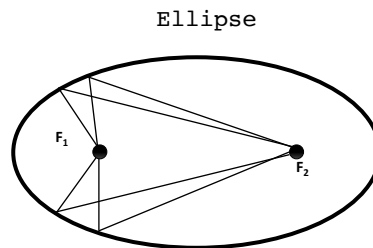
Figure 6.1: Rays from one focus point will terminate at the second. This property allows elliptical quadrics to be utilized for image formation.

Oliker's method of supporting quadrics follows a well-prescribed algorithm combining quadric surfaces, and taking the envelope of those surfaces as the resultant optic. A single quadric, as seen in Fig. 6.2, takes incoming parallel light and directs it to a single point on a target screen a distance, d from the optic. More quadrics can be combined. For example, two quadric surfaces as shown in Fig. 6.3, can take incoming parallel light and each surface will focus to a point such that there are now two points. This can be done iteratively, building more and more quadrics to increase the number of points of light arriving to a target screen. The end result is an image or illumination distribution consisting of points of light in combination.
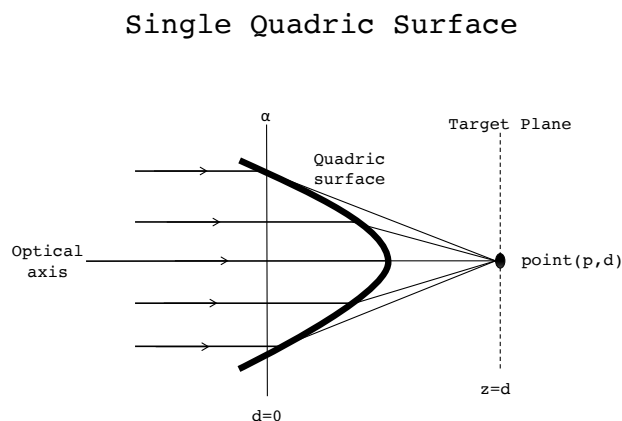
Figure 6.2: Single quadric takes incoming parallel light and focuses it onto a target.
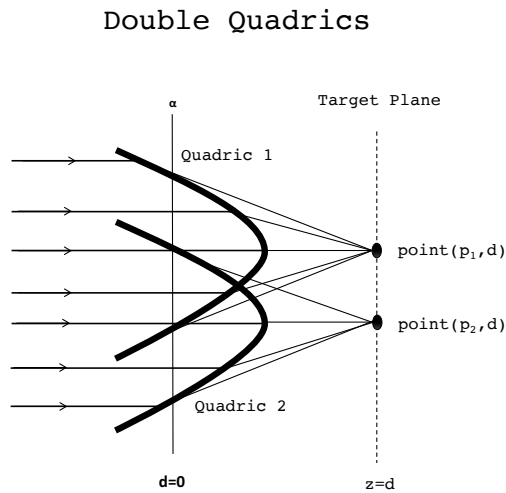
Double Quadrics

Figure 6.3: Double quadrics combine two individual quadrics, each responsible for sending light to a specific point location on a target screen a distance, d from the optic.

To find the envelope of quadrics, which becomes the resultant lens, consider Fig. 6.4 with two quadrics. This is the simplest case. Because each individual quadric is responsible for sending light to a respective point, one can subtract away the innermost (R_1) or outermost (R_2) portions of the two quadrics; light striking each remaining portion will still behave appropriately, traveling to its respective point.

Envelope
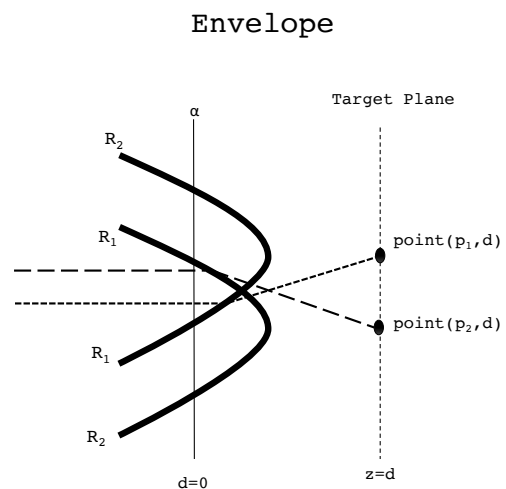
Figure 6.4: Double quadrics combine two individual quadrics, each responsible for sending light to a specific point location on a target screen a distance, d from the optic.

The surface quadric is defined by the "z-sag" function[73]. The z-sag is a way to represent

a conic surface as a graph of a function. For the SQM, the z-sag defines each individual quadric, and is represented as,

$$z_{p,f} = \frac{-\sqrt{f^2 + (n^2 - 1)(x - p)^2} + fn}{n^2 - 1} + d, \tag{6.1}$$

where $x = (x_1, x_2)$ is a point in the plane of α, n is the index of refraction, point $p = (p_1, p_2)$ is located on the target plane, and set by the user, and distance d is the distance between the quadric plane $\alpha = 0$ and the target plane.

Using Eq.6.1, one can control the exact light pattern based on the arrangement of quadrics and their foci. In a prescribed irradiance problem, a target can consist of thousands of points across a surface. The resulting lens for this problem, consisting of the envelope of quadrics, would then consist of corresponding number of quadrics. The ratio of quadrics to pixels is then 1:1.

This method allows for the generation of a range of illumination patterns, such as uniform distributions, or complicated images possessing no symmetry, the latter of which weighing in as the most complicated. Creating an image of high quality requires large pixel counts, thus, the resultant lens can require tens of thousands of quadrics.

Ray trace software is a popular way to design and test freeform surfaces. However, geometrical optics does not account for diffraction, a realistic problem encountered when light passes through any aperture approaching the order of a wavelength. Thus, when moving to the prototyping process, diffraction should be well understood to eliminate any errors in the optical design.

For optical surfaces in SQM, each section of quadric in the envelope is an aperture that sends incoming light to a target (for example, R_1 consists of two apertures). If they are small enough, these apertures may create diffraction effects on each pixel, which can affect the final irradiance distribution. The larger an aperture is, the more light it will allow to pass through. The smaller apertures will let less light through. The gradient of light and dark can be used to create grayscale images like the ones shown in Oliker's work [74]. Finally, smaller apertures are more prone to diffraction effects. It is important to understand to what degree these effects take place, and how much they will detract from the irradiance distribution.

The most likely case where diffraction effects can occur is for a small lens possessing large quadric counts. Splitting a lens into a few quadrics generally implies that these quadrics will be larger (depending on the overall lens size). It stands to reason that as the quadric count increases, the size of these quadrics decrease. Thus, diffraction effects are more likely. In the following section, a simpler case of five quadrics covering a lens of diameter $2\ mm$ will be evaluated.

6.3   Diffraction

The diffraction analysis begins by making approximations using Fraunhofer's diffraction model for a circular aperture; the quadrics used in SQM are estimated as circular apertures

due to their curved surfaces. In this case, determining the diffraction of an optic as a single aperture (ie. convex lens) cannot be done, because diffraction effects will occur for each quadric present. And, the effects seen for various quadrics may be different as quadric apertures are not necessarily the same area. Thus, the problem is approached by looking first at diffraction patterns on individual quadrics, and then determining how these diffraction patterns combine to affect the overall distribution accuracy.

Within each optic, there exists a number i of quadric apertures in possession of varying surface areas A_i. These can be summed together such that these areas comprise the total surface area of the optic,

$$\Sigma_{i=1}^{n}(A_i) = A_0 \tag{6.2}$$

We are able to set the total area (size of optic) as a design parameter, but we do not necessarily know the individual areas of each quadric. Thus, we need a way to determine diffraction without knowing A_i. For simple uniform distributions, as is common in uniform illumination problems, all quadrics will possess an equal area. In this case, an equation can be written to determine individual aperture area:

$$\frac{A_0}{N} = A_i \tag{6.3}$$

Where N is the number of quadrics present in the optic.

The area of an aperture is directly proportional to the intensity of light we see on the target screen; an aperture with a larger area will transmit more light to the target than an aperture half its size. In fact, the apertures in possession of larger areas will steal light from those with smaller areas. Thus the light at the screen will differ depending on the supporting aperture. These various intensities I_i are part of the total intensity of incident light given to the optic I_0, thus,

$$\Sigma_{i=1}^{n}(I_i) = I_0 \tag{6.4}$$

The total intensity is set by the user, and the various intensities seen at the target screen are also known. Thus we can solve for the first ring/fringe of the diffraction pattern based on wavelength, the intensities, total surface area, and distance between aperture and target screen. These are all known quantities.

The ratio of areas to intensities is as follows,

$$\frac{A_0}{A_i} = \frac{I_0}{I_i} \tag{6.5}$$

For a quadric aperture of radius r_i and diameter d_i we have,

$$A_i = \pi[d_i/2]^2 \tag{6.6}$$

$$d_i = \sqrt{4A_i/\pi} = C\sqrt{A_i} \cong \sqrt{A_i} \tag{6.7}$$

C $\sim$ 1 so we see that the diameter of an aperture is approximately the square root of the surface area of each quadric. As mentioned previously, for distributions with varying intensities, we do not know A_i; this is the nature of how the envelope is generated. We do know the original size of the quadric. To get around the use of individual quadric areas, we relate A_i to intensity I_0 and I_i using Eq.6.5.

$$\frac{A_0 I_i}{I_0} = A_i \tag{6.8}$$

We denote the diffraction radius for the first minima as y, as seen in Fig. 6.5.

$$\sin \theta = \frac{m\lambda}{d} \qquad \frac{y}{D} = \tan \theta \cong \theta$$
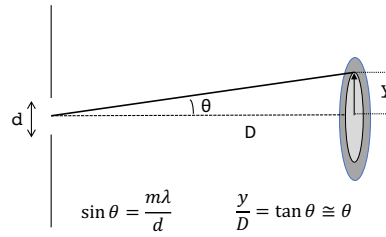
Figure 6.5: Fraunhofer diffraction for a circular aperture. The first ring, with radius y is surrounded by a secondary ring with a lower intensity, possessing a larger radius.

Using equations from Fig. 6.5, we eliminate θ and find,

$$\frac{y_i}{D} = \frac{m\lambda}{d_i} \tag{6.9}$$

For our case, the distance D depends on the quadric in question. This is because the apertures in the lens are at varying distances from the target screen. The difference in D is small but should be accounted for. Thus the distance, D, becomes D_i, and our final equation for the diffraction ring radius is,

$$y_i = \frac{m\lambda}{\sqrt{\frac{A_0 I_i}{I_0}}} D_i \tag{6.10}$$

Also, if the distribution is uniform (all intensities are equal) this equation can be written as,

$$y_i = \frac{m\lambda}{\sqrt{A_i}} D_i = \frac{m\lambda}{\sqrt{\frac{A_0}{N}}} D_i \tag{6.11}$$

For this paper, we will be using equation 9 to evaluate a nonuniform distribution.

When a quadric aperture sends light to its designated target point, that point, or "pixel" of light, will have a large maxima surrounded by less intense rings. The brightest spot in

the middle of this diffraction pattern - the maxima, will have some area. We use the value $m = 1.22$, the first minima point on our distribution to determine the first maxima ring radius.

It is important to note that both the SQM and this model of Fraunhofer diffraction rely on incoming light to be parallel (ie. point source at infinity). In reality, these optics will be used with LEDs, which are neither point source nor at infinity, but rather, extended sources of a lambertian distribution.

Using MATLAB, we implement this final equation into an algorithm that takes our known input data and generates values for the first diffraction fringe seen at each intensity value. The data is assessed to determine the severity of effects for various distributions.

$$\Sigma_{i=1}^{n}(I_i) = I_0 = 575\frac{W}{mm^2} \tag{6.12}$$

For the purpose of simplicity, a lens of five quadrics is used. This lens has a diameter of 2 mm. The illumination distribution consists of five points of light, each at a specific (x, y, z) location on the target screen as shown in Fig. 6.6. The distribution of light intensity is not uniform. The total intensity given to the system is the sum of each individual intensity (assuming no light loss).

$$\Sigma_{i=1}^{n}(I_i) = I_0 = 575\frac{W}{mm^2} \tag{6.13}$$
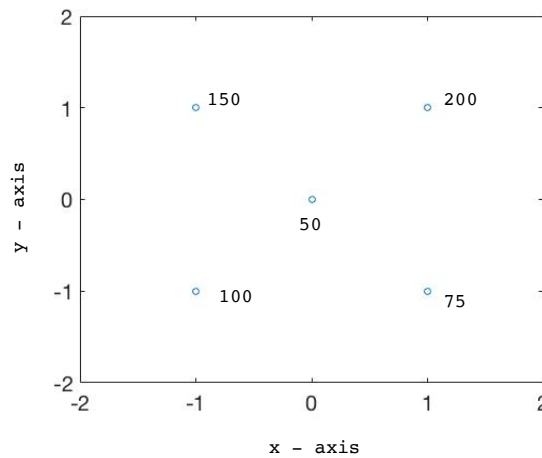
Figure 6.6: Intensity distribution in cartesian coordinates on the z=d=150 plane measured in W/mm^2

For these calculations, red light ($700nm$), and blue light ($500nm$) are used. The results can be seen in the tables below. Figure 6.6 shows the location on the target screen for each point and how much intensity it possesses.

$x-position$	$y-position$	$z=d$	$Intensity[W/mm^2]$	$Ringradius[mm]$
0	0	150	50	0.051
−1	−1	150	100	0.036
−1	1	150	150	0.030
1	1	150	200	0.026
1	−1	150	75	0.042

Table 6.1: Supporting Quadrics Method: diffraction ring radius values for $\lambda = 500nm$.

$x-position$	$y-position$	$z=d$	$Intensity[W/mm^2]$	$Ringradius[mm]$
0	0	150	50	0.072
−1	−1	150	100	0.051
−1	1	150	150	0.041
1	1	150	200	0.036
1	−1	150	75	0.059

Table 6.2: Supporting Quadrics Method: diffraction ring radius values for $\lambda = 700nm$.

The analysis shows the diffraction effects to be negligible. The pixel locations are at least 1.4 mm in distance from the nearest neighbors. The size they will have gained from the first diffraction peak will not cause them to significantly overlap, or merge with each other. Because there are only five apertures in the lens, the ratio of aperture areas to the total lens area means the apertures are much larger than the case where the lens has many apertures. In that case, shrinking aperture sizes will lead to larger diffraction patterns. Secondary rings have not yet been accounted for in this analysis.

6.4   Conclusion

The emerging field of freeform optics is solving many illumination problems that conventional optics has been unsuccessful with. There are many new methods including the supporting quadrics method. For instances where a lens is small (on the order of millimeters), and possesses numerous quadrics, diffraction effects are a viable concern. Given that ray trace

software will not analyze these effects, a more direct measurement method should be performed. Here, an example of such a method was given for five quadrics in a millimeter sized lens. Because the number of quadrics was low in this case, diffraction effects were deemed negligible. But that may not be the case for more complicated scenarios, where thousands of quadrics are used. To produce some of the complicated images such as the ones Oliker has been successful with, a more thorough analysis should be completed to ensure that image blur does not occur. However, for uniform illumination problems, diffraction effects, should there be any, can be ignored.

Chapter 7

Summary and Conclusions

The need for controlled illumination arises from emerging efficiency standards and increasing light pollution. Various applications of lighting such as street lighting, automotive lighting, indoor lighting (industrial, commercial, and residential), and outdoor lighting can all be improved by utilizing concepts in freeform and nonimaging optics. Initially, imaging optics dominated the lighting industry. When the illumination sector diverged from imaging optics finding solutions in nonimaging optics, the field of illumination engineering greatly evolved.

Nonimaging optics was initially developed for solar concentration in the 1960s. During the 1990s, it quickly because useful for illumination, revolutionizing the industry. Illumination optics can now minimize light waste, improve light quality, and enhance light aesthetics. And because illumination optics is concerned with the transferring of light, fundamental concepts in nonimaging optics lead to solutions without imposing the constraints found in imaging optics.

This dissertation is largely concerned with nonimaging optics pertaining to illumination. An overview of this field is given, addressing fundamental topics such as edge-ray theory, strings method, étendue, phase space, angular space, thermodynamics, and flow lines. These concepts are applied to new advances, specifically, the theoretical advances pertaining to the asymmetric compound parabolic concentrator (ACPC). Although similar to the compound parabolic concentrator, this optic has differing acceptance angles, making it versatile for both the fields of solar concentration and illumination. For solar concentration, its asymmetry can be utilized for areas of the world far from the equator, where more extreme seasons are experienced. Also, in regards to illumination, the ACPC offers more specialized control in non-symmetric instances.

A method to determine the acceptance angles for the ACPC based on its design angles is provided. It is important to note that the acceptance angle(s) for any nonimaging optic is always taken with respect to the optic's aperture, which is not always apparent when dealing with symmetric optics. The étendue, phase space, and angular acceptance for the ACPC is derived and discussed, where graphs are provided in parallel with ray trace results to illustrate a more visual way of studying the results. Two cases for each of these results is acknowledged. A way to predict these cases is also discussed.

Finally, flow lines for this asymmetric design are derived, pushing the boundaries of this relatively new nonimaging optics topic. Flow lines are emerging to be useful in both fields of illumination and concentration. Determining the flow line generator for the ACPC helps to round out the overall understanding of how flow lines arise from geometry.

Because of its asymmetry, the ACPC could potentially help in reducing light pollution. Light pollution is a growing problem worldwide. The valley in Yosemite National Park is one example of a place in need of lighting reform. Nonimaging optics offers ways to improve the light quality there. Using a wedge design as a primary optic to transform phase space for a compound parabolic concentrator (CPC), illumination for an equipment yard was controlled to reduce stray light. This nonimaging optics solution was both quick and inexpensive to produce. Furthermore, its small size allowed for retrofitting, which is an ideal way to fix the lights in Yosemite.

Another optic discussed in this dissertation utilizes total internal reflection (TIR) to control illumination. Nicknamed "The Jellyfish" for its shape, this novel aplanatic lens is one of a kind. Impressively, the Jellyfish can be used as either an illuminator or a solar concentrator because its optics work in both forward and reverse scenarios. When designed on a small scale, this optic becomes useful for micro-optic scale concentrating photovoltaic (CPV) solutions. As a light source, its adjustable size, acceptance angle, and thickness can be increased to meet various lighting standards. When designed for ideal cases, emerging rays exit the surface nearly parallel to one another. In fact, high efficiencies are seen for rays to within two degrees of the optical axis. This is due in large part to the design method, which is carried out using the concepts first developed by Ernest Abbe. The Abbe Sphere offers a starting point, after which, ideas of reflection and refraction can be utilized at front and back surfaces to guid light via TIR to its exit points.

Work documented here takes the Jellyfish and optimizes it for illumination solutions. It is adjusted to operate with an extended source (LED) and meet MR-16 standards. Design and simulation processes are given, along with prototyping results.

Finally, design methods in freeform optics offers solutions that can be tailored for even the most complicated illumination distributions. One method discussed in this dissertation, the Supporting Quadrics Method (SQM), takes light rays and directs them to designated locations on a target. The quadrics used for these designs can be ellipsoids, hyperbolids, or parabolids. Numbers of them can used in conjunction with one another to create a desired distribution, after which an envelope is taken to generate a final surface. When the number of these quadrics increase, they must become smaller to accommodate the overall size of the lens. This leads to the question of diffraction effects. Because each quadric is its own aperture, does diffraction play a role in disrupting what should be a precise distribution? Preliminary analysis is done to address this question. Results show that the smaller the aperture, the darker the pixel it creates. This means that as the aperture shrinks in size the pixel intensity decreases. It was determined that for simple distributions, diffraction effects can be neglected.

All the work completed within this dissertation falls into nonimaging optics for illumination. With the growing prevalence of energy standards, optical design is important for

controlling the light emitted from LEDs. This relatively new field provides the fundamental concepts necessary to design solutions for preventing light pollution, creating prescribed distributions, and achieving high efficiencies. Many of the concepts discussed can be taken further for development.

Future Work

Illumination optics is constantly evolving. Not only are new, more advanced LED technologies rapidly developing, but the optics used in conjunction must also keep up. Thus, significant change in solid state lighting has been witnessed in the last decade and will continue to rapidly change [75, 76]. In this section, some directions of future work will be discussed for the topics covered in this dissertation.

Nonimaging Optics and Flow lines

Progress in nonimaging optics theory will focus on flow lines. Flow lines are the newest topic, offering great potential. Future work should address ways to use flow line generators to develop new flow line designs. Because current flow line designs are restricted to existing concentrator geometry, the best way to to deviate from this constraint is to determine a method for constructing flow lines using generators that are designed from desired target distributions. This way, a user can generate a flow line design based solely on the desired illumination distribution and distance from target to optic.

Some of the newest work in this field has demonstrated that with any set of flow lines, new optics can be realized. This is revolutionary. If new sets of flow lines can be obtained without constraints of existing optics, flow lines will offer a new method of design for prescribed illumination problems.

Aplanatic Jellyfish Optic

The optical efficiencies obtained by the Jellyfish show that there is still more work to be done. An efficiency of 80% is ideal for illumination optics. Several areas within manufacturing should be evaluated to find potential areas for improvement, namely, the molding process and materials used.

The Jellyfish optic has the potential to reach high optical efficiencies and cheaper fabrication costs. One way to achieve reduced cost is by eliminating the reflective coating on the back. Utilizing the idea that v-grooves can work as reflectors, which was shown in [77] where TIR devices were designed using v-groove back surfaces in lieu of reflective back surfaces. To implement this idea for the Jellyfish, the back surface curve must be used to create a groove with sides of the same profile, and this design must be replicated radially. By avoiding reflective coating on the back surface, the optic could be molded with the grooves and eliminate part of the fabrication process and cost. However, manufacturing capabilities should be addressed.

Supporting Quadrics Method

The supporting quadrics method is designed to work for parallel light (point sources) as shown in Chapter 6. Its ideas work great for mathematical and theoretical solutions. However, point sources are unrealistic. LEDs are extended sources and cannot be treated like point sources for use with this method. For the SQM to work with real sources, its theory requires adjustment. Several groups have already started using various optimizing tools to utilize methods similar to SQM for work with LED sources. The steps to allow the SQM to work with LEDs will be vital for any future work regarding illumination optics implementation. However, once the ideas have been adjusted, this method offers great potential for prescribed irradiance problems.

Conclusion

Illumination optics with its rapidly growing technology is addressing new concepts to further push the field. Healthy lighting, color mixing, new optical designs, organic LED implementation, and integration of lighting into smart systems, are all areas under intense research and improvement. The market has become a frenzied race to develop the latest and most state of the art technology, whilst boosting improvements in efficiencies, aesthetics, and performance. Yet, at the base of it all are concepts rooted in nonimaging optics and the ability to transport electromagnetic radiation from one surface to another. These fundamentals such as étendue and edge ray theory, will continue to carry the industry, helping to push the boundaries of new designs and concepts for years to come.

Bibliography

[1]   J.C. Miñano R. Winston and P. Benítez. *Nonimaging Optics*. Burlington, MA: Elsevier Academic Press, 2005.

[2]   J. Chaves. *Introduction to Nonimaging Optics*. Second. Boca Raton, FL: CRC Press, 2015.

[3]   Warren J. Smith. *Modern Optical Engineering*. Second. McGraw-Hill, Inc., 1990.

[4]   W. T. Welford and R. Winston. "The ellipsoid paradox in thermodynamics". In: *Journal of Statistical Physics* 28.3 (1982), pp. 603–606. ISSN: 1572-9613. DOI: 10.1007/BF01008327. URL: http://dx.doi.org/10.1007/BF01008327.

[5]   Bennett Widyolar, Lun Jiang, and Roland Winston. "Thermodynamics and the segmented compound parabolic concentrator". In: *Journal of Photonics for Energy* 7.2 (2017), p. 028002. DOI: 10.1117/1.JPE.7.028002. URL: http://dx.doi.org/10.1117/1.JPE.7.028002.

[6]   W.H. McAdams. *Hottel, H.C.: Radiant Heat Transmission, in Heat Transmission*. Third Ed. Chapter 3. McGraw-Hill, 1954.

[7]   Michael F Modest. "Chapter 4 - View Factors". In: *Radiative Heat Transfer (Third Edition)*. Ed. by Michael F Modest. Third Edition. Boston: Academic Press, 2013, pp. 129 –159. ISBN: 978-0-12-386944-9. DOI: https://doi.org/10.1016/B978-0-12-386944-9.50004-2. URL: http://www.sciencedirect.com/science/article/pii/B9780123869449500042.

[8]   Boe Colabewala, Lun Jiang, and Roland Winston. *String method of nonimaging optics from a radiation theory perspective*. 2015. DOI: 10.1117/12.2191946. URL: http://dx.doi.org/10.1117/12.2191946.

[9]   Lun Jiang and Roland Winston. *Asymmetric design for Compound Elliptical Concentrators (CEC) and its geometric flux implications*. 2015. DOI: 10.1117/12.2191948. URL: http://dx.doi.org/10.1117/12.2191948.

[10] R. Winston W.T. Welford. *High Collection Nonimaging Optics*. San Diego: Academic Press, INC., 1991.

[11] R. Winston and W. T. Welford. "Geometrical vector flux and some new nonimaging concentrators". In: *J. Opt. Soc. Am.* 69.4 (1979), pp. 532–536. DOI: 10.1364/JOSA.69.000532. URL: http://www.osapublishing.org/abstract.cfm?URI=josa-69-4-532.

[12] R. Winston. "Light Collection within the Framework of Geometrical Optics*". In: *J. Opt. Soc. Am.* 60.2 (1970), pp. 245–247. DOI: `10.1364/JOSA.60.000245`. URL: `http://www.osapublishing.org/abstract.cfm?URI=josa-60-2-245`.

[13] R.J. Koshel. *Illumination Engineering: Design with Nonimaging Optics*. Wiley, 2012. ISBN: 9781118462492. URL: `https://books.google.com/books?id=WpL2SQoJCXYC`.

[14] E. Fermi. *Notes on thermodynamics and statistics*. Phoenix science series. University of Chicago Press, 1966. URL: `https://books.google.com/books?id=ufF8wYXwiqUC`.

[15] Roland Winston. "Principles of solar concentrators of a novel design". In: *Solar Energy* 16.2 (1974), pp. 89 –95. ISSN: 0038-092X. DOI: `http://dx.doi.org/10.1016/0038-092X(74)90004-8`. URL: `http://www.sciencedirect.com/science/article/pii/0038092X74900048`.

[16] P. Moon and D. Spencer. *The Photic Field*. Cambridge, MA: MIT Press, 1981.

[17] Marc Levoy and Pat Hanrahan. "Light Field Rendering". In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '96. New York, NY, USA: ACM, 1996, pp. 31–42. ISBN: 0-89791-746-4. DOI: `10.1145/237170.237199`. URL: `http://doi.acm.org/10.1145/237170.237199`.

[18] A. Gershun. "The Light Field". In: *Journal of Mathematics and Physics* 18.1-4 (1939), pp. 51–151. ISSN: 1467-9590. DOI: `10.1002/sapm193918151`. URL: `http://dx.doi.org/10.1002/sapm193918151`.

[19] Parry Moon and Domina Eberle Spencer. "Theory of the photic field". In: *Journal of the Franklin Institute* 255.1 (1953), pp. 33 –50. ISSN: 0016-0032. DOI: `http://dx.doi.org/10.1016/0016-0032(53)90727-3`. URL: `http://www.sciencedirect.com/science/article/pii/0016003253907273`.

[20] Roland Winston. "Thermodynamically efficient solar concentrators". In: *Journal of Photonics for Energy* 2.1 (2012), pp. 025501–1–025501–6. DOI: `10.1117/1.JPE.2.025501`. URL: `http://dx.doi.org/10.1117/1.JPE.2.025501`.

[21] H. Hinterberger and R. Winston. "Efficient Light Coupler for Threshold Čerenkov Counters". In: *Review of Scientific Instruments* 37.8 (1966), pp. 1094–1095. DOI: `10.1063/1.1720428`. eprint: `http://dx.doi.org/10.1063/1.1720428`. URL: `http://dx.doi.org/10.1063/1.1720428`.

[22] H. Hinterberger and R. Winston. "Use of a Solid Light Funnel to Increase Phototube Aperture without Restricting Angular Acceptance". In: *Review of Scientific Instruments* 39.8 (1968), pp. 1217–1218. DOI: `10.1063/1.1683629`. eprint: `http://dx.doi.org/10.1063/1.1683629`. URL: `http://dx.doi.org/10.1063/1.1683629`.

[23] H. Hinterberger et al. "The Design and Performance of a Gas Čerenkov Counter with Large PhaseÂêSpace Acceptance". In: *Review of Scientific Instruments* 41.3 (1970), pp. 413–418. DOI: `10.1063/1.1684531`. eprint: `http://dx.doi.org/10.1063/1.1684531`. URL: `http://dx.doi.org/10.1063/1.1684531`.

[24]   A. Rabl. *Active Solar Collectors and their Applications*. New York: Oxford University Press, 1985.

[25]   R. Winston W.T. Welford. *The Optics of Nonimaging Concentrators - Light and Solar Energy*. New York: Academic Press, INC., 1978.

[26]   Ari Rabl and Roland Winston. "Ideal concentrators for finite sources and restricted exit angles". In: *Appl. Opt.* 15.11 (1976), pp. 2880–2883. DOI: 10.1364/AO.15.002880. URL: http://ao.osa.org/abstract.cfm?URI=ao-15-11-2880.

[27]   David R. Mills and John E. Giutronich. "Asymmetrical non-imaging cylindrical solar concentrators". In: *Solar Energy* 20.1 (1978), pp. 45 –55. ISSN: 0038-092X. DOI: http://dx.doi.org/10.1016/0038-092X(78)90140-8. URL: http://www.sciencedirect.com/science/article/pii/0038092X78901408.

[28]   D.R. Mills and J.E. Giutronich. "New ideal concentrators for distant radiation sources". In: *Solar Energy* 23.1 (1979), pp. 85 –87. ISSN: 0038-092X. DOI: http://dx.doi.org/10.1016/0038-092X(79)90048-3. URL: http://www.sciencedirect.com/science/article/pii/0038092X79900483.

[29]   Tapas K. Mallick, Philip C. Eames, and Brian Norton. "Non-concentrating and asymmetric compound parabolic concentrating building façade integrated photovoltaics: An experimental comparison". In: *Solar Energy* 80.7 (2006), pp. 834 –849. ISSN: 0038-092X. DOI: http://dx.doi.org/10.1016/j.solener.2005.05.011. URL: http://www.sciencedirect.com/science/article/pii/S0038092X0500188X.

[30]   A. Harmim et al. "Performance study of a box-type solar cooker employing an asymmetric compound parabolic concentrator". In: *Energy* 47.1 (2012). Asia-Pacific Forum on Renewable Energy 2011, pp. 471 –480. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2012.09.037. URL: http://www.sciencedirect.com/science/article/pii/S0360544212007153.

[31]   Siti Hawa Abu-Bakar et al. "Rotationally asymmetrical compound parabolic concentrator for concentrating photovoltaic applications". In: *Applied Energy* 136 (2014), pp. 363 –372. ISSN: 0306-2619. DOI: http://dx.doi.org/10.1016/j.apenergy.2014.09.053. URL: http://www.sciencedirect.com/science/article/pii/S0306261914010010.

[32]   D.R. Mills and J.E. Giutronich. "Symmetrical and asymmetrical ideal cylindrical radiation transformers and concentrators". In: *J. Opt. Soc. Am.* 69.2 (1979).

[33]   Melissa N. Ricketts, Roland Winston, and Jonathan Ferry. "Étendue and angular acceptance of the asymmetric compound parabolic concentrator". In: *Journal of Photonics for Energy* 7.2 (2017), p. 028003. DOI: 10.1117/1.JPE.7.028003. URL: http://dx.doi.org/10.1117/1.JPE.7.028003.

[34]   Lun Jiang and Roland Winston. *Asymmetric design for Compound Elliptical Concentrators (CEC) and its geometric flux implications*. 2015. DOI: 10.1117/12.2191948. URL: http://dx.doi.org/10.1117/12.2191948.

[35] Angel García-Botella et al. "Hyperparabolic concentrators". In: *Appl. Opt.* 48.4 (2009), pp. 712–715. DOI: 10.1364/AO.48.000712. URL: http://ao.osa.org/abstract.cfm?URI=ao-48-4-712.

[36] Lun Jiang and Roland Winston. "Thermodynamic origin of nonimaging optics". In: *Journal of Photonics for Energy* 6.4 (2016), p. 047003. DOI: 10.1117/1.JPE.6.047003. URL: http://dx.doi.org/10.1117/1.JPE.6.047003.

[37] R. Winston and W. T. Welford. "Ideal flux concentrators as shapes that do not disturb the geometrical vector flux field: A new derivation of the compound parabolic concentrator". In: *J. Opt. Soc. Am.* 69.4 (1979), pp. 536–539. DOI: 10.1364/JOSA.69.000536. URL: http://www.osapublishing.org/abstract.cfm?URI=josa-69-4-536.

[38] 2016. URL: https://www.nps.gov/subjects/nightskies/index.htm.

[39] Manuel Collares-Pereira et al. *Redirecting concentrated radiation*. 1995. DOI: 10.1117/12.216975. URL: http://dx.doi.org/10.1117/12.216975.

[40] Melissa Ricketts et al. *Nonimaging optics in lighting to reduce light pollution*. 2016. DOI: 10.1117/12.2239506. URL: http://dx.doi.org/10.1117/12.2239506.

[41] O.S. America. *Handbook of Optics*. v. 4. Mcgraw-hill, 2000. ISBN: 9780071364560. URL: https://books.google.com/books?id=vWAWVJ3vjPcC.

[42] 2014. URL: http://camera-wiki.org/wiki/Steinheil.

[43] A K Head. "The Two-Mirror Aplanat". In: *Proceedings of the Physical Society. Section B* 70.10 (1957), p. 945. URL: http://stacks.iop.org/0370-1301/70/i=10/a=304.

[44] D. Lynden-Bell. "Exact optics: a unification of optical telescope design". In: *Mon. Not. R. Astron. Soc.* 334 (2002), pp. 787 –796.

[45] R. V. Willstrop and Lynden-Bell. ""Exact optics - II. Exploration of designs on- and off- axis". In: *Monthly Notice of the Royal Astronomical Society* 342.1 (2003), pp. 33 –49.

[46] Jeffrey M. Gordon, Daniel Feuermann, and Pete Young. "Unfolded aplanats for high-concentration photovoltaics". In: *Opt. Lett.* 33.10 (2008), pp. 1114–1116. DOI: 10.1364/OL.33.001114. URL: http://ol.osa.org/abstract.cfm?URI=ol-33-10-1114.

[47] Jeffrey M. Gordon. "Aplanatic optics for solar concentration". In: *Opt. Express* 18.S1 (2010), A41–A52. DOI: 10.1364/OE.18.000A41. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-101-A41.

[48] Alex Goldstein et al. "Nested aplanats for practical maximum-performance solar concentration". In: *Opt. Lett.* 36.15 (2011), pp. 2836–2838. DOI: 10.1364/OL.36.002836. URL: http://ol.osa.org/abstract.cfm?URI=ol-36-15-2836.

[49] Heylal Mashaal, Daniel Feuermann, and Jeffrey M. Gordon. "Aplanatic Fresnel optics". In: *Opt. Express* 25.8 (2017), A274–A282. DOI: 10.1364/OE.25.00A274. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-25-8-A274.

[50] R. Winston and W. Zhang. "Novel aplanatic designs". In: *Opt. Lett.* 34.19 (2009), pp. 3018–3019. DOI: 10.1364/OL.34.003018. URL: http://ol.osa.org/abstract.cfm?URI=ol-34-19-3018.

[51] R. Winston and W. Zhang. *Light collection apparatus, system and method.* US Patent 8,355,214. 2013. URL: https://www.google.com/patents/US8355214.

[52] Roland Winston and Weiya Zhang. *Aplanatism and nonimaging optics.* 2008. DOI: 10.1117/12.813936. URL: http://dx.doi.org/10.1117/12.813936.

[53] M. Born and E. Wolf. *Principles of Optics.* Pergamon Press, 1964.

[54] Roland Winston and Wenjun Ge. *Hamilton optics: transformational theory of optics.* 2013. DOI: 10.1117/12.2027837. URL: http://dx.doi.org/10.1117/12.2027837.

[55] Eugene Hecht. *Optics.* 4th. Boston: Pearson, 2016.

[56] Jannick P. Rolland and Kevin P. Tompson. "Freeform optics: Evolution? No, revolution!" In: *SPIE.* (2012). DOI: 10.1117/2.1201207.004309.

[57] Kevin P. Thompson and Jannick P. Rolland. "Freeform Optical Surfaces: A Revolution in Imaging Optical Design". In: *Opt. Photon. News* 23.6 (2012), pp. 30–35. DOI: 10.1364/OPN.23.6.000030. URL: http://www.osa-opn.org/abstract.cfm?URI=opn-23-6-30.

[58] Groot Gregory, Craig Olson, and Florian Fournier. "Special Section Guest Editorial: Freeform Optics". In: *Optical Engineering* 53.3 (2014), p. 031301. DOI: 10.1117/1.OE.53.3.031301. URL: http://dx.doi.org/10.1117/1.OE.53.3.031301.

[59] Roland Winston and Harald Ries. "Nonimaging reflectors as functionals of the desired irradiance". In: *J. Opt. Soc. Am. A* 10.9 (1993), pp. 1902–1908. DOI: 10.1364/JOSAA.10.001902. URL: http://josaa.osa.org/abstract.cfm?URI=josaa-10-9-1902.

[60] Roland Winston. *Nonimaging optics: optical design at the thermodynamic limit.* 1991. DOI: 10.1117/12.49125. URL: http://dx.doi.org/10.1117/12.49125.

[61] Vladimir Oliker. "Mathematical Aspects of Design of Beam Shaping Surfaces in Geometrical Optics". In: *Trends in Nonlinear Analysis.* Ed. by Markus Kirkilionis et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 193–224. ISBN: 978-3-662-05281-5. DOI: 10.1007/978-3-662-05281-5_4. URL: https://doi.org/10.1007/978-3-662-05281-5_4.

[62] Harald Ries and Julius Muschaweck. "Tailored freeform optical surfaces". In: *J. Opt. Soc. Am. A* 19.3 (2002), pp. 590–595. DOI: 10.1364/JOSAA.19.000590. URL: http://josaa.osa.org/abstract.cfm?URI=josaa-19-3-590.

[63] Kai Wang et al. "Freeform LED lens for rectangularly prescribed illumination". In: *Journal of Optics A: Pure and Applied Optics* 11.10 (2009), p. 105501. URL: http://stacks.iop.org/1464-4258/11/i=10/a=105501.

[64] Florian R. Fournier, William J. Cassarly, and Jannick P. Rolland. "Fast freeform reflector generation using source-target maps". In: *Opt. Express* 18.5 (2010), pp. 5295–5304. DOI: 10.1364/OE.18.005295. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-5-5295.

[65] Cristina Canavesi, William J. Cassarly, and Jannick P. Rolland. "Target flux estimation by calculating intersections between neighboring conic reflector patches". In: *Opt. Lett.* 38.23 (2013), pp. 5012–5015. DOI: 10.1364/OL.38.005012. URL: http://ol.osa.org/abstract.cfm?URI=ol-38-23-5012.

[66] Pablo Benítez et al. "Simultaneous multiple surface optical design method in three dimensions". In: *Optical Engineering* 43.7 (2004), pp. 1489–1502. DOI: 10.1117/1.1752918. URL: http://dx.doi.org/10.1117/1.1752918.

[67] Pablo Benítez et al. "Simultaneous multiple surface optical design method in three dimensions". In: *Optical Engineering* 43.7 (2004), pp. 1489–1502. DOI: 10.1117/1.1752918. URL: http://dx.doi.org/10.1117/1.1752918.

[68] Juan C. Miñano, Pablo Benítez, and Asunción Santamaría. "Free-form optics for illumination". In: *Optical Review* 16.2 (2009), pp. 99 –102.

[69] Juan Miñano and Pablo Benítez. "Method designs free-form optical devices". In: *SPIE Newsroom* (2009).

[70] Bo Yang. "Automating design of free-form optics for LED lighting". In: *SPIE Newsroom* (2008). DOI: 10.1117/2.1200802.1064.

[71] F.Z. Fang et al. "Manufacturing and measurement of freeform optics". In: *CIRP Annals - Manufacturing Technology* 62.2 (2013), pp. 823 –846. ISSN: 0007-8506. DOI: http://dx.doi.org/10.1016/j.cirp.2013.05.003. URL: http://www.sciencedirect.com/science/article/pii/S0007850613001935.

[72] L. A. Romero and F. M. Dickey. "Lossless laser beam shaping". In: *J. Opt. Soc. Am. A* 13.4 (1996), pp. 751–760. DOI: 10.1364/JOSAA.13.000751. URL: http://josaa.osa.org/abstract.cfm?URI=josaa-13-4-751.

[73] Vladimir I. Oliker. "Differential equations for design of a freeform single lens with prescribed irradiance properties". In: *Optical Engineering* 53 (2013), pp. 53 –53 –11. DOI: 10.1117/1.OE.53.3.031302. URL: http://dx.doi.org/10.1117/1.OE.53.3.031302.

[74] Vladimir I. Oliker and Boris Cherkasskiy. "Controlling light with freeform optics: recent progress in computational methods for optical design of freeform lenses with prescribed irradiance properties". In: *Proc. SPIE* 9191 (2014), pp. 9191 –9191 –7. DOI: 10.1117/12.2063355. URL: http://dx.doi.org/10.1117/12.2063355.

[75] M. H. Crawford. "LEDs for Solid-State Lighting: Performance Challenges and Recent Advances". In: *IEEE Journal of Selected Topics in Quantum Electronics* 15.4 (2009), pp. 1028–1040. ISSN: 1077-260X. DOI: 10.1109/JSTQE.2009.2013476.

[76] M. R. Krames et al. "Status and Future of High-Power Light-Emitting Diodes for Solid-State Lighting". In: *Journal of Display Technology* 3.2 (2007), pp. 160–175. ISSN: 1551-319X. DOI: 10.1109/JDT.2007.895339.

[77] Pablo BenÌtez Jess LÛpez Juan Vilaplana Guillermo Biot Marina Buljan Dejan Grabovi?ki? Juan C. MiÒano. *Metal-less V-groove RXI collimator.* 2011. DOI: 10.1117/12.896783. URL: http://dx.doi.org/10.1117/12.896783.