Mathematical Methods for Image Synthesis

Thesis by
Min Chen

Advisor
James R. Arvo

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Science

California Institute of Technology
Pasadena, California

(Defended April 2001)

November 2, 2001

Acknowledgements

There are many people I wish to thank. First and foremost, I offer my sincere thanks to my advisor, James Arvo for his steadfast confidence in me and getting me admitted into Caltech. Without his unhesitating support and encouragement, this thesis work would never have been completed. He has been a wonderful influence and a guide in my research life. From working with him, I tasted the joy of serious research work and developed strong interest in solving challenging graphics problems. I am grateful to him for his kindness and unreservedly sharing his brilliant insights with me.

I am also deeply indebted to my colleague, Anil Hirani. Talking with him always gives me some strength and courage to face some mathematical challenges I have encountered, and helps me to re-capture my self-confidence. Besides that, his broad and deep mathematical knowledge is always a great help on my way to completion.

I also wish to thank my other committee members, Peter Schröder, whose consistent confidence and support are always a great cure for my doubts. Many thanks to him for a lot of new and related reference he provided, which not only broadened my view, but also are a great help to my work. Besides, it is always a pleasure to exchange personal feelings with him. Al Barr, who generously offered his time and expertise to solve my problem, proof-read my papers and give me very constructive suggestions.

Special thanks to my officemate Mark Meyer for stimulating discussions and offering detailed comments, as well as nice hugs and talking during my tough time. Thanks to Mathieu Desbrun for both technical and casual advises on finishing PhD and pursuing personal careers. Thanks to Eitan Grinspun for proofreading my papers. Thanks to Dave Felt for his help with my videotape. Thanks also to my women friends at Caltech, Zoë

Wood and Eve Schooler, who taught me to be robust and brave in an intense research environment. I am also very grateful for various help and support from our secretaries: Louise Foucher and Jeri Chittum.

The person to whom I am most indebted is my dear husband, Hui, whose love and understanding are truly without bound. No matter what happened, he always provides me a cozy harbor to cry, yell, calm down and then go on. It is difficult to imagine how I could have completed this work without his support at every step along the way.

The work on path perturbation was supported by the Army Research Office Young Investigator Program (DAAH04-96-100077), the NSF Science and Technology Center for Computer Graphics and Scientific Visualization, and the Alfred P. Sloan Foundation. The research on non-uniform luminaires was supported by a Microsoft Research Fellowship and NSF Career Award CCR9876332.

Mathematical Methods for Image Synthesis

by

Min Chen

In Partial Fulfillment of the

Requirements for the Degree of

Doctor of Science

Abstract

This thesis presents the application of some advanced mathematical methods to image synthesis. The mainstream of our work is to formulate and analyze some rendering problems in terms of mathematical concepts, and develop some new mathematical machineries to pursue analytical solutions or nearly analytical approximations to them. An enhanced Taylor expansion formula is derived for the perturbation of a *general* ray-traced path and new theoretical results are presented for *spatially-varying luminaires*. On top of them, new deterministic algorithms are presented for simulating direct lighting and other scattering effects involving a wide range of non-diffuse surfaces and spatially-varying luminaires. Our work greatly extends the repertoire of non-Lambertian effects that can be handled in a deterministic fashion.

First, my previous work on "Perturbation Methods for Image Synthesis" is extended here in several ways: 1) I propose a coherent framework using closed-form path Jacobians and path Hessians to perturb a general ray-traced path involving both specular reflections and refractions, and an algorithm similar to that used for interactive specular reflections is employed to simulate lens effects. 2) The original path Jacobian formula is simplified by means of matrix manipulations. 3) Path Jacobians and Hessians are extended to parametric surfaces which may not have an implicit definition. 4) Theoretical comparisons and connections are made with related work including *pencil tracing* and *ray differentials*. 5) Identify potential applications of perturbation methods of this nature in rendering and computer vision.

Next, a closed-form solution is derived for the irradiance at a point on a surface due to an arbitrary polygonal Lambertian luminaire with linearly-varying radiant exitance. The solution consists of elementary functions and a single well-behaved special function known as the Clausen integral. The expression is derived from the Taylor expansion and a recurrence formula derived for an extension of double-axis moments, and then verified by Stokes' theorem and Monte Carlo simulation. The study of linearly-varying luminaires introduces much of the machinery needed to derive closed-form solutions for the general case of luminaires with radiance distributions that vary polynomially in both position and direction.

Finally, the concept of irradiance tensors is generalized to account for inhomogeneous radiant exitance distributions from luminaires. These tensors are comprised of scalar elements that consist of *constrained* rational polynomials integrated over regions of the sphere, which arise frequently in simulating some rendering effects due to polynomially-varying luminaires. Several recurrence relations are derived for generalized irradiance tensors and their scalar elements, which reduce the surface integrals associated with spatially-varying luminaires to one-dimensional boundary integrals, leading to closed-form solutions in polyhedral environments. These formulas extend the range of illumination and non-Lambertian effects that can be computed deterministically, which includes illumination from polynomially-varying luminaires, reflections from and transmissions through glossy surfaces due to these emitters. Particularly, we have derived a general tensor formula for the irradiance due to a luminaire whose radiant exitance varies according to a monomial of any order, which subsumes Lambert's formula and expresses the solution for higher order monomials in terms of those for lower-order cases.

Table of Contents

List of Figures

Nomenclature

Symbol	*Page*	*Meaning*
$\mathbf{q}(q_1, q_2, q_3)$	16	The fixed endpoint of a path
$\mathbf{p}(p_1, p_2, p_3)$	16	The varying endpoint of a path
$\mathbf{p}'$	16	The new position of the varying endpoint
$\mathbf{x}(x_1, x_2, x_3)$	16	The bouncing point of a path
$\mathbf{x}'$	16	The new bouncing point while perturbing $\mathbf{p}$ to $\mathbf{p}'$
Ψ	16	Path function mapping $\mathbf{p}$ to the bouncing points, $\mathbb{R}^3 \to \mathbb{R}^3$
$\Delta\mathbf{p}$	17	Perturbation vector from $\mathbf{p}$, i.e, $\mathbf{p}' - \mathbf{p}$
$\mathbf{J}$	17	Path Jacobian matrix
$\mathbf{H}$	18	Path Hessian tensor, consisting of three matrices $\mathbf{H}^1, \mathbf{H}^2, \mathbf{H}^3$
$\mathbf{G}$	22	Surfaces encountered in a ray-traced path
g	22	Implicit function defining a surface
η	18	Refractive index of the medium
d	19	Optical path length
∇	23	Gradient operator
F_{η_1, g, η_2}	23	The non-linear system from Lagrange multiplier method
f	24	Explicit function ($\mathbb{R}^3 \to \mathbb{R}^4$) determined from $F_{\eta_1, g, \eta_2} = \mathbf{0}$
sub	24	Operator to drop the last row of a 4×3 matrix
aug	29	Append a 4×3 matrix by a zero column
$\mathbf{J}^*$	26	Bounce Jacobian in a multiple-bounce path
$\mathbf{H}^*$	31	Bounce Hessian in a multiple-bounce path
$\hat{F}$	33	The non-linear system formed by the first three equations of F
ξ	41	Image plane point in $\mathbb{R}^3$
$\hat{\xi}$	41	Image plane pixel, $\mathbb{R}^2$

Symbol	*Page*	*Meaning*
L	61	Planar luminaire
$\mathbf{x}$	61	Point on the luminaire
$\phi(\mathbf{x})$	61	Radiant exitance distribution function of a luminaire
Φ	61	Irradiance
$\mathbf{o}$	61	The receiver point at the origin
$\mathbf{q}$	63	Arbitrary receiver point
A	61	Regions on the sphere
$\mathcal{S}^2$	68	Unit sphere
$\mathbf{u}$	62	Unit vector in $\mathcal{S}^2$
$\mathbf{r}$	68	A point in $\mathbb{R}^3$
r	86	The Euclidean length of the vector $\mathbf{r}$
σ	62	Measurement for area on the sphere
$\mathbf{w}$	62	The unit vector perpendicular to the luminaire
h	62	Distance from the receiver point to the luminaire
$\mathbf{a}$	63	Vector encoding the linear variation of L
$\mathbf{b}$	63	Receiver normal
$\langle \cdot, \cdot \rangle$	63	Standard inner product
$f(\mathbf{r}, \mathbf{u})$	68	Monochromatic radiance function
$f(\mathbf{u})$	68	Monochromatic radiance distribution at a point
$\mathbf{T}^n(A)$	69	The nth order irradiance tensor of A
δ_{ij}	69	Kronecker delta
ε_{ijk}	66	Permutation symbol
∂A	69	Boundary curve of a spherical region A
$\mathbf{n}$	69	Outward-pointing normal to the boundary curve on the sphere
I	70	The multi-index $(i_1, i_2, \ldots, i_n)$
$I \backslash k$	70	The multi-index $(i_1, cdots, i_{k-1}, i_{k+1}, \cdots, i_n)$

Symbol	*Page*	*Meaning*
$\bar{\tau}$	70	nth order axial moment
$\bar{\bar{\tau}}$	70	Double axis moment
$\bar{\bar{\bar{\tau}}}$	75	Triple axis moment
$\mathbf{M}$	84	The tensor quantity in the irradiance formula
η	84	The logarithmic function $\ln \langle \mathbf{w}, \mathbf{u} \rangle / (1 - \langle \mathbf{w}, \mathbf{u} \rangle^2)$
$\mathbf{\Phi}$	85	Vector-valued irradiance
P	93	A polygonal luminaire
ζ	93	An edge of a spherical polygon
$B^{m,l}$	123	Line integral defined in (4.22)
B^*	95	Line integral of η
$\tilde{B}$	133	A special class of $B^{m,l}$ with repeated numerator axes
Λ	96	The two-parameter special function given in (3.115)
$\mathrm{Li}_2(z)$	99	Dilogarithm
$\mathrm{Cl}_2(x)$	99	Clausen's integral
Υ	99	The two-parameter special function given in (3.122)
${}_pF_q$	104	Hypergeometric function
$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$	111	An orthogonal coordinate system at the origin
$\mathbf{T}^{n;q}$	112	Generalized irradiance tensors
$\tau^{n;q}$	119	Rational angular moments
T	128	Time complexity
$\tilde{\tau}$	130	Moments corresponding to Phong distributions
F	133	the integral of power of cosine
Φ_n	152	the irradiance due to a monomially-varying lumiaires of nth order

Chapter 1

Introduction

A major problem in computer graphics is *physically-based rendering*, which is concerned with creating physically accurate synthetic images of a given model by simulating the light transport process inside the environment. The input of the rendering algorithm is a description of a scene by way of its geometry and the optical characteristics of objects. In physically-based rendering, one is interested in the light and complex interreflection of light with the environment, known as *global illumination*, which is governed by the rendering equation [62]. This Fredholm integral equation of the second kind determines the radiance function contributing to the perceived image. While the global illumination problem is easy to formulate mathematically, it is hard to simulate due to the computational cost.

The broad application of image synthesis in various fields, including mechanical engineering, entertainment industry and illumination engineering, etc., places two contradictory requirements on global illumination algorithms. On the one hand, the rendering process should be as fast as possible, allowing users to explore the illuminated scene and perceive that at least some degree of global illumination effects get updated interactively. On the other hand an increasing number of applications demand a high degree of realism, such as the movie business, lighting design, scientific visualization, etc. Ideally, the images should truly exhibit global illumination effects, such as lighting, shadows, specular and glossy reflections and refractions, etc, which are a computationally expensive

procedures. Because these two goals conflict with each other, all rendering algorithms have to make some tradeoff, in one way or another. Research on image synthesis seeks to push the boundary between speed and realism.

This thesis addresses both aspects of image synthesis. We will stress the use of advanced mathematical methods in formulating some rendering problems and developing deterministic techniques. The well-founded mathematical results not only serve as an initial step to the scientific and predictive approach to the problems, but also give rise to better numerical approximations feasible in practice.

1.1 Global Illumination

The limitations of current global illumination algorithms stem from the computational aspects of following physical laws of transport. In studying the computational aspects of global illumination, it is often convenient to partition illumination into two components: *direct* and *indirect* [8]. By direct illumination we mean the process of light emitted from the luminaire (area light source), propagated through the scene, and subsequently scattered at a second surface. By indirect illumination we mean light that undergoes multiple scattering and interactions with the environment. We shall address these two issues in the next sections.

1.1.1 Basic Techniques

Both direct and indirect illumination require us to faithfully represent the distribution of light scattering from real surface materials, which is characterized in computer graphics by various reflection models. Due to computational constrains, global illumination algorithms frequently idealize the surface materials as Lambertian (ideal diffuse), ideal specular or refractive, or a combination of the two. The assumptions of these two extremes lead respectively to two well-understood global illumination procedures: *ray tracing* and *radiosity*.

The first example of ray tracing in computer graphics is due to Appel [6] and then significantly extended by Whitted [119]. Ray tracing succeeds in producing brilliant

specular reflection and transparent effects as well as caustics and shadows. Two fundamental operations are involved in ray tracing: ray casting, which is used to follow optical paths through a simulated environment, and shading, which is applied at the points where rays intersect objects. Of the two operations, generating the optical paths is by far the most expensive and typically the chief obstacle to interactive ray tracing. Consequently, finding some means of quickly generating or reusing ray paths becomes a critical challenge for interactive ray tracing.

The radiosity method, first applied to image synthesis by Goral et al. [47, 33], is a global illumination model of choice for polygonal, diffuse environments. It accommodates multiple diffuse interreflections using a *finite element method* and is well suited for subtle color bleeding and penumbra effects. Furthermore, its view-independent solutions make it ideal for architectural walk-through. In general, three components are involved in a finite element approach [8]: 1) discretization [56, 53], 2) computing element interactions, and 3) solving a linear system. Among them, computing element interactions is perhaps the most costly component, involving the computation of a large number of multi-variable integrals, called *form factors*, over irregular domains. These integrals may be evaluated in closed form under certain configurations and piecewise diffuse environments [58, 97]. However, there are virtually no formulas available for analytical integration when non-diffuse emission or reflection is involved, where Monte Carlo approximation [32, 120, 72, 111] is still the basic technique.

Both ray tracing and radiosity only partially simulate the indirect illumination process occurring in a complex scene. Classical ray tracing simplifies this process by neglecting all multiple scatterings except those from mirror-like surfaces, while radiosity does not accommodate mirror surfaces. To account for more complex surface reflections such as glossy surfaces [102], Monte Carlo methods have been applied in combination with ray tracing and radiosity to simulate a variety of global illumination effects [62, 36, 35, 101, 112].

1.1.2 Direct Illumination and Light Source Description

Although indirect illumination is important to consider for realistically rendering a complex environment, such as a room, direct illumination is a prerequisite for indirect illumination. For example, direct illumination plays an important role in computing element interaction in finite element methods mentioned above. A challenging topic in direct illumination is to model the distribution of light emitted from luminaires [115, 8, 40, 38, 42, 39] and simulate direct lighting and scattering effects due to various light emitters (maybe partially occluded), including computing irradiance and its features, modeling surface reflections, computing reflected radiance and subsurface scattering [52, 86, 60], and so on. Handling these effects is also necessary for providing artisticly rich and subtle illumination effects in lighting design [44, 43].

To correctly describe the physical characteristic of light sources, three attributes must be modeled [115]:

Geometry: Light sources have well-defined, finite geometries that greatly affect the emission distribution. Sources may vary dramatically in shape and size, ranging from the sun and the skylight to a window pane and a pixel element. The physical geometry of light sources can be easily depicted by any standard 3D modeling technique. However, from emission geometry, we may categorize light sources as three types: *point* sources (zero dimensional) [87, 82], *linear* or *curve* sources (one dimensional) [82, 90, 88, 14, 15] and *area* sources (two dimensional) [83, 110, 80, 81]. Among them, point and polygonal light sources are the most common ones encountered in graphics applications.

Spectral distribution: The emission, reflection and transmission characteristics of light sources are wavelength-dependent functions. Similarly, our human perceptual response is also dependent on wavelength, sensitive to stimuli over the range of the visible spectrum. To avoid aliasing artifacts due to calculation based on RGB colorimetry, we sometimes consider the spectral distribution of the light source over the visible wavelength range.

Radiant exitance function: Perhaps the most important characteristic of a luminaire

that must be included to accurately model its emission distribution is the *radiant exitance function*. Although the rendering algorithms developed so far have been primarily restricted to *uniform* luminaires, most luminaires do not emit light equally in all directions or positions. These *non-uniform* luminaires may be specified by an exitance function $\phi(\mathbf{x}, \mathbf{u})$ defined at all points $\mathbf{x}$ and in all directions $\mathbf{u}$, which may either vary with respect to the position $\mathbf{x}$ or the direction $\mathbf{u}$. Some literature in illumination engineering characterizes these two types of variations respectively as *non-diffuse* (or non-Lambertian) and *inhomogeneous*, as shown in Figure 1.1a and Figure 1.1b. More intuitively, we refer to them as *directionally-varying* and *spatially-varying* emitters, respectively. For directional point light sources, this radiant exitance function simplifies to a *luminous intensity distribution*, which may be measured by a goniometer and represented by goniometric diagrams [115, 42].

In the remainder of this thesis, we shall direct our attention to polygonal *diffuse* and *gray* luminaires whose radiant exitance is directionally uniform and varies spatially with respect to the position over the emitter plane. By diffuse (or Lambertian) we mean that it is directionally uniform. The notion of "gray" removes the dependency on wavelength. Spatially-varying luminaires constitute an important class of light sources with immediate application in higher-order finite element methods for global illumination, both for direct lighting [57] and final gathers from coarse global solutions [93, 103, 19]. Few tools currently exist for handling this type of luminaire aside from Monte Carlo or numerical quadrature [39].

1.2 Prior Art

1.2.1 Path Reuse in Ray Tracing

With the invention of ray tracing, many acceleration techniques were developed. A good summary of these techniques may be found in Arvo and Kirk [10]. To date, most research has focused on speeding up the process of ray casting. The rationale for considering path reuse instead is that updating a path can be far less expensive than recomputing

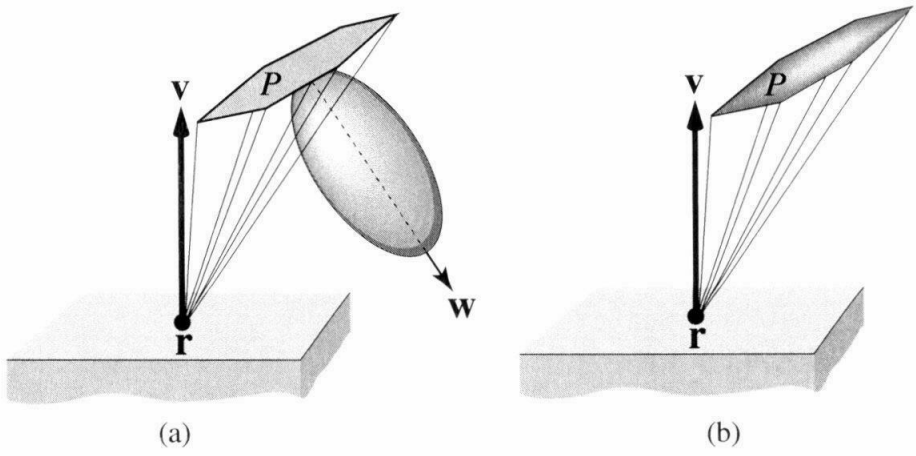


Figure 1.1: *Two types of non-uniform luminaires: (a) Directionally-varying luminaires. (b) Spatially-varying luminaires.*

it. Moreover, this approach leads to new algorithms that exploit coherence both within a single image and among similar images.

Many aspects of path reuse have been investigated for interactive rendering. Typically, a ray tree or equivalent data structure is used to retain all ray-object intersections computed during the rendering of each pixel, then used in the generation of a nearby image. Cook [34], and Séquin and Smyrl [99] handled material changes in this way by re-traversing the retained ray trees with modified parameters. Murakami and Hirota [77], and Jevans [61] handled changes in geometry by indexing each ray path with the list of voxels it traverses; any change to the scene affects some subset of the voxels, which in turn determines the potentially affected rays. Briere [23] employed a ray tree and a color tree to preserve the path information and shading expressions separately, and proposed a novel data structure to efficiently detect and recompute the exact portion of the image that has changed after an arbitrary manipulation of a scene. Although these techniques are very effective for geometric and material changes, they are ineffective for viewpoint motion, where even a small eye movement may change the ray paths associated with most pixels. To compensate for the viewpoint movement, Badt [13], and Adelson and Hodges [2] reprojected the old intersections to the new view position by

applying 3D warping, which exploits perspective coherence. This 3D warp amounts to an image space reprojection (2D warp) in stereoscopic ray tracing [3]. Unfortunately, reprojection only works for the first level rays and is applicable only to diffuse scenes. Chapman et al. [25, 26] computed "continuous intersection" information of rays with the scene through time by using the trajectory of the viewpoint through the scene, but their method is restricted to a polygonal scene model and predefined viewpoint paths.

Some researchers have attempted to exploit the path coherence in ray tracing analytically. Shinya et al [100] employed paraxial ray theory to linearly approximate the propagation of paraxial rays by a system matrix, called *pencil tracing*. Igehy [59] computed ray differentials, partial derivatives of a ray with respect to the position on the image plane to estimate the footprint of a ray. These ray differentials give an idea of the distance to neighboring rays, which proves to be very effective for anti-aliasing texture by filtering locally over the footprint.

1.2.2 Form Factors

The computation of radiant energy transfers is an essential component of physically-based rendering algorithms, both for local and global illumination. In particular, the determination of radiant interchange among discrete surface elements arises frequently in finite element methods for global illumination, and in simulating direct illumination effects [9, 54, 107]. Consideration has been given to finite, *diffuse* and *gray* surfaces with uniform radiation properties, where the *form factor*, alternatively known as the *angle factor*, or *configuration factor*, is used to characterize such radiant transfer [46, 105]. It represents the fraction of energy that leaves element i and arrives directly at element j. In calculating these form factors, it has been standard practice to make two basic assumptions about the participating surface: (1) The angular distribution of the radiant flux leaving it is *diffuse*; (2) The magnitude of the radiant flux is *uniform* over all parts of the surface. We refer to them as *homogeneous Lambertian surfaces*. The essential simplification associated with these assumptions is that the form factor is purely geometrical, independent of surface properties, leading to an assortment of analytic solutions for some special geometric configurations [58], including Lambert's

formula [18, 97] and the form factor between two polygons [97] well known in the graphics community.

The irradiance at a point on a surface due to a polygonal luminaire of uniform brightness can be computed from a formula attributed to Lambert. Suppose P is a simple planar polygon in $\mathbb{R}^3$ with vertices $\mathbf{v}_1, \ldots, \mathbf{v}_n$. If P is a diffuse light source with constant radiant exitance M, then the irradiance at a point $\mathbf{q}$ due to P is given by

$$\Phi(\mathbf{q}) = \frac{M}{2\pi} \sum_{i=1}^{n} \beta_i \cos \eta_i,$$

where β_i is the angle subtended by the edge $\mathbf{v}_i\mathbf{v}_{i+1}$ as seen from the receiver point $\mathbf{q}$, and η_i is the angle between the plane containing $\mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{q}$ and the surface normal at $\mathbf{q}$.

The form factor between two arbitrary diffuse polygons can be computed exactly using a formula due to Schröder and Hanrahan [97]. One complication that arises in computing this polygon-to-polygon form factor is the appearance of a special function studied by Powell in 1943, and subsequently named the *dilogarithm* [67]. Powell showed that it is equivalent to the *Airey radiation integrals* that arise in both astrophysics and quantum mechanics.

1.2.3 Directional and Spatial Luminaires

Methods for simulating the illumination due to diffuse area sources [84] and directional point sources [115, 42] are well-known. To represent an arbitrary luminous intensity distribution of a point light source, Dobashi et al. [42] decomposed it as combinations of basis functions, which were chosen to be spherical harmonics [76]. By calculating in advance the intensity impinging on each surface lit by the luminaires whose luminous intensity distributions are the same as spherical harmonic functions, they proposed a quick rendering method for interactive lighting design. By contrast, non-uniform area sources are problematic for deterministic methods. Handling the rendering effects involving this type of emitters still relies heavily on Monte Carlo method or numerical quadrature [42, 40, 38, 39].

Arvo [8, 9] presented the first direct method to compute the irradiance at a point due

to a class of directional luminaires whose radiant exitance varies directionally according to a Phong distribution, that is, as a cosine to a power [87], shown in Figure 1.1a. By formulating this illumination problem as a double-axis moment, the irradiance due to this type of polygonal directional emitters may be computed in closed form. In addition, Arvo predicted that the irradiance integrals due to spatially-varying area light sources generally cannot be evaluated in terms of elementary functions only and must depend on at least a special function.

By applying Stokes' theorem to surface integrals formulated in a spherical coordinate system, DiLaura et al. derived a single contour integral expression and a double contour integral expression for calculating the irradiance received at a point [40] and an area [38] due to an arbitrary directional luminaire. Furthermore, they extended the similar procedure from non-diffuse homogeneous area sources to inhomogeneous sources, attaining a similar boundary integral formula for the irradiance at a point due to spatially-varying luminaires [39]. Their new formulations reduce the original two-dimensional area integrals to one-dimensional integrals, whose evaluations, however, still resort to numerical quadratures. With a similar technique but applied in a Cartesian coordinate system, Holzschuch and Sillion [57] were only able to decompose the irradiance integral at a point due to a linearly-varying emitter into boundary integral terms and a simpler surface integral, which were then evaluated using numerical approximation. Besides these analytical efforts, Troutman and Max [72, 111] used the hemicube algorithm [32] to approximate generalized form factors involving surface elements with linear radiosity variations as summations of vertex-to-vertex form factors, taking advantage of hardware shading features to do the necessary interpolation.

1.3 Contributions

In this section the original contributions of this thesis, both theoretical and practical, are sketched in an abbreviated, informal way.

Derivatives of Ray-Traced Paths Closed-form expressions are derived for the *path Jacobian* (the first-order derivative) and *path Hessian* (the second-order derivative)

that hold for a general ray-traced path mixing both specular reflections and refractions. The perturbation formula based on these expressions allows us to generate a family of closely-related optical paths by expanding the given path into a high-dimensional Taylor series, which may find many potential applications, such as sterescopic ray tracing [3], image-based rendering of non-diffuse scenes [68], accelerating caustic generations, guiding importance sampling and improving existing Monte Carlo methods [112], and so on.

Generalized Irradiance Tensors A natural generalization of irradiance tensors is presented to account for the polynomial spatial variation of the emission distribution. Several useful formulas are derived from generalized Stokes' theorem for these new tensors, which provides a powerful tool to compute different moments of the radiance distribution function that are expressed as a family of rational polynomials over the sphere.

Extended Lambert's Formula The generalized irradiance tensors associated with the irradiance due to a monomially-varying luminaire are shown to satisfy a recurrence relation with respect to the order of the monomial function, which gives rise to a natural extension of Lambert's formula for uniform luminaires. The new formula relates the solution to higher-order polynomials to those for lower-order cases, and is verified independently by comparison with Monte Carlo and Finite Element methods.

Closed-Form Solutions for Linearly- and Quadratically-Varying Luminaires Closed-form expressions for linearly-varying and quadratically-varying luminaires follow easily from the extended Lambert's formula. An analytical algorithm based on them is developed for simulating direct lighting effects due to such emitters.

A New Class of Closed-Form Expressions From recurrence formulas derived for generalized irradiance tensors, we present several expressions for moments of the radiance distribution from polynomially-varying luminaires. These expressions may be evaluated in closed form for polyhedral environments, except for a single special function that can be expressed in terms of the Clausen integral [67]. We also show

that this special function is inevitable for such irradiance problems, consistent with Arvo's prediction [8] and the form factor formula between two polygons.

Rational Moments for Non-Lambertian Simulations Closed-form solutions are presented for moments of the radiance from polygonal luminaires with polynomially-varying radiant exitance, along with efficient algorithms for the evaluation of a subclass of these rational moments corresponding to Phong distributions. These expressions and procedures are demonstrated by simulating some non-Lambertian phenomena involving polynomially-varying luminaires and non-diffuse surfaces, such as view-dependent glossy reflections and glossy transmissions. Our algorithms are validated by corresponding Monte Carlo simulations.

1.4 Thesis Overview

The remaining chapters of this dissertation are organized as follows:

Chapter 2 builds upon our previous work on perturbing a specular reflection path and develops a coherent mathematical framework for a general ray-traced path mixing both reflections and refractions. Moreover, we simplify the path Jacobian formula using matrix operations and propose an idea to extend the results to other surface representations, such as parametric surfaces. Finally, the concept of path Jacobian is compared with other linear approximations.

Chapter 3 serves as the first step to the main topic of the thesis. A closed-form solution for the irradiance at a point due to a linearly-varying luminaire is derived using Taylor expansion and *triple-axis moments*, a natural generalization of double-axis moments. Although the same result is derived in a much simpler way in Chapter 5, the machinery used in this chapter is the initial motivation for the mathematical tools presented in Chapter 4. Most importantly, a two-parameter special function appears for the first time. This chapter shows its connection with the Clausen integral and discusses various evaluation methods in practice.

Chapter 4 introduces the most powerful mathematical tools of this thesis. By char-

acterizing the radiance distribution caused by a polynomially-varying luminaire, we propose a natural generalization of irradiance tensors. The expressions derived in this chapter for these new tensors and associated rational moments lead to a number of new techniques for image synthesis involving spatially-varying luminaires. Efficient algorithms are also presented for evaluating a subclass of rational polynomials integrated over the sphere.

Chapter 5 demonstrates the applications of generalized irradiance tensors in image synthesis, which include computing the irradiance due to a polynomially-varying luminaire, simulating non-Lambertian scattering effects involving spatial luminaires and non-diffuse surfaces, combining directional variations as well as spatial variations. Particularly for the irradiance problem, we have presented an extended Lambert's formula for any nth-order monomially-varying luminaire, which is independently verified by comparing our analytical results with Monte Carlo estimates and Finite Element solutions.

Chapter 6 summarizes the thesis work and indicates possible future research directions.

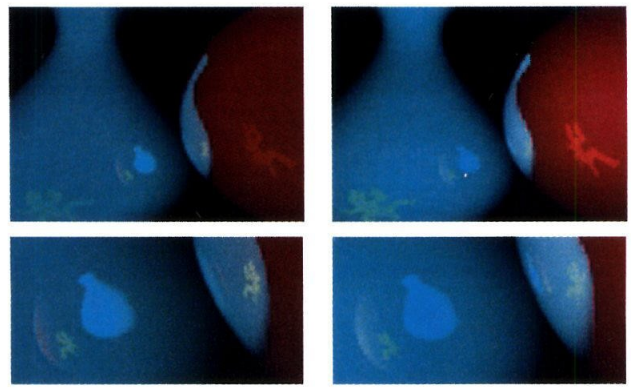
Appendix attaches some additional proofs and derivations for the formulas appearing in this thesis.

Chapter 2

Perturbation of a Ray-Traced Path

In previous work [27], we applied perturbation methods to the problem of computing specular reflections in arbitrary implicitly-defined curved surfaces. Our path perturbation technique is motivated by the observation that specular reflections in smooth surfaces vary continuously as a function of object position (see Figure 2.2a), except when boundary or visibility conditions intervene. The key idea is to generate families of closely related reflection paths by expanding a given path into a high-dimensional Taylor series. The result is an analytical perturbation formula that holds for a general multiple-bounce reflection path, which is based on closed-form expressions for linear and higher-order approximations of the ray path, called path Jacobians and path Hessians, respectively. This perturbation formula provides a mathematical foundation for exploiting path coherence and reuse in ray tracing acceleration techniques and incremental rendering [28]. As illustrated by the side-by-side comparison in Figure 2.1, our perturbation approach gains a great speedup over conventional ray tracing in approximating specular reflections on curved surfaces, with little sacrifice in accuracy. More importantly, by taking advantage of coherent reflections from frame to frame, the perturbed reflections can be updated interactively in less than one second while moving the diffuse object around.

The present work in this chapter enriches this path perturbation theory from the



Perturbation (9.6s/0.7s)

Ray Tracing (67s/67s)

Figure 2.1: *Side-by-side comparison of multiple-bounce reflection images generated by the perturbation method (left) and ray tracing (right). The closeup images on the bottom show that the result from path perturbation is almost indistinguishable from that generated by ray tracing, but with a great speedup (9.6s vs. 67s). However, by exploiting path coherence from frame to frame, the perturbed reflections can be updated in less than one second (0.7s) while moving the lizard around.*

following theoretical aspects:

1. Present a unified mathematical framework to derive closed-form expressions for path Jacobians and path Hessians that hold for a general ray-traced path involving both specular reflections and refractions.

2. Simplify the old path Jacobian formula by means of matrix operations.

3. Extend the path perturbation theory from implicit surfaces to other representations such as parametric surfaces.

4. Compare our path Jacobian analysis with other linear approximations explored thus far, which include *system matrix* from pencil tracing [100] and *ray differentials* [59].

While there are direct applications of this extended work, such as the simulation of lens

effects demonstrated in section 2.6, it is mainly of theoretical importance as it provides a complete perturbation formula in the ray-tracing context and clarifies possible confusions from different linear approximations. We shall elaborate these generalizations in the remainder of this chapter. In the end, we will discuss some potential applications of this mathematical tool.

2.1 Related Work

In essence, our approach of path perturbation aims at computing the first- and second-order derivatives of the intersection points along a path. This differential point of view has been pursued by many researchers. As a basic mathematical tool to characterize function differentiability, various kinds of derivatives have been studied in computer graphics in order to accurately, quantitatively model features such as discontinuity, variations, etc. Derivative information has played an important role in mesh construction, guiding importance sampling, improving function interpolation and texture filtering.

In previous work, Heckbert [56] and Lischinski et al. [69] identified derivative discontinuities in irradiance and used them to construct effective surface meshes. The irradiance gradient has been used to get higher order approximations of irradiance functions. For example, Ward et al. [118] and Vedel [113, 114] estimated irradiance gradients by Monte Carlo path tracing and used them to improve the interpolation accuracy. Salesin et al. [94] and Bastos et al. [17] employed gradients to construct higher-order interpolants for irradiance functions. Drettakis et al. [45] estimated gradients as well as isolux contours from a collection of discrete samples and used them to guide subsequent sampling, placing more samples where the curvature of the isolux contours was large. Arvo [7] extended the derivative to a vector function and derived the closed-form expression for the *irradiance Jacobian*, defined as the derivative of the vector representation of irradiance with respect to the receiver point. The expression holds for any number of polyhedral blockers, and was used for direct computation of isolux contours, finding local irradiance extrema and iso-meshing. A similar idea was also employed by him to generate the iso-contour in volume rendering [11]. Holzschuch and Sillion [57] computed the radiosity

gradient for constant and linear emitters to control the discretization error in a radiosity calculation. Igehy [59] differentiated a ray-traced path with respect to the image plane position to estimate the footprint of a ray for texture filtering.

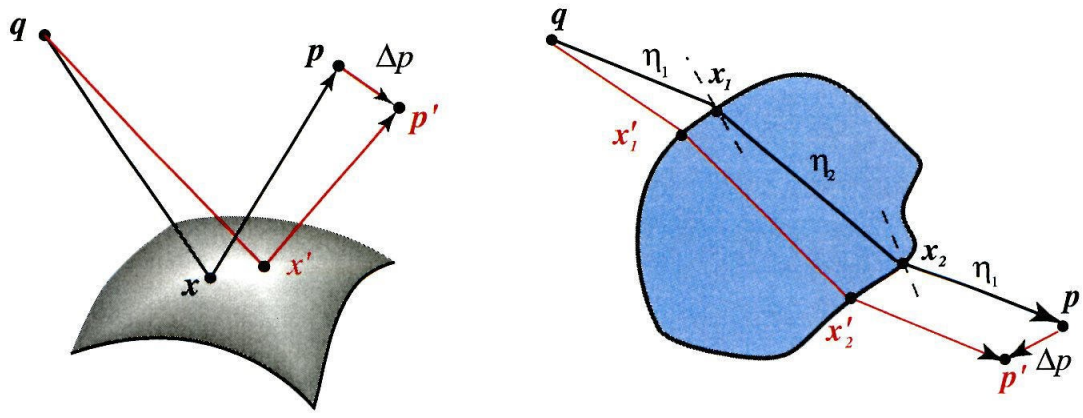
2.2 Path Perturbation Theory

In this section, we shall first review the idea of applying perturbation method to a ray path and then present a perturbation formula for a general ray-traced path involving both specular reflections and refractions. Similar to the case of reflection paths, this formula is also based on closed-form expressions for path Jacobians and path Hessians, which are, however, derived from a more general formulation with the index of refraction taken into consideration, using a procedure quite similar to that performed for a specular reflection path. For the sake of self-containedness, we present the complete derivation here, which may have some overlaps with the previous report [27], but in a more general sense.

2.2.1 Path Coherence and Perturbation Formula

Reflections and refractions tend to have a great degree of coherence; in general, as two objects $\mathbf{p}$ and $\mathbf{p}'$ grow nearer to each other, so will their reflections or refractions $\mathbf{x}$ and $\mathbf{x}'$, as shown respectively in Figure 2.2a and Figure 2.2b. Here coherence means local smoothness of specular reflections or refractions. That is, given a known path from a fixed point $\mathbf{q}$ to $\mathbf{p}$ via a smooth reflecting (or refracting) surface, it is likely that the bouncing point $\mathbf{x}$ will vary continuously as a function of the position of $\mathbf{p}$, that is, $\mathbf{x} = \Psi(\mathbf{p})$, where the smooth function $\Psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is called the *path function*. The essence of path perturbation is to approximate families of closely related ray-traced paths using a high-dimensional Taylor expansion of Ψ around a given path. In this way, the reflection or refraction from $\mathbf{q}$ to a new point $\mathbf{p}'$ in the neighborhood of $\mathbf{p}$ can be obtained by perturbing $\mathbf{x}$ analytically.

For a vector-valued function Ψ with three components Ψ_1, Ψ_2 and Ψ_3, the second-



(a) (b)

Figure 2.2: *Reflection (left) and refraction (right) tend to have a great degree of coherence, the corresponding reflection (or refraction) points* $\mathbf{x}$ *and* $\mathbf{x}'$ *will converge as* $\mathbf{p}$ *approaches* $\mathbf{p}'$.

order Taylor expansion can be expressed as:

$$\Psi(\mathbf{p}+\Delta\mathbf{p}) = \Psi(\mathbf{p}) + \mathbf{J}\Delta\mathbf{p} + \frac{1}{2}\begin{bmatrix} \Delta\mathbf{p}^{\mathrm{T}}\mathbf{H}^1\Delta\mathbf{p} \\ \Delta\mathbf{p}^{\mathrm{T}}\mathbf{H}^2\Delta\mathbf{p} \\ \Delta\mathbf{p}^{\mathrm{T}}\mathbf{H}^3\Delta\mathbf{p} \end{bmatrix} + \mathrm{O}(||\,\Delta\mathbf{p}\,||^3), \tag{2.1}$$

which is the *perturbation formula* to update a given path through $\mathbf{p}$ to a new path reaching the neighboring point $\mathbf{p}+\Delta\mathbf{p}$. Here, the first-order derivative $\mathbf{J}$ is a 3×3 matrix defined by

$$\mathbf{J} \equiv \frac{\partial\Psi(\mathbf{p})}{\partial\mathbf{p}} = \begin{bmatrix} \Psi_{1,1} & \Psi_{1,2} & \Psi_{1,3} \\ \Psi_{2,1} & \Psi_{2,2} & \Psi_{2,3} \\ \Psi_{3,1} & \Psi_{3,2} & \Psi_{3,3} \end{bmatrix}, \tag{2.2}$$

where $\Psi_{i,j} \equiv \partial\Psi_i/\partial\mathbf{p}_j$. The second-order derivative $\mathbf{H}$ is a third-order tensor, which consists of three Hessian matrices $\mathbf{H}^1, \mathbf{H}^2, \mathbf{H}^3$ corresponding respectively to coordinates

x, y and z, given by

$$\mathbf{H}^i \equiv \frac{\partial^2 \Psi_i(\mathbf{p})}{\partial \mathbf{p}^2} = \begin{bmatrix} \Psi_{i,11} & \Psi_{i,12} & \Psi_{i,13} \\ \Psi_{i,21} & \Psi_{i,22} & \Psi_{i,23} \\ \Psi_{i,31} & \Psi_{i,32} & \Psi_{i,33} \end{bmatrix} \tag{2.3}$$

for $i = 1, 2, 3$, where $\Psi_{i,jk} \equiv \partial^2 \Psi_i / \partial \mathbf{p}_j \partial \mathbf{p}_k$. We call $\mathbf{J}$ and $\mathbf{H}$ the *path Jacobian* and the *path Hessian*, which provide the first- and second-order approximations to the path function, respectively. We shall derive closed-form expressions for them next.

2.2.2 The Path Jacobian

In this section we shall first derive the expression for the path Jacobian for a single-bounce path, and then extend the result to a multiple-bounce case. The term "bounce" here means one interaction of a ray hitting a surface. Depending on the property of the surface encountered, this event may be reflection from or refraction through the surface, as demonstrated in Figure 2.3.

Preliminaries

The closed-form formula for the path Jacobian is derived from the *Fermat principle* [22], *Lagrange Multiplier Theorem* [70], and the *Implicit Function Theorem* [71], which will be briefly reviewed next.

Fermat's Principle: In geometrical optics, the propagation of the light obeys *Fermat's principle*, also known as the principle of the *shortest optical path*. This principle asserts that the optical length of an actual light ray between any two points P_1 and P_2 is a local extremum among all paths between these points, within a small neighborhood [22, pp. 128–129]. The *optical length* from one point P_1 on a ray to another point P_2 is defined as the geometric path length weighted by the refractive index of the medium, η [22]:

$$\int_\gamma \eta \, ds, \tag{2.4}$$

where $\gamma(s)$ is the parametric path from P_1 to P_2 and s is arc length. In a ray tracing setting where regions of constant refractive index are separated by smooth boundaries, rays will travel along piecewise straight paths, reflecting or refracting at boundaries. The optical length function d between two arbitrary points $\mathbf{p}$ and $\mathbf{q}$ is simply the sum of the segment lengths weighted by their corresponding refractive indices. That is,

$$d(\mathbf{x}_0, \ldots, \mathbf{x}_{N+1}) = \sum_{i=0}^{N} \eta_i \, || \, \mathbf{x}_i - \mathbf{x}_{i+1} \, || \, . \tag{2.5}$$

where $\mathbf{x}_0 = \mathbf{p}$ and $\mathbf{x}_{N+1} = \mathbf{q}$. For a specular reflection path in a homogeneous medium, equation (2.5) can be further simplified to

$$d(\mathbf{x}_0, \ldots, \mathbf{x}_{N+1}) = \sum_{i=0}^{N} || \, \mathbf{x}_i - \mathbf{x}_{i+1} \, || \, , \tag{2.6}$$

where all the refractive indices are assumed to be 1 since they are constant. This is exactly the assumption we used to derive the path Jacobian formula for a specular reflection path. However, this assumption breaks down for a general ray-traced path involving refractions into a different medium, where equation (2.5) has to be used instead.

Lagrange Multiplier Theorem: The Fermat path problem represents a class of optimization problems with equality constraints, for which the *Lagrange Multiplier Theorem* [5, p. 315] provides a first-order necessary condition for the local extrema.

Theorem 1 (Lagrange Multiplier Theorem) Let $\mathbf{x}^*$ be a local extremal point of $f : \mathbb{R}^n \to \mathbb{R}$, subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, where $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, and $\mathbf{x}^*$ is a regular point. Then, there exists $\lambda^* \in \mathbb{R}^{\mathbf{m}}$ such that

$$\nabla f(\mathbf{x}^*) + \lambda^{*\mathrm{T}} D\mathbf{h}(\mathbf{x}^*) = \mathbf{0}^{\mathrm{T}}, \tag{2.7}$$

where

$$D\mathbf{h}(\mathbf{x}^*) = \begin{bmatrix} \nabla h_1(\mathbf{x}^*) \\ \vdots \\ \nabla h_m(\mathbf{x}^*) \end{bmatrix}$$

is the Jacobian matrix of $\mathbf{h} = [h_1, \ldots, h_m]^{\mathrm{T}}$ at $\mathbf{x}^*$.

We refer to the vector λ^* in the above theorem as the *Lagrange multiplier vector.*

Implicit Function Theorem: In general, it is impossible to extract a closed-form solution for the non-linear system (2.7) obtained by means of Lagrange Multiplier Theorem. Fortunately, the *Implicit Function Theorem* (IFT) [71, pp. 211–213] provides a method for explicitly computing the *derivative* of such an implicitly-defined function without finding the function itself; it is this tool that allows us to derive a closed-form expression for the path Jacobian.

Theorem 2 (Implicit Function Theorem) Let $A \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set and let $F : A \to \mathbb{R}^m$ be a function of class C^p. Suppose $(\mathbf{x}_0, \mathbf{y}_0) \in A$ such that $F(\mathbf{x}_0, \mathbf{y}_0) = 0$ and

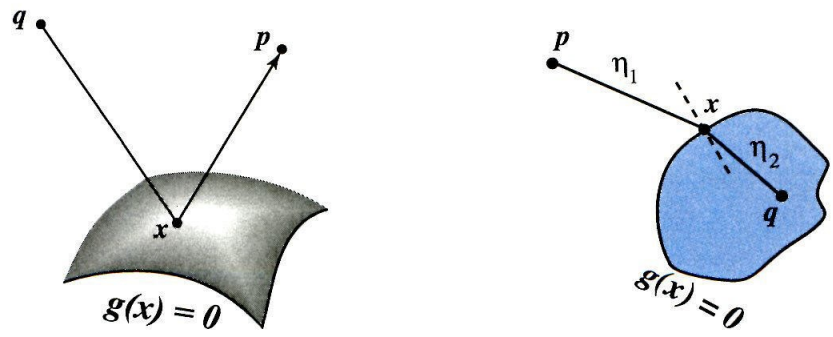
$$\det \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix} \neq 0, \tag{2.8}$$

where $F = (F_1, \ldots, F_m)$, and the Jacobian matrix is evaluated at the point $(\mathbf{x}_0, \mathbf{y}_0)$. Then there exists an open neighborhood $\mathbf{x}_0 \in U \subset \mathbb{R}^n$, a neighborhood $\mathbf{y}_0 \in V \subset \mathbb{R}^m$, and a unique function $f : U \to V$ such that

$$F(\mathbf{x}, f(\mathbf{x})) = 0 \tag{2.9}$$

for all $\mathbf{x} \in U$. Furthermore, $f \in C^p$.

By differentiating both sides of equation (2.9) with respect to the independent variable $\mathbf{x} \in \mathbb{R}^n$, we obtain the following well-known corollary:



(a) (b)

Figure 2.3: *A path from* **p** *to* **q** *via a bounce on a surface. The index of refraction is labelled over each ray segment. (a) Specular reflection. (b) Refraction.*

Corollary 1 The Jacobian matrix of the implicit function f in Theorem 2 is given by

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{bmatrix}. \quad (2.10)$$

Note that the inverse on the right hand side of equation (2.10) is guaranteed to exist since the determinant is necessarily nonzero.

Using these mathematical tools, we are ready to derive the formulas for the path Jacobian.

Single-Bounce Path

Given a surface and two points **p** and **q** in space, the problem of finding a point **x** on the surface where a ray is reflected or refracted from **p** to **q** has a long history in optics. For a spherical mirror, this problem is known as *Alhazen's problem* [78]. Even for simple convex shapes, such as a sphere, it is difficult to find **x** analytically. For non-convex surfaces the problem is even more difficult, as there may be many such reflecting (refraction) paths

connecting two points. By computing the path Jacobian $\mathbf{J} = \partial\mathbf{x}/\partial\mathbf{p}$, we can attack the problem of finding these intersection points in a completely different manner; by approximating the reflection (or refraction) $\mathbf{x}'$ of a point $\mathbf{p}'$ using a known reflection (or refraction) $\mathbf{x}$ of a nearby point $\mathbf{p}$. For example,

$$\mathbf{x}' = \mathbf{x} + \mathbf{J}(\mathbf{p}' - \mathbf{p}) \tag{2.11}$$

approximates the new intersection $\mathbf{x}'$ to first-order accuracy by perturbing the known intersection $\mathbf{x}$.

Consider a ray-traced path involving a single-surface interaction, where a ray cast from $\mathbf{p}$ travels through a medium of refractive index η_1, scatters at a point $\mathbf{x}$ on a smooth surface $\mathbf{G}$ implicitly defined by $g(\mathbf{x}) = 0$, and then reaches a fixed point $\mathbf{q}$ located in a medium with refractive index η_2. The interacting surface may be a specular reflector (Figure 2.3a), or a transparent refractive surface (Figure 2.3b). In the former case, we have $\eta_1 = \eta_2$. It follows from Fermat's principle that the optical path length assumes a local extremum. In addition, the bouncing point $\mathbf{x}$ is required to lie on the surface $\mathbf{G}$. Therefore, we may recast the problem of computing $\mathbf{x}$ as a *constrained optimization problem*, that is,

$$\text{minimize or maximize } \; d(\mathbf{p}, \mathbf{x}, \mathbf{q}) \quad \text{subject to } \; \mathbf{g(x)} = \mathbf{0}, \tag{2.12}$$

where $d(\mathbf{p}, \mathbf{x}, \mathbf{q})$ is the optical length of the path $\mathbf{p} - \mathbf{x} - \mathbf{q}$, expressed from equation (2.5) as

$$d(\mathbf{p}, \mathbf{x}, \mathbf{q}) = \eta_1 \,||\, \mathbf{p} - \mathbf{x} \,|| + \eta_2 \,||\, \mathbf{q} - \mathbf{x} \,||\,. \tag{2.13}$$

Applying the method of Lagrange multipliers, in the same fashion demonstrated by Mitchell and Hanrahan [74], we obtain

$$\begin{aligned} \nabla d(\mathbf{p}, \mathbf{x}, \mathbf{q}) + \lambda \nabla g(\mathbf{x}) &= \mathbf{0} \\ g(\mathbf{x}) &= 0, \end{aligned} \tag{2.14}$$

where λ is a Lagrange multiplier, and $d(\mathbf{p}, \mathbf{x}, \mathbf{q})$ is given by equation (2.13). By fixing one endpoint of the path, $\mathbf{q}$, and allowing $\mathbf{p}$ to vary, we may rewrite equation (2.14) as an implicit equation that relates $\mathbf{p}$, $\mathbf{x}$ and λ by

$$F_{\eta_1,g,\eta_2}(\mathbf{p}, \mathbf{x}, \lambda) = \mathbf{0}, \tag{2.15}$$

where $F_{\eta_1,g,\eta_2} : \mathbb{R}^3 \times \mathbb{R}^4 \rightarrow \mathbb{R}^4$, and $\mathbf{p}$ is regarded as the independent variable. As a convention, η_1 and η_2 in the subscript denote the refractive index of the medium where the ray travels before and after hitting the surface $\mathbf{G}$, respectively. We refer to equation (2.15) as the *Fermat equation*. The explicit form of the Fermat equation can be derived by expanding the ∇ operator in equation (2.14), yielding

$$\begin{aligned} F_i(\mathbf{p}, \mathbf{x}, \lambda) &= -\eta_1 \frac{(p_i - x_i)}{||\,\mathbf{p} - \mathbf{x}\,||} - \eta_2 \frac{(q_i - x_i)}{||\,\mathbf{q} - \mathbf{x}\,||} + \lambda \frac{\partial g(\mathbf{x})}{\partial x_i} \\ F_4(\mathbf{p}, \mathbf{x}, \lambda) &= g(\mathbf{x}), \end{aligned} \tag{2.16}$$

where $i = 1, 2, 3$.

As we mentioned before, it is generally impossible to solve these non-linear equations for the intersection point $\mathbf{x}$ in a close form, even for trivial functions $g(\mathbf{x})$ [78]. Mitchell and Hanrahan [74] dealt with this problem by applying an iterative root-finding technique known as the interval Newton method to solve for $\mathbf{x}$. In contrast, the *derivative* of the intersection point can be computed directly, without knowing the explicit functional relationship between $\mathbf{x}$ and the end points $\mathbf{q}$ and $\mathbf{p}$, by means of the Implicit Function Theorem.

Since this given path provides a solution $(\tilde{\mathbf{p}}, \tilde{\mathbf{x}}, \tilde{\lambda})$ to the Fermat equation (2.15), it follows from the condition (2.8) in Theorem 2 that if[1]

$$\det\left[\frac{\partial F_{\eta_1,g,\eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial(\mathbf{x}, \lambda)}\right] = \det\left[\frac{\partial F_{\eta_1,g,\eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial \mathbf{x}} \quad \frac{\partial F_{\eta_1,g,\eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial \lambda}\right] \neq 0$$

[1]The condition is violated when a ray is tangent to the surface. This degenerate case is considered as direct transfer instead of reflection or refraction.

at $(\tilde{\mathbf{p}}, \tilde{\mathbf{x}}, \tilde{\lambda})$, there exists a function $f : \mathbb{R}^3 \to \mathbb{R}^4$ such that

$$f(\mathbf{p}) = (\mathbf{x}, \lambda) \tag{2.17}$$

and

$$F_{\eta_1, g, \eta_2}(\mathbf{p}, f(\mathbf{p})) \;=\; \mathbf{0}$$

for all $\mathbf{p}$ sufficiently close to $\tilde{\mathbf{p}}$. That is, f solves the ray intersection problem in a neighborhood of the known ray-traced path. More importantly, according to Corollary 1, we can solve for $\partial f / \partial \mathbf{p}$ in this neighborhood using equation (2.10), yielding

$$\left[\frac{\partial f(\mathbf{p})}{\partial \mathbf{p}}\right]_{4\times 3} = -\left[\frac{\partial F_{\eta_1, g, \eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial (\mathbf{x}, \lambda)}\right]^{-1}_{4\times 4} \left[\frac{\partial F_{\eta_1, g, \eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial (\mathbf{p})}\right]_{4\times 3.} \tag{2.18}$$

For clarity, we shall frequently indicate the matrix dimensions with subscripts, as we have done above. Equation (2.17) shows that the path function $\Psi : \mathbf{p} \to \mathbf{x}$ is easily obtained from f by discarding its last component λ. By introducing an operator $\mathbf{sub} : \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^4) \to \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^3)$, which drops the last row of a 4×3 matrix, the 3×3 path Jacobian $\mathbf{J}$ can be expressed as

$$\boxed{\mathbf{J} \;=\; \left[\frac{\partial \Psi}{\partial \mathbf{p}}\right]_{3\times 3} \;=\; \mathbf{sub}\left(-\left[\frac{\partial F_{\eta_1, g, \eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial (\mathbf{x}, \lambda)}\right]^{-1}_{4\times 4} \left[\frac{\partial F_{\eta_1, g, \eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial \mathbf{p}}\right]_{4\times 3}\right),} \tag{2.19}$$

which characterizes the variation in $\mathbf{x}$ with respect to $\mathbf{p}$.

N-Bounce Path

Now Let us consider a general ray-traced path from a varying point $\mathbf{p}$ to a fixed point $\mathbf{q}$ via N surface interactions. We order all the intersection points from $\mathbf{p}$ to $\mathbf{q}$ as $\mathbf{x}_1, \ldots, \mathbf{x}_N$, with $\mathbf{x}_0 = \mathbf{p}$ and $\mathbf{x}_{N+1} = \mathbf{q}$. All the interacting surfaces $\mathbf{G}_i$ and their corresponding implicit definitions g_i are ordered accordingly, as shown in Figure 2.4, for a three-bounce

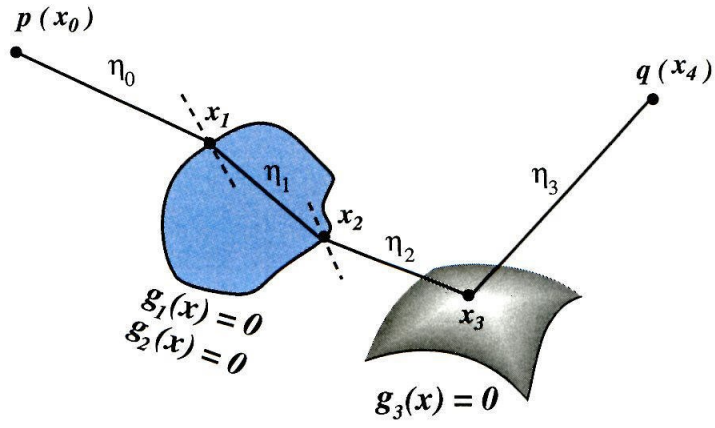


Figure 2.4: *A generic three-bounce ray-traced path from* $\mathbf{p}$ *to* $\mathbf{q}$ *which mixes refraction and specular reflection.* $\mathbf{x}_1$, $\mathbf{x}_2$ *and* $\mathbf{x}_3$ *are three intersecting points with implicitly-defined surfaces* $\mathbf{G}_1$, $\mathbf{G}_1$ *and* $\mathbf{G}_2$. *The indices of refraction of the medium where the path travels are labeled in each segment.*

path. We shall always consider the varying endpoint $\mathbf{p} = \mathbf{x}_0$ as the starting point and the fixed endpoint $\mathbf{q} = \mathbf{x}_{N+1}$ as the ending point of an N-bounce path. Note that we view a ray through a transparent solid surface as two surface interactions, with the same implicit function counted twice, as demonstrated in Figure 2.4 by $g_1(\mathbf{x}) = g_2(\mathbf{x})$. In addition, we associate an index of refraction η_i $(i = 1, \ldots, N)$ with each ray segment from $\mathbf{x}_{i-1}$ to $\mathbf{x}_i$ to specify the medium in which it is immersed. By viewing the position of each bouncing point $\mathbf{x}_i$ as a function Ψ_i of the endpoint $\mathbf{p}$, we may define the path Jacobian $\mathbf{J}_i$ at $\mathbf{x}_i$ as the derivative of Ψ_i with respect to $\mathbf{p}$; this Jacobian characterizes how $\mathbf{x}_i$ changes with respect to perturbations in $\mathbf{p}$. Accordingly, the N reflection points in the path can be updated to the first order using

$$\mathbf{x}_i' = \mathbf{x}_i + \mathbf{J}_i \Delta\mathbf{p} \tag{2.20}$$

for $i = 1, 2, \ldots, N$, where the N path Jacobians $\mathbf{J}_i$s can be computed recursively, rather than calculating $\mathbf{J}_i$s all at once by applying Lagrange Multiplier Methods to the whole path, which involves inverting an $N \times N$ matrix that increases quadratically [27, 29].

By applying the chain rule to the definition of $\mathbf{J}_i$, it follows that we may express $\mathbf{J}_i$ as a product of i Jacobian matrices. That is,

$$\mathbf{J}_i = \frac{\partial \mathbf{x}_i}{\partial \mathbf{p}} = \frac{\partial \mathbf{x}_i}{\partial \mathbf{x}_{i-1}} \cdot \frac{\partial \mathbf{x}_{i-1}}{\partial \mathbf{x}_{i-2}} \cdots \frac{\partial \mathbf{x}_1}{\partial \mathbf{p}} \tag{2.21}$$

for $i = 1, 2, \ldots, N$. Note that each factor on the right hand side of equation (2.21) is a 3×3 Jacobian matrix, which denotes the derivative of the position of the ith intersection point with respect to its previous point $\mathbf{x}_{i-1}$. Let us define

$$\mathbf{J}_i^* \equiv \frac{\partial \mathbf{x}_i}{\partial \mathbf{x}_{i-1}} \tag{2.22}$$

at each intersection point $\mathbf{x}_i$ $(i = 1, 2, \ldots, N)$. Then equation (2.21) can be written as

$$\mathbf{J}_i = \mathbf{J}_i^* \cdot \mathbf{J}_{i-1}^* \cdots \mathbf{J}_1^*,$$

or, equivalently, by the recurrence relation

$$\mathbf{J}_i = \mathbf{J}_i^* \cdot \mathbf{J}_{i-1}, \tag{2.23}$$

where $\mathbf{J}_1 = \mathbf{J}_1^*$. We refer to the Jacobian matrix defined in equation (2.22) as the *bounce Jacobian*. Using equation (2.23), we have transformed the problem of computing N path Jacobians $\mathbf{J}_i$ to an equivalent problem of computing N bounce Jacobians $\mathbf{J}_i^*$. We now describe how the bounce Jacobians can be computed.

Observe that for the last bounce $(\mathbf{x}_{N-1}, \mathbf{x}_N, \mathbf{q})$, where $\mathbf{q}$ is the fixed endpoint, the problem reduces to a simple one-bounce path. As discussed in the previous section, we have

$$F_{\eta_{N-1}, g_N, \eta_N}(\mathbf{x}_{N-1}, \mathbf{x}_N, \lambda_N) = \mathbf{0}, \tag{2.24}$$

where $\mathbf{x}_{N-1}$ is considered as the independent variable. The Implicit Function Theorem

implies that if

$$\det\left(\frac{\partial F_{\eta_{N-1},g_N,\eta_N}(\mathbf{x}_{N-1},\mathbf{x}_N,\lambda_N)}{\partial(\mathbf{x}_N,\lambda_N)}\right)\neq 0,$$

there exists a function $f_N : \mathbf{x}_{N-1} \to (\mathbf{x}_N, \lambda_N)$. The function f_N can be decomposed into two components: $f_N(\mathbf{x}) = (f_{N1}(\mathbf{x}), f_{N2}(\mathbf{x}))$ where $f_{N1} : \mathbb{R}^3 \to \mathbb{R}^3$ and $f_{N2} : \mathbb{R}^3 \to \mathbb{R}$ map $\mathbf{x}_{N-1}$ to $\mathbf{x}_N$ and λ_N respectively. With the existence of f_N, the last bounce Jacobian $\mathbf{J}_N^*$ has been derived in equation (2.19). As a segment of an N-bounce path, the ith bounce still observes Fermat's principle. Applying Lagrange multipliers to $(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})$, where $i \neq N$, we obtain a similar Fermat equation:

$$F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},\lambda_i)=\mathbf{0}, \tag{2.25}$$

which differs from equation (2.15) by having two varying end points $\mathbf{x}_{i-1}$ and $\mathbf{x}_{i+1}$. By processing the N bounces from the ending point $\mathbf{q}$ to the starting point $\mathbf{p}$, the functions f_is mapping $\mathbf{x}_{i-1}$ to $(\mathbf{x}_i, \lambda_i)$ $(i = N, \ldots, 1)$ can be implicitly determined from the Implicit Function Theorem in sequence. Thus when we reach the ith bounce, the function $f_{i+1} : \mathbb{R}^3 \to \mathbb{R}^4$ mapping $\mathbf{x}_i$ to $(\mathbf{x}_{i+1}, \lambda_{i+1})$ has been constructed implicitly from the $(i+1)$-th bounce. By decomposing f_{i+1} into $f_{(i+1)1} : \mathbf{x}_i \to \mathbf{x}_{i+1}$ and $f_{(i+1)2} : \mathbf{x}_i \to \lambda_{i+1}$, we can express $\mathbf{x}_{i+1}$ in equation (2.25) in terms of $f_{(i+1)1}$, obtaining

$$F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\mathbf{x}_i,f_{(i+1)1}(\mathbf{x}_i),\lambda_i)=\mathbf{0}, \tag{2.26}$$

which can be treated as an implicit equation of three variables $\mathbf{x}_{i-1}$, $\mathbf{x}_i$, and λ_i, with $\mathbf{x}_{i-1}$ as the independent variable. That is,

$$H_i(\mathbf{x}_{i-1},\mathbf{x}_i,\lambda_i)=\mathbf{0} \tag{2.27}$$

where

$$H_i(\mathbf{x}_{i-1},\mathbf{x}_i,\lambda_i)=F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\mathbf{x}_i,f_{(i+1)1}(\mathbf{x}_i),\lambda_i).$$

Applying the Implicit Function Theorem to equation (2.27), the ith explicit function $f_i : \mathbf{x}_{i-1} \to (\mathbf{x}_i, \lambda_i)$ exists under the condition that

$$\det\left(\frac{\partial H_i(\mathbf{x}_{i-1}, \mathbf{x}_i, \lambda_i)}{\partial(\mathbf{x}_i, \lambda_i)}\right) = \det\left(\frac{\partial F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1}, \mathbf{x}_i, f_{(i+1)1}(\mathbf{x}_i), \lambda_i)}{\partial(\mathbf{x}_i, \lambda_i)}\right) \neq 0. \qquad (2.28)$$

Note that the existence of the function f_i makes the bounce Jacobian $\mathbf{J}_i^*$ in (2.22) well-defined. In terms of f_i, equation (2.26) can be reformulated as

$$F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\ f_{i1}(\mathbf{x}_{i-1}),\ f_{(i+1)1}(f_{i1}(\mathbf{x}_{i-1})),\ f_{i2}(\mathbf{x}_{i-1})) = \mathbf{0}. \qquad (2.29)$$

Differentiating both sides of equation (2.29) with respect to $\mathbf{x}_{i-1}$, we get

$$\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i-1}} + \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_i} \cdot \frac{\partial f_{i1}}{\partial \mathbf{x}_{i-1}} + \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i+1}} \cdot \frac{\partial f_{(i+1)1}}{\partial \mathbf{x}_i} \cdot \frac{\partial f_{i1}}{\partial \mathbf{x}_{i-1}} + \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \lambda_i} \cdot \frac{\partial f_{i2}}{\partial \mathbf{x}_{i-1}} = \mathbf{0}. \qquad (2.30)$$

Rearranging the terms in equation (2.30) and substituting $\mathbf{J}_{i+1}^*$ computed previously for $\dfrac{\partial f_{(i+1)1}}{\partial \mathbf{x}_i}$, we obtain

$$\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i-1}} = -\left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}_{i+1}^* + \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_i} \quad \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \lambda_i}\right] \begin{bmatrix} \dfrac{\partial f_{i1}(\mathbf{x}_{i-1})}{\partial \mathbf{x}_{i-1}} \\ \\ \dfrac{\partial f_{i2}(\mathbf{x}_{i-1})}{\partial \mathbf{x}_{i-1}} \end{bmatrix}.$$

By inverting the matrix, we solve for $\partial f_i / \partial \mathbf{x}_{i-1}$:

$$\left[\frac{\partial f_i}{\partial \mathbf{x}_{i-1}}\right]_{4\times 3} = -\left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}_{i+1}^* + \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_i} \quad \frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \lambda_i}\right]^{-1} \left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i-1}}\right], \qquad (2.31)$$

where the existence of the inverse is guaranteed by the condition in (2.28). Introducing an operator $\mathbf{aug} : \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^4) \to \mathrm{Hom}(\mathbb{R}^4, \mathbb{R}^4)$, which expands a 4×3 matrix by

appending a zero column, we may rewrite equation (2.31) as

$$\left[\frac{\partial f_i}{\partial \mathbf{x}_{i-1}}\right]_{4\times 3} = -\left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial(\mathbf{x}_i,\lambda_i)} + \mathbf{aug}\left(\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i+1}}\cdot \mathbf{J}^*_{i+1}\right)\right]^{-1}\left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}}{\partial \mathbf{x}_{i-1}}\right]. \quad (2.32)$$

Finally, taking the submatrix of the left hand side of equation (2.32), we have derived a recurrence relation for the ith bounce Jacobian $\mathbf{J}^*_i$:

$$\boxed{\mathbf{J}^*_i \;=\; \mathbf{sub}\left(-\left[A_i + \mathbf{aug}(B_i\cdot \mathbf{J}^*_{i+1})\right]^{-1} T_i\right), \quad (2.33)}$$

where $i = N-1,\ldots,1$ and

$$\begin{aligned} A_i &= \left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},\lambda_i)}{\partial(\mathbf{x}_i,\lambda_i)}\right]_{4\times 4} \\ B_i &= \left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},\lambda_i)}{\partial \mathbf{x}_{i+1}}\right]_{4\times 3} \\ T_i &= \left[\frac{\partial F_{\eta_{i-1},g_i,\eta_i}(\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},\lambda_i)}{\partial \mathbf{x}_{i-1}}\right]_{4\times 3}. \end{aligned} \quad (2.34)$$

Equation (2.33) suggests that the bounce Jacobians $\mathbf{J}^*_i$ for an N-bounce path are computed *backward*, from the fixed point $\mathbf{q}$ towards the perturbed point $\mathbf{p}$. The starting Jacobian $\mathbf{J}^*_N$ for the last intersection point $\mathbf{x}_N$, which is calculated first, can be viewed as a special case of this recurrence relation where $B_N = 0$ and $\mathbf{J}^*_{N+1} = \mathbf{0}$.

In order to use these bounce Jacobians $\mathbf{J}^*_i$ to perturb a given N-bounce path, a naive approach is to compute each path Jacobian $\mathbf{J}_i$ ($i = 1,2,\ldots,N$) from the $\mathbf{J}^*_i$s with equation (2.21), and then update the corresponding bouncing point $\mathbf{x}_i$ to first order accuracy using equation (2.20). However, motivated by the observation that the term ultimately required for linear perturbation in equation (2.20) is $\mathbf{J}_i\Delta\mathbf{p}$, we may optimize the naive algorithm by perturbing the N intersection points incrementally in a prescribed order. Let $\Delta\mathbf{x}_i = \mathbf{J}_i\Delta\mathbf{p}$ ($i = 1,2,\ldots,N$), which denotes the perturbation of each intersection

point in this linear approximation. Then the perturbation formula (2.20) becomes

$$\mathbf{x}_i' \;=\; \mathbf{x}_i + \Delta\mathbf{x}_i. \tag{2.35}$$

Expressing $\Delta\mathbf{x}_i$ in terms of $\mathbf{J}_i^*$s using equation (2.23), we have

$$\Delta\mathbf{x}_i \;=\; \mathbf{J}_i^* \cdot \mathbf{J}_{i-1}\Delta\mathbf{p} \;=\; \mathbf{J}_i^*\Delta\mathbf{x}_{i-1}$$

for $i = 1, 2, \ldots, N$. Consequently, equation (2.35) can be expressed as

$$\mathbf{x}_i' \;=\; \mathbf{x}_i + \mathbf{J}_i^*\Delta\mathbf{x}_{i-1}, \tag{2.36}$$

where $i = 1, 2, \ldots, N$, $\Delta\mathbf{x}_i = \mathbf{x}_i' - \mathbf{x}_i$ and $\Delta\mathbf{x}_0 = \Delta\mathbf{p}$. We can interpret equation (2.36) as the first-order Taylor approximation of f_{i1} around its previous point $\mathbf{x}_{i-1}$, which allows us to perturb the intersection points in an N-bounce path incrementally, from the starting point $\mathbf{p}$ until the ending point $\mathbf{q}$.

Equations (2.19) and (2.33) suggest that the path Jacobian formulas for a general ray-traced path have the same form as those derived for a specular reflection path [27], which actually corresponds to the special case of $\eta_1 = \eta_2 = 1$.

2.2.3 The Path Hessian

Based on the expression for path Jacobians, path Hessians can be computed via tensor differentiation. Furthermore, the Hessians for multiple bounces can also be computed iteratively from the ending point $\mathbf{q}$ to the starting point $\mathbf{p}$, in a similar fashion as the Jacobians. Consequently, for an N-bounce path, a second-order perturbation formula similar to equation (2.36) can be applied from the starting point $\mathbf{p}$ to the ending point $\mathbf{q}$ as:

$$\mathbf{x}_i' = \mathbf{x}_i + \mathbf{J}_i^*\Delta\mathbf{x}_{i-1} + \frac{1}{2}(\Delta\mathbf{x}_{i-1})^{\mathrm{T}}\mathbf{H}_i^*\Delta\mathbf{x}_{i-1}, \tag{2.37}$$

where $i = 1, 2, 3, \ldots, N$, $\Delta\mathbf{x}_{i-1} = \mathbf{x}_{i-1}' - \mathbf{x}_{i-1}$, and $\Delta\mathbf{x}_0 = \Delta\mathbf{p}$. The Hessian $\mathbf{H}_i^*$ is defined as the second-order derivative $\partial^2\mathbf{x}_i/\partial^2\mathbf{x}_{i-1}$, called *bounce Hessian* to be consistent with

$\mathbf{J}_i^*$.

Starting with nearly the same formula (2.33), the derivation of the path Hessian parallels that for a specular reflection path [27, 29], resulting in a similar expression for a general ray-traced path. The only difference is that the definitions for A_i, B_i, T_i are replaced by equation (2.34), involving $F_{\eta_{i-1},g_i,\eta_i}$, instead of F_i before [27]. The outline of this process can be stated briefly as follows:

Letting $\Gamma_i = [A_i + \mathbf{aug}(B_i \cdot \mathbf{J}_{i+1}^*)]$ and $\mathbf{D}_i = -\Gamma_i^{-1} T_i$, we may rewrite equation (2.33) as

$$T \;=\; -\Gamma \cdot \mathbf{D}, \tag{2.38}$$

where we have omitted the subscript i for simplicity. Henceforth, we assume that all the variables with no subscripts are associated with a particular bouncing point $\mathbf{x}_i$. In order to compute $\mathbf{H}_i^*$, we need to compute the gradient of the matrix $\mathbf{D}$, denoted by $\nabla\mathbf{D}$ [98, p.62]. The gradient $\nabla\mathbf{D}$ is a third-order tensor and can be described with several different formalisms. It can be viewed as: the scalar components $(\nabla\mathbf{D})_{mij}$, the vector entries $\nabla(\mathbf{D}_{ij})$, or the matrix layers $\nabla_m\mathbf{D}$, defined respectively as

$$(\nabla\mathbf{D})_{mij} \;\equiv\; \frac{\partial \mathbf{D}_{ij}}{\partial(\mathbf{x}_{i-1})_m}, \;\; \nabla(\mathbf{D}_{ij}) \;\equiv\; \frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{x}_{i-1}}, \;\; \nabla_m\mathbf{D} \;\equiv\; \left\{ \frac{\partial \mathbf{D}_{ij}}{\partial(\mathbf{x}_{i-1})_m} \right\},$$

where $i = 1,2,3,4$ and $m,j = 1,2,3$. By differentiating both sides of equation (2.38) with respect to $\mathbf{x}_{i-1}$ using Cartesian tensor notation, we obtain

$$\boxed{\nabla_m\mathbf{D} = -\Gamma^{-1}\left(\nabla_m T + \nabla_m\Gamma \cdot \mathbf{D}\right)} \tag{2.39}$$

for $m = 1,2,3$. The formulas for ∇T and $\nabla\Gamma$ are given by

$$\nabla(T_{jk}) \;=\; \frac{\partial T_{jk}}{\partial \mathbf{x}_{i-1}} + \frac{\partial T_{jk}}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}_{i+1}^* \cdot \mathbf{J}_i^* + \frac{\partial T_{jk}}{\partial(\mathbf{x}_i, \lambda_i)} \cdot \mathbf{D} \tag{2.40}$$

$$(\nabla\Gamma)_{mjl} = \begin{cases} (\nabla_m A + \nabla_m B \cdot \mathbf{J}^*_{i+1} + B \cdot \nabla_m C)_{jl} & (l = 1,2,3) \\ (\nabla A)_{mjl} & (l = 4) \end{cases} \tag{2.41}$$

where

$$\begin{aligned} \nabla(A_{jr}) &= \frac{\partial A_{jr}}{\partial \mathbf{x}_{i-1}} + \frac{\partial A_{jr}}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}^*_{i+1} \cdot \mathbf{J}^*_i + \frac{\partial A_{jr}}{\partial(\mathbf{x}_i, \lambda_i)} \cdot \mathbf{D} \\ \nabla(B_{jk}) &= \frac{\partial B_{jk}}{\partial \mathbf{x}_{i-1}} + \frac{\partial B_{jk}}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}^*_{i+1} \cdot \mathbf{J}^*_i + \frac{\partial B_{jk}}{\partial(\mathbf{x}_i, \lambda_i)} \cdot \mathbf{D} \\ \nabla(C_{kl}) &= \nabla\left((\mathbf{J}^*_{i+1})_{kl}\right) \cdot \mathbf{J}^*_i \end{aligned} \tag{2.42}$$

for $j, r = 1,2,3,4$ and $m, k, l = 1,2,3$. It follows from equation (2.3) and $\mathbf{J}^* = \mathbf{sub}(\mathbf{D})$ that the Hessian tensor can be formed from $\nabla\mathbf{D}$ by

$$\mathbf{H}^*_{jkm} = \frac{\partial \mathbf{x}_j}{\partial \mathbf{p}_k \mathbf{p}_m} = \frac{\partial \mathbf{J}^*_{jk}}{\partial \mathbf{p}_m} = (\nabla \mathbf{J}^*)_{mjk} = (\nabla \mathbf{D})_{mjk}. \tag{2.43}$$

for $j, k, m = 1,2,3$. The dependence of the ith bounce Hessian on $\mathbf{J}^*_i$, $\mathbf{J}^*_{i+1}$ and $\nabla\mathbf{J}^*_{i+1}$ implied by equation (2.42) suggests that the Jacobians and Hessians for an N-bounce path can be computed together, from the fixed point propagated to the perturbed point.

2.3 Simplified Path Jacobian

In equation (2.19) of section 2.2.2, we have derived the formula for the path Jacobian as

$$\mathbf{J} = \mathbf{sub}\left(-\Gamma^{-1} T\right), \tag{2.44}$$

where

$$\Gamma = \frac{\partial F_{\eta_1, g, \eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial(\mathbf{x}, \lambda)}, \quad T = \frac{\partial F_{\eta_1, g, \eta_2}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial \mathbf{p}} \tag{2.45}$$

and the operator **sub** is introduced to remove the effect of the lagrange multiplier λ. Observing the special structures of the Fermat equation (2.16) and the resulting two matrices Γ and T given in equation (2.45), we may simplify equation (2.44) in such a way that the path Jacobian $\mathbf{J}$ can be computed directly, without inverting a 4×4 matrix in certain circumstances. Consequently, we can get rid of the operator **sub**. For simplicity, in the remaining section we shall omit the subscripts and use symbols F and F_i to denote F_{η_1,g,η_2} and $F_{\eta_{i-1},g_i,\eta_i}$, respectively.

Let $\hat{F}(\mathbf{p}, \mathbf{x}, \lambda) = \mathbf{0}$ represent the non-linear system formed by the first three equations of F shown in equation (2.16), that is,

$$\begin{aligned} \hat{F}(\mathbf{p}, \mathbf{x}, \lambda) &= \nabla d(\mathbf{p}, \mathbf{x}, \mathbf{q}) + \lambda \nabla g(\mathbf{x}) \\ &= -\eta_1 \frac{\mathbf{p} - \mathbf{x}}{||\,\mathbf{p} - \mathbf{x}\,||} - \eta_2 \frac{\mathbf{q} - \mathbf{x}}{||\,\mathbf{q} - \mathbf{x}\,||} + \lambda \mathbf{h}, \end{aligned} \tag{2.46}$$

where the vector $\mathbf{h} = \nabla g(\mathbf{x}) = \partial g(\mathbf{x})/\partial \mathbf{x}$. We observe that

1. The matrix Γ is symmetric, and can be expressed as a block matrix

$$\begin{bmatrix} \hat{\Gamma}_{3\times 3} & \mathbf{h}_{3\times 1} \\ \mathbf{h}^{\mathrm{T}}_{1\times 3} & 0 \end{bmatrix}, \tag{2.47}$$

   where the symmetry of the 3×3 matrix $\hat{\Gamma} = \dfrac{\partial \hat{F}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial \mathbf{x}}$ follows from the fact that it can be viewed as the Hessian matrix of a smooth function $d(\mathbf{p}, \mathbf{x}, \mathbf{q}) + \lambda g(\mathbf{x})$ with respect to $\mathbf{x}$.

2. The 4×3 matrix T has a form

$$\begin{bmatrix} \hat{T}_{3\times 3} \\ \mathbf{0}_{1\times 3} \end{bmatrix}, \tag{2.48}$$

   where $\hat{T} = \dfrac{\partial \hat{F}(\mathbf{p}, \mathbf{x}, \lambda)}{\partial \mathbf{p}}$.

It follows from linear algebra that the inverse Γ^{-1} of matrix (2.47) must also be sym-

metric, which may be expressed in a form of block matrix as

$$\begin{bmatrix} N_{3\times3} & \mathbf{v}_{3\times1} \\ \mathbf{v}_{1\times3}^{\mathrm{T}} & d \end{bmatrix}, \tag{2.49}$$

where the scalar d, the vector $\mathbf{v}$ and the 3×3 symmetric submatrix N are to be determined. On the other hand, as an inverse of Γ, the matrix (2.49) should satisfy the following equation

$$\begin{bmatrix} \hat{\Gamma}_{3\times3} & \mathbf{h}_{3\times1} \\ \mathbf{h}_{1\times3}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} N_{3\times3} & \mathbf{v}_{3\times1} \\ \mathbf{v}_{1\times3}^{\mathrm{T}} & d \end{bmatrix} = \begin{bmatrix} \hat{\Gamma}N + \mathbf{h}\mathbf{v}^{\mathrm{T}} & \hat{\Gamma}\mathbf{v} + d\mathbf{h} \\ \mathbf{h}^{\mathrm{T}}N & \mathbf{h}^{\mathrm{T}}\mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{3\times3} & 0 \\ 0 & 1 \end{bmatrix}, \tag{2.50}$$

where $\mathbf{I}$ denotes the 3×3 identity matrix. Solving the linear system (2.50) for N, we obtain

$$N = \hat{\Gamma}^{-1}\left(\mathbf{I} - \frac{\mathbf{h}\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}}{\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}\mathbf{h}}\right) \tag{2.51}$$

for non-singular $\hat{\Gamma}$. See Appendix A.1 for details. According to the definition of **sub** in equation (2.44), the path Jacobian $\mathbf{J}$ we are interested in is composed of the first three rows of $\begin{bmatrix} N_{3\times3} & \mathbf{v}_{3\times1} \\ \mathbf{v}_{1\times3}^{\mathrm{T}} & d \end{bmatrix} \begin{bmatrix} \hat{T}_{3\times3} \\ \mathbf{0}_{1\times3} \end{bmatrix}$. Thus, we may directly evaluate it as the product of two 3×3 matrices,

$$\mathbf{J} \;=\; -N \times \hat{T}. \tag{2.52}$$

Equation (2.52) provides a more efficient way to evaluate $\mathbf{J}$ in practice since the matrix to be inverted is 3×3 instead of the 4×4 matrix in equation (2.44).

For an N-bounce path, we can do the similar optimization on the recursive formula (2.31), which can be expressed as

$$\begin{aligned} \mathbf{J}_i^* &= \mathbf{sub}\left(-\left[\frac{\partial F_i}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}_{i+1}^* + \frac{\partial F_i}{\partial \mathbf{x}_i} \;\; \frac{\partial F_i}{\partial \lambda_i}\right]^{-1} \left[\frac{\partial F_i}{\partial \mathbf{x}_{i-1}}\right]\right), \\ &= \mathbf{sub}\left({\Gamma_i}^{-1} T_i\right), \end{aligned} \tag{2.53}$$

where

$$\Gamma_i = \left[\frac{\partial F_i}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}^*_{i+1} + \frac{\partial F_i}{\partial \mathbf{x}_i} \quad \frac{\partial F_i}{\partial \lambda_i} \right], \quad T_i = \left[\frac{\partial F_i}{\partial \mathbf{x}_{i-1}} \right]. \tag{2.54}$$

Let $\hat{F}_i(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}) = \nabla d(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}) + \lambda_i \nabla g_i(\mathbf{x}_i) = \mathbf{0}$ denote the non-linear system formed from the first three rows of F_i and $\mathbf{h}_i = \nabla g_i$. Similarly, we may represent Γ_i and T_i in equation (2.54) as block matrices, as we did for Γ and T in equations (2.47) and (2.48). That is,

$$\Gamma_i = \begin{bmatrix} \hat{\Gamma}_i & \mathbf{h}_i \\ \mathbf{h}_i{}^{\mathrm{T}} & 0 \end{bmatrix}, \quad T_i = \begin{bmatrix} \hat{T}_i \\ \mathbf{0} \end{bmatrix},$$

where

$$\begin{aligned} \hat{\Gamma}_i &= \frac{\partial \hat{F}_i(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})}{\partial \mathbf{x}_{i+1}} \cdot \mathbf{J}^*_{i+1} + \frac{\partial \hat{F}_i(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})}{\partial \mathbf{x}_i}, \\ \hat{T}_i &= \frac{\partial \hat{F}_i(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})}{\partial \mathbf{x}_{i-1}}. \end{aligned}$$

However, Γ_i may not be symmetric here. Accordingly, we have to denote its inverse as

$$\begin{bmatrix} (N_i)_{3\times 3} & \mathbf{v}_{3\times 1} \\ \mathbf{w}^{\mathrm{T}}_{1\times 3} & d \end{bmatrix}.$$

To compute the upper 3×3 corner of this inverse matrix, we have to solve a larger linear system, consisting of 8 equations from the conditions:

$$\begin{bmatrix} \hat{\Gamma}_i & \mathbf{h}_i \\ \mathbf{h}_i{}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} N_i & \mathbf{v} \\ \mathbf{w}^{\mathrm{T}} & d \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{3\times 3} & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$\begin{bmatrix} N_i & \mathbf{v} \\ \mathbf{w}^{\mathrm{T}} & d \end{bmatrix} \begin{bmatrix} \hat{\Gamma}_i & \mathbf{h}_i \\ \mathbf{h}_i{}^{\mathrm{T}} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{3\times 3} & 0 \\ 0 & 1 \end{bmatrix}.$$

Using the similar technique shown in Appendix A.1, we attain an analogous formula for $\mathbf{J}_i^*$ to simplify equation (2.53), given by

$$\mathbf{J}_i^* = -N_i \times \hat{T}_i, \tag{2.55}$$

where

$$N_i = \hat{\Gamma}_i^{-1}\left(\mathbf{I} - \frac{\mathbf{h}_i{\mathbf{h}_i}^{\mathrm{T}}\hat{\Gamma}_i^{-1}}{{\mathbf{h}_i}^{\mathrm{T}}\hat{\Gamma}_i^{-1}\mathbf{h}_i}\right), \tag{2.56}$$

Discussion: In the above simplification, we have assumed that the inverse of $\hat{\Gamma}$ or $\hat{\Gamma}_i$ in equations (2.51) and (2.56) exists. Unfortunately, given that Γ is nonsingular (guaranteed by the Implicit Function Theorem), we cannot determine whether $\hat{\Gamma}$ is singular or not, which may be singular. Here is an example. Let

$$\hat{\Gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and $\mathbf{h}^{\mathrm{T}} = [1\,1\,1]$, then $\hat{\Gamma}$ is singular, but

$$\Gamma = \begin{bmatrix} \hat{\Gamma} & \mathbf{h} \\ \mathbf{h}^{\mathrm{T}} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

is nonsingular.

2.4 Parametric Surfaces

It can be seen from equations (2.19) and (2.34) that the first-order partial derivatives of F are required to compute the path Jacobians. Since the four explicit equations of the non-linear system F shown in equation (2.16) are based on the gradient of the implicit function, the local shape of the interacting surface is captured to the second order

necessary to get the first-order approximation to the scattering effects. Consequently, our path Jacobian theory could be adapted to handle more general surface representations, as long as we can estimate the first- and second-order properties of the underlying surfaces, such as parametric surfaces, triangular meshes [37], etc. Next we shall examine the widely-used parametric representation.

Specifically, since the path Jacobian depends on the gradient and Hessian of the implicit function representing a surface, in order to extend this first-order approximation to parametric surfaces, we need to replace these two terms with their equivalent geometric quantities available from the parametric representation of the same surface. Since we are interested in analyzing the surface locally, we can always consider the corresponding *Monge representation* [30] of the surface. That is, the surface can be described by the graph of a function $z = G(x, y)$. In particular, if we introduce a local coordinate system $(\mathbf{S}, \mathbf{T}, \mathbf{H})$ centered at a bouncing point $\mathbf{x}$, where $\mathbf{H}$ is chosen to be the surface normal at $\mathbf{x}$, and $\mathbf{S}$ and $\mathbf{T}$ are two orthonormal vectors in the tangent plane at $\mathbf{x}$, and $\mathbf{H} = \mathbf{S} \times \mathbf{T}$, the equation of the surface can be written in a *special Monge form* from the Taylor expansion:

$$z = \frac{1}{2}(ax^2 + 2cxy + by^2) + \frac{1}{3!}(ex^3 + 3fx^2y + 3gxy^2 + hy^3) + \cdots. \tag{2.57}$$

Figure 2.5 shows a Monge patch expressed as the height field over the tangent plane at one point. Therefore, locally the mirror surface can be implicitly described by an equation $z - G(x, y) = 0$, denoted by $g(x, y, z) = 0$. You may notice that this implicit function is defined with respect to the local reference frame at $\mathbf{x}$, but it can be easily expressed in a given global coordinate system $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ by a tranformation. It then follows that the first-order derivative of g in equation (2.57), i.e, ∇g, is coincident with the unit normal $\mathbf{n}$ at $\mathbf{x}$. Consequently, the Hessian of g just corresponds to $\frac{\partial \mathbf{n}}{\partial \mathbf{x}}$, which is the matrix associated with the *shape operator* $\mathbf{S}$. In differential geometry, the shape operator at a surface point $\mathbf{p}$ defines a linear map from a direction $\mathbf{b}$ tangent to the surface to the negative derivative of the unit normal $\mathbf{n}$ along $\mathbf{b}$ at $\mathbf{p}$. This operator completely describes a differentially small area on a surface. Given a parametric surface

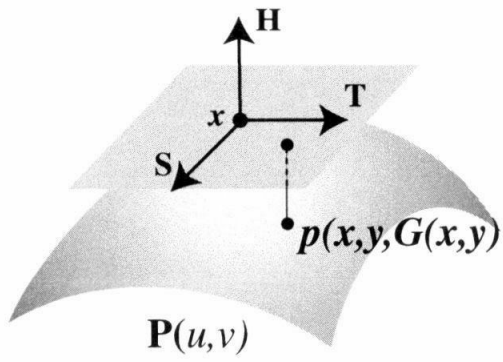


Figure 2.5: *Locally, a parametric surface can be written as the graph of the function:* $z = G(x, y)$.

$\mathbf{P}(u, v)$, the matrix of this linear map $\mathbf{S}$ may be computed as $I^{-1}II$ [30, p.31], where

$$I = \begin{bmatrix} \langle \mathbf{P}_u, \mathbf{P}_u \rangle & \langle \mathbf{P}_u, \mathbf{P}_v \rangle \\ \langle \mathbf{P}_v, \mathbf{P}_u \rangle & \langle \mathbf{P}_v, \mathbf{P}_v \rangle \end{bmatrix},$$

and

$$II = \begin{bmatrix} \langle \mathbf{P}_{uu}, \mathbf{n} \rangle & \langle \mathbf{P}_{uv}, \mathbf{n} \rangle \\ \langle \mathbf{P}_{vu}, \mathbf{n} \rangle & \langle \mathbf{P}_{vv}, \mathbf{n} \rangle \end{bmatrix}$$

are the *first fundamental form* and the *second fundamental form*, respectively. These formulae may be found in classical textbooks on differential geometry (e.g. [108, 41]).

2.5 Comparisons with Related Work

Similar algorithms that linearly approximate a perturbed ray path have been developed for use in ray tracers to calculate a bundle of rays at the cost of tracing a single ray. In this section we shall make some comparisons and connections between our path Jacobians and two closely-related previous work, that is, pencil tracing [100] and ray differentials [59].

2.5.1 Pencil Tracing

Paraxial ray theory [22] is an approximation technique originally designed for lens design, and its application to ray tracing is known as *pencil tracing*. The paraxial approximation theory provides a linear approximation for ray changes due to refraction, reflection, and transfer (propagation in a homogeneous medium). In order to linearize each phenomenon, pencil tracing makes a distinct set of simplifying assumptions on each optical event (transfer, reflection and refraction). In pencil tracing, paraxial rays to an axial ray are parameterized by point-vector pairs on a plane perpendicular to the axial ray. A linear ray change from the axial ray can be represented by a 4×4 matrix called *system matrix*, and thus, the propagation of these paraxial rays is approximated linearly by a product of system matrices.

The approach taken by ray differentials and our path Jacobians is to formulate the propagation of a ray as a single ray function that combines a series of surface interactions occurring during the process of ray tracing, and then compute the first-order Taylor approximation to this function by means of analytical differentiation. As we can see, instead of performing a linear approximation at each intersection as in pencil tracing, linearization only happens once and is applied to the whole path for both approaches based on Taylor expansion, thus leading to more accurate results.

2.5.2 Ray Differentials

Tracing ray differentials formulates a ray from the view point as a function parameterized in terms of its image coordinates on the image plane. The ray propagation is represented by a series of function compositions and the ray differential, which is the derivative of the resulting composite function with respect to the image plane, is traced using the chain rule. This derivative information may be used for estimating a texture filtering kernel [59].

Tracing ray differentials is similar to our work in that we both compute a first-order Taylor approximation to a ray path function and essentially obtain a Jacobian matrix characterizing the perturbation of intersection points. Moreover, as described in section 2.4, path Jacobians and ray differentials are based on the same geometric

quantities of the scattering surface, that is, normal and shape operator. The most significant point of departure is that the path functions involved in these two frameworks are parameterized differently. Specifically, instead of limiting the path to be a ray cast from a fixed eye point, we parameterize a path in terms of the positions of its end points, and compute the path Jacobian as its first-order derivative with respect to the varying end point. Our path function works for any ray segment with a fixed end. There are no particular physical meanings about both endpoints, which allows us to perturb either end of a ray traced path flexibly, leading to approximate analytical methods applicable in a variety of contexts, such as moving the view point, locating the reflection/refraction point of a varying scene point, etc. In addition, we also obtain a second order approximation for a ray by computing path Hessians. On the one hand, different perturbation variables of these two approaches lead to their different derivation strategies and application contexts. However, on the other hand, it is very interesting that the ray differential and the path Jacobian computed for the same path are closely related by an *inverse* relationship, which will be described in detail next. As a shorthand, we shall use FP to denote our approach of computing the path Jacobian from the Fermat Principle and the Implicit Function Theorem, and use TRD for tracing ray differentials.

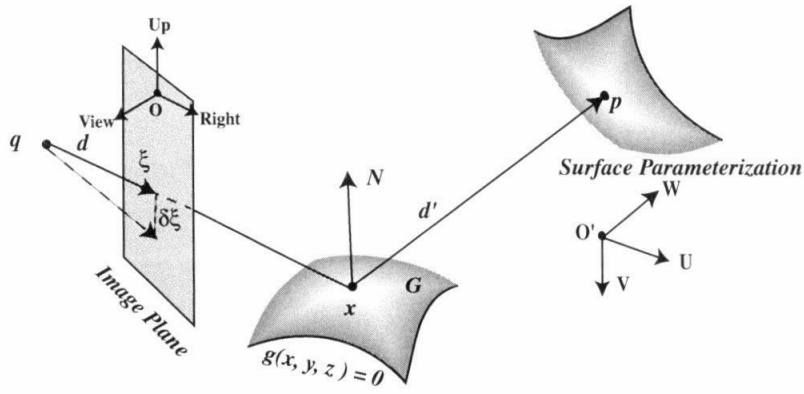


Figure 2.6: *An One-bounce Path and Local Coordinate Systems*

Consider an one-bounce specular reflection path $(\mathbf{q}, \mathbf{x}, \mathbf{p})$ shown in Figure 2.5.2. We should mention that TRD and FP have two subtle differences:

- All the surfaces involved in TRD are parametric patches, while original FP assumes

the reflecting surface $\mathbf{G}$ is implicitly defined by $g(\mathbf{x}) = 0$.

- TRD is specially designed for a traditional ray-tracing scenario, where the ray is cast from the eye point $\mathbf{q}$, and $\mathbf{p}$ must lie on a diffuse surface to obtain a function representation to apply differentiation. Therefore, in order to apply TRD to a path with an end point $\mathbf{p}$ free in 3D, which is allowed in the FP framework, we need to introduce a virtual plane passing through $\mathbf{p}$ and perpendicular to the ray direction $\mathbf{d}'$.

For comparison purpose, we assume that $\mathbf{G}$ has both a parametric representation and an implicit definition and restrict $\mathbf{p}$ on a surface and $\mathbf{q}$ as a fixed view point. Next we shall show that TRD and FP yield two 2×2 Jacobian matrices ($\mathbf{J}_{trd}$ and $\mathbf{J}_{fp}$, respectively), which are inverse matrices with each other.

TRD framework casts a ray from $\mathbf{q}$ to an image plane point ξ. While tracing each step of ray propagation, we express $\mathbf{p}$ as a composite function in terms of ξ using Snell's law or direct transfer equation. By differentiating the resulting function with the chain rule, TRD essentially provides a mechanism to compute a Jacobian matrix $\dfrac{\partial \mathbf{p}}{\partial \xi}$, which captures how the intersection point $\mathbf{p}$ on the path changes with respect to the perturbation in the image plane point ξ. If we think of ξ and $\mathbf{p}$ as points in $\mathbb{R}^3$, $\dfrac{\partial \mathbf{p}}{\partial \xi}$ is a 3×3 Jacobian matrix. However, in the ray-tracing setting ξ only moves on the image plane, while $\mathbf{p}$ on a surface is associated with a 2D parameterization (u, v).[2] Therefore, ξ and $\mathbf{p}$ only have two degrees of freedom. For clarity, we use the symbols $\hat{\xi}(x, y)$ and $\hat{\mathbf{p}}(u, v)$ to represent ξ and $\mathbf{p}$ viewed in $\mathbb{R}^2$, respectively. Then TRD actually computes a 2×2 Jacobian matrix $\mathbf{J}_{trd} = \dfrac{\partial \hat{\mathbf{p}}}{\partial \hat{\xi}}$.

In the *FP framework*, we formulate an implicit system satisfied by $\mathbf{q}$, $\mathbf{x}$ and $\mathbf{p}$ from the Fermat Principle and Lagrange Multiplier method. By applying the Implicit Function Theorem to this implicit system, we can directly compute a 3×3 Jacobian matrix $\mathbf{J} = \dfrac{\partial \mathbf{x}}{\partial \mathbf{p}}$, which characterizes how the intersection point $\mathbf{x}$ changes with respect to the perturbation in the other endpoint $\mathbf{p}$. In terms of the path Jacobian $\mathbf{J}$, we can easily compute a 2×2 matrix $\dfrac{\partial \hat{\xi}}{\partial \hat{\mathbf{p}}}$, which is the inverse of $\mathbf{J}_{trd}$ and states how the image coordinates change

[2] For the application context of texture filtering [59], (u, v) are the texture coordinates.

with respect to the movement of the surface point. This conversion can be carried out by following two steps:

1. Since $\hat{\xi}$ is uniquely determined by intersecting the line connecting $\mathbf{q}$ and $\mathbf{x}$ with the image plane, we can easily derive

$$\left[\frac{\partial \hat{\xi}}{\partial \mathbf{p}}\right]_{2\times 3} = \left[\frac{\partial \hat{\xi}}{\partial \mathbf{x}}\right]_{2\times 3} \cdot \mathbf{J} \tag{2.58}$$

   from our path Jacobian $\mathbf{J}$ with the chain rule.

2. The restriction of $\mathbf{p}$ moving on the surface allows us to compute $\dfrac{\partial \mathbf{p}}{\partial \hat{\mathbf{p}}}$ depending on the surface parameterization chosen. Consequently, it follows from the chain rule that

$$\mathbf{J}_{fp} = \left[\frac{\partial \hat{\xi}}{\partial \hat{\mathbf{p}}}\right]_{2\times 2} = \left[\frac{\partial \hat{\xi}}{\partial \mathbf{x}}\right]_{2\times 3} \cdot \mathbf{J} \cdot \left[\frac{\partial \mathbf{p}}{\partial \hat{\mathbf{p}}}\right]_{3\times 2}. \tag{2.59}$$

It then follows from $\mathbf{J}_{trd} = {\mathbf{J}_{fp}}^{-1}$ that the ray differential $\mathbf{J}_{trd}$ is inversely related to the path Jacobian $\mathbf{J}$, indirectly through equation (2.59).

This inverse connection can be stated more formally by defining several mappings among $\xi, \hat{\xi}, \mathbf{p}, \hat{\mathbf{p}}, \mathbf{x}$, as shown in the diagram 2.7. Here $\hat{\xi}$, $\hat{\mathbf{p}}$ are 2D pixel coordinates and surface parameterization. ξ, $\mathbf{p}$ represent the same points in $\mathbb{R}^3$. The existence of f is guaranteed from IFT (the Implicit Function Theorem). By means of function composition, we attain two mappings that correspond to *forward ray tracing* and *backward ray tracing* respectively, that is:

$$\begin{aligned} \mathbf{T}_{forward} &= l_2 \circ h \circ {l_1}^{-1} \quad : \quad \hat{\xi} \to \hat{\mathbf{p}} \\ \mathbf{T}_{backward} &= l_1 \circ g \circ f \circ {l_2}^{-1} \; : \; \hat{\mathbf{p}} \to \hat{\xi}, \end{aligned}$$

where the inverse ${l_1}^{-1}$ and ${l_2}^{-1}$ are guaranteed by assuming that l_1 and l_2 are bijective mappings. Accordingly, we may define two identity maps from $\mathbf{T}_{forward}$ and $\mathbf{T}_{backward}$

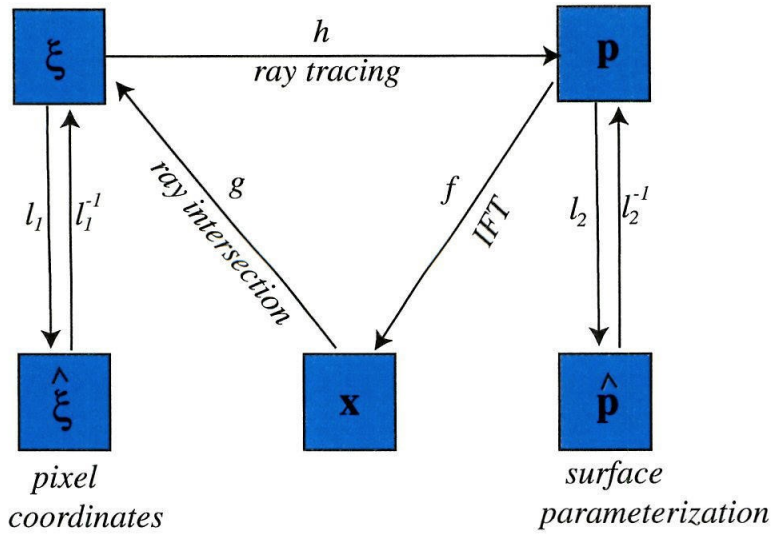


Figure 2.7: *Mappings used for computing* $\mathbf{J}_{trd}$ *and* $\mathbf{J}_{fp}$. *The label associated with each mapping suggests the corresponding procedure represented.*

as follows:

$$\begin{aligned} \mathbf{I}_1 &= \mathbf{T}_{forward} \circ \mathbf{T}_{backward} \quad : \quad \hat{\mathbf{p}} \rightarrow \hat{\mathbf{p}} \quad (1) \\ \mathbf{I}_2 &= \mathbf{T}_{backward} \circ \mathbf{T}_{forward} \quad : \quad \hat{\xi} \rightarrow \hat{\xi} \quad (2). \end{aligned}$$

Differentiating both sides of (1) ((2)) above with respect to $\hat{\mathbf{p}}$ ($\hat{\xi}$), we have

$$\begin{aligned} \frac{\partial \mathbf{T}_{forward}}{\partial \hat{\xi}} \cdot \frac{\partial \mathbf{T}_{backward}}{\partial \hat{\mathbf{p}}} &= \mathbf{I}, \\ \frac{\partial \mathbf{T}_{backward}}{\partial \hat{\mathbf{p}}} \cdot \frac{\partial \mathbf{T}_{forward}}{\partial \hat{\xi}} &= \mathbf{I}, \end{aligned}$$

which have shown that

$$\left[\frac{\partial \mathbf{T}_{forward}}{\partial \hat{\xi}}\right] = \left[\frac{\partial \mathbf{T}_{backward}}{\partial \hat{\mathbf{p}}}\right]^{-1}.$$

Here the matrices on both sides are related to the Jacobians from TRD and FP by

$$\begin{aligned} \frac{\partial \mathbf{T}_{forward}}{\partial \hat{\xi}} &= \mathbf{J}_{trd} \\ \frac{\partial \mathbf{T}_{backward}}{\partial \hat{\mathbf{p}}} &= \mathbf{J}_{fp} = l_1' \cdot g' \cdot \mathbf{J} \cdot l_2'^{-1}. \end{aligned}$$

In fact, ray differentials and path Jacobians are computed by differentiating a given path with respect to the perturbations at two opposite ends. Particularly, ray differentials provide a first-order approximation to the family of rays due to perturbations of the image plane point ξ and thus ray direction $\mathbf{d}$, which are closer to the fixed eye point. Alternatively, path Jacobians approximates a bundle of rays due to perturbations of the other end that is away from the eye. These two mechanisms are not designed for replacing one with the other. Instead, different motivations underlying them find them appropriate applications. For example, ray differentials were computed to estimate the footprint of texture mapping for anti-aliasing [59], while path Jacobians were combined with the *virtual object method* for interactive specular reflection [27, 28]. In either of these contexts, it is not natural to change one to the other.

2.6 Simulation of Lens Effects

To illustrate the power of our specular path perturbation formula, we designed an algorithm of approximating highly accurate and dynamic specular reflections on curved surfaces for scenes consisting of only diffuse and specular surfaces [28]. Using the formulas (2.19), (2.33) and (2.36) that we have derived for a general ray-traced path in section 2.2, we may extend the nearly same perturbation-based algorithm to scenes with refractive surfaces. In this section we shall examine a practical application of the machinery for perturbing a refraction path, that is, simulating the lens effect through a refractive glass. As a recap, we shall briefly review the algorithm developed for interactive specular reflections, and explain some adaptations we made to accommodate refraction. A more detailed description of our previous algorithm can be found in the paper [28].

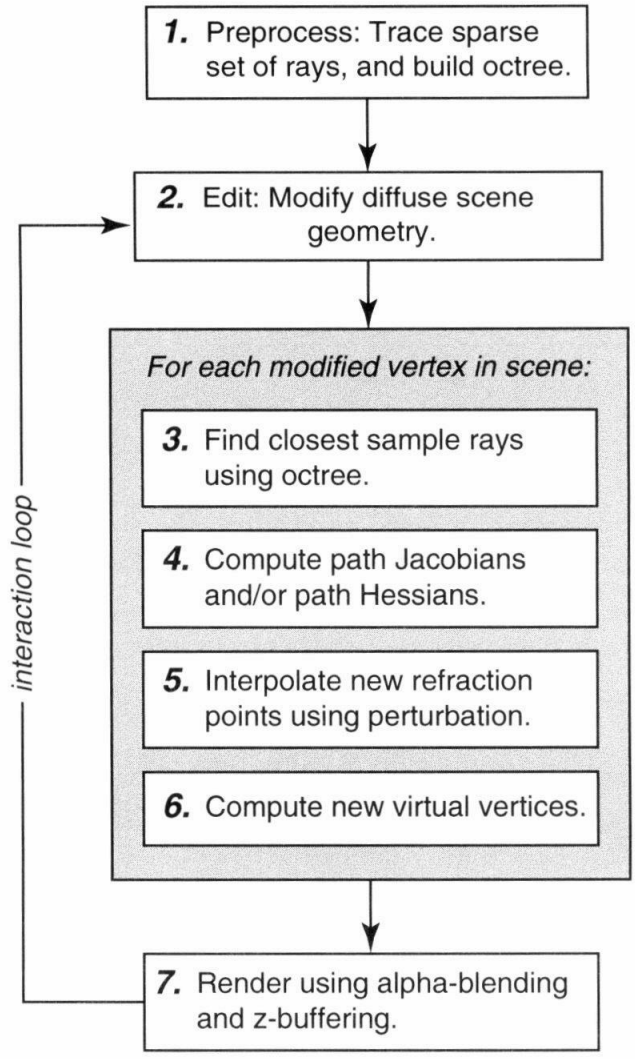


Figure 2.8: *The flowchart of the algorithm for simulating lens effects using perturbation.*

2.6.1 Algorithm Description

The algorithm depicted here is quite similar to that for fast computation of specular reflections, with some minor changes to adapt a refracting surface rather than a mirror surface. As usual, each surface must be tessellated into polygons, and each refracting surface must be equipped with a corresponding implicit equation and a ray intersection procedure. The outline of the algorithm is shown in Figure 2.8. We now summarize the

major steps.

Using standard ray tracing, a sparse set of rays with respect to a given vantage point is traced through an environment consisting of only the static refractive surfaces found in the original scene. The resulting ray paths are stored in a hierarchical data structure (such as an octree) that enables fast searching later. Then, based on these pre-computed refraction paths, the perturbation formulae are employed to interpolate new refraction paths reaching each vertex of each diffuse object found in the original scene. That is to say, the refraction of each object vertex in each refractor is approximated by perturbing the nearest known sample path. For a solid refractor, two refraction events occur at its boundary, thus the formula (2.33) for a multiple-bounce path is employed to compute path Jacobians. Here, only the linear approximation is computed for efficiency. Of course, quadratic approximation using path Hessians can be used for higher accuracy.

Using these perturbed refraction paths associated with object vertices, the algorithm approximates refractions close to the eye at the object level, rather than forming them pixel by pixel as in ray tracing. For each refracted object, its image in a curved refractor is rendered by constructing the associated *virtual object*. We shall explain the formula to compute virtual images for refractions in the next section. From the refraction point interpolated for each object vertex, a virtual vertex is computed and then the virtual object corresponding to each refraction is created by connecting the virtual vertices. The virtual object thus constructed may be positioned between the refractor and the diffuse object or behind the object, depending on the curvature of the refracting surface and the ratio of refractive indices $\eta = \eta_1/\eta_2$, where η_1, η_2 denote the refractive index for the refractor material and the air, respectively. Finally, the entire scene, consisting of both real and virtual objects, is rendered in a single pass through a standard graphics pipeline, where refractivity is simulated using alpha-blended transparency to "merge" the refraction onto real objects. By preserving the optical path length of virtual objects, hidden surface removal and relative visibility are correctly handled by z-buffering. This virtual object method is similar to the approach described by Ofek and Rappoport [85], where virtual objects with respect to reflections were computed.

One complication that arises with this approach is that the shading of diffuse virtual

objects must be computed with respect to the original positions of the refractions, not their virtual coordinates. This is because the orientation of the virtual objects with respect to the light sources will in general be quite different from that of the actual object. Consequently, the shading must be pre-computed, when the virtual vertices are constructed, and not shaded by the graphics pipeline.

By exploiting path coherence from frame to frame, this perturbation approach shows a great advantage over ray tracing while dynamically moving the diffuse object.

2.6.2 Virtual Image from Refraction

As we demonstrated in the example of interactive specular reflection, the reflection of an object can be rendered by constructing its virtual image with respect to the reflective surface. A similar idea can be applied to the refraction image. For reflection, the relation between the original object and its virtual object is quite simple, a virtual vertex is placed at a distance from the eye that is equal to the original optical path length from the eye to the vertex. It becomes a little complicated for refraction, where the location of the virtual image is dependent on the refractive angle.

Consider a situation that an observer B in a medium of index of refraction η_2 looks through a plane surface OX at an object point A situated at a distance Y below OX in an adjacent medium of index η_1. Given a refraction path from A to B, hitting the surface OX at R, we would like to derive the location of the virtual image corresponding to R. See Figure 2.9. A full analysis of this situation shows that the image seen by the observer is actually astigmatic [16]. Imagine that the ray AR is the central ray along the axis of a small cone of rays leaving the object A, the location of the virtual image of A is found by determining the intersection of a pair of rays from the object that reach the eye of an observer B. The virtual image formed by light leaving an object point with *vertical divergence* will be located at a different position than the virtual image formed by light leaving an object point with *horizontal divergence*. Pairs of rays with vertical divergence refer to those included in the small cone that have the same incident angle as AR, while pairs of rays with horizontal divergence refer to rays in the same plane of incidence as AR and characterized by slightly different values of the incident angle θ_1

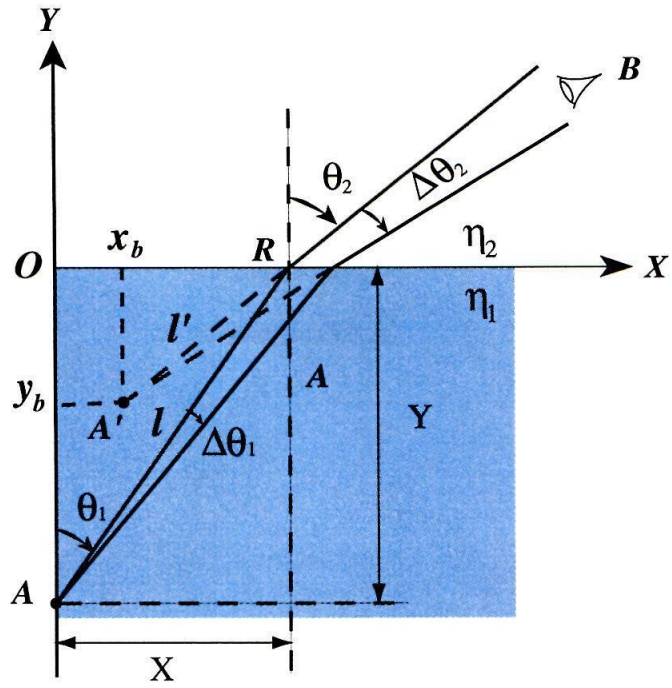


Figure 2.9: *The coordinate geometry of virtual image due to refraction. The X axis represents the interface between two medias with the index of refraction, η_1 and η_2. A point object A is located at a distance Y from the interface, rays of light from A travel through the medium and reach the observer at B, with a virtual image at A' with coordinates (x_b, y_b).*

of AR. Here, we are constructing the virtual image corresponding to the latter pairs of rays. Our derivation follows that proposed by Bartlett [16], but we are using a different quantity to characterize the position of the virtual image.

We examine two rays leaving the object A and striking the interface at angles of incidence θ_1 and $\Delta\theta_1$, passing into medium 2 with angle of refraction θ_2 and $\Delta\theta_2$, respectively. Extending backward these two rays will converge at a site A' identified as the virtual image of A. The position of A' can be uniquely specified by the length $l' = |A'R|$. Let $l = |AR|$, next we shall show that the ratio between l' and l satisfies an identity depending on the refractive angle θ_2 and the refractive indices η_1, η_2.

According to the coordinate geometry in Figure 2.9, the refracted ray RB is a straight

line, which can be represented as

$$y = mx + b,$$

where

$$m = \cot\theta_2, \tag{2.60}$$

$$b = -mX = -X\cot\theta_2, \tag{2.61}$$

and X is the distance from R to the origin O. Thus, the equation of the ray RB is

$$y = \frac{x - X}{\tan\theta_2}. \tag{2.62}$$

Distinguishing the ray RB and its deviated ray by the subscript e and f, we can compute their intersection $A'(x_b, y_b)$ from equation (2.62) as

$$\begin{aligned} x_b &= -\frac{b_e - b_f}{m_e - m_f} \\ y_b &= \frac{x_b - X}{\tan\theta_2}. \end{aligned} \tag{2.63}$$

In the limiting case of small deviations from RA, i.e. small changes in m, b, θ_2, x_b can be written as

$$x_b = -\frac{db}{dm} = -\frac{db}{d\theta_2} \Big/ \frac{dm}{d\theta_2}. \tag{2.64}$$

It follows from equation (2.60) that

$$\frac{dm}{d\theta_2} = -\csc^2\theta_2. \tag{2.65}$$

Note that $X = Y\tan\theta_1$ from Figure 2.9, we have

$$b = -Y\frac{\tan\theta_1}{\tan\theta_2}.$$

By means of Snell's law of refraction, that is,

$$\eta_1 \sin\theta_1 \;=\; \eta_2 \sin\theta_2,$$

we can express b in terms of the refractive angle θ_2

$$b \;=\; -Y\frac{\cos\theta_2}{\sqrt{\eta^2-\sin^2\theta_2}}, \tag{2.66}$$

where $\eta = \eta_1/\eta_2$. Differentiating equation (2.66) with respect to θ_2, we attain

$$\frac{db}{d\theta_2} \;=\; \frac{Y(\eta^2-1)\sin\theta_2}{\sqrt{(\eta^2-\sin^2\theta_2)^3}}. \tag{2.67}$$

Combining equations (2.64), (2.65) and (2.67), we get

$$\frac{x_b}{Y} \;=\; \frac{(\eta^2-1)\sin^3\theta_2}{\sqrt{(\eta^2-\sin^2\theta_2)^3}}. \tag{2.68}$$

Equations (2.68) and (2.63) yield

$$\frac{y_b}{Y} \;=\; -\frac{\eta^2\cos^3\theta_2}{\sqrt{(\eta^2-\sin^2\theta_2)^3}}. \tag{2.69}$$

It follows from the geometry shown in Figure 2.9 that

$$\begin{aligned}\frac{|y_b|}{l'} &= \cos\theta_2\\ \frac{Y}{l} &= \cos\theta_1.\end{aligned}$$

Therefore, we have the ratio

$$\begin{aligned}\frac{l'}{l} &= \frac{|y_b|}{Y}\cdot\frac{\cos\theta_1}{\cos\theta_2}\\ &= \frac{\eta^2\cos^3\theta_2}{\sqrt{(\eta^2-\sin^2\theta_2)^3}}\cdot\frac{\sqrt{\eta^2-\sin^2\theta_2}}{\eta\cos\theta_2}\end{aligned}$$

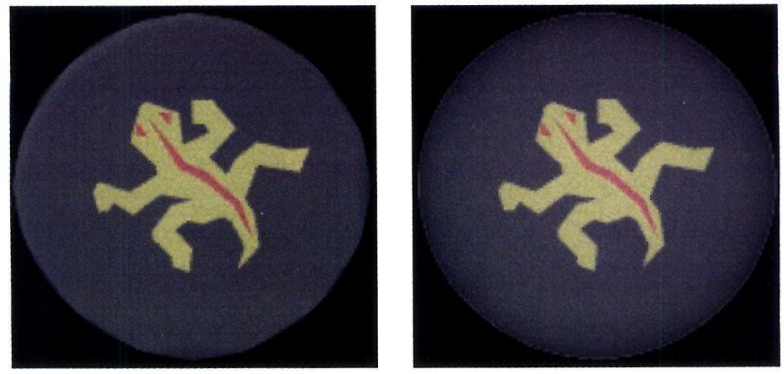


Figure 2.10: *Side-by-side comparison of the refraction images of the lizard through a glass lens, generated by perturbation method (left) and ray tracing (right).*

$$= \frac{\eta \cos^2 \theta_2}{\eta^2 - \sin^2 \theta_2}. \tag{2.70}$$

Equation (2.70) suggests that the virtual point A' can be found by extending backward the ray RB by a distance $\frac{\eta \cos^2 \theta_2}{\eta^2 - \sin^2 \theta_2} l$.

2.6.3 Results and Discussions

Based on the new path perturbation formula, in combination with the method for constructing virtual images from refractions discussed in the previous section, we have implemented our perturbation algorithm shown in Figure 2.8 using Open Inventor on a SGI Indigo2 to simulate the refractive effect through a lens. In our experimentation, the scene is composed of a glass lens and a diffuse lizard-shaped polygon underneath, and the lens is approximated by a squashed sphere of radius 1.5 (scaled by 0.3 in the direction perpendicular to the paper). That is, an ellipsoid with an implicit definition $g(x, y, z) = 0$, where

$$g(x, y, z) = \frac{x^2}{1.5^2} + \frac{y^2}{1.5^2} + \frac{z^2}{0.45^2} - 1.$$

Figure 2.10 shows a side-by-side comparison of the refraction images of the lizard seen through the glass lens, generated respectively by our perturbation method and ray tracing (using the commercial software PovRay). The left image is generated by perturbing the

Perturbed (error) Ray-traced

Figure 2.11: *A sequence of images generated by perturbation method while moving the lizard to the right. Error occurs when it is close to the lens boundary.*

sample rays cast every 8th pixel along each horizontal and vertical scanline, and the results are almost identical.

In Figure 2.11 we show a sequence of refraction images generated by our perturbation method while we dynamically move the lizard to the right. Because refraction transformation is quite non-linear, the linear perturbation based on refractive path Jacobians will cause error when the lizard is close to the lens boundary, as demonstrated in the bottom row of Figure 2.11. The error is also caused by the way we used to construct the virtual object. As seen from the ray-traced image in the bottom, the vertices of the lizard leg tip may have multiple refractions on the lens, while other vertices only generate one refraction. However, when we construct the virtual object by connecting virtual vertices corresponding to each object vertex, we actually have made an assumption that all the object vertices should have the same number of refractions on the lens. Other-

wise, the current algorithm doesn't know how to correctly separate different refractions corresponding to a single vertex, thus will fail by mixing one refraction with another. That is the case here. To solve this problem, we have to break the diffuse object into different parts, where vertices belonging to one part should generate a fixed number of refractions on the lens. Then the challenge becomes: how could we determine in advance the number of refractions that an object point will generate on a lens? Answering this question requires a detailed analysis from optics, which is the future direction we would like to pursue.

2.7 Potential Applications

Our path perturbation theory, as well as the work of tracing ray differentials and its extension to path differentials by Suykens and Willems [109], provides mathematical foundations for several perturbation methods in image synthesis. Perturbation methods of this nature can find many other potential applications in image generation. Moreover, they may also give some insights to practical solutions to some inverse modeling and rendering problems known in the computer vision community. In this section we shall point out such possibilities, which are future directions worthy of further investigation.

Accelerating Geometry-Based Rendering: An important application area of path perturbation is incremental rendering to accelerate some existing geometry-based rendering algorithms. As we mentioned in Section 1.2, perspective coherence between our left and right eye views has been exploited for stereoscopic rendering [3], using the technique of image space reprojection through a 2D warping equation, which, however, cannot handle diffuse environments. Due to the proximity of human eyes, strong path coherence exist between stereoscopic image pairs, thus our path perturbation formula can be directly applied to calculate both images at the cost of tracing from a single eye position. This optimization is very useful in the context of responsive workbench [65]. The idea of path perturbation may also be used to speed up the computation of caustics in the approach proposed by Mitchell and Hanrahan [74]. Instead of computing the reflecting path for every pixel on the

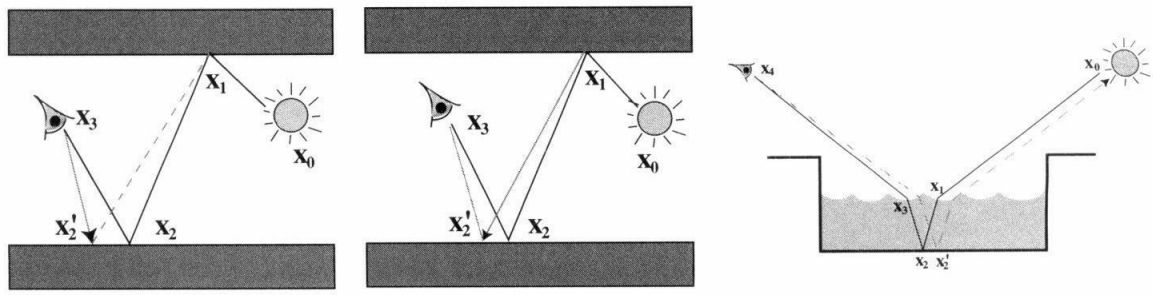
floor using costly interval analysis and automatic differentiation, we could sparsely sample it and interpolate the remaining paths using perturbation.

Image-Based Rendering: To circumvent the difficulties of traditional modeling and intimidating computational cost suffered by geometry-based method, image-based modeling and rendering generate model or new images from other images rather than geometric primitives. However, normal image warping and view interpolation will cause artifacts with the existence of specular or glossy surfaces. Currently, image-based rendering of non-diffuse environments is still an open question. As demonstrated by Gortler et. al [48], the accuracy of image-based approaches can be greatly enhanced by incorporating approximate knowledge of the underlying geometry. An obvious extension of this idea is to augment the geometric information with first-order approximations of prominent optical effects, such as coherent reflections. From the closed-form expression derived for the path Jacobian, we may consider computing the *image Jacobian*, which approximates how pixels moves with respect to the view change and can be obtained by projecting the path Jacobian into the image plane. Consequently, by differentiating an image with respect to the viewpoint, one can approximate a neighborhood of similar views to first-order accuracy. The remaining hurdles include the computation of image Jacobians from approximate geometry, gracefully handling image discontinuities, and exploring accuracy/efficiency tradeoff, etc. This image differentiation approach could be applied to image warping where specular reflection/refraction effects are prominent, as in the work of Lischinski and Rappoport [68].

Monte Carlo Methods: A new approach to solving the light transport problem, called *Metropolis Light Transport* (MLT), was recently developed by Veach and Guibas [112], inspired by the Metropolis sampling method in computational physics [73]. The basic idea behind MLT is that a sequence of light-carrying paths through the scene is computed, with each path generated by mutating the previous path in some manner. Each mutation is accepted or rejected with some carefully chosen probability, to ensure that the overall distribution of sampled paths in the

scene is proportional to their contribution to the image being generated.

Compared to previous unbiased Monte Carlo approaches to image synthesis, MLT has two important advantages embodied in its mutation strategies: *path reuse*, where paths are often constructed using some of the segments of the previous one; and *local exploration*, by favoring mutations that make small changes to the current found path that make large contributions to the final image. It is these key features that make MLT stand out for its efficiency and robustness, amortizing the expense of finding difficult paths over many samples. More importantly, MLT places relatively few restrictions on the types of mutations that can be applied and thus the mutation set is easily extended. It is often possible to handle difficult lighting problems efficiently by designing a specialized mutation by exploiting various kinds of coherence in the scene. In the original paper of MLT [112], the authors proposed three perturbations to handle some lighting situations where a small fraction of paths contribute much more than average, for example, due to caustics. They are respectively *lens perturbation*, *caustics perturbation*, and *multi-chain perturbation*, as shown in Figure 2.12. Due to specular scatterings, these random perturbations may cause an excessive number of rejections before finding a valid path. Take the two-chain perturbation shown in Figure 2.12c for an example. First, the lens edge $\mathbf{x}_3\mathbf{x}_4$ is perturbed to generate a point $\mathbf{x}_2'$ on the pool bottom. Then, the direction from the original point $\mathbf{x}_2$ toward the light source is perturbed, and a ray is cast from $\mathbf{x}_2'$ in this direction. Because of the specular refraction on the water surface, very frequently, the new ray cast in the perturbed direction will miss a point (or small area) light source $\mathbf{x}_0$, and get rejected. The consequence is the slow convergence rate and high variance due to repeated samples in the Markov chain. Another drawback of these random perturbations is that the intersection points on the new path must be obtained through ray casting, which may be costly. However, these situations may be greatly optimized by using analytical path perturbations, which are characterized by such properties as high acceptance probability, low cost, and easy changes of the image locations, and so on. The framework of path Jacobians, as well as ray differentials and path differentials, provide mathematical tools



(a)Lens Perturbation (b)Caustic Perturbation (c)Multi-chain Perturbation

Figure 2.12: *Three perturbations proposed by Veach and Guibas: both lens perturbation (left) and caustics perturbation (middle) make small changes to subpaths including the lens edge, but from different ends. The former casts a ray at a new image location to generate* $\mathbf{x}_2'$*, good for subpaths of the form* $(L|D)DS^*E$*, while the later perturbs the direction of the ray segment away from the eye to get* $\mathbf{x}_2'$*, and works for subpaths of the form* $(D|L)S^*DE$*. Multi-chain perturbation (right) handles other paths of the form* $(D|L)DS^*DS^*DE$ *by perturbing the path through more than one specular chain, such as caustics seen through a specular surface. (Courtesy of Veach and Guibas [112])*

to enhance existing perturbation strategies and design new low-variance mutation strategies in the context of MLT.

In detail, MLT may benefit from our and others' work on path perturbation in the following aspects:

- Improve the efficiency of current mutations by regenerating the new path analytically instead of casting rays. For example, to replace the lens subpath $\mathbf{x}_t \cdots \mathbf{x}_k$ of the form $(L|D)DS^*E$[3], lens perturbation perturbs the old image location by moving it a random distance in a random direction and then traces the subpath from the new image location through additional specular bounces to be the same length as the original. Using our path Jacobians (if we require that $\mathbf{x}_{t+1}$ is fixed) or ray differentials, the new subpath can be determined

[3]This is the regular notation for paths, where S, D, E, and L stand for specular, non-specular (diffuse or glossy), lens and light vertices respectively.

immediately to the first order accuracy, without tracing any further paths. Similar speedup may be obtained for caustic and multi-chain perturbations using path differentials for glossy surfaces.

- Design new perturbation strategies and machinery. Currently, caustic perturbations and multi-chain perturbations are carried out by choosing a new direction at each non-specular vertex by perturbing the corresponding direction of the original subpath. Inspired by the motivation underlying our path perturbation formula, we may easily extend these mutations to perturb the positions of non-specular vertices as well, which allows us to better explore the local space around an important path that is hard to find. The goal of designing better and more efficient MLT mutation strategies also motivates us to develop new mathematical models to exploit a wide range of coherence in the path space. For example, to accommodate caustic and multi-chain perturbation, we are required to formulate the other kind of coherent reflection, as illustrated in Figure 2.13, where the family of similar paths are parameterized by the deviation angle θ. If we could work out the differential relationship between the location of the reflection/refraction points and θ, some mutations may be performed very efficiently.

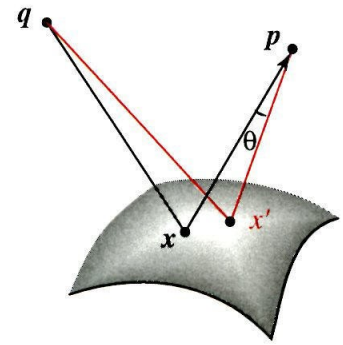


Figure 2.13: *Another kind of specular path coherence, where both endpoints are fixed and the path gets changed while the ray direction from one end is slightly perturbed. The corresponding reflection points* $\mathbf{x}'$ *and* $\mathbf{x}$ *will converge when the angle* θ *is small.*

- Increase the acceptance probability. As we mentioned before, random perturbation leads to frequently rejections in some cases (i.e.,missing the light or the lens), while staying in the same state for many samples in a row causes increased variance. Due to the constraint of a fixed endpoint required by our path function formulation, we may be able to generate new perturbed paths with more chances to be accepted in certain circumstances. For instance, in the second step of the two-chain perturbation example shown in Figure 2.12c, we may require that the light source $\mathbf{x}_0$ is fixed and use path Jacobians computed for the original path to locate the new refraction point $\mathbf{x}_1'$ on the water surface. Compared to perturbing the direction from $\mathbf{x}_2'$, this new mutation definitely yields paths with higher acceptance chances.

Computer Vision: One of the main tasks in computer vision is computing the shape of objects [21]. A number of cues, notably stereoscopic disparity, texture gradient, motion parallax, contours and shading, have been shown to carry valuable information on surface shape. However, the study of shading has been mostly restricted to the Lambertian case, and specularities have been mostly ignored with the exception of the analysis of specular highlight [55, 51]. The machinery of path perturbation brings up a possibility of recovering information on the shape of a surface from the specular component of its reflectance function, which is being explored recently by Savarese and Perona [95]. They proposed an analysis of the relationship between the image of a calibrated scene composed of patterned lines through a point, and the geometry of a curved mirror surface on which the scene is reflected. Their main result is that it is possible to reconstruct the position and shape of the mirror surface up to the second order from the orientation and curvature of the images of two intersecting lines. The key idea is to study the orientation and curvatures of the "distorted" curves of the patterned lines on the image plane, which are then used as constraints to solve for the surface parameters. During their analysis and derivation, our theory of specular path perturbation plays an important role.

Chapter 3

Irradiance due to Linearly-Varying Luminaires

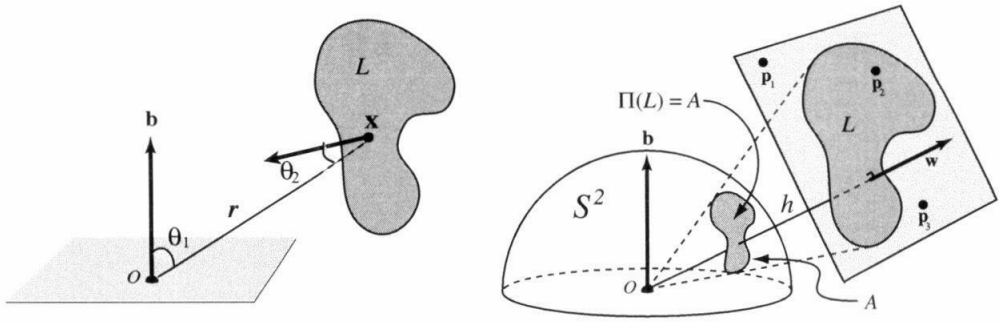
The determination of radiant interchange between a finite surface and an infinitesimal area is one of the most ubiquitous computings in radiative heat transfer. As mentioned in Chapter 1, these transfers correspond to point-to-area form-factors [18] in *diffuse piecewise-constant* environments. The present investigation is concerned with calculating the illumination of an infinitesimal area due to emittance from an *inhomogeneous* Lambertian surface, which may be used to exactly compute element interactions in the context of higher-order finite element methods [111].

In this chapter we present a closed-form expression for the irradiance at a differential surface element due to an arbitrary polygonal Lambertian luminaire in which the radiant exitance is *linearly* varying with position over the surface of the luminaire. The solution consists of elementary functions and a single well-behaved special function that can be either approximated directly or computed exactly in terms of classical special functions such as Clausen's integral or the closely related dilogarithm. We first provide a general boundary integral that applies to planar luminaires of arbitrary shape and linearly-varying radiant exitance, and then derive the *closed-form* expression for the restricted case of arbitrary polygons, which is the result most relevant for global illumination. Our approach is to express the problem as an integral of a simple class of rational functions

over regions of the sphere, and to convert the surface integral to a boundary integral using a recurrence formula derived for a generalization of axial moments known as *triple-axis moments*.

The closed-form expression presented in this chapter extends the class of available closed-form expressions for computing direct radiative transfer from finite areas to differential areas. One direct application of our result is the exact computation of transfers from linear polygonal elements to differential areas in the context of a collocation-based radiosity method [111, 72]. However, the theoretical contribution of this work is somewhat broader as it introduces much of the machinery needed to derive closed-form solutions for the general case of luminaires with radiance distributions that vary polynomially in both position and direction. As we shall demonstrate in chapter 5, all cases of polynomial variation admit similar closed-form solutions, which rest on the same special function that emerges in this linear case.

The remainder of this chapter is organized as follows. Section 3.1 formulates the problem of computing the irradiance at a point due to a linearly-varying Lambertian luminaire as integrating a simple class of rational functions over spherical regions. To convert the resulting surface integral to a boundary integral representation, we extend *axial moments* reviewed in section 3.2 to accommodate this type of integral. Based on the recurrence formula derived for this generalization, we derive a general boundary integral formula for the irradiance integral using Taylor expansion and a variety of binomial identities in section 3.3. Finally, by restricting this general formula that applies to planar linearly-varying luminaires with arbitrary shapes to polygons, we obtain a closed-form expression for the irradiance at a point due to a polygonal Lambertian luminaire with linearly-varying radiant exitance in section 3.4.



(a) (b)

Figure 3.1: *(a) Geometrical configuration for computing the irradiance from an area to a point. (b) The irradiance at* **o** *due to a planar figure L with linearly varying radiant exitance is computed by a boundary integral. The radiant exitance variation is uniquely determined by any three non-collinear points* $\mathbf{p}_1$, $\mathbf{p}_2$, *and* $\mathbf{p}_3$ *on the luminaire plane.*

3.1 Luminaires with Linearly-Varying Radiant Exitance

The irradiance Φ impinging on a surface at the point $\mathbf{o}$ due to a planar luminaire L is given by

$$\Phi(L) = \frac{1}{\pi} \int_L \phi(\mathbf{x}) \frac{\cos\theta_1 \cos\theta_2}{r^2} \, d\mathbf{x}, \tag{3.1}$$

where $\mathbf{x}$ is a point on L, r is its distance from $\mathbf{o}$, and $\phi(\mathbf{x})$ denotes the *radiant exitance* of the luminaire at the point $\mathbf{x}$; thus, $\phi(\mathbf{x})/\pi$ is the radiance of the surface at $\mathbf{x}$, which is constant in all directions. Here θ_1 denotes the angle between $\mathbf{x}-\mathbf{o}$ and the receiver normal $\mathbf{b}$, and θ_2 denotes the angle between $\mathbf{x} - \mathbf{o}$ and the luminaire normal. See Figure 3.1a. Without loss of generality, we shall assume that $\mathbf{o}$ is the origin throughout the remainder of the chapter. If we integrate over the spherical projection of L, denoted by $A = \prod(L)$, we have

$$\begin{aligned} \Phi(A) &= \frac{1}{\pi} \int_A \phi(\mathbf{x}) \cos\theta_1 \, d\sigma(\mathbf{u}) \\ &= \frac{1}{\pi} \int_A \phi(\mathbf{x}) \, \langle \mathbf{b}, \mathbf{u} \rangle \, d\sigma(\mathbf{u}), \end{aligned} \tag{3.2}$$

where $\mathbf{u}$ is the unit vector on the sphere, and $\mathbf{x}$ is the point on L in the direction of $\mathbf{u}$. The measure σ denotes area on the sphere.

The class of luminaires that we are considering here are those that are planar and with linearly-varying radiant exitance, or equivalently, for which the function $\phi : \mathbb{R}^3 \to \mathbb{R}$ is linear. Thus, if $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ are three non-collinear points on L with radiant exitance w_1, w_2 and w_3, respectively, then $\phi(\mathbf{x})$ can be expressed as a linear combination of w_1, w_2 and w_3:

$$\phi(\mathbf{x}) = [w_1\ w_2\ w_3]\,\hat{\mathbf{x}}, \tag{3.3}$$

where $\hat{\mathbf{x}}$ is barycentric coordinate vector of $\mathbf{x}$ with respect to $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$; that is,

$$\hat{\mathbf{x}} = [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]^{-1}\,\mathbf{x},$$

where $[\mathbf{p}_1\ \ \mathbf{p}_2\ \ \mathbf{p}_3]$ is the matrix with $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ as its columns. Thus, equation (3.3) becomes

$$\phi(\mathbf{x}) = [w_1\ w_2\ w_3]\,[\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]^{-1}\,\mathbf{x}. \tag{3.4}$$

Let $\mathbf{w}$ denote the unit vector orthogonal to the planar luminaire, and let h denote the distance from the origin to the plane containing the luminaire, as shown in Figure 3.1b. By expressing $\mathbf{x}$ in equation (3.4) in terms of the unit vector $\mathbf{u}$,

$$\mathbf{x} = \frac{h}{\langle \mathbf{w}, \mathbf{u} \rangle}\mathbf{u}, \tag{3.5}$$

we have

$$\phi(\mathbf{u}) = \frac{\langle \mathbf{a}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle},$$

where the vector $\mathbf{a}$ is given by

$$\mathbf{a} = h \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{bmatrix}^{-1}. \tag{3.6}$$

Consequently, the irradiance formula (3.2) becomes the integral of a simple rational function over A. Specifically,

$$\Phi(A) = \frac{1}{\pi} \int_A \frac{\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \, d\sigma(\mathbf{u}), \tag{3.7}$$

where $\mathbf{w}$ is the unit vector orthogonal to the planar luminaire, the constant vector $\mathbf{a}$ given in equation (3.6) encodes the linear variation of radiant exitance, and $\mathbf{b}$ is the receiver normal at the origin. Here, $\langle \cdot, \cdot \rangle$ denotes the standard inner product.

As for a receiver point $\mathbf{q}$ other than the origin $\mathbf{o}$, the vector $\mathbf{a}$ can be computed by translating the luminaire by $-\mathbf{q}$, that is

$$\mathbf{a}(\mathbf{q}) = h \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 - \mathbf{q} & \mathbf{p}_2 - \mathbf{q} & \mathbf{p}_3 - \mathbf{q} \end{bmatrix}^{-1}. \tag{3.8}$$

However, this translation method is inefficient since it requires inverting a different matrix at each receiver point. In fact, the vector argument $\mathbf{a}$ corresponding to an arbitrary receiver point $\mathbf{q}$ can be updated immediately from that computed for the origin. Let

$$\mathbf{a}_0 = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{bmatrix}^{-1},$$

it follows from equation (3.4) that the radiant exitance distribution function may be expressed as

$$\phi(\mathbf{x}) = \langle \mathbf{a}_0, \mathbf{x} \rangle \tag{3.9}$$

for a point $\mathbf{x} \in L$, which is denoted by its coordinates with respect to the coordinate system located at the origin $\mathbf{o}$. By rewriting $\mathbf{x} = \mathbf{q} + (\mathbf{x} - \mathbf{q})$, we have

$$\phi(\mathbf{x}) = \langle \mathbf{a}_0, \mathbf{q} \rangle + \langle \mathbf{a}_0, \mathbf{x} - \mathbf{q} \rangle .$$

It then follows from equation (3.1) that the irridance at $\mathbf{q}$ due to the linearly-varying

luminaire is given by

$$\begin{aligned}\Phi(L,\mathbf{q}) &= \frac{1}{\pi}\int_L \phi(\mathbf{x})\frac{\cos\theta_1\cos\theta_2}{r^2}\,d\mathbf{x} \\ &= \frac{1}{\pi}\left[\int_L \langle \mathbf{a}_0,\mathbf{q}\rangle\,\frac{\cos\theta_1\cos\theta_2}{r^2}\,d\mathbf{x} + \int_L \langle \mathbf{a}_0,\mathbf{x}-\mathbf{q}\rangle\,\frac{\cos\theta_1\cos\theta_2}{r^2}\,d\mathbf{x}\right]. \quad (3.10)\end{aligned}$$

Now we integrate the two terms on the right hand side of equation (3.10) over the projection $A = \prod(L)$ of L onto the hemisphere around $\mathbf{q}$. Since $\mathbf{x}-\mathbf{q}$ represents the position vector on the luminaire that is in the same direction $\mathbf{u}$ with respect to $\mathbf{q}$, the identity (3.5) should hold for $\mathbf{x}-\mathbf{q}$ and $\mathbf{u}$. Consequently, equation (3.10) becomes

$$\begin{aligned}\Phi(A,\mathbf{q}) &= \frac{1}{\pi}\left[\langle \mathbf{a}_0,\mathbf{q}\rangle\int_A \langle \mathbf{b},\mathbf{u}\rangle\;d\sigma(\mathbf{u}) + \int_A \left\langle \mathbf{a}_0,\frac{h(\mathbf{q})}{\langle \mathbf{w},\mathbf{u}\rangle}\mathbf{u}\right\rangle \langle \mathbf{b},\mathbf{u}\rangle\;d\sigma(\mathbf{u})\right] \\ &= \frac{1}{\pi}\left[\langle \mathbf{a}_0,\mathbf{q}\rangle\int_A \langle \mathbf{b},\mathbf{u}\rangle\;d\sigma(\mathbf{u}) + h(\mathbf{q})\int_A \frac{\langle \mathbf{a}_0,\mathbf{u}\rangle\,\langle \mathbf{b},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\,d\sigma(\mathbf{u})\right], \quad (3.11)\end{aligned}$$

where $h(\mathbf{q})$ represents the distance from $\mathbf{q}$ to the luminaire plane. By matching equation (3.11) with equation (3.7), we have shown that the vector $\mathbf{a}(\mathbf{q})$ in equation (3.8) associated with an arbitrary receiver point $\mathbf{q}$ can be equivalently expressed as

$$\mathbf{a}(\mathbf{q}) = \langle \mathbf{a}_0,\mathbf{q}\rangle\;\mathbf{w} + h(\mathbf{q})\,\mathbf{a}_0, \quad (3.12)$$

Compared to equation (3.8), the update formula (3.12) is more efficient since only one matrix inverse is required for a given linearly-varying luminaire.

Equation (3.7) suggests that the irradiance at a point due to a linearly-varying luminaire can be expressed as a simple class of rational functions integrated over regions on the sphere. Common to all methods for computing surface integrals of this nature is Stokes' theorem [106], by which surface integrals can be converted into boundary integrals. Such reductions are generally advantageous, as boundary integrals are better suited for computation [39] and frequently lead to closed-form solutions [9, 97]. We perform this reduction with an approach similar to that used by Arvo [9]; specifically, we extend the concept of axial moments to accommodate the rational functions appearing in the irradiance integral (3.7) and then derive a recurrence relation for this generalization,

which is then used to reduce the surface integral (3.7) to a one-dimensional integral. Furthermore, our recurrence formula for the generalization of axial moments is based on the fundamental expression for *irradiance tensors*. Therefore, as a background, we need to review the concepts of irradiance tensors and angular moments proposed by Arvo [9].

3.2 Background

3.2.1 Tensor Calculus and Differential Forms

To study irradiance tensors and angular moments, we shall be familiar with some elementary machinery from *tensor calculus* [64] and *differential forms* [96]. In this section we summarize some relevant facts that will be needed in the remainder of this thesis.

Tensors are an important tool in studying higher dimensional quantities. Scalars, vectors and matrices are just special tensors of order 0, 1 and 2. Tensors of higher order can be constructed from lower-order tensors by using tensor product operator, denoted by $\otimes$. The components of an n-tensor can be denoted by using either n subscripts or a single multi-index explained later. For notational convenience, Einstein *summation convention* is adopted to express tensor quantities, where summation is implicit over every index occurring twice in a term, including repeated multi-index. For example, given two vectors $\mathbf{A}, \mathbf{B} \in \mathbb{R}^3$, we have

$$\mathbf{A}_i \mathbf{B}_i = \sum_{i=1}^{3} \mathbf{A}_i \mathbf{B}_i. \tag{3.13}$$

This convention will not apply to sub-indices, however. By default, we shall assume this convention throughout this thesis.

Two of the fundamental tensors employed in tensor calculus are

1. The *Kronecker delta* δ_{ij} defined by

$$\delta_{ij} \equiv \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

2. The *Levi-Civita symbol* or *permutation symbol* ε_{ijk}, defined by

$$\varepsilon_{ijk} \equiv \begin{cases} 1 & \text{if (i,j,k) is an even permutation} \\ -1 & \text{if (i,j,k) is an odd permutation} \\ 0 & \text{if (i,j,k) is not a permutation} \end{cases} . \tag{3.15}$$

There exist several useful identities involving these two tensors, which will be employed frequently in the proofs that follows. They are:

$$\delta_{ij}\delta_{ij} = 3 \tag{3.16}$$

$$\delta_{ijk}\delta_{ljk} = 2\delta_{il} \tag{3.17}$$

$$\varepsilon_{pjl}\varepsilon_{kml} = \delta_{pk}\delta_{jm} - \delta_{pm}\delta_{jk} \tag{3.18}$$

$$\delta_{ij}\mathbf{A}_i\mathbf{B}_j = \langle \mathbf{A}, \mathbf{B} \rangle \tag{3.19}$$

$$\varepsilon_{ijk}\mathbf{A}_j\mathbf{B}_k = (\mathbf{A} \times \mathbf{B})_i, \tag{3.20}$$

where $\times$ means cross product of two vectors.

Differential forms provide a powerful formalism for representing multi-dimensional integrals over manifolds, it subsumes the classical differential operators in vector analysis, such as gradient, divergence, curl, etc. Higher-order forms can be formed from lower-order forms using an *exterior multiplication* operator $\wedge$, also called *wedge product*. The classical notation $dx_i dx_j$ for integration is replaced by $dx_i \wedge dx_j$, which is *anti-commutative*,

$$dx_i \wedge dx_j = -dx_j \wedge dx_i. \tag{3.21}$$

and

$$dx_i \wedge dx_i = 0. \tag{3.22}$$

Differential forms are associated with a collection of algebraic rules known as *exterior calculus* that relies mainly on an *exterior derivative* operator d, which is a linear mapping

that takes a p-form to a $(p+1)$-form, independent of coordinate systems. It is related to classical derivatives by

$$d(f d\omega) = df \wedge d\omega, \tag{3.23}$$

$$df = \sum \frac{\partial f}{\partial x_i} dx_i, \tag{3.24}$$

where f is a function.

A r-form ω is *closed* if $d\omega = 0$ and *exact* if there exists a $(r-1)$-form γ such that $\omega = d\gamma$. Particularly, integration of an exact r-form can be reduced to an integral of lower dimensions by generalized Stokes' theorem [96, pp. 87-88].

Theorem 3 *If M is an oriented r-manifold in $\mathbb{R}^k$ with boundary ∂M, and ω is an $(r-1)$ form, $r = 1, 2, \ldots, k$, then*

$$\int_{\partial M} \omega = \int_M d\omega, \tag{3.25}$$

where d on the right is exterior derivative operator.

The advantage of differential forms is that they allow us to unify several well-known variants of Stokes' theorem, such as Curl Theorem, Green Theorem, and Divergent Theorem, into one elegant and compact identity in (3.25).

To illustrate the use of these identities and properties about tensors, $\wedge$ and d, we end this section with the following example:

$$\begin{aligned} d\mathbf{r}_{\mathrm{m}} \wedge d\mathbf{r}_{\mathrm{l}} &= \frac{d\mathbf{r}_{\mathrm{m}} \wedge d\mathbf{r}_{\mathrm{l}} - d\mathbf{r}_{\mathrm{l}} \wedge d\mathbf{r}_{\mathrm{m}}}{2} \\ &= \left[\frac{\delta_{sm}\delta_{tl} - \delta_{sl}\delta_{tm}}{2}\right] d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}} \\ &= \varepsilon_{kml}\left[\frac{\varepsilon_{kst}\, d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}}}{2}\right], \end{aligned} \tag{3.26}$$

where we have used equation (3.18) and the anti-commutativity (3.21) of the wedge product $\wedge$. This transformation is generally employed to complete the 2-form $d\omega$, given

by [106, p.131]

$$d\omega \equiv -\frac{\varepsilon_{qlm}\mathbf{r}_q \, d\mathbf{r}_l \wedge d\mathbf{r}_m}{2r^3}. \tag{3.27}$$

3.2.2 Irradiance Tensors

Let $f(\mathbf{r}, \mathbf{u})$ denote a *radiance function* defined at all points $\mathbf{r} \in \mathbb{R}^3$ and all directions $\mathbf{u} \in \mathcal{S}^2$, the set of all unit vectors in $\mathbb{R}^3$. For a fixed $\mathbf{r}$, the function simplifies to a *radiance distribution* $f(\mathbf{u})$. The notion of irradiance tensors is motivated by several fundamental radiometric quantities that are defined in terms of weighted integrals of this radiance function. In particular, three such quantities lead naturally to it. They are (up to some scalar constants) [92]:

radiation energy density:

$$\int_{\mathcal{S}^2} f(\mathbf{r}, \mathbf{u}) \, d\sigma(\mathbf{u}), \tag{3.28}$$

vector irradiance:

$$\int_{\mathcal{S}^2} \mathbf{u} f(\mathbf{r}, \mathbf{u}) \, d\sigma(\mathbf{u}), \tag{3.29}$$

radiation pressure tensor:

$$\int_{\mathcal{S}^2} \mathbf{u}\mathbf{u}^{\mathrm{T}} f(\mathbf{r}, \mathbf{u}) \, d\sigma(\mathbf{u}), \tag{3.30}$$

where $f(\mathbf{r}, \mathbf{u})$ reduces to $f(\mathbf{u})$ at a given point $\mathbf{r}$. We can extend them to higher orders by characterizing them as lower-order instances of a multi-linear function, given by the form

$$\int_{\mathcal{S}^2} \mathbf{u} \otimes \cdots \otimes \mathbf{u} f(\mathbf{u}) \, d\sigma(\mathbf{u}). \tag{3.31}$$

By restricting $f(\mathbf{u})$ to be piecewise constant or piecewise polynomial over the sphere, Arvo introduced a tensor form, defined at a spherical region A and for an integer $n \geq 0$

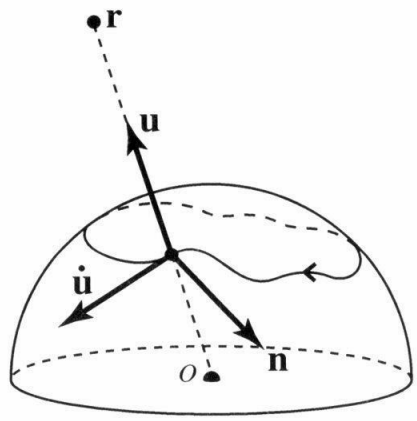


Figure 3.2: *The outward normal* $\mathbf{n} = \mathbf{u} \times \dot{\mathbf{u}}$ *is tangent to the sphere and orthogonal to the boundary of the spherical projection. (Courtesy of James Arvo [8])*

by

$$\mathbf{T}^n(A) \equiv \int_A \underbrace{\mathbf{u} \otimes \cdots \otimes \mathbf{u}}_{n \text{ factors}} \, d\sigma(\mathbf{u}), \tag{3.32}$$

called *irradiance tensors* of nth order, which satisfies the recurrence relation

$$\mathbf{T}^n_{Ij}(A) = \frac{1}{n+1} \left(\sum_{k=1}^{n-1} \delta_{jI_k} \mathbf{T}^{n-2}_{I\backslash k}(A) - \int_{\partial A} \mathbf{u}_I^{n-1} \mathbf{n}_j \, ds \right), \tag{3.33}$$

where ds denotes integration with respect to arclength and $\mathbf{n}$ denotes the unit vector normal to the boundary curve ∂A of a spherical region A, directed outward and tangent to the sphere, as shown in Figure 3.2. δ_{ij} is the Kronecker delta given by equation (3.14) and $\mathbf{u}_I^{n-1} = (\mathbf{u} \otimes \cdots \otimes \mathbf{u})_I$.

To conveniently index the elements of $\mathbf{T}^n$, we have made up some special notations about multi-index in equation (3.33). As a convention throughout this thesis, we shall use the n-index $I \equiv \{i_1, i_2, \ldots, i_n\}$, where $i_k \in \{1, 2, 3\}$ for $1 \leq k \leq n$. We define I_k as the kth sub-index, $I\backslash k$ as the $(n-1)$-index obtained by deleting the kth sub-index from

I, Ij as the $(n+1)$-index formed by appending j to I. That is,

$$\begin{aligned} I\backslash k &= (i_1, \cdots, i_{k-1}, i_{k+1}, \cdots, i_n) \\ Ij &= (i_1, i_2, \cdots, i_n, j) \end{aligned}$$

3.2.3 Angular Moments

Irradiance tensors are comprised of *angular moments* as their elements, which are weighted integrals of radiance with respect to directions [89]. It is this scalar quantities that are of primary interest in computer graphics. The angular moments of $\mathbf{T}^n$ given in equation (3.32) represent polynomials integrated over the sphere, which are useful in simulating a variety of non-diffuse phenomena. Using the fundamental formula (3.33) for irradiance tensors, we may directly derive new closed-form expressions for a subclass of polynomials that corresponds to Phong distributions [87].

Given an arbitrary subset $A \subset \mathcal{S}^2$ and unit vectors $\mathbf{r}$ and $\mathbf{s}$, Arvo defined two simple case of moments as follows:

axial moment:

$$\bar{\tau}^n(A, \mathbf{r}) \equiv \int_A \langle \mathbf{r}, \mathbf{u} \rangle^n \, d\sigma(\mathbf{u}). \tag{3.34}$$

double-axis moment:

$$\bar{\bar{\tau}}^{n,1}(A, \mathbf{r}, \mathbf{s}) \equiv \int_A \langle \mathbf{r}, \mathbf{u} \rangle^n \langle \mathbf{s}, \mathbf{u} \rangle \, d\sigma(\mathbf{u}), \tag{3.35}$$

where only the first axis is raised to the power n. Double-axis moments can be used to formulate the irradiance due to directionally-varying luminaires, reflected (or transmitted) radiance from glossy surfaces [9].

By means of tensor composition, we may express them respectively as

$$\bar{\tau}^n(A, \mathbf{r}) = \mathbf{T}^n(A)(\mathbf{r} \otimes \cdots \otimes \mathbf{r})$$

$$\bar{\bar{\tau}}^{n,1}(A,\mathbf{r},\mathbf{s}) \;=\; \mathbf{T}^{n+1}(A)(\mathbf{r}\otimes\cdots\otimes\mathbf{r}\otimes\mathbf{s}). \tag{3.36}$$

It then follows from equation (3.33) that we obtain

$$(n+1)\bar{\tau}^{n}(A,\mathbf{r}) \;=\; (n-1)\bar{\tau}^{n-2}(A,\mathbf{r}) - \int_{\partial A}\langle\mathbf{r},\mathbf{u}\rangle^{n-1}\langle\mathbf{r},\mathbf{n}\rangle\,ds, \tag{3.37}$$

$$(n+2)\bar{\bar{\tau}}^{n,1}(A,\mathbf{r},\mathbf{s}) \;=\; n\langle\mathbf{r},\mathbf{s}\rangle\,\bar{\tau}^{n-1}(A,\mathbf{r}) - \int_{\partial A}\langle\mathbf{r},\mathbf{u}\rangle^{n}\langle\mathbf{s},\mathbf{n}\rangle\,ds. \tag{3.38}$$

Equation (3.37) leads further to

$$(n+1)\bar{\tau}^{n}(A,\mathbf{r}) \;=\; \bar{\tau}^{q} - \int_{\partial A}\left[\langle\mathbf{r},\mathbf{u}\rangle^{n-1} + \langle\mathbf{r},\mathbf{u}\rangle^{n-3} + \cdots + \langle\mathbf{r},\mathbf{u}\rangle^{q+1}\right]\langle\mathbf{r},\mathbf{n}\rangle\,ds, \tag{3.39}$$

where $q = 0$ if n is even, and $q = -1$ if n is odd. Moreover, $\bar{\tau}^{-1} = 0$ and $\bar{\tau}^{0}(A) = \sigma(A)$, solid angle extended by A.

Up to this point, we have reviewed the results for angular moments derived by Arvo. Next we shall demonstrate a simplification of equation (3.39), which follows from the partial sum formula for a geometric series. That is

$$\langle\mathbf{r},\mathbf{u}\rangle^{n-1} + \langle\mathbf{r},\mathbf{u}\rangle^{n-3} + \cdots + \langle\mathbf{r},\mathbf{u}\rangle^{q+1} = \begin{cases} \dfrac{1-\langle\mathbf{r},\mathbf{u}\rangle^{n+1}}{1-\langle\mathbf{r},\mathbf{u}\rangle^{2}} - \dfrac{1}{1+\langle\mathbf{r},\mathbf{u}\rangle} & \text{even } n \\ \dfrac{1-\langle\mathbf{r},\mathbf{u}\rangle^{n+1}}{1-\langle\mathbf{r},\mathbf{u}\rangle^{2}} & \text{odd } n \end{cases}$$

Consequently, the axial moment in equation (3.39) may be expressed by

$$(n+1)\bar{\tau}^{n}(A,\mathbf{r}) = \begin{cases} \left(\sigma(A) + \displaystyle\int_{\partial A}\frac{\langle\mathbf{r},\mathbf{n}\rangle}{1+\langle\mathbf{r},\mathbf{u}\rangle}\,ds\right) - \int_{\partial A}\frac{1-\langle\mathbf{r},\mathbf{u}\rangle^{n+1}}{1-\langle\mathbf{r},\mathbf{u}\rangle^{2}}\langle\mathbf{r},\mathbf{n}\rangle\,ds & \text{even } n \\ -\displaystyle\int_{\partial A}\frac{1-\langle\mathbf{r},\mathbf{u}\rangle^{n+1}}{1-\langle\mathbf{r},\mathbf{u}\rangle^{2}}\langle\mathbf{r},\mathbf{n}\rangle\,ds & \text{odd } n \end{cases} \tag{3.40}$$

On the other hand, given a spherical region $A \subset \mathcal{S}^2$ and an arbitrary unit vector $\mathbf{w}$ such that $-\mathbf{w} \notin A$, we have proved in Appendix A.2 that the solid angle $\sigma(A)$ can be

expressed as a boundary integral

$$-\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{1 + \langle \mathbf{w}, \mathbf{u} \rangle} \, ds \; = \; \sigma(A), \tag{3.41}$$

where $1 + \langle \mathbf{w}, \mathbf{u} \rangle$ is guaranteed to be nonzero over A by the condition of $-\mathbf{w} \notin A$. Equation (3.41) can be viewed as a generalization of Girard's formula [20] for the area of spherical triangles to arbitrary regions on the sphere, as will be demonstrated in section 4.3.2. More importantly, equation (3.41) tells us that although the 2-form $d\omega$ is not *exact* over the whole sphere S^2 [8, p.75][106], it is *locally exact* over $S^2 - \mathbf{p}$, that is, the sphere minus one point. With $d\omega$ as an *exterior derivative* of a 1-form, $\sigma(A)$ thus can be represented as a boundary integral over any spherical region A other than S^2.

Using the identity (3.41), equation (3.40) finally simplifies to

$$(n+1)\bar{\tau}^n(A, \mathbf{r}) \; = \; -\int_{\partial A} \frac{1 - \langle \mathbf{r}, \mathbf{u} \rangle^{n+1}}{1 - \langle \mathbf{r}, \mathbf{u} \rangle^2} \langle \mathbf{r}, \mathbf{n} \rangle \, ds \tag{3.42}$$

for any unit vector $\mathbf{r}$ such that $-\mathbf{r} \notin A$. The formula (3.42) provides us a more compact expression for axial moments $\bar{\tau}$ than equation (3.39) as it gets away with the solid angle difference between even and odd n. In addition, it also suggests alternate approaches to evaluate or approximate axial moments, for example, in terms of hypergeometric functions [4], or using numerical quadratures.

3.3 A General Boundary Integral for Irradiance

In this section we shall derive a boundary integral representation for the irradiance integral (3.7) due to a linearly-varying luminaire. The result is general in that it applies to planar luminaires of arbitrary shapes. This reduction to a one-dimensional integral is performed by using Taylor expansion and generalizing axial moments described in section 3.2.3, and can be stated as three steps detailed in the following sections.

3.3.1 Taylor Expansion

The formula (3.33) for irradiance tensors and equations (3.37) and (3.38) presented in section 3.2 allow us to transform the integral of polynomial functions over a spherical region to a boundary integral around its curve. In order to use these formulas to reduce the rational function in the irradiance integral (3.7) to a one-dimensional integral, we need to first change this rational function to a polynomial form, and the mathematical tool is *Taylor expansion.*

We proceed by expanding the denominator $1/\langle \mathbf{w}, \mathbf{u} \rangle$ in equation (3.7) into an infinite series according to Taylor's theorem and the binomial theorem. Let $y = \langle \mathbf{w}, \mathbf{u} \rangle$, then the series expansion of $1/y$ at y_0 can be computed as:

$$\begin{aligned}
\frac{1}{y} &= \frac{1}{y_0 + (y - y_0)} \\
&= \frac{1}{y_0}\left[\sum_{n=0}^{\infty}\left(-\frac{y - y_0}{y_0}\right)^n\right] \\
&= \frac{1}{y_0}\left[\sum_{n=0}^{\infty}\left(1 - \frac{y}{y_0}\right)^n\right] \\
&= \frac{1}{y_0}\left[\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \binom{n}{k}\left(\frac{y}{y_0}\right)^k\right] \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(-1)^k}{y_0^{k+1}}\binom{n}{k} y^k,
\end{aligned} \tag{3.43}$$

where we have used the fact that

$$(1 + x)^{-1} = \sum_{n=0}^{\infty}(-x)^n, \quad (|x| < 1). \tag{3.44}$$

In order to make the expansion (3.43) converge for all $\mathbf{u}$ over the spherical projection A, it follows from equation (3.44) that we have to choose a scalar y_0, or equivalently a unit direction $\mathbf{u}_0 \in \mathcal{S}^2$, such that

$$|\langle \mathbf{w}, \mathbf{u} \rangle - \langle \mathbf{w}, \mathbf{u}_0 \rangle| < \langle \mathbf{w}, \mathbf{u}_0 \rangle \tag{3.45}$$

for any unit direction $\mathbf{u} \in A$. As we defined in section 3.1, $\mathbf{w}$ is the unit vector from the

origin orthogonal to the planar luminaire. Let φ denote the angle between any unit vector $\mathbf{u} \in A = \Pi(L)$ and the perpendicular vector $\mathbf{w}$, then φ falls in the range $[0, \pi/2)$ as A is on the illuminated hemisphere. By choosing $\mathbf{u}_0 = \mathbf{w}$, the convergence condition (3.45) is satisfied as $|\cos\varphi - 1| < 1$. Consequently, the Taylor expansion of $1/\langle \mathbf{w}, \mathbf{u} \rangle$ around $\mathbf{u}_0 = \mathbf{w}$ is simplified by $y_0 = 1$, yielding

$$\frac{1}{\langle \mathbf{w}, \mathbf{u} \rangle} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \langle \mathbf{w}, \mathbf{u} \rangle^k .$$

As a result, the integrand $\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle / \langle \mathbf{w}, \mathbf{u} \rangle$ can also be expanded into an infinite series, obtaining

$$\Phi(A) = \frac{1}{\pi} \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \int_A \langle \mathbf{w}, \mathbf{u} \rangle^k \langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle \, d\omega. \tag{3.46}$$

Equation (3.46) represents the irradiance $\Phi(A)$ as an infinite sum, comprised of terms of the form

$$\int_A \langle \mathbf{w}, \mathbf{u} \rangle^k \langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle \, d\omega, \tag{3.47}$$

which motivates us to extend the idea of axial moments and double-axis moments described in section 3.2.3 to allow for three axes, resulting in the concept of *triple-axis moments*.

3.3.2 Triple-axis Moments

By analogy with the definitions (3.34) and (3.35), we define the *triple-axis moment* of a region $A \subset S^2$ with respect to three axes $\mathbf{r}$, $\mathbf{s}$ and $\mathbf{t}$ by

$$\overset{\equiv}{\tau}^{i,j,k}(A, \mathbf{r}, \mathbf{s}, \mathbf{t}) \equiv \int_A \langle \mathbf{r}, \mathbf{u} \rangle^i \langle \mathbf{s}, \mathbf{u} \rangle^j \langle \mathbf{t}, \mathbf{u} \rangle^k \, d\omega,$$

for integers i, j, k. Specifically, to accommodate the terms (3.47) resulting from Taylor expansion, we shall only consider a special case where $j = k = 1$; that is

$$\overset{\equiv}{\tau}^{n,1,1}(A, \mathbf{r}, \mathbf{s}, \mathbf{t}) \equiv \int_A \langle \mathbf{r}, \mathbf{u} \rangle^n \langle \mathbf{s}, \mathbf{u} \rangle \langle \mathbf{t}, \mathbf{u} \rangle \, d\omega. \tag{3.48}$$

The irradiance in (3.46) thus becomes an infinite sum of triple-axis moments, given by

$$\Phi(A) = \frac{1}{\pi} \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \overset{\equiv}{\tau}^{k,1,1}(A, \mathbf{w}, \mathbf{a}, \mathbf{b}), \tag{3.49}$$

which may be expressed as a one-dimensional integral by deriving a boundary integral representation for the moment $\overset{\equiv}{\tau}^{k,1,1}(A, \mathbf{w}, \mathbf{a}, \mathbf{b})$.

By expressing the integrand as a tensor composition

$$\langle \mathbf{r}, \mathbf{u} \rangle^n \langle \mathbf{s}, \mathbf{u} \rangle \langle \mathbf{t}, \mathbf{u} \rangle = (\mathbf{u} \otimes \cdots \otimes \mathbf{u})^{n+2}_{Iij} (\mathbf{r} \otimes \cdots \otimes \mathbf{r})^n_I \mathbf{s}_i \mathbf{t}_j,$$

where the summation convention is implied for all repeated subscripts, including the multi-index, we may represent the triple-axis moment $\overset{\equiv}{\tau}^{n,1,1}$ defined in (3.48) in terms of irradiance tensors given in equation (3.32), obtaining

$$\overset{\equiv}{\tau}^{n,1,1}(A, \mathbf{r}, \mathbf{s}, \mathbf{t}) = \mathbf{T}^{n+2}(A)(\mathbf{r} \otimes \cdots \mathbf{r} \otimes \mathbf{s} \otimes \mathbf{t}). \tag{3.50}$$

From the recurrence relation (3.33), we get

$$\mathbf{T}^{n+2}_{Iij}(A) = \frac{1}{n+3} \left[\sum_{k=1}^{n} \delta_{jI_k} \mathbf{T}^n_{(I \setminus k)i}(A) + \delta_{ji} \mathbf{T}^n_I(A) - \int_{\partial A} \mathbf{u}^{n+1}_{Ii} \mathbf{n}_j ds \right]. \tag{3.51}$$

Substituting equation (3.51) into (3.50), we have

$$\begin{aligned}
(n+3)\overset{\equiv}{\tau}^{n,1,1}(A, \mathbf{r}, \mathbf{s}, \mathbf{t}) &= (n+3)\mathbf{T}^{n+2}_{Iij}(A)(\mathbf{r} \otimes \cdots \otimes \mathbf{r})_I \mathbf{s}_i \mathbf{t}_j \\
&= \left[\sum_{k=1}^{n} \delta_{jI_k} \mathbf{T}^n_{(I \setminus k)i}(A) + \delta_{ji} \mathbf{T}^n_I(A) - \int_{\partial A} \mathbf{u}^{n+1}_{Ii} \mathbf{n}_j ds \right] (\mathbf{r} \otimes \cdots \otimes \mathbf{r})_I \mathbf{s}_i \mathbf{t}_j \\
&= \sum_{k=1}^{n} \delta_{jI_k} \mathbf{t}_j \mathbf{r}_{I_k} \mathbf{T}^n_{(I \setminus k)i}(A)(\mathbf{r} \otimes \cdots \otimes \mathbf{r})_{I \setminus k} \mathbf{s}_i +
\end{aligned}$$

$$\delta_{ji}\mathbf{s}_i\mathbf{t}_j\mathbf{T}^n_I(A)(\mathbf{r}\otimes\cdots\otimes\mathbf{r})_I - \int_{\partial A}\mathbf{u}^{n+1}_{Ii}(\mathbf{r}\otimes\cdots\otimes\mathbf{r})_I\mathbf{s}_i\mathbf{t}_j\mathbf{n}_j ds.$$

Using equation (3.36) to replace the first two terms on the right, we have derived a recurrence formula for $\bar{\bar{\tau}}^{n,1,1}$, given by

$$\begin{aligned}(n+3)\bar{\bar{\tau}}^{n,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) &= n\langle\mathbf{r},\mathbf{t}\rangle\,\bar{\bar{\tau}}^{n-1,1}(A,\mathbf{r},\mathbf{s}) + \langle\mathbf{s},\mathbf{t}\rangle\,\bar{\tau}^n(A,\mathbf{r}) - \\ &\quad \int_{\partial A}\langle\mathbf{r},\mathbf{u}\rangle^n\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle\,ds \qquad (3.52)\end{aligned}$$

for any integer $n \geq 0$, where we define $\bar{\bar{\tau}}^{0,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) = \bar{\bar{\tau}}^{1,1}(A,\mathbf{s},\mathbf{t})$, $\bar{\bar{\tau}}^{-1,1}(A,\mathbf{r},\mathbf{s}) = 0$ and $\bar{\bar{\tau}}^{0,1}(A,\mathbf{r},\mathbf{s}) = \bar{\tau}^1(A,\mathbf{s})$.

Comprised of the axial moment $\bar{\tau}$ and double-axis moment $\bar{\bar{\tau}}$ as its components in equation (3.52), we may obtain an expression for triple-axis moments in terms of boundary integrals, by expanding these components using equations (3.38), (3.37) and (3.42). From equation (3.38), we may replace $\bar{\bar{\tau}}^{n-1,1}$ in equation (3.52) by

$$\bar{\bar{\tau}}^{n-1,1}(A,\mathbf{r},\mathbf{s}) = \frac{1}{n+1}\left((n-1)\langle\mathbf{r},\mathbf{s}\rangle\,\bar{\tau}^{n-2}(A,\mathbf{r}) - \int_{\partial A}\langle\mathbf{r},\mathbf{u}\rangle^{n-1}\langle\mathbf{s},\mathbf{n}\rangle\,ds\right),$$

resulting in

$$\begin{aligned}\bar{\bar{\tau}}^{n,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) &= \frac{n(n-1)}{(n+3)(n+1)}\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{s}\rangle\,\bar{\tau}^{n-2}(A,\mathbf{r}) + \frac{1}{n+3}\langle\mathbf{s},\mathbf{t}\rangle\,\bar{\tau}^n(A,\mathbf{r}) - \\ &\quad \int_{\partial A}\left[\frac{1}{n+3}\langle\mathbf{r},\mathbf{u}\rangle^n\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle + \right. \\ &\quad \left.\frac{n}{(n+3)(n+1)}\langle\mathbf{r},\mathbf{u}\rangle^{n-1}\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{s},\mathbf{n}\rangle\right]ds. \qquad (3.53)\end{aligned}$$

Then by applying the recurrence formula (3.37) for $\bar{\tau}^n$, we attain

$$\begin{aligned}\bar{\bar{\tau}}^{n,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) &= \left(\frac{n-1}{(n+3)(n+1)}\left(n\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{s}\rangle + \langle\mathbf{s},\mathbf{t}\rangle\right)\right)\bar{\tau}^{n-2}(A,\mathbf{r}) - \\ &\quad \int_{\partial A}\left[\frac{1}{n+3}\langle\mathbf{r},\mathbf{u}\rangle^n\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle + \right. \\ &\quad \left.\langle\mathbf{r},\mathbf{u}\rangle^{n-1}\left(\frac{n\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{s},\mathbf{n}\rangle + \langle\mathbf{s},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{n}\rangle}{(n+3)(n+1)}\right)\right]ds. \qquad (3.54)\end{aligned}$$

It then follows from equation (3.42) that $\overline{\overline{\tau}}^{n,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t})$ can be expressed as a boundary integral, that is

$$\begin{aligned}
\overline{\overline{\tau}}^{n,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) &= -\frac{1}{n+3}\left(\int_{\partial A}\left[\langle\mathbf{r},\mathbf{u}\rangle^{n}\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle + \langle\mathbf{r},\mathbf{u}\rangle^{n-1}\frac{n\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{s},\mathbf{n}\rangle + \langle\mathbf{s},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{n}\rangle}{n+1} + \right.\right. \\
&\quad \left.\left.\frac{1-\langle\mathbf{r},\mathbf{u}\rangle^{n-1}}{1-\langle\mathbf{r},\mathbf{u}\rangle^{2}}\left(\frac{n\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{s}\rangle + \langle\mathbf{s},\mathbf{t}\rangle}{n+1}\langle\mathbf{r},\mathbf{n}\rangle\right)\right] ds\right) - \\
&= -\frac{1}{n+3}\left(\int_{\partial A}\left[\langle\mathbf{r},\mathbf{u}\rangle^{n}\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle + \right.\right. \\
&\quad \frac{n}{n+1}\langle\mathbf{r},\mathbf{u}\rangle^{n-1}\langle\mathbf{r},\mathbf{t}\rangle\left(\langle\mathbf{s},\mathbf{n}\rangle - \langle\mathbf{r},\mathbf{s}\rangle\langle\mathbf{r},\mathbf{n}\rangle\right) + \\
&\quad \left.\left.\frac{1-\langle\mathbf{r},\mathbf{u}\rangle^{n+1}}{1-\langle\mathbf{r},\mathbf{u}\rangle^{2}}\left(\frac{n\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{s}\rangle + \langle\mathbf{s},\mathbf{t}\rangle}{n+1}\langle\mathbf{r},\mathbf{n}\rangle\right)\right] ds\right), \qquad (3.55)
\end{aligned}$$

where $-\mathbf{r} \notin A$.

Although we have assumed $n \geq 2$ in the step shown in equation (3.54), equation (3.55) still holds for all $n \geq 0$, which can be easily verified by examining the following two base cases:

$\underline{n=0:}$ $\overline{\overline{\tau}}$ reduces to a double-axis moment, thus

$$\begin{aligned}
\overline{\overline{\tau}}^{0,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) &= \bar{\tau}^{1,1}(A,\mathbf{s},\mathbf{t}) \\
&= \frac{1}{3}\langle\mathbf{s},\mathbf{t}\rangle\sigma(A) - \int_{\partial A}\frac{\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle}{3}ds \\
&= -\frac{1}{3}\int_{\partial A}\left(\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle + \frac{\langle\mathbf{s},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{n}\rangle}{1+\langle\mathbf{r},\mathbf{u}\rangle}\right) ds. \qquad (3.56)
\end{aligned}$$

$\underline{n=1:}$ It follows from equation (3.52) that

$$\begin{aligned}
\overline{\overline{\tau}}^{1,1,1}(A,\mathbf{r},\mathbf{s},\mathbf{t}) &= \frac{1}{4}\left(\langle\mathbf{r},\mathbf{t}\rangle\bar{\tau}(A,\mathbf{s}) + \langle\mathbf{s},\mathbf{t}\rangle\bar{\tau}(A,\mathbf{r}) + \int_{\partial A}\langle\mathbf{r},\mathbf{u}\rangle\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle\, ds\right) \\
&= -\frac{1}{4}\int_{\partial A}\left[\langle\mathbf{r},\mathbf{u}\rangle\langle\mathbf{s},\mathbf{u}\rangle\langle\mathbf{t},\mathbf{n}\rangle + \right. \\
&\qquad \left.\frac{\langle\mathbf{r},\mathbf{t}\rangle\langle\mathbf{s},\mathbf{n}\rangle + \langle\mathbf{s},\mathbf{t}\rangle\langle\mathbf{r},\mathbf{n}\rangle}{2}\right] ds. \qquad (3.57)
\end{aligned}$$

Both results (3.56) and (3.57) are consistent with those from equation (3.55).

Incorporating equations (3.55) and (3.49), we are able to reduce the surface integral

$\Phi(A)$ given in (3.49) to an infinite summation of boundary integrals, which may then be further simplified down to finite terms. However, this step is quite tedious, resting upon assorted binomial summation identities summarized in the next section.

3.3.3 Assorted Summation Identities

In the simplification that follows we employ a number of useful power series and algebraic transformations listed below to absorb the infinite sums.

$$\frac{1}{(n+3)(n+2)(n+1)} = \frac{1}{2}\left(\frac{1}{n+1} - \frac{2}{n+2} + \frac{1}{n+3}\right) \tag{3.58}$$

$$\sum_{n=0}^{\infty}\left(\frac{1}{n+1} - \frac{2}{n+2} + \frac{1}{n+3}\right) = \frac{1}{2} \tag{3.59}$$

$$\sum_{n=0}^{\infty}\frac{(1-x)^{n+1}}{n+1} = -\ln x \tag{3.60}$$

$$\sum_{n=0}^{\infty}\frac{(1-x)^{n+2}}{n+2} = x - 1 - \ln x \tag{3.61}$$

$$\sum_{n=0}^{\infty}\frac{(1-x)^{n+3}}{n+3} = -\frac{(x-3)(x-1)}{2} - \ln x \tag{3.62}$$

Here we assume $0 < x \leq 2$ to guarantee the convergence of the Taylor series of $\ln x$.

In order to get rid of the binomial summation in (3.49), we shall also require some identities regarding binomial coefficients, which are described here without proof. The detailed derivations are attached in Appendix A.3.

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^k}{k+1} = \frac{1-(1-x)^{n+1}}{(n+1)x} \tag{3.63}$$

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^k}{k+3} = \frac{1}{x^3}\left(\frac{2}{(n+1)(n+2)(n+3)} - \frac{(1-x)^{n+3}}{n+3} + \frac{2(1-x)^{n+2}}{n+2} - \frac{(1-x)^{n+1}}{n+1}\right) \tag{3.64}$$

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1}{k+1} = \frac{1}{n+1} \tag{3.65}$$

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1}{k+3} = \frac{2}{(n+3)(n+2)(n+1)} \tag{3.66}$$

By applying the infinite summation over the index n to these binomial identities and simplifying the results using identities (3.58) to (3.62), we can easily derive the following double summation identities involving binomial coefficients as follows:

$$\begin{aligned}
\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1-x^{k+1}}{k+1} &= \sum_{n=0}^{\infty}\left(\frac{1}{n+1}-\frac{1-(1-x)^{n+1}}{n+1}\right)\\
&= \sum_{n=0}^{\infty}\frac{(1-x)^{n+1}}{n+1}\\
&= -\ln x \qquad (3.67)\\
\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^k}{k+3} &= \frac{1}{x^3}\left(\sum_{n=0}^{\infty}\frac{2}{(n+3)(n+2)(n+1)}-\sum_{n=0}^{\infty}\frac{(1-x)^{n+3}}{n+3}+\right.\\
&\quad \left.\sum_{n=0}^{\infty}\frac{2(1-x)^{n+2}}{n+2}-\sum_{n=0}^{\infty}\frac{(1-x)^{n+1}}{n+1}\right)\\
&= \frac{1}{x^3}\left(\frac{1}{2}+\frac{(x-3)(x-1)}{2}+\ln x-2(1-x)-2\ln x+\ln x\right)\\
&= \frac{1}{2x} \qquad (3.68)\\
\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1}{k+3} &= \sum_{n=0}^{\infty}\frac{2}{(n+1)(n+2)(n+3)}\\
&= \sum_{n=0}^{\infty}\left(\frac{1}{n+1}-\frac{2}{n+2}+\frac{1}{n+3}\right)\\
&= \frac{1}{2}. \qquad (3.69)\\
\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1-x^{k+1}}{k+3} &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1}{k+3}-x\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^k}{k+3}\\
&= 0 \qquad (3.70)
\end{aligned}$$

3.3.4 Simplifying Infinite Sum to Finite Terms

The simplification of the infinite sum shown in (3.49) is performed by applying assorted binomial identities discussed in the previous section. In this section we illustrate each step in detail.

With $\mathbf{w}$ as a unit vector orthogonal to the planar luminaires L, we have $-\mathbf{w} \notin A = \Pi(L)$. It then follows from equation (3.55) that

$$\begin{aligned}
\overset{\equiv}{\tau}^{k,1,1}(A,\mathbf{w},\mathbf{a},\mathbf{b}) &= -\frac{1}{k+3}\left(\int_{\partial A}\left[\ \langle \mathbf{w},\mathbf{u}\rangle^{k}\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{b},\mathbf{n}\rangle\ +\right.\right. \\
&\quad \frac{k}{k+1}\langle \mathbf{w},\mathbf{u}\rangle^{k-1}\langle \mathbf{w},\mathbf{b}\rangle\left(\langle \mathbf{a},\mathbf{n}\rangle - \langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle\right) + \\
&\quad \left.\left.\frac{1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{1-\langle \mathbf{w},\mathbf{u}\rangle^{2}}\left(\frac{k\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle + \langle \mathbf{a},\mathbf{b}\rangle}{k+1}\langle \mathbf{w},\mathbf{n}\rangle\right)\right]\ ds\right). \qquad (3.71)
\end{aligned}$$

Let us define

$$\begin{aligned}
T_1 &= \int_{\partial A}\frac{\langle \mathbf{w},\mathbf{u}\rangle^{k}}{k+3}\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{b},\mathbf{n}\rangle\ ds, \\
T_2 &= \int_{\partial A}\frac{k}{(k+3)(k+1)}\langle \mathbf{w},\mathbf{u}\rangle^{k-1}\langle \mathbf{w},\mathbf{b}\rangle\left(\langle \mathbf{a},\mathbf{n}\rangle - \langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle\right)\ ds, \\
T_3 &= \int_{\partial A}\frac{\langle \mathbf{w},\mathbf{n}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^{2}}\frac{1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{(k+3)(k+1)}\left(k\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle + \langle \mathbf{a},\mathbf{b}\rangle\right)\ ds.
\end{aligned}$$

Equation (3.71) then reduces to

$$\overset{\equiv}{\tau}^{k,1,1}(A,\mathbf{w},\mathbf{a},\mathbf{b}) = -\left(T_1+T_2+T_3\right),$$

Consequently, the irradiance Φ becomes

$$\Phi(A) = -\frac{1}{\pi}\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^{k}\binom{n}{k}\left(T_1+T_2+T_3\right). \qquad (3.72)$$

Using those double summation identities (3.67) to (3.70) derived in section 3.3.3, we may reduce the three resulting infinite summation terms on the right of equation (3.72) to finite terms.

(1) Term One It follows from equation (3.68) that

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^{k}\binom{n}{k}\ T_1 = \int_{\partial A}\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^{k}\binom{n}{k}\frac{\langle \mathbf{w},\mathbf{u}\rangle^{k}}{k+3}\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{b},\mathbf{n}\rangle\ ds$$

$$= \frac{1}{2}\int_{\partial A} \frac{\langle \mathbf{a}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \langle \mathbf{b}, \mathbf{n} \rangle \, ds. \tag{3.73}$$

(2) Term Two

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \binom{n}{k} T_2 \\
= &\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \binom{n}{k} \int_{\partial A} \langle \mathbf{w}, \mathbf{b} \rangle \left(\langle \mathbf{a}, \mathbf{n} \rangle - \langle \mathbf{w}, \mathbf{a} \rangle \langle \mathbf{w}, \mathbf{n} \rangle \right) \times \\
&\qquad \frac{k}{(k+3)(k+1)} \langle \mathbf{w}, \mathbf{u} \rangle^{k-1} \, ds \\
= &\frac{1}{2}\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \binom{n}{k} \int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{b} \rangle \left(\langle \mathbf{a}, \mathbf{n} \rangle - \langle \mathbf{w}, \mathbf{a} \rangle \langle \mathbf{w}, \mathbf{n} \rangle \right)}{\langle \mathbf{w}, \mathbf{u} \rangle^2} \times \\
&\qquad \left(\frac{3}{k+3} - \frac{1}{k+1} \right) \langle \mathbf{w}, \mathbf{u} \rangle^{k+1} \, ds \\
= &\frac{1}{2}\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \binom{n}{k} \int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{b} \rangle \left(\langle \mathbf{a}, \mathbf{n} \rangle - \langle \mathbf{w}, \mathbf{a} \rangle \langle \mathbf{w}, \mathbf{n} \rangle \right)}{\langle \mathbf{w}, \mathbf{u} \rangle^2} \times \\
&\qquad \left(-\frac{1}{k+1} + \frac{3\langle \mathbf{w}, \mathbf{u} \rangle^{k+1}}{k+3} + \frac{1 - \langle \mathbf{w}, \mathbf{u} \rangle^{k+1}}{k+1} \right) ds.
\end{aligned} \tag{3.74}$$

In addition, it follows from Stokes' theorem that

$$\int_{\partial A} \frac{\langle \mathbf{v}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2} \, ds = \langle \mathbf{w}, \mathbf{v} \rangle \int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2} \, ds \tag{3.75}$$

for any arbitrary vector $\mathbf{v}$ and unit vector $\mathbf{w}$ in S^2. See Appendix A.4 for its proof. Letting $\mathbf{v} = \mathbf{a}$ in equation (3.75), we have

$$\int_{\partial A} \frac{\langle \mathbf{a}, \mathbf{n} \rangle - \langle \mathbf{w}, \mathbf{a} \rangle \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2} = 0.$$

Consequently, combined with identities (3.67) and (3.68) that equation (3.74) simplifies to

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \binom{n}{k} T_2$$

$$= \frac{1}{2}\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\int_{\partial A}\frac{\langle \mathbf{w},\mathbf{b}\rangle\left(\langle \mathbf{a},\mathbf{n}\rangle - \langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle\right)}{\langle \mathbf{w},\mathbf{u}\rangle^2}\times$$

$$\left(\frac{3\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{k+3}+\frac{1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{k+1}\right)ds$$

$$= \frac{1}{2}\int_{\partial A}\frac{\langle \mathbf{w},\mathbf{b}\rangle\left(\langle \mathbf{a},\mathbf{n}\rangle - \langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle\right)}{\langle \mathbf{w},\mathbf{u}\rangle^2}\times$$

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\left(\frac{3\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{k+3}+\frac{1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{k+1}\right)ds$$

$$= \frac{1}{2}\int_{\partial A}\frac{\langle \mathbf{w},\mathbf{b}\rangle\left(\langle \mathbf{a},\mathbf{n}\rangle - \langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle\right)}{\langle \mathbf{w},\mathbf{u}\rangle^2}\left(\frac{3}{2}-\ln\langle \mathbf{w},\mathbf{u}\rangle\right)ds$$

$$= \frac{\langle \mathbf{w},\mathbf{b}\rangle}{2}\int_{\partial A}\left(\langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle - \langle \mathbf{a},\mathbf{n}\rangle\right)\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle^2}\,ds. \tag{3.76}$$

(3) Term Three By applying the transformation

$$\frac{1}{(k+3)(k+1)} = \frac{1}{2}\left(\frac{1}{k+1}-\frac{1}{k+3}\right),$$

we have

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}T_3$$

$$= \int_{\partial A}\frac{\langle \mathbf{w},\mathbf{n}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\left[\left(1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}\right)\times\right.$$

$$\left.\left(\frac{k\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle}{(k+3)(k+1)}+\frac{\langle \mathbf{a},\mathbf{b}\rangle}{(k+3)(k+1)}\right)\right]ds$$

$$= \frac{1}{2}\int_{\partial A}\frac{\langle \mathbf{w},\mathbf{n}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\left[\left(3\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle - \langle \mathbf{a},\mathbf{b}\rangle\right)\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{k+3}+\right.$$

$$\left.\left(\langle \mathbf{a},\mathbf{b}\rangle - \langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle\right)\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1-\langle \mathbf{w},\mathbf{u}\rangle^{k+1}}{k+1}\right]ds$$

$$= \frac{\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle - \langle \mathbf{a},\mathbf{b}\rangle}{2}\int_{\partial A}\langle \mathbf{w},\mathbf{n}\rangle\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\,ds, \tag{3.77}$$

where the last step follows from identities (3.67) and (3.70)

Incorporating equations (3.73), (3.76) and (3.77), equation (3.72) finally reduces to finite terms, given by

$$\begin{aligned}\Phi(A) &= -\frac{1}{2\pi}\int_{\partial A}\left(\frac{\langle \mathbf{a},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\langle \mathbf{b},\mathbf{n}\rangle + \left[\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle - \langle \mathbf{a},\mathbf{b}\rangle\right]\langle \mathbf{w},\mathbf{n}\rangle\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\right.\\ &\quad\left. + \left[\langle \mathbf{w},\mathbf{a}\rangle\langle \mathbf{w},\mathbf{n}\rangle - \langle \mathbf{a},\mathbf{n}\rangle\right]\langle \mathbf{w},\mathbf{b}\rangle\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle^2}\right)ds \qquad (3.78)\end{aligned}$$

Noticing that another identity

$$\int_{\partial A}\left[\langle \mathbf{a},\mathbf{n}\rangle - \frac{\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{w},\mathbf{n}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\right]ds = \int_{\partial A}\left[\langle \mathbf{a},\mathbf{w}\rangle\langle \mathbf{w},\mathbf{n}\rangle - \langle \mathbf{a},\mathbf{n}\rangle\right]\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle^2}ds \qquad (3.79)$$

holds for an arbitrary vector $\mathbf{a}$ and a unit vector $\mathbf{w}$ (See Appendix A.5 for details), we may further simplify equation (3.78) as follows:

$$\begin{aligned}\Phi(A) &= -\frac{1}{2\pi}\int_{\partial A}\left(\frac{\langle \mathbf{a},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\langle \mathbf{b},\mathbf{n}\rangle + \left[\langle \mathbf{w},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{a}\rangle - \langle \mathbf{a},\mathbf{b}\rangle\right]\langle \mathbf{w},\mathbf{n}\rangle\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\right.\\ &\qquad\left. + \langle \mathbf{b},\mathbf{w}\rangle\left[\langle \mathbf{a},\mathbf{n}\rangle - \frac{\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{w},\mathbf{n}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\right]\right)ds\\ &= -\frac{1}{2\pi}\int_{\partial A}\left(\langle \mathbf{b},\mathbf{w}\rangle\langle \mathbf{a},\mathbf{n}\rangle + \left[\langle \mathbf{b},\mathbf{n}\rangle - \langle \mathbf{b},\mathbf{w}\rangle\langle \mathbf{w},\mathbf{n}\rangle\right]\frac{\langle \mathbf{a},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\right.\\ &\qquad\left. + \left[\langle \mathbf{a},\mathbf{w}\rangle\langle \mathbf{b},\mathbf{w}\rangle - \langle \mathbf{a},\mathbf{b}\rangle\right]\langle \mathbf{w},\mathbf{n}\rangle\frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\right)ds. \qquad (3.80)\end{aligned}$$

Moreover, using the fact that

$$\begin{aligned}\langle \mathbf{b},\mathbf{w}\rangle\langle \mathbf{w},\mathbf{n}\rangle - \langle \mathbf{b},\mathbf{n}\rangle &= \mathbf{b}^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{n} - \mathbf{b}^{\mathrm{T}}\mathbf{n}\\ &= \mathbf{b}^{\mathrm{T}}\left(\mathbf{w}\mathbf{w}^{\mathrm{T}} - \mathbf{I}\right)\mathbf{n},\end{aligned}$$

and

$$\begin{aligned}\langle \mathbf{a},\mathbf{w}\rangle\langle \mathbf{b},\mathbf{w}\rangle - \langle \mathbf{a},\mathbf{b}\rangle &= \mathbf{b}^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{a} - \mathbf{b}^{\mathrm{T}}\mathbf{a}\\ &= \mathbf{b}^{\mathrm{T}}\left(\mathbf{w}\mathbf{w}^{\mathrm{T}} - \mathbf{I}\right)\mathbf{a},\end{aligned}$$

equation (3.80) becomes

$$-\frac{1}{2\pi}\int_{\partial A}\left(\langle \mathbf{b},\mathbf{w}\rangle\,\langle \mathbf{a},\mathbf{n}\rangle + \mathbf{b}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}}\right)\left[\frac{\langle \mathbf{a},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\mathbf{n} - \eta(\mathbf{w},\mathbf{u})\ \langle \mathbf{w},\mathbf{n}\rangle\,\mathbf{a}\right]\right) ds, \tag{3.81}$$

where the scalar-valued function η is dependent on $\mathbf{w}$ and $\mathbf{u}$ and given by

$$\eta(\mathbf{w},\mathbf{u}) \;\equiv\; \frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}, \tag{3.82}$$

which is well-defined since $\langle \mathbf{w},\mathbf{u}\rangle > 0$ for the illuminated hemisphere and $\eta = 0$ at the singular point $\langle \mathbf{w},\mathbf{u}\rangle = 1$.

To see the connection between equation (3.81) and Lambert's formula for uniform luminaires, we are intended to express the irradiance formula in (3.81) in a tensor form. By using equation (3.19), we rewrite the integrand in equation (3.81) as

$$\begin{aligned}
&\langle \mathbf{b},\mathbf{w}\rangle\,\langle \mathbf{a},\mathbf{n}\rangle + \mathbf{b}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}}\right)\left[\frac{\langle \mathbf{a},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\mathbf{n} - \eta\ \langle \mathbf{w},\mathbf{n}\rangle\,\mathbf{a}\right] \\
= \;& \delta_{ik}\mathbf{a}_i\mathbf{n}_k\mathbf{b}_j\mathbf{w}_j + \mathbf{b}_j\left(\delta_{jm}-\mathbf{w}_j\mathbf{w}_m\right)\left[\frac{\mathbf{a}_i\mathbf{u}_i}{\langle \mathbf{w},\mathbf{u}\rangle}\mathbf{n}_m - \eta\ \mathbf{w}_k\mathbf{n}_k\mathbf{a}_m\right] \\
= \;& \left[\delta_{ik}\mathbf{w}_j + \left(\delta_{jm}-\mathbf{w}_j\mathbf{w}_m\right)\left(\frac{\delta_{km}\mathbf{u}_i}{\langle \mathbf{w},\mathbf{u}\rangle} - \delta_{im}\mathbf{w}_k\,\eta\right)\right]\mathbf{a}_i\mathbf{b}_j\mathbf{n}_k.
\end{aligned} \tag{3.83}$$

We shall define a 3-tensor $\mathbf{M}$ depending on unit vectors $\mathbf{w}$ and $\mathbf{u}$ by

$$\mathbf{M}_{ijk}(\mathbf{w},\mathbf{u}) \;\equiv\; \delta_{ik}\mathbf{w}_j + \left(\frac{\delta_{km}\mathbf{u}_i}{\langle \mathbf{w},\mathbf{u}\rangle} - \delta_{im}\mathbf{w}_k\,\eta\right)\left(\delta_{jm}-\mathbf{w}_j\mathbf{w}_m\right) \tag{3.84}$$

with η given in equation (3.82). It then follows from equations (3.81) and (3.83) that the irradiance at the origin due to a Lambertian luminaire L with linearly-varying radiant exitance is given by the boundary integral

$$\boxed{\Phi(A) \;=\; -\frac{1}{2\pi}\int_{\partial A}\mathbf{M}_{ijk}(\mathbf{w},\mathbf{u})\,\mathbf{a}_i\,\mathbf{b}_j\,\mathbf{n}_k\,ds,} \tag{3.85}$$

where $A \subset S^2$ is the spherical projection of L and it denotes a region on the sphere with a *rectifiable* boundary ∂A; that is, a boundary for which an outward normal $\mathbf{n}$ is defined almost everywhere. See Figure 3.2. Here, ds denotes integration with respect to arclength.

Another advantage of expressing the irradiance $\Phi(A)$ as a tensor form shown in (3.85) is that we can directly derive various *vector-valued* irradiances from it, such as *vector irradiance* defined in (3.29), which actually corresponds to

$$\frac{1}{\pi}\int_A \frac{\mathbf{u}\,\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle}\, d\sigma(\mathbf{u}),$$

independent of the receiver normal $\mathbf{b}$. However, we are more interested in another irradiance vector that removes the parameter $\mathbf{a}$ from $\Phi(A)$. That is

$$\mathbf{\Phi}(A) \;=\; \frac{1}{\pi}\int_A \frac{\mathbf{u}\,\langle \mathbf{b}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle}\, d\sigma(\mathbf{u}), \tag{3.86}$$

which may be computed from equation (3.85) by

$$\mathbf{\Phi}(A) \;=\; -\frac{1}{2\pi}\int_{\partial A} \mathbf{M}_{ijk}(\mathbf{w}, \mathbf{u})\, \mathbf{b}_j\, \mathbf{n}_k\, ds. \tag{3.87}$$

We denote this vector-valued irradiance by $\mathbf{\Phi}(A)$. It is related to the scalar-valued irradiance by

$$\Phi(A, \mathbf{w}, \mathbf{a}, \mathbf{b}) \;=\; \langle \mathbf{a}, \mathbf{\Phi}(A, \mathbf{w}, \mathbf{b})\rangle\,. \tag{3.88}$$

By separating the vector $\mathbf{a}$ encoding the linear variation of the luminaire from the computation of the irradiance, the vector quantity $\mathbf{\Phi}(A)$ is very useful in certain circumstances. For example, suppose that we want to compute the irradiance at a receiver point due to a luminaire with more than one linear variation superimposed over the plane. Since different variations only differ in the vector argument $\mathbf{a}$, computing this vector form of irradiance allows us to amortize the computation cost among these variations for high efficiency.

3.3.5 Verifying the Boundary Integral (3.85)

It is generally quite difficult to express a surface integral into integrals over the boundaries[1], which is equivalent to solving a *form differential equation* [8, p.72]. Interestingly, once the solution is found with whatever methods, we can, however, easily verify its correctness from Stokes' theorem. In this section we shall provide such a direct proof of our general boundary integral (3.85) for the irradiance due to a linearly-varying luminaire derived in the previous section. The proof strategy is quite general and applies to a class of identities that relates boundary integrals to surface integrals, which will be encountered frequently in the next chapter.

Combining equations (3.7) and (3.85), we phrase the identity to be proved as the following theorem.

Theorem 4 *Let A be the spherical projection of a planar luminaire L onto the hemisphere at the origin, $\mathbf{u} \in \mathcal{S}^2$ and the vectors $\mathbf{a}, \mathbf{b}, \mathbf{w}$ as defined before. Then we have*

$$\int_A \frac{\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \, d\omega \;=\; -\frac{1}{2} \int_{\partial A} \mathbf{M}_{ijk} \, \mathbf{a}_i \, \mathbf{b}_j \, \mathbf{n}_k \, ds, \tag{3.89}$$

where the tensor $\mathbf{M}$ is given in equation (3.84).

Proof: The proof is done by applying theorem 3 on the right hand side and changing it into a surface integral, which can be stated in the following steps:

Step 1: Let $\mathbf{r}$ be the position vector of the luminaire point and $r = ||\,\mathbf{r}\,||$. Using the fact that

$$\mathbf{n} \, ds \;=\; \frac{\mathbf{r} \times d\mathbf{r}}{r^2}, \quad \mathbf{u} \;=\; \frac{\mathbf{r}}{r}, \tag{3.90}$$

as shown in Figure 3.2, we may rewrite the integrand on the right hand of (3.89) in terms of the position vector $\mathbf{r}$ and its derivatives:

$$\int_{\partial A} \mathbf{M}_{ijk} \, \mathbf{a}_i \, \mathbf{b}_j \, \mathbf{n}_k \, ds \;=\; \int_{\partial A} \mathbf{M}_{ijk} \, \mathbf{a}_i \, \mathbf{b}_j \frac{\varepsilon_{kpl} \mathbf{r}_p \, d\mathbf{r}_l}{r^2}$$

[1] This conversion may not exist when the 2-form corresponding to the surface integral is not exact.

$$= \int_{\partial A} B_l \, d\mathbf{r}_l, \tag{3.91}$$

where have introduced a vector B_l, given by

$$B_l = \frac{\varepsilon_{kpl}\mathbf{M}_{ijk}\mathbf{a}_i\mathbf{b}_j\mathbf{r}_p}{r^2}.$$

It follows from equation (3.84) that we may write $\mathbf{M}$ as

$$\mathbf{M}_{ijk} = \delta_{ik}\mathbf{w}_j + \frac{\delta_{jk}-\mathbf{w}_j\mathbf{w}_k}{\langle \mathbf{w}, \mathbf{u}\rangle}\mathbf{u}_i - (\delta_{ij}\mathbf{w}_k - \mathbf{w}_i\mathbf{w}_j\mathbf{w}_k)\eta. \tag{3.92}$$

Step 2: To convert the boundary integral in equation (3.91) into a surface integral using Stokes' theorem (3.25), we must compute the partial derivative of B_l with respect to $\mathbf{r}_m$, that is,

$$B_{l,m} = \varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\mathbf{M}_{ijk}\mathbf{r}_p}{r^2}\right]. \tag{3.93}$$

Notice that

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\mathbf{r}_p}{r^2}\right] &= \frac{\delta_{pm}r^2 - 2\mathbf{r}_p\mathbf{r}_m}{r^4} \\
\frac{\partial}{\partial \mathbf{r}_m}\left[\mathbf{M}_{ijk}\right] &= (\delta_{jk} - \mathbf{w}_j\mathbf{w}_k)\frac{\delta_{im}\langle \mathbf{w}, \mathbf{r}\rangle - \mathbf{r}_i\mathbf{w}_m}{\langle \mathbf{w}, \mathbf{r}\rangle^2} - \eta_m(\delta_{ij}\mathbf{w}_k - \mathbf{w}_i\mathbf{w}_j\mathbf{w}_k),
\end{aligned}$$

where η_m denotes the partial derivatives of η with respect to $\mathbf{r}_m$, which can be computed as follows:

$$\begin{aligned}
&\eta_m \left(= \frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\ln\langle \mathbf{w}, \mathbf{u}\rangle}{1 - \langle \mathbf{w}, \mathbf{u}\rangle^2}\right]\right) \\
= \; &\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{(\ln\langle \mathbf{w}, \mathbf{r}\rangle - \ln r)r^2}{r^2 - \langle \mathbf{w}, \mathbf{r}\rangle^2}\right] \\
= \; &\frac{\left(r^2\mathbf{w}_m/\langle \mathbf{w}, \mathbf{r}\rangle - \mathbf{r}_m + 2\mathbf{r}_m\ln\langle \mathbf{w}, \mathbf{u}\rangle\right)\left(r^2 - \langle \mathbf{w}, \mathbf{r}\rangle^2\right) - \left(2\mathbf{r}_m - 2\langle \mathbf{w}, \mathbf{r}\rangle \mathbf{w}_m\right)r^2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\left(r^2 - \langle \mathbf{w}, \mathbf{r}\rangle^2\right)^2}
\end{aligned}$$

$$= \frac{r^2\mathbf{w}_m\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2 + 2\langle\mathbf{w},\mathbf{r}\rangle^2 \ln\langle\mathbf{w},\mathbf{u}\rangle\right)/\langle\mathbf{w},\mathbf{r}\rangle - \mathbf{r}_m\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2 + 2\langle\mathbf{w},\mathbf{r}\rangle^2 \ln\langle\mathbf{w},\mathbf{u}\rangle\right)}{\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2\right)^2}$$

$$= \frac{r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2 + 2\langle\mathbf{w},\mathbf{r}\rangle^2 \ln\langle\mathbf{w},\mathbf{u}\rangle}{\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2\right)^2}\left(\frac{r^2\mathbf{w}_m}{\langle\mathbf{w},\mathbf{r}\rangle} - \mathbf{r}_m\right). \tag{3.94}$$

It follows from the chain rule and equation (3.93) that

$$\begin{aligned} B_{l,m} &= \varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j\left[\mathbf{M}_{ijk}\frac{\delta_{pm}r^2 - 2\mathbf{r}_p\mathbf{r}_m}{r^4} + (\delta_{jk} - \mathbf{w}_j\mathbf{w}_k)\frac{\delta_{im}\mathbf{r}_p\langle\mathbf{w},\mathbf{r}\rangle - \mathbf{r}_i\mathbf{r}_p\mathbf{w}_m}{r^2\langle\mathbf{w},\mathbf{r}\rangle^2}\right. \\ &\qquad \left. - \frac{\mathbf{r}_p}{r^2}\eta_m(\delta_{ij}\mathbf{w}_k - \mathbf{w}_i\mathbf{w}_j\mathbf{w}_k)\right] \\ &= \varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j\left[A_1 + A_2 + A_3\right], \end{aligned} \tag{3.95}$$

where A_1, A_2, A_3 are given by

$$\begin{aligned} A_1 &= \mathbf{M}_{ijk}\frac{\delta_{pm}r^2 - 2\mathbf{r}_p\mathbf{r}_m}{r^4} \\ A_2 &= (\delta_{jk} - \mathbf{w}_j\mathbf{w}_k)\frac{\delta_{im}\mathbf{r}_p\langle\mathbf{w},\mathbf{r}\rangle - \mathbf{r}_i\mathbf{r}_p\mathbf{w}_m}{r^2\langle\mathbf{w},\mathbf{r}\rangle^2} \\ A_3 &= -\frac{\mathbf{r}_p}{r^2}\eta_m(\delta_{ij}\mathbf{w}_k - \mathbf{w}_i\mathbf{w}_j\mathbf{w}_k). \end{aligned}$$

and $\mathbf{M}$, η_m are given by equations (3.92) and (3.94), respectively. Thus, by applying theorem 3, we have

$$\begin{aligned} \int_{\partial A} B_l\, d\mathbf{r}_l &= \int_A B_{l,m}\, d\mathbf{r}_m \wedge d\mathbf{r}_l \\ &= \int_A \varepsilon_{qml}\varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j\left[A_1 + A_2 + A_3\right]\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right]. \end{aligned} \tag{3.96}$$

where the transformation (3.26) is applied.

Step 3: Applying the tensor identity (3.18) to the three terms in equation (3.96), we get

$$\varepsilon_{qml}\varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j A_1 = (\delta_{qk}\delta_{pm} - \delta_{pq}\delta_{km})\frac{\delta_{pm}r^2 - 2\mathbf{r}_p\mathbf{r}_m}{r^4}\mathbf{a}_i\mathbf{b}_j\mathbf{M}_{ijk}$$

$$
\begin{aligned}
&= \frac{\mathbf{a}_i\mathbf{b}_j\mathbf{M}_{ijk}}{r^4}\left(3\delta_{qk}r^2 - 2\delta_{qk}r^2 - \delta_{qk}r^2 + 2\mathbf{r}_q\mathbf{r}_k\right) \\
&= \frac{2\mathbf{r}_q}{r^4}\left(\mathbf{a}_i\mathbf{b}_j\mathbf{r}_k\mathbf{M}_{ijk}\right) \\
&= \frac{2\mathbf{r}_q}{r^4}\left(\langle\mathbf{a},\mathbf{r}\rangle\langle\mathbf{b},\mathbf{w}\rangle + \frac{\langle\mathbf{b},\mathbf{r}\rangle - \langle\mathbf{w},\mathbf{b}\rangle\langle\mathbf{w},\mathbf{r}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\langle\mathbf{a},\mathbf{u}\rangle\right. \\
&\qquad \left.-\eta\left(\langle\mathbf{a},\mathbf{b}\rangle\langle\mathbf{w},\mathbf{r}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\langle\mathbf{r},\mathbf{w}\rangle\right)\right) \\
&= \frac{2\mathbf{r}_q}{r^3}\left[\frac{\langle\mathbf{a},\mathbf{u}\rangle\langle\mathbf{b},\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle} - \eta\left(\langle\mathbf{a},\mathbf{b}\rangle\langle\mathbf{w},\mathbf{u}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\langle\mathbf{w},\mathbf{u}\rangle\right)\right], \\
\varepsilon_{qml}\varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j A_2 &= \frac{\delta_{qk}\delta_{pm} - \delta_{pq}\delta_{km}}{r^2\langle\mathbf{w},\mathbf{r}\rangle^2}\left[(\mathbf{b}_k - \mathbf{w}_k\langle\mathbf{b},\mathbf{w}\rangle)(\mathbf{a}_m\mathbf{r}_p\langle\mathbf{w},\mathbf{r}\rangle - \mathbf{w}_m\mathbf{r}_p\langle\mathbf{a},\mathbf{r}\rangle)\right] \\
&= \frac{\mathbf{r}_q}{r^3}\left[-\frac{\langle\mathbf{a},\mathbf{b}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle} + \frac{\langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\right], \\
\varepsilon_{qml}\varepsilon_{kpl}\mathbf{a}_i\mathbf{b}_j A_3 &= -\frac{\varepsilon_{qml}\varepsilon_{kpl}}{r^2}\eta_m\left(\mathbf{r}_p\mathbf{w}_k\langle\mathbf{a},\mathbf{b}\rangle - \mathbf{r}_p\mathbf{w}_k\langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\right) \\
&= -\frac{\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2 + 2\langle\mathbf{w},\mathbf{r}\rangle^2\ln\langle\mathbf{w},\mathbf{u}\rangle\right)\left(\langle\mathbf{a},\mathbf{b}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\right)}{r^2\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2\right)^2}\times \\
&\qquad \left[\left(\delta_{qk}\delta_{pm} - \delta_{pq}\delta_{km}\right)\mathbf{r}_p\mathbf{w}_k\left(\frac{r^2\mathbf{w}_m}{\langle\mathbf{w},\mathbf{r}\rangle} - \mathbf{r}_m\right)\right] \\
&= -\frac{\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2 + 2\langle\mathbf{w},\mathbf{r}\rangle^2\ln\langle\mathbf{w},\mathbf{u}\rangle\right)\left(\langle\mathbf{a},\mathbf{b}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\right)}{r^2\left(r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2\right)^2}\times \\
&\qquad \left[\mathbf{r}_q\langle\mathbf{w},\mathbf{r}\rangle - \mathbf{r}_q\frac{r^2}{\langle\mathbf{w},\mathbf{r}\rangle}\right] \\
&= \frac{\left(1 - \langle\mathbf{w},\mathbf{u}\rangle^2 + 2\langle\mathbf{w},\mathbf{u}\rangle^2\ln\langle\mathbf{w},\mathbf{u}\rangle\right)\left(\langle\mathbf{a},\mathbf{b}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\right)}{r^2 - \langle\mathbf{w},\mathbf{r}\rangle^2}\cdot\frac{\mathbf{r}_q}{\langle\mathbf{w},\mathbf{r}\rangle} \\
&= \frac{1 - \langle\mathbf{w},\mathbf{u}\rangle^2 + 2\langle\mathbf{w},\mathbf{u}\rangle^2\ln\langle\mathbf{w},\mathbf{u}\rangle}{r^3\left(1 - \langle\mathbf{w},\mathbf{u}\rangle^2\right)\langle\mathbf{w},\mathbf{u}\rangle}\mathbf{r}_q\left[\langle\mathbf{a},\mathbf{b}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\right] \\
&= \frac{\mathbf{r}_q}{r^3}\left[\left(\frac{1}{\langle\mathbf{w},\mathbf{u}\rangle} + 2\eta\langle\mathbf{w},\mathbf{u}\rangle\right)\left(\langle\mathbf{a},\mathbf{b}\rangle - \langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle\right)\right] \\
&= \frac{\mathbf{r}_q}{r^3}\left[\frac{\langle\mathbf{a},\mathbf{b}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle} - \frac{\langle\mathbf{a},\mathbf{w}\rangle\langle\mathbf{b},\mathbf{w}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle} + \right.
\end{aligned}
$$

$$2\eta\left(\langle \mathbf{a},\mathbf{b}\rangle\langle \mathbf{w},\mathbf{u}\rangle - \langle \mathbf{a},\mathbf{w}\rangle\langle \mathbf{b},\mathbf{w}\rangle\langle \mathbf{w},\mathbf{u}\rangle\right)\Big].$$

It follows that

$$\varepsilon_{qml}\varepsilon_{kql}\mathbf{a}_i\mathbf{b}_j\left[A_1+A_2+A_3\right] \;=\; \frac{2\mathbf{r}_q}{r^3}\left[\frac{\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{b},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\right].$$

Consequently, equation (3.96) simplifies to

$$\int_{\partial A} B_l\, d\mathbf{r}_l \;=\; 2\int_A \left[\frac{\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{b},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}\right]\frac{\mathbf{r}_q}{r^3}\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s\wedge d\mathbf{r}_t}{2}\right]. \tag{3.97}$$

Step 4: Expressing the surface integral in equation (3.97) in terms of solid angle using the definition (3.27), we get

$$-2\int_A \frac{\langle \mathbf{a},\mathbf{u}\rangle\langle \mathbf{b},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}d\omega.$$

From equations (3.91) and (3.97), we have proved equation (3.89). □□

In deriving the general boundary integral (3.85) in section 3.3.4, we have exchanged the order of infinite summation and integration without proof. The direct proof presented here serves as an indirect verification of the feasibility of these operations.

As another verification, we shall demonstrate that the formula (3.85) for linearly-varying luminaires subsumes Lambert's formula for uniform luminaires. As a special case of linearly-varying luminaires, we can specify a luminaire with constant radiant exitance c everywhere by three non-collinear points $\mathbf{p}_1(x_1,y_1,z_1)$, $\mathbf{p}_2(x_2,y_2,z_2)$ and $\mathbf{p}_3(x_3,y_3,z_3)$ on the luminaire and their associated radiant exitance values $w_1=w_2=w_3=c$. Then equation (3.6) gives us

$$\begin{aligned}\mathbf{a} &= c\,h\,[1\;1\;1]\begin{bmatrix} x_1 & x_2 & x_3\\ y_1 & y_2 & y_3\\ z_1 & z_2 & z_3\end{bmatrix}^{-1}\\ &= \frac{c\,h}{\triangle}\,(\triangle_1,\triangle_2,\triangle_3),\end{aligned} \tag{3.98}$$

where $\triangle_1, \triangle_2, \triangle_3$ and $\triangle$ denote determinants, given by

$$\triangle = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$$

and

$$\triangle_1 = \begin{vmatrix} 1 & 1 & 1 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} \quad \triangle_2 = -\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ z_1 & z_2 & z_3 \end{vmatrix} \quad \triangle_3 = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}. \tag{3.99}$$

Here, we have used Cramer's rule, that is, considering

$$[1\ 1\ 1] \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}^{-1}$$

as the solution to a linear system $Ax = b$. Let V denote the volume of the tetrahedron defined by $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ and the origin $\mathbf{o}$, then

$$V = \frac{1}{3} S\, h, \tag{3.100}$$

where S is the area of $\triangle \mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$, and h is the perpendicular distance from $\mathbf{o}$ to the luminaire plane determined by $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$. On the other hand, by considering $\triangle \mathbf{p}_2\mathbf{p}_3\mathbf{o}$ as the bottom of this tetrahedron, we can also express V as [20]

$$V = \frac{1}{6} (\mathbf{p}_1 \cdot (\mathbf{p}_2 \times \mathbf{p}_3)) = \frac{1}{6} \triangle. \tag{3.101}$$

It follows from equations (3.100) and (3.101) that $\triangle = 2Sh$. From equation (3.98), we have

$$\mathbf{a} = \frac{c}{2S} (\triangle_1, \triangle_2, \triangle_3). \tag{3.102}$$

By expressing the vector $(\triangle_1, \triangle_2, \triangle_3)$ in terms of coordinates of $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ using equation (3.99), we observe that[2]

$$(\triangle_1, \triangle_2, \triangle_3) \;=\; (\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_3 - \mathbf{p}_1).$$

Since the cross product $(\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_3 - \mathbf{p}_1)$ represents a vector that is perpendicular to two contiguous edges of $\triangle \mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$ and has a magnitude equivalent to twice the area of the triangle, we have

$$(\triangle_1, \triangle_2, \triangle_3) \;=\; 2S\,\mathbf{w}.$$

Consequently, equation (3.102) gives us

$$\mathbf{a} \;=\; c\,\mathbf{w}.$$

As a result, equation (3.85) combined with equation (3.84) yields

$$\Phi(A) \;=\; -\frac{c}{2\pi} \int_{\partial A} \langle \mathbf{b}, \mathbf{n} \rangle \; ds, \tag{3.103}$$

which is Lambert's well-known formula for the irradiance at a point due to a uniform luminaire with radiant exitance c, where $\mathbf{b}$ is the normal at the receiver point.

3.4 Closed-form Solution for Polygonal Luminaires

The general boundary integral in equation (3.85) reduces the area integral (3.7) corresponding to the irradiance due to linearly-varying luminaires to a one-dimensional integral. Thus far, no restrictions have been placed on the shape of the integration region $A \subset \mathcal{S}^2$. This section describes how this boundary integral formula leads to a closed-form solution for polygonal luminaires.

[2]Depending on the orientation of $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$, the order of the cross product may be reversed.

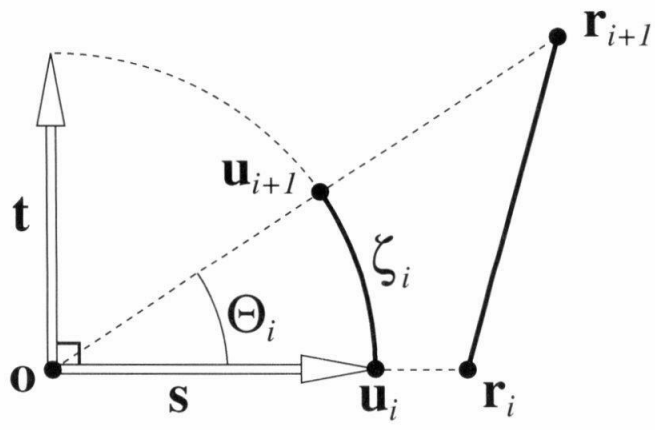


Figure 3.3: *The vectors used to parameterize an arc* ζ_i *(*$\mathbf{u}_i\mathbf{u}_{i+1}$*) defined by a polygon edge* $\mathbf{r}_i\mathbf{r}_{i+1}$.

3.4.1 Spherical Polygon

Given a polygonal luminaire P, the resulting projection A on the sphere is a *spherical polygon*, whose edges are great arcs, that is, segments of great circles.

When A is a spherical polygon, the boundary integral (3.85) can be evaluated along each edge ζ of A, resulting in a summation of line integrals, which are greatly simplified in two ways. First and most importantly, the outward normal $\mathbf{n}$ remains constant along ζ, which allows us to move the factors involving $\mathbf{n}$ outside the integrals. Another simplification emerges from a certain parameterization for the edge, as we shall show below.

Let ζ_i denote a great arc connecting two points $\mathbf{u}_i$ and $\mathbf{u}_{i+1}$ on a unit sphere, which are spherical projections of an edge $\mathbf{r}_i\mathbf{r}_{i+1}$ of the polygonal luminaire P. Then the outward normal $\mathbf{n}$ is in the direction of $\mathbf{u}_i \times \mathbf{u}_{i+1}$, perpendicular to the plane formed by $\mathbf{u}_i$ and $\mathbf{u}_{i+1}$. We may parameterize this great arc ζ_i as

$$\mathbf{u}(\theta) = \mathbf{s}\cos\theta + \mathbf{t}\sin\theta, \tag{3.104}$$

where $\mathbf{s}$ and $\mathbf{t}$ are two orthonormal vectors in the plane containing the edge and the origin $\mathbf{o}$, with $\mathbf{s}$ directed toward the first vertex of the edge, and θ is the rotation angle of $\mathbf{u}$ from $\mathbf{s}$. See Figure 3.3. This parameterization allows us to compute the line integral

of the form

$$\int_{\zeta} \langle \mathbf{v}, \mathbf{u} \rangle \, ds \tag{3.105}$$

by an integral of angle θ, where $\mathbf{v}$ is an arbitrary vector. Let

$$a \equiv \langle \mathbf{v}, \mathbf{s} \rangle, \quad b \equiv \langle \mathbf{v}, \mathbf{t} \rangle,$$

we define

$$c \equiv \sqrt{a^2 + b^2}, \tag{3.106}$$

and an angle $\phi \in [-\pi, \pi]$ such that

$$(\cos\phi, \ \sin\phi) = \left(\frac{a}{c}, \ \frac{b}{c} \right). \tag{3.107}$$

It follows from the trigonometric identity

$$\cos(\theta_1 + \theta_2) \ = \ \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2$$

that we may evaluate the type of line integral shown in (3.105) by

$$\begin{aligned} \int_{\zeta} \langle \mathbf{v}, \mathbf{u} \rangle \ ds &= \int_0^{\Theta} [a\cos\theta + b\sin\theta] \ d\theta \\ &= c \int_0^{\Theta} \cos(\theta - \phi) \, d\theta, \end{aligned} \tag{3.108}$$

where Θ is the angle subtended by the arc ζ_i. Note that these variables a, b, c, ϕ depend on the given vector $\mathbf{v}$. For a unit vector $\mathbf{v}$, we have $0 \leq c \leq 1$. As a convention, later on we shall use a subscript to denote this dependency, such as, $c_{\mathbf{v}}$, $\phi_{\mathbf{v}}$, and so on.

3.4.2 Restriction to Polygons

We now specialize the result given in the previous section to spherical polygons, which results from projecting simple planar polygons onto the sphere. The resulting formula

will allow us to compute the exact irradiance due to a polygonal luminaire with linearly-varying radiant exitance.

For a polygonal luminaire P with s edges, we evaluate the boundary integrals in equation (3.85) along each great arc ζ_i defined by a polygon edge, obtaining

$$\Phi(P) \;=\; -\frac{1}{2\pi}\sum_{m=1}^{s}\left[\int_{\zeta_m}\mathbf{M}_{ijk}\,\mathbf{a}_i\,\mathbf{b}_j\;ds\right]\,\mathbf{n}_k(m), \tag{3.109}$$

where the normal $\mathbf{n}$ that depends on the edge ζ_m has been moved outside the integral due to the property of a great arc. Consequently, for each m, the line integral in equation (3.109) simplifies to

$$\begin{aligned}
&\int_{\zeta}\mathbf{M}_{ijk}\,\mathbf{a}_i\,\mathbf{b}_j\,\mathbf{n}_k\;ds\\
=\;&\int_{\zeta}\left[\delta_{ik}\mathbf{a}_i\mathbf{n}_k\mathbf{b}_j\mathbf{w}_j+\left(\frac{\delta_{km}\mathbf{u}_i}{\langle\mathbf{w},\mathbf{u}\rangle}-\delta_{im}\mathbf{w}_k\;\eta\right)\mathbf{a}_i\mathbf{n}_k\mathbf{b}_j\;(\delta_{jm}-\mathbf{w}_j\mathbf{w}_m)\right]\;ds\\
=\;&\int_{\zeta}\left[\langle\mathbf{a},\mathbf{n}\rangle\,\langle\mathbf{b},\mathbf{w}\rangle+\left(\frac{\langle\mathbf{a},\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\mathbf{n}_m-\eta\;\langle\mathbf{w},\mathbf{n}\rangle\,\mathbf{a}_m\right)\mathbf{b}_j\;(\delta_{jm}-\mathbf{w}_j\mathbf{w}_m)\right]\;ds\\
=\;&\int_{\zeta}\left[\langle\mathbf{a},\mathbf{n}\rangle\,\langle\mathbf{b},\mathbf{w}\rangle+\mathbf{b}^{\mathrm{T}}(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\left(\frac{\langle\mathbf{a},\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\mathbf{n}-\eta\;\langle\mathbf{w},\mathbf{n}\rangle\,\mathbf{a}\right)\right]\;ds\\
=\;&\langle\mathbf{a},\mathbf{n}\rangle\,\langle\mathbf{b},\mathbf{w}\rangle\,\Theta+\mathbf{b}^{\mathrm{T}}(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\left[B^{1,1}(\mathbf{a})\,\mathbf{n}-B^{*}\;\langle\mathbf{w},\mathbf{n}\rangle\,\mathbf{a}\right],
\end{aligned}\tag{3.110}$$

where the boundary integrals are defined by

$$B^{1,1}(\mathbf{a}) \;\equiv\; \int_{\zeta}\frac{\langle\mathbf{a},\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\;ds, \tag{3.111}$$

$$B^{*} \;\equiv\; \int_{\zeta}\eta(\mathbf{w},\mathbf{u})\;ds. \tag{3.112}$$

By parameterizing ζ as we described in section 3.4.1, it follows from equation (3.108) that we can evaluate the line integrals in equations (3.111) and (3.112) as follows:

$$\begin{aligned}
\overline{B}^{1,1} \;&=\; \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}}\int_{0}^{\Theta}\frac{\cos(\theta-\phi_{\mathbf{a}})}{\cos(\theta-\phi_{\mathbf{w}})}\;d\theta\\
&=\; \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}}\int_{-\phi_{\mathbf{w}}}^{\Theta-\phi_{\mathbf{w}}}\frac{\cos(\theta+\phi_{\mathbf{w}}-\phi_{\mathbf{a}})}{\cos\theta}\;d\theta
\end{aligned}$$

$$
\begin{aligned}
&= \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}} \int_{-\phi_{\mathbf{w}}}^{\Theta-\phi_{\mathbf{w}}} \frac{\cos\theta\cos(\phi_{\mathbf{w}}-\phi_{\mathbf{a}}) - \sin\theta\sin(\phi_{\mathbf{w}}-\phi_{\mathbf{a}})}{\cos\theta}\, d\theta \\
&= \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}} \int_{-\phi_{\mathbf{w}}}^{\Theta-\phi_{\mathbf{w}}} \left[\cos(\phi_{\mathbf{w}}-\phi_{\mathbf{a}}) - \sin(\phi_{\mathbf{w}}-\phi_{\mathbf{a}})\frac{\sin\theta}{\cos\theta}\right] d\theta \\
&= \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}} \left[\cos(\phi_{\mathbf{w}}-\phi_{\mathbf{a}})\Theta + \sin(\phi_{\mathbf{w}}-\phi_{\mathbf{a}}) \left[\ln\cos\theta\right]_{-\phi_{\mathbf{w}}}^{\Theta-\phi_{\mathbf{w}}}\right] \\
&= \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}} \left[\cos(\phi_{\mathbf{w}}-\phi_{\mathbf{a}})\Theta + \sin(\phi_{\mathbf{w}}-\phi_{\mathbf{a}}) \ln\frac{\cos(\Theta-\phi_{\mathbf{w}})}{\cos\phi_{\mathbf{w}}}\right]
\end{aligned}
\tag{3.113}
$$

and

$$
\begin{aligned}
\overline{B^*} &= \int_0^{\Theta} \frac{\ln\left(c_{\mathbf{w}}\cos(\theta-\phi_{\mathbf{w}})\right)}{1-\left(c_{\mathbf{w}}\cos(\theta-\phi_{\mathbf{w}})\right)^2}\, d\theta \\
&= \int_{-\phi_{\mathbf{w}}}^{\Theta-\phi_{\mathbf{w}}} \frac{\ln\left(c_{\mathbf{w}}\cos\theta\right)}{1-\left(c_{\mathbf{w}}\cos\theta\right)^2}\, d\theta \\
&= \Lambda(c_{\mathbf{w}}, \Theta-\phi_{\mathbf{w}}) - \Lambda(c_{\mathbf{w}}, -\phi_{\mathbf{w}}),
\end{aligned}
\tag{3.114}
$$

where Θ is the angle subtended by the edge ζ, and the variables c's and ϕ's are defined by equation (3.106) and equation (3.107), respectively. The two-parameter function Λ appearing in equation (3.114) is given by

$$
\boxed{\Lambda(\alpha,\beta) \equiv \int_0^{\beta} \frac{\ln(\alpha\cos\theta)}{1-(\alpha\cos\theta)^2}\, d\theta,} \tag{3.115}
$$

where $0 < \alpha \leq 1, -\pi/2 \leq \beta \leq \pi/2$. As we will show in the next two chapters, this same integral arises in computing the irradiance from all polynomially-varying Lambertian luminaires, and is therefore of practical importance beyond the linear problem we are addressing here. We shall discuss this integral more fully in the next section.

Combining Equations (3.109), (3.110), (3.113) and (3.114), we arrive at a "closed-form" solution for the irradiance at a point due to a polygonal luminaire P with s edges,

given by

$$\boxed{\Phi(P) = -\frac{1}{2\pi} \sum_{m=1}^{s} \left[\langle \mathbf{a}, \mathbf{n} \rangle \langle \mathbf{b}, \mathbf{w} \rangle \Theta + \mathbf{b}^{\mathrm{T}} (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}}) \left(\overline{B}^{1,1}(\mathbf{a})\, \mathbf{n} - \overline{B^*}\, \langle \mathbf{w}, \mathbf{n} \rangle\, \mathbf{a} \right) \right],} \tag{3.116}$$

where $\overline{B}^{1,1}$, $\overline{B^*}$ and $\mathbf{n}$ vary over each edge ζ_m. One complication that arises in our solution is the appearance of a *special function* that has no finite representation in terms of elementary functions.

To compute the irradiance vector $\mathbf{\Phi}(P)$ described on Page 85, we are required to evaluate a vector-valued boundary integral that corresponds to $B^{1,1}$ in equation (3.111), defined by

$$\mathbf{B}^{1,1} \equiv \int_{\zeta} \frac{\mathbf{u}}{\langle \mathbf{w}, \mathbf{u} \rangle}\, ds,$$

denoted by $\mathbf{B}^{1,1}$. Similar to equation (3.113), we use the parameterization given in equation (3.104) and evaluate $\mathbf{B}^{1,1}$ as follows:

$$\begin{aligned}
\overline{\mathbf{B}}^{1,1} &= \frac{1}{c_{\mathbf{w}}} \int_0^{\Theta} \frac{\mathbf{s}\cos\theta + \mathbf{t}\sin\theta}{\cos(\theta - \phi_{\mathbf{w}})}\, d\theta \\
&= \frac{1}{c_{\mathbf{w}}} \int_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}} \frac{\mathbf{s}\cos(\theta + \phi_{\mathbf{w}}) + \mathbf{t}\sin(\theta + \phi_{\mathbf{w}})}{\cos\theta}\, d\theta \\
&= \frac{1}{c_{\mathbf{w}}} \int_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}} \frac{(\mathbf{s}\cos\phi_{\mathbf{w}} + \mathbf{t}\sin\phi_{\mathbf{w}})\cos\theta - (\mathbf{s}\sin\phi_{\mathbf{w}} - \mathbf{t}\cos\phi_{\mathbf{w}})\sin\theta}{\cos\theta}\, d\theta \\
&= \frac{1}{c_{\mathbf{w}}} \int_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}} \left[(\mathbf{s}\cos\phi_{\mathbf{w}} + \mathbf{t}\sin\phi_{\mathbf{w}}) - (\mathbf{s}\sin\phi_{\mathbf{w}} - \mathbf{t}\cos\phi_{\mathbf{w}}) \frac{\sin\theta}{\cos\theta} \right] d\theta \\
&= \frac{1}{c_{\mathbf{w}}} \left[(\mathbf{s}\cos\phi_{\mathbf{w}} + \mathbf{t}\sin\phi_{\mathbf{w}})\Theta + (\mathbf{s}\sin\phi_{\mathbf{w}} - \mathbf{t}\cos\phi_{\mathbf{w}}) \left[\ln\cos\theta \right]_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}} \right] \\
&= \frac{1}{c_{\mathbf{w}}} \left[(\mathbf{s}\cos\phi_{\mathbf{w}} + \mathbf{t}\sin\phi_{\mathbf{w}})\Theta + (\mathbf{s}\sin\phi_{\mathbf{w}} - \mathbf{t}\cos\phi_{\mathbf{w}}) \ln \frac{\cos(\Theta - \phi_{\mathbf{w}})}{\cos\phi_{\mathbf{w}}} \right].
\end{aligned}$$

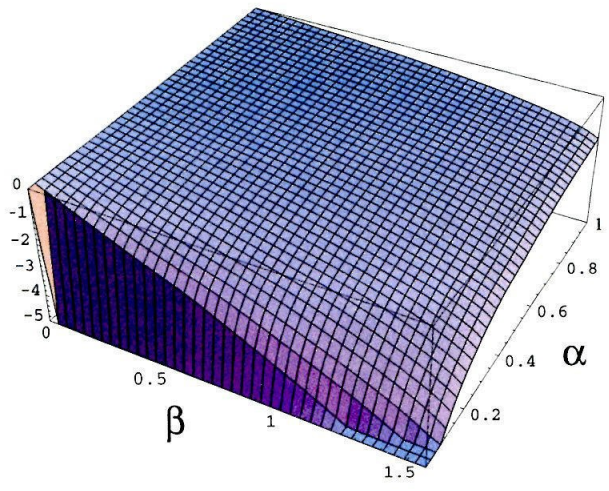


Figure 3.4: *The special function* $\Lambda(\alpha, \beta)$ *defined in equation (3.115), over the range* $0 < \alpha \leq 1$, *and* $0 \leq \beta \leq \pi/2$. *When* $\alpha \to 0$, $\Lambda(\alpha, \beta)$ *tends to* $-\infty$.

Thus, it follows from equation (3.88) and equation (3.116), we have

$$\mathbf{\Phi}(P) \;=\; -\frac{1}{2\pi} \sum_{m=1}^{s} \left[\mathbf{n} \left\langle \mathbf{b}, \mathbf{w} \right\rangle \Theta + \mathbf{b}^{\mathrm{T}} (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}}) \left(\mathbf{n}\, \overline{\mathbf{B}}^{1,1} - \overline{B^*} \left\langle \mathbf{w}, \mathbf{n} \right\rangle \right) \right]. \qquad (3.117)$$

3.4.3 The Special Function Λ

To our knowledge, the two-parameter function $\Lambda(\alpha, \beta)$ defined in (3.115) has no finite representation in terms of elementary exponential and logarithmic functions, and therefore qualifies as a *higher transcendental function*, also known as a *special function*. Figure 3.4 shows $\Lambda(\alpha, \beta)$ over its required domain, $(0, 1] \times [0, \pi/2]$. For certain values of the parameters such as $\alpha = 1$, $\Lambda(\alpha, \beta)$ simplifies immediately to a known definite integral with exact solution. For example,

$$\begin{aligned} \Lambda(1, \beta) \;&=\; \int_0^{\beta} \frac{\ln(\cos\theta)}{\sin^2\theta}\, d\theta \\ &=\; -\int_0^{\beta} \ln(\cos\theta)\; d(\cot\theta) \\ &=\; -\beta - \cot\beta \ln(\cos\beta). \end{aligned} \qquad (3.118)$$

However, to evaluate $\Lambda(\alpha, \beta)$ in general, we reduce it to a well-known special function that can be evaluated to high accuracy, such as the *dilogarithm* $\mathrm{Li}_2(z)$, given by [91, 67]

$$\mathrm{Li}_2(z) \;=\; -\int_0^z \frac{\ln(1-t)}{t}\,dt \;=\; \sum_{n=1}^{\infty} \frac{z^n}{n^2} \tag{3.119}$$

for complex argument z with $|z| \leq 1$; or a somewhat simpler function *Clausen integral* $\mathrm{Cl}_2(x)$ with real arguments, defined as [1, 12, 67]

$$\mathrm{Cl}_2(x) \;=\; -\int_0^x \ln\left|2\sin\frac{\theta}{2}\right|\,d\theta \;=\; -\frac{1}{2}\int_0^x \ln\left(4\sin^2\frac{\theta}{2}\right)\,d\theta \;=\; \sum_{n=1}^{\infty} \frac{\sin nx}{n^2}. \tag{3.120}$$

These two special functions are closely related, as $\mathrm{Cl}_2(\theta)$ is simply the imaginary part of $\mathrm{Li}_2(e^{i\theta})$ [67, pp.91], which can be seen from their corresponding series definitions (3.119) and (3.120).

Next we shall derive a formula to express $\Lambda(\alpha, \beta)$ in terms of the Clausen integrals. This reduction can be stated in three steps.

First, by the change of variable

$$\tan\theta \;=\; \sqrt{1-\alpha^2}\tan t, \tag{3.121}$$

we relate $\Lambda(\alpha, \beta)$ to an intermediate two-parameter function $\Upsilon(\mu, \nu)$ introduced by Arvo, defined by [8, pp.112],

$$\Upsilon(\mu, \nu) \;\equiv\; \int_0^{\mu} \ln\left(1+\nu^2\sec^2\theta\right)\,d\theta. \tag{3.122}$$

The functional relationship is

$$\Lambda(\alpha, \beta) \;=\; -\frac{1}{2\sqrt{1-\alpha^2}}\,\Upsilon\left(\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right),\, \frac{\sqrt{1-\alpha^2}}{\alpha}\right). \tag{3.123}$$

The detail of this step is provided in Appendix A.6. The change of variable shown in equation (3.121) that leads to equation (3.123) may seem not intuitive at all, and the readers may wonder how we came up with this formula. In fact, the connection between equation (3.121) and equation (3.123) was reversed in our actual derivation.

That is, equation (3.121) is obtained from relationship (3.123) using reverse engineering. We first proved (3.123) independently by formulating $\Lambda(\alpha, \beta)$ as the irradiance due to special quadratically-varying luminaires, which is known to have an expression in terms of $\Upsilon(\mu, \nu)$. We shall describe this derivation in chapter 5.

Then, we express $\Upsilon(\mu, \nu)$ in terms of Clausen integrals. The key to finding the exact connection between Υ and the Clausen integral is an identity involving the Clausen integral derived by Newman [79, pp.88] in the 19th century. We have shown a similar proof in Appendix A.7. Following a similar procedure employed by Arvo [8], we have derived an expression for $\Upsilon(\mu, \nu)$ in terms of three Clausen integrals, given by

$$\Upsilon(\mu, \nu) = -\frac{1}{2}[2\mathrm{Cl}_2(2\mu) - \mathrm{Cl}_2(4\mu - 2\eta) - \mathrm{Cl}_2(2\eta) + 2(\eta - \mu)\ln\gamma], \tag{3.124}$$

where

$$\begin{aligned} \gamma &= (\nu - \sqrt{\nu^2 + 1})^2, \\ \eta &= \tan^{-1}\left(\frac{\sin(2\mu)}{\gamma + \cos(2\mu)}\right). \end{aligned} \tag{3.125}$$

Equation (3.124) is in a much simpler form than the original formula from Arvo [8, p.197]. The complete derivation can be found in Appendix A.7.

Finally, incorporating equations (3.123) and (3.124), we have shown that $\Lambda(\alpha, \beta)$ can be expressed in terms of three Clausen integrals. Specifically,

$$\Lambda(\alpha, \beta) = \frac{2(\eta - \mu)\ln\gamma + 2\mathrm{Cl}_2(2\mu) - \mathrm{Cl}_2(4\mu - 2\eta) - \mathrm{Cl}_2(2\eta)}{4\sqrt{1 - \alpha^2}}, \tag{3.126}$$

where $0 < \alpha < 1, 0 \leq \beta \leq \pi/2$ and

$$\begin{aligned} \mu(\alpha, \beta) &= \tan^{-1}\left(\frac{\tan\beta}{\sqrt{1 - \alpha^2}}\right), \\ \gamma(\alpha) &= \left(\frac{1 - \sqrt{1 - \alpha^2}}{\alpha}\right)^2, \\ \eta(\alpha, \beta) &= \tan^{-1}\left(\frac{\sin(2\mu)}{\gamma + \cos(2\mu)}\right). \end{aligned}$$

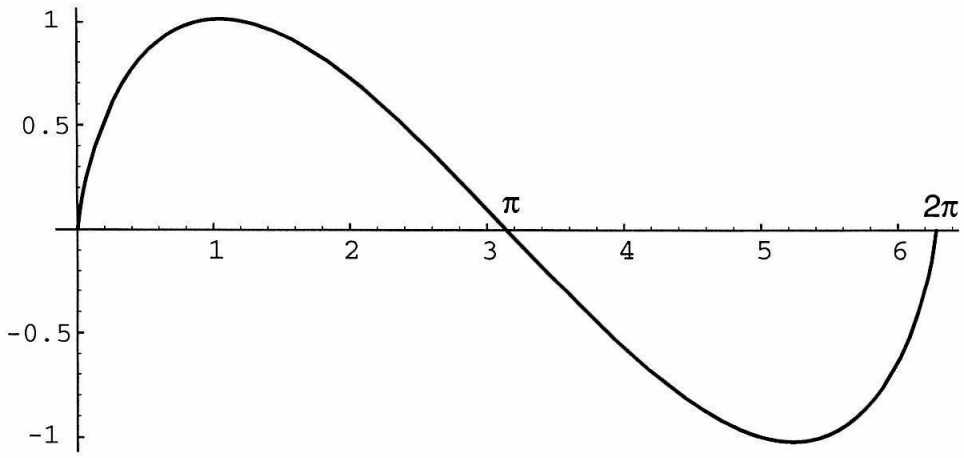


Figure 3.5: *The Clausen integral* $\mathrm{Cl}_2(\theta)$ *over the domain* $[0, 2\pi]$.

The case of $\alpha = 1$ has been handled by equation (3.118) and the values of Λ when $\beta \in [-\pi/2, 0]$ can be easily computed by using the fact that

$$\Lambda(\alpha, -\beta) = -\Lambda(\alpha, \beta). \tag{3.127}$$

It is not surprising that a special function like $\Lambda(\alpha, \beta)$ appears in the problem we are addressing; integrals of a similar nature have been discovered and studied in connection with radiative transfer for half a century [91]. It is interesting to note that the dilogarithms and related functions like the Clausen integral also appear in computing the form factor between two arbitrary polygons, as shown by Schröder and Hanrahan [97]. In addition, Powell [91] showed that dilogarithms are equivalent to the *Airey radiation integrals* that arise in both astrophysics and quantum mechanics.

By using equation (3.126), $\Lambda(\alpha, \beta)$ can be evaluated directly. The representation of $\Lambda(\alpha, \beta)$ in terms of Clausen integrals has some advantages, since $\mathrm{Cl}_2(\theta)$ is well studied and has some nice properties, such as being cyclic with respect to 2π. A plot of this function over the domain $[0, 2\pi]$ is shown in Figure 3.5. The Clausen integral can be easily approximated to high accuracy. Kölbig [63] presents an approximation of $\mathrm{Cl}_2(\theta)$ by series of even Chebyshev polynomials [104], which are defined separately for the

intervals $[0, \pi/2]$ and $[\pi/2, \pi]$. That is,

$$\mathrm{Cl}_2(\theta) \;=\; \theta - \theta \ln|\theta| + \frac{1}{2}\theta^3 \sum_{n=0}^{\infty} a_n \, T_{2n}\left(\frac{2\theta}{\pi}\right) \quad \left(0 \le \theta \le \frac{\pi}{2}\right), \tag{3.128}$$

and

$$\mathrm{Cl}_2(\theta) \;=\; (\pi - \theta) \sum_{n=0}^{\infty} b_n \, T_{2n}\left(\frac{2\theta}{\pi} - 2\right) \quad \left(\frac{\pi}{2} \le \theta \le \pi\right), \tag{3.129}$$

where the Chebyshev polynomial $T_n(x)$ is defined for $x \in [-1, 1]$ as

$$T_n(x) \;\equiv \cos(n\theta), \qquad x = \cos\theta.$$

Due to the property of smooth spreading out the error, the truncated Chebyshev expansion is very nearly the same polynomial as the *minimax polynomial*, the optimal polynomial approximation that minimizes the maximum deviation from the function to be approximated. Therefore, Chebyshev approximation is widely used in approximating mathematical functions [31]. With 20-digit Chebyshev coefficients given in Table 3.4.3, the Chebyshev series in formulas (3.128) and (3.129) converges very fast. Truncation error is negligible due to rapidly decreasing coefficients a_n or b_n. For an accuracy of 20 digits, we only need to evaluate the Chebyshev series in equation (3.128) (or (3.129)) with 10 terms (or 16 terms). A procedure for evaluating a sum $S = \sum_{n=0}^{N} c_n \, T_{2n}(x)$ of even Chebyshev polynomials is given in the pseudocode *ChebySum* , which is based on the recurrence relation attributed to Clenshaw [31]. It should be pointed out that this method of summation has no worse round-off characteristics than ordinary polynomial evaluation and use the same number of multiplications [104, p.35].

ChebySum(**real** $c[\,]$, **real** x, **integer** N)

$h \leftarrow 2x^2 - 1$

$\alpha \leftarrow 2h$

$t_1 \leftarrow 0$

$t_2 \leftarrow 0$

for $n \leftarrow N, N-1, \ldots, 0$ **do**

$t_0 \leftarrow c[n] + \alpha * t_1 - t_2$

$t_2 \leftarrow t_1$

$t_1 \leftarrow t_0$

endfor

return $t_0 - h * t_2$

Another approximation formula, due to Grosjean [50], consists of elementary real functions only and is given by

$$\begin{aligned} \mathrm{Cl}_2(\theta) \;\approx\; & -\theta \ln \sin \frac{1}{2}\theta + \frac{1}{2880}\theta(\pi^2 - \theta^2)(120 - 7\pi^2 + 3\theta^2) + \left(2\ln 2 - \frac{5}{4}\right)\sin\theta \\ & - \left(\frac{89}{128} - \ln 2\right)\sin 2\theta + \left(\frac{2}{3}\ln 2 - \frac{449}{972}\right)\sin 3\theta - \left(\frac{4259}{12288} - \frac{1}{2}\ln 2\right)\sin 4\theta \\ & + \left(\frac{2}{5}\ln 2 - \frac{10397}{37500}\right)\sin 5\theta, \quad 0 \le \theta \le \pi. \end{aligned} \tag{3.130}$$

Experimentation shows that equation (3.130) can approximate $\mathrm{Cl}_2(x)$ to a relative error less than 0.003%, which may be enough for graphics applications. This procedure is shown in the pseudo-code *ApproxClausen* in Figure 3.6. If most time is spent in evaluating trigonometric functions in equation (3.130), optimization can be performed such that only one call to sine and cosine function is necessary. The idea is to compute the sine and cosine of a multiple angle using iterated rotations rather than direct call to trigonometric functions. For example,

$$\begin{bmatrix} \cos(n+1)\theta \\ \sin(n+1)\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos(n\theta) \\ \sin(n\theta) \end{bmatrix}.$$

n	a_n	b_n
0	$2.795283197357566135E-02$	$6.390970888572653413 1E-01$
1	$1.7630887438981157E-04$	$-5.498056930185171564E-02$
2	$1.26627414611565E-06$	$-9.6126194595060643E-04$
3	$1.171718181344E-08$	$-3.205468682255048E-05$
4	$1.2300641288E-10$	$-1.32946169542555E-06$
5	$1.39527290E-12$	$-6.209360182440E-08$
6	$1.669078E-14$	$-3.12960065639E-09$
7	$2.0761E-16$	$-1.6635195382E-10$
8	$2.66E-18$	$-9.19652725E-12$
9	$3E-20$	$-5.2400377E-13$
10		$-3.058038E-14$
11		$-1.81969E-15$
12		$-1.1004E-16$
13		$-6.75E-18$
14		$-4.2E-19$
15		$-3.0E-20$

Table 3.1: Chebyshev coefficients for $\mathrm{Cl}_2(\theta)$.

Or alternatively, we may compute the Clausen integral using a hypergeometric representation [4, p.102], given by

$$\mathrm{Cl}_2(\theta) \;=\; \mathrm{Im}(\mathrm{Li}_2(e^{i\theta})) \;=\; \mathrm{Im}\left(e^{i\theta} {}_3F_2 \left[\begin{matrix} 1\,,\,1\,,\,1 \\ 2\,,\,2 \end{matrix} \;;\; e^{i\theta} \right] \right),$$

where $\mathrm{Im}(z)$ is the imaginary part of the complex number z.

However, as a more efficient alternative, one can tabulate the function $\Lambda(\alpha, \beta)$ on a regular grid, and retrieve values by direct indexing and bilinear interpolation. Unfortunately, it can be seen from the definition (3.115) that

$$\lim_{\alpha \to 0} \Lambda(\alpha, \beta) \;=\; -\infty,$$

ApproxClausen(**real** x)

$c_1 \leftarrow \quad 3.472222222E-04$

$c_2 \leftarrow \quad 9.869604401E+00$

$c_3 \leftarrow \quad 5.091276919E+01$

$c_4 \leftarrow \quad 1.362943611E-01$

$c_5 \leftarrow -2.165319440E-03$

$c_6 \leftarrow \quad 1.639639947E-04$

$c_7 \leftarrow -2.471701169E-05$

$c_8 \leftarrow \quad 5.538890645E-06$

if $x == 0$

    return 0.0

endif

if $x > \pi$

    return -ApproxClausen($2\pi - x$)

endif

return

$c_1 * x * (c_2 - x^2) * (c_3 + 3x^2)$

$+c_4 * \sin(x) + c_5 * \sin(2x)$

$+c_6 * \sin(3x) + c_7 * \sin(4x)$

$+c_8 * \sin(5x)$

$-x * \ln(\sin(x/2))$

end

Figure 3.6: *Pseudo-code for Grosjean's approximation of Clausen's integral for* $x \in [0, 2\pi]$.

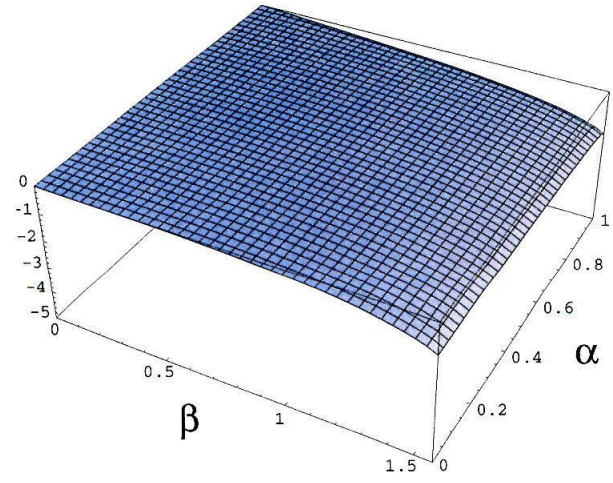


Figure 3.7: *The special function* $\Lambda'(\alpha, \beta)$ *defined in equation (3.132), over the range* $[0, 1] \times [0, \pi/2]$.

(also see Figure 3.4), which makes it impossible to evaluate Λ over its entire domain using bilinear interpolation. To remedy this, noticing from Figure 3.4 that $\Lambda(\alpha, \beta)$ varies smoothly in most places except the neighborhood of $\alpha = 0$, we may separate a small strip $[0, \alpha_\epsilon] \times [0, \pi/2]$ from the domain, and evaluate $\Lambda(\alpha, \beta)$ using formula (3.126) inside the strip and use bilinear interpolation over the rest of the domain. In our experimentation, α_ϵ is chosen as 0.01. After creating a table on a regular 300×300 grid over the domain $[0.01, 1] \times [0, \pi/2]$, we estimate the error by evaluating Λ over 3000×3000 regular grid points. It is shown that such a hybrid approach achieves a maximum relative error 0.2%, which is accurate enough for graphics applications. For simplicity, we only need to tabulate the value of Λ with the parameter β over the range $[0, \pi/2]$, since $\beta \in [-\pi/2, 0]$ can be easily converted to the required domain by equation (3.127).

Another approach is to split $\Lambda(\alpha, \beta)$ $(\alpha \neq 1)$ into two parts by replacing $\ln(\alpha \cos\theta)$ in the numerator by $\ln(\alpha) + \ln(\cos\theta)$. This results in

$$\begin{aligned} \Lambda(\alpha, \beta) &= (\ln \alpha) \int_0^\beta \frac{1}{1 - (\alpha \cos\theta)^2} \, d\theta + \int_0^\beta \frac{\ln(\cos\theta)}{1 - (\alpha \cos\theta)^2} \, d\theta \\ &= \lambda(\alpha, \beta) + \Lambda'(\alpha, \beta), \end{aligned} \tag{3.131}$$

where we define

$$\Lambda'(\alpha,\beta) \equiv \int_0^\beta \frac{\ln(\cos\theta)}{1-(\alpha\cos\theta)^2}\,d\theta, \tag{3.132}$$

and λ can be evaluated exactly as [49, p.185]

$$\lambda(\alpha,\beta) = \frac{\ln\alpha}{\sqrt{1-\alpha^2}}\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right). \tag{3.133}$$

A plot of the new function $\Lambda'(\alpha,\beta)$ for $0 \le \alpha \le 1$ and $0 \le \beta \le \pi/2$ is shown in Figure 3.7. Since Λ' is smooth and bounded, it can be easily computed using numerical quadrature, or in terms of Clausen integrals from equations (3.126), (3.131) and (3.133). Furthermore, by tabulating $\Lambda'(\alpha,\beta)$ we can approximate Λ over its entire domain via interpolation.

To aid in other possible interpolation using B-splines, we provide convenient expressions for the four boundary curves of $\Lambda'(\alpha,\beta)$. Except the boundary $\Lambda'(0,\beta)$, given by

$$\Lambda'(0,\beta) = \int_0^\beta \ln(\cos\theta)\,d\theta, \tag{3.134}$$

all the other three boundaries can be expressed in closed-form. That is

$$\Lambda'(1,\beta) = \Lambda(1,\beta) = -\beta - \cot\beta\ln(\cos\beta), \tag{3.135}$$

$$\Lambda'(\alpha,0) = 0,$$

$$\Lambda'(\alpha,\frac{\pi}{2}) = -\frac{\pi}{2\sqrt{1-\alpha^2}}\ln\left(1+\sqrt{1-\alpha^2}\right), \tag{3.136}$$

where equation (3.136) is derived in Appendix A.8. The boundary defined in (3.134) is called *Lobachevsky's integral* [49, pp. 941], denoted by $L(x)$, which can be expressed in terms of the Clausen integral as

$$L(x) = \frac{1}{2}\mathrm{Cl}_2(\pi-2x) - x\ln 2, \tag{3.137}$$

See Appendix A.7. Lobachevsky's integral is a well-known special function which cannot be evaluated in finite terms [67]. Since this special function Λ' contains Lobachevsky's integral [49] as a special case of $\alpha = 0$, we thus claim in the beginning of this section that $\Lambda(\alpha, \beta)$ can not be evaluated in terms of a finite number of elementary functions.

Chapter 4

Generalized Irradiance Tensors

In this chapter we generalize the concept of irradiance tensors to accommodate the emission features manifested by planar luminaires with polynomially-varying radiant exitance. These tensors, called *generalized irradiance tensors*, are natural generalizations of the *radiation pressure tensor* of a radiation field due to a polynomially-varying luminaire. The elements of these tensors are moments of the radiance distribution function $f(\mathbf{u})$ that corresponds to a polynomial radiant exitance function, which are expressed as a class of constrained rational polynomial functions integrated over regions of the sphere, called *rational (angular) moments*. A variety of non-Lambertian rendering effects involving spatially-varying luminaires can be formulated as combinations of these moments, as we shall demonstrate in the next chapter.

Generalized irradiance tensors provide us a powerful vehicle to concisely represent a wide range of rational polynomials over the sphere. Based on several useful recurrence formulas identified for these tensors, we obtain a number of new expressions for their individual moments that are relevant to applications in computer graphics. Particularly, we derive closed-form expressions for a limited subclass of rational moments corresponding to Phong distributions [87] in polyhedral environments, and present complete algorithms for their efficient evaluation.

We organize the remainder of this chapter as follows. Section 4.1 characterizes that the radiance distribution at the origin due to a polynomially-varying luminaire is ex-

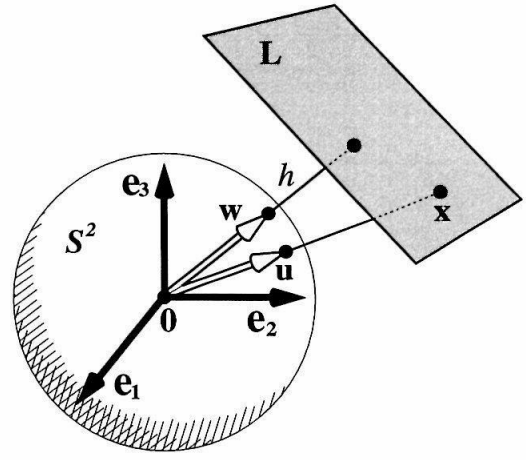


Figure 4.1: *A polynomial radiant exitance function over the luminaire plane corresponds to a rational polynomial function over the sphere.*

pressed as a class of constrained rational polynomial over the sphere, which motivates our generalization of irradiance tensors in section 4.2. These generalized irradiance tensors satisfy several useful recurrence relations developed in section 4.2. In section 4.3 we use these tensor formulas to derive expressions for rational angular moments, and discuss the exact evaluation of these expressions. In section 4.4 we derive simplified closed-form expressions for a subclass of rational moments corresponding to Phong distributions, and present algorithms for their efficient evaluation over polygonal domains, which will be applied to image synthesis in the next chapter.

4.1 Radiance from Polynomially-Varying Luminaires

As we mentioned in section 3.1, a spatially-varying planar luminaire is specified by a radiant exitance function $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$, mapping a point $\mathbf{x}$ on the luminaire to its radiant exitance value. The goal of this section is to characterize the radiance distribution function $f(\mathbf{u})$ at the origin due to a polynomially-varying luminaire, which turns out to be a rational polynomial function over the sphere.

To see how rational polynomials arise in this context, consider the radiation field at the origin caused by a polygonal luminaire $L \subset \mathbb{R}^3$ not containing the origin $\mathbf{o}$.

Suppose that the radiant exitance of L varies according to a polynomial $\phi(\mathbf{x})$ at each point $\mathbf{x} = (x_1, x_2, x_3) \in L$, where x_1, x_2, x_3 are the coordinates of $\mathbf{x}$ with respect to an orthogonal coordinate system $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ at the origin. As usual, we define $\mathbf{w}$ as the unit vector orthogonal to the luminaire plane and h as the distance from the origin to the plane, as shown in Figure 4.1.

Since any nth order polynomial can be expressed as linear combinations of monomials of order n or less, it is only necessary to consider these monomials, given by

$$\phi(\mathbf{x}) \;=\; x_1^p\, x_2^q\, x_3^r \;=\; \langle \mathbf{x}, \mathbf{e}_1 \rangle^p \, \langle \mathbf{x}, \mathbf{e}_2 \rangle^q \, \langle \mathbf{x}, \mathbf{e}_3 \rangle^r \,, \tag{4.1}$$

where $p + q + r = n$ for a nth order monomial. By using the identity (3.5) between the position vector $\mathbf{x}$ and its unit direction $\mathbf{u}$, we may express the radiance distribution $f(\mathbf{u})$ at the origin emitted from the Lambertian planar luminaire L in terms of ϕ, yielding

$$f(\mathbf{u}) \;=\; \frac{1}{\pi}\, \phi\left(h\, \frac{\mathbf{u}}{\langle \mathbf{w}, \mathbf{u} \rangle} \right), \tag{4.2}$$

where the radiant exitance keeps constant in all directions. Consequently, a Lambertian luminaire whose radiant exitance varies according to a nth order monomial given in equation (4.1) generates a radiation field at the origin distributed directionally by

$$f(\mathbf{u}) \;=\; \frac{h^n}{\pi} \frac{\langle \mathbf{e}_1, \mathbf{u} \rangle^p \, \langle \mathbf{e}_2, \mathbf{u} \rangle^q \, \langle \mathbf{e}_3, \mathbf{u} \rangle^r}{\langle \mathbf{w}, \mathbf{u} \rangle^n}, \tag{4.3}$$

which is a rational polynomial over the sphere. However, the form of (4.1) is *constrained*, as only powers of $\langle \mathbf{w}, \mathbf{u} \rangle$ appear in the denominator, which simplifies our formalism discussed next. More precisely, equation (4.3) is a *characteristic function* of A, the spherical projection of L at the origin. For $\mathbf{u} \notin A$, $f(\mathbf{u}) = 0$.

4.2 Generalizing Irradiance Tensors

The radiance distribution function from polynomially-varying luminaires can be accommodated by a simple generalization of irradiance tensors, which provides us a means of formulating illuminating and scattering effects involving this type of emitters.

4.2.1 Definition

In section 3.2.2, we have shown that the notion of irradiance tensors is motivated by extending three related physical quantities, i.e, the radiation energy density, the vector irradiance and the radiation pressure tensor, to higher orders. One assumption was made in defining $\mathbf{T}^n$ by equation (3.32), that is, $f(\mathbf{u})$ is piecewise constant or piecewise polynomial over the sphere, which no longer holds for polynomially-varying luminaires, as shown in equation (4.3). However, this type of variation can be easily accommodated by a simple generalization of irradiance tensors of the form (3.32).

By restricting $f(\mathbf{u})$ in equation (3.31) to be a class of constrained rational polynomials given in equation (4.3), we define

$$\mathbf{T}^{n;q}(A, \mathbf{w}) \equiv \int_A \frac{\mathbf{u} \otimes \cdots \otimes \mathbf{u}}{\langle \mathbf{w}, \mathbf{u} \rangle^q} \, d\sigma(\mathbf{u}) \tag{4.4}$$

for integers $n, q \geq 0$, where $A \subset \mathcal{S}^2$ and we restrict $\mathbf{w}$ as a unit vector such that $\langle \mathbf{w}, \mathbf{u} \rangle > 0$ for any $\mathbf{u} \in A$, and $\otimes$ denotes a tensor product. In the context of simulating polynomially-varying luminaires, A is the spherical projection of a luminaire L at the origin, and $\mathbf{w}$ is the perpendicular vector to the luminaire plane. With a tensor product of n factors in the numerator, the integrand of such a tensor contains all rational polynomials of the form $x^i y^j z^k / \langle \mathbf{w}, \mathbf{u} \rangle^q$ where $(x, y, z) \equiv \mathbf{u} \in \mathcal{S}^2$ and $i + j + k = n$. We call the integrals in (4.4) *generalized irradiance tensors* and denote them by $\mathbf{T}^{n;q}(A, \mathbf{w})$, distinct from irradiance tensors $\mathbf{T}^n(A)$. As a convention, the numerator order n and the denominator order q for the integrand of a generalized irradiance tensor is indicated in the superscript of the notation $\mathbf{T}^{n;q}$, separated by a semicolon. Obviously, when $q = 0$, $\mathbf{T}^{n;q}$ reduces to the irradiance tensor of nth order. Therefore, generalized irradiance tensors $\mathbf{T}^{n;q}$ play a more powerful role than $\mathbf{T}^n$, allowing us to perform some additional computations.

4.2.2 Recurrence Relations

Generalized irradiance tensors of all orders are defined by a surface integral given in (4.4), which may be reduced to one-dimensional integrals by means of several recurrence formulas presented in this section. This reformulation generally simplifies the analytical

evaluation of original integrals and may lead to closed-form solutions for certain geometries. This has been a common strategy to compute integrals of this nature, such as form factors [97, 105], vector irradiance [75], and so on.

Using the notation for the multi-indices introduced in section 3.2.2, we now state and prove several fundamental formulas about $\mathbf{T}^{n;q}$ in the following theorems. The basic mathematical tool used in the foregoing analysis is generalized Stokes' theorem, that is, Theorem 3 in section 3.2.1.

Theorem 5 *Let $n \geq 0$ and $q \geq 2$ be integers, then the tensor $\mathbf{T}^{n;q}(A, \mathbf{w})$ satisfies the recurrence relation*

$$\mathbf{T}_I^{n;q}(A, \mathbf{w}) = \frac{1}{q-1}\left[\sum_{k=1}^{n} \mathbf{w}_{I_k} \mathbf{T}_{I\backslash k}^{n-1;q-1} - (n-q+3)\, \mathbf{T}_I^{n;q-2} - \int_{\partial A} \frac{\mathbf{u}_I^n \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^{q-1}}\, ds\right], \quad (4.5)$$

where I is an n-index, ds denotes integration with respect to arclength, and $\mathbf{n}$ is the outward normal to the curve ∂A.

Proof: The proof is performed by considering an equivalent formula

$$(q-1)\mathbf{T}_I^{n;q} - \sum_{k=1}^{n} \mathbf{w}_{I_k} \mathbf{T}_{I\backslash k}^{n-1;q-1} + (n-q+3)\, \mathbf{T}_I^{n;q-2} = -\int_{\partial A} \frac{\mathbf{u}_I^n \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^{q-1}}\, ds. \quad (4.6)$$

Equation (4.6) reformulates equation (4.5) as an identity that equates a boundary integral on the right with an expression involving surface integrals on the left, which can be proved using the same strategy as that employed in the proof of Theorem 4. We only show the difference here.

Corresponding to equations (3.91) and (3.96), we have

$$\begin{aligned} \int_{\partial A} \frac{\mathbf{u}_I^n \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^{q-1}}\, ds &= \int_{\partial A} \left[\frac{r^{q-1}}{\langle \mathbf{w}, \mathbf{r} \rangle^{q-1}}\right] \left[\frac{\mathbf{r}_I^n}{r^n}\right] \left[\frac{\varepsilon_{kpl} \mathbf{w}_k \mathbf{r}_p\, d\mathbf{r}_l}{r^2}\right] \\ &= \int_{\partial A} B_{Il}^{n+1}\, d\mathbf{r}_l \\ &= \int_A \varepsilon_{qml} B_{Il,m} \left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right], \end{aligned} \quad (4.7)$$

where we have introduced $(n+1)$-tensor B, given by

$$B_{Il} = \varepsilon_{kpl}\mathbf{w}_k \left[\frac{\mathbf{r}_p \mathbf{r}_I^n}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1} r^{n-q+3}}\right].$$

and its derivative $B_{Il,m}$ is computed from the chain rule by

$$B_{Il,m} = \varepsilon_{kpl}\mathbf{w}_k \left(\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{1}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1}}\right]\left[\frac{\mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+3}}\right] + \left[\frac{1}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1}}\right]\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+3}}\right]\right), \quad (4.8)$$

where

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{1}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1}}\right] &= -\frac{(q-1)\mathbf{w}_m}{\langle \mathbf{w}, \mathbf{r}\rangle^{q}}, \\
\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+3}}\right] &= \frac{\delta_{pm}\mathbf{r}_I^n + \mathbf{r}_p \mathbf{r}_{I,m}^n}{r^{n-q+3}} - (n-q+3)\frac{\mathbf{r}_m \mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+5}},
\end{aligned}$$

Applying identity (3.18) to the two terms in equation (4.8), we get

$$\begin{aligned}
&\varepsilon_{qml}\varepsilon_{kpl}\mathbf{w}_k \frac{\partial}{\partial \mathbf{r}_m}\left[\frac{1}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1}}\right]\left[\frac{\mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+3}}\right] \\
&= -\frac{q-1}{\langle \mathbf{w}, \mathbf{r}\rangle^{q} r^{n-q+3}} \left[\delta_{qk}\delta_{pm} - \delta_{pq}\delta_{km}\right] \mathbf{w}_k \mathbf{w}_m \mathbf{r}_p \mathbf{r}_I^n \\
&= (q-1)\frac{\mathbf{r}_I^n}{\langle \mathbf{w}, \mathbf{r}\rangle^{q} r^{n-q}}\left(\frac{\mathbf{r}_q}{r^3}\right) - (q-1)\frac{\mathbf{w}_q \mathbf{r}_I^n}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1} r^{n-q+3}},
\end{aligned}$$

and

$$\begin{aligned}
&\varepsilon_{qml}\varepsilon_{kpl}\mathbf{w}_k \left[\frac{1}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1}}\right]\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+3}}\right] \\
&= \frac{\mathbf{w}_q \delta_{pm} - \mathbf{w}_m \delta_{pq}}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1}}\left[\frac{\delta_{pm}\mathbf{r}_I^n + \mathbf{r}_p \mathbf{r}_{I,m}^n}{r^{n-q+3}} - (n-q+3)\frac{\mathbf{r}_m \mathbf{r}_p \mathbf{r}_I^n}{r^{n-q+5}}\right] \\
&= (q-1)\frac{\mathbf{w}_q \mathbf{r}_I^n}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1} r^{n-q+3}} + \left[(n-q+3)\frac{\mathbf{r}_I^n}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-2} r^{n-q+2}} - \frac{\mathbf{w}_m \mathbf{r}_{I,m}^n}{\langle \mathbf{w}, \mathbf{r}\rangle^{q-1} r^{n-q}}\right]\left(\frac{\mathbf{r}_q}{r^3}\right),
\end{aligned}$$

where we have used the fact that

$$\mathbf{r}_m \mathbf{r}^n_{I,m} \;=\; \mathbf{r}_m \sum_{k=1}^{n} \delta_{mI_k} \mathbf{r}^{n-1}_{I\backslash k} \;=\; \sum_{k=1}^{n} \mathbf{r}_{I_k} \mathbf{r}^{n-1}_{I\backslash k} \;=\; n\, \mathbf{r}^n_I. \tag{4.9}$$

Consequently, in terms of solid angle defined in equation (3.27), equation (4.7) simplifies to

$$-\int_A \left[(q-1) \frac{\mathbf{r}^n_I}{\langle \mathbf{w}, \mathbf{r} \rangle^q \, r^{n-q}} - \frac{\mathbf{w}_m \mathbf{r}^n_{I,m}}{\langle \mathbf{w}, \mathbf{r} \rangle^{q-1} \, r^{n-q}} + (n-q+3) \frac{\mathbf{r}^n_I}{\langle \mathbf{w}, \mathbf{r} \rangle^{q-2} \, r^{n-q+2}} \right] d\omega \tag{4.10}$$

According to the definition (4.4), we express equation (4.10) in terms of generalized irradiance tensors, obtaining

$$\int_{\partial A} B^{n+1}_{Il} \, d\mathbf{r}_l \;=\; -(q-1)\mathbf{T}^{n;q}_I - (n-q+3)\mathbf{T}^{n;q-2}_I + \sum_{k=1}^{n} \mathbf{w}_{I_k} \mathbf{T}^{n-1;q-1}_{I\backslash k}, \tag{4.11}$$

where $\mathbf{r}^n_{I,m}$ is expanded, as we did in equation (4.9). Finally, equation (4.6) follows easily from equations (4.7) and (4.11). □□

Theorem 5 reduces the generalized irradiance tensor $\mathbf{T}^{n;q}$ of higher denominator order q to tensors of lower q, finally resulting in a formula in terms of $\mathbf{T}^{m;0}$ (i.e, irradiance tensors $\mathbf{T}^m$) and $\mathbf{T}^{m;1}$ $(m \leq n)$ only. Irradiance tensors have been addressed in section 3.2.2 and can be computed by equation (3.32), while a recurrence relation for $\mathbf{T}^{n;1}$ can be derived from another useful formula about $\mathbf{T}^{n;q}$ given below, which was originally identified by Arvo [8, p.112].

Theorem 6 *Let $n \geq 1$ and $q \geq 0$ be integers. If $q \neq n+1$, then the tensor $\mathbf{T}^{n;q}(A, \mathbf{w})$ satisfies the recurrence relation*

$$\mathbf{T}^{n;q}_{Ij} \;=\; \frac{1}{n-q+1} \left(\sum_{k=1}^{n-1} \delta_{jI_k} \mathbf{T}^{n-2;q}_{I\backslash k} - q\, \mathbf{w}_j \mathbf{T}^{n-1;q+1}_I - \int_{\partial A} \frac{\mathbf{u}^{n-1}_I \mathbf{n}_j}{\langle \mathbf{w}, \mathbf{u} \rangle^q} \, ds \right), \tag{4.12}$$

where I is a $(n-1)$-index, ds and $\mathbf{n}$ are the same as those in Theorem 5.

Proof: The proof parallel that of Theorem 5. See Appendix A.9. □□

From Theorem 5 and Theorem 6, we are able to derive a formula for $\mathbf{T}^{n;1}$. For integer $n > 0$, we may apply equation (4.12) to express $\mathbf{T}^{n;1}$ as

$$\mathbf{T}^{n;1}_{Ij} = \frac{1}{n}\left(\sum_{k=1}^{n-1}\delta_{jI_k}\mathbf{T}^{n-2;1}_{I/k} - \mathbf{w}_j\mathbf{T}^{n-1;2}_{I} - \int_{\partial A}\frac{\mathbf{u}^{n-1}_I\mathbf{n}_j}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds\right).$$

Replacing $\mathbf{T}^{n-1;2}$ above by equation (4.5), that is,

$$\mathbf{T}^{n-1;2}_{I} = \sum_{k=1}^{n-1}\mathbf{w}_{I_k}\mathbf{T}^{n-2;1}_{I/k} - n\mathbf{T}^{n-1;0}_{I} - \int_{\partial A}\frac{\mathbf{u}^{n-1}_I\langle\mathbf{w},\mathbf{n}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds,$$

we get

$$\mathbf{T}^{n,1}_{Ij} = \mathbf{w}_j\mathbf{T}^{n-1}_{I} + \frac{1}{n}\left[\sum_{k=1}^{n-1}\left(\delta_{jI_k} - \mathbf{w}_j\mathbf{w}_{I_k}\right)\mathbf{T}^{n-2,1}_{I/k} + \int_{\partial A}\frac{\mathbf{u}^{n-1}_I\left(\mathbf{w}_j\langle\mathbf{w},\mathbf{n}\rangle - \mathbf{n}_j\right)}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds\right].$$

Noticing that

$$\begin{aligned}\delta_{jk} - \mathbf{w}_j\mathbf{w}_k &= \left(\delta_{ji} - \mathbf{w}_j\mathbf{w}_i\right)\delta_{ik},\\ \mathbf{n}_j - \mathbf{w}_j\langle\mathbf{w},\mathbf{n}\rangle &= \left(\delta_{ji} - \mathbf{w}_j\mathbf{w}_i\right)\mathbf{n}_i,\end{aligned}$$

we arrive at the following theorem:

Theorem 7 *The tensor $\mathbf{T}^{n;1}$ $(n > 0)$ satisfies the recurrence relation*

$$\mathbf{T}^{n;1}_{Ij}(A,\mathbf{w}) = \mathbf{w}_j\mathbf{T}^{n-1}_{I} + \frac{1}{n}\left(\delta_{jm} - \mathbf{w}_j\mathbf{w}_m\right)\left[\sum_{k=1}^{n-1}\delta_{mI_k}\mathbf{T}^{n-2;1}_{I\backslash k} - \int_{\partial A}\frac{\mathbf{u}^{n-1}_I\mathbf{n}_m}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds\right], \quad (4.13)$$

where I is an $(n-1)$-index.

Proof: See Appendix A.10. □□

To complete the recurrence shown in (4.13), we end this section with a theorem that expresses the base case $\mathbf{T}^{0;1}$ in terms of a boundary integral.

Theorem 8 *The base case* $\mathbf{T}^{0;1}(A, \mathbf{w})$ *is given by*

$$\mathbf{T}^{0;1}(A, \mathbf{w}) = \int_{\partial A} \frac{\ln \langle \mathbf{w}, \mathbf{u} \rangle}{1 - \langle \mathbf{w}, \mathbf{u} \rangle^2} \langle \mathbf{w}, \mathbf{n} \rangle \, ds = \int_{\partial A} \eta(\mathbf{w}, \mathbf{u}) \langle \mathbf{w}, \mathbf{n} \rangle \, ds, \tag{4.14}$$

where η *is given in equation (3.82).*

Proof: See Appendix A.11. □□

It is not surprising that the function $\eta(\mathbf{w}, \mathbf{u})$ defined on page 84 arises here, as equation (4.14) was derived using a similar expansion approach employed in chapter 3. According to the definition (4.4), $\mathbf{T}^{0;1}$ is given by a surface integral, that is

$$\mathbf{T}^{0;1}(A, \mathbf{w}) = \int_A \frac{1}{\langle \mathbf{w}, \mathbf{u} \rangle} \, d\sigma(\mathbf{u}).$$

Expanding the integrand as a Taylor series, as we did in section 3.3.1, we can derive equation (4.14) as follows:

$$\begin{aligned}
\mathbf{T}^{0,1}(A, \mathbf{w}) &= \int_A \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \langle \mathbf{w}, \mathbf{u} \rangle^k \, d\sigma(\mathbf{u}) \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \int_A \langle \mathbf{w}, \mathbf{u} \rangle^k \, d\sigma(\mathbf{u}) \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \bar{\tau}^k(A, \mathbf{w}) \\
&= -\sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1}{k+1} \int_{\partial A} \frac{1 - \langle \mathbf{w}, \mathbf{u} \rangle^{k+1}}{1 - \langle \mathbf{w}, \mathbf{u} \rangle^2} \langle \mathbf{w}, \mathbf{n} \rangle \, ds \\
&= -\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{1 - \langle \mathbf{w}, \mathbf{u} \rangle^2} \left(\sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1 - \langle \mathbf{w}, \mathbf{u} \rangle^{k+1}}{k+1} \right) ds \\
&= \int_{\partial A} \frac{\ln \langle \mathbf{w}, \mathbf{u} \rangle}{1 - \langle \mathbf{w}, \mathbf{u} \rangle^2} \langle \mathbf{w}, \mathbf{n} \rangle \, ds,
\end{aligned}$$

where we have used the formula (3.42) for the axial moment $\bar{\tau}^k$ and the double summation identity (3.67).

Considering the connection between η and the special function Λ arising in computing

the irradiance due to a linearly-varying luminaire, Theorem 8 suggests that the special function Λ may essentially come from $\mathbf{T}^{0;1}$. Furthermore, on observing that $\mathbf{T}^{0;1}$ is the base case of generalized irradiance tensors $\mathbf{T}^{n;q}$, which are closely related to the irradiance problem due to polynomially-varying luminaires, we may speculate that the irradiance due to all polynomially-varying Lambertian luminaires probably depend on the same special function, which will be verified in the next chapter.

To sum up, for all non-negative integers n and q, we may obtain a contour integral representation for generalized irradiance tensors $\mathbf{T}^{n;q}$ by recursively applying the three formulas stated respectively in theorem 5, theorem 7 and theorem 8, as well as the recurrence formula for irradiance tensors given in equation (3.32).

4.3 Rational Angular Moments

In simulating some illuminating and scattering effects involving polynomially-varying luminaires, we are required to compute moments of the radiance distribution function $f(\mathbf{u})$ from such emitters. With $f(\mathbf{u})$ given in (4.3), these moments are expressed as rational polynomials integrated over the sphere. To study them, we generalize the concept of axial moments in section 3.2.3 to accommodate a rational component coming from the spatial variation of the emitters. Given an arbitrary subset $A \subset S^2$, a unit vector $\mathbf{w} \in \mathcal{S}^2$ and n vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$ which may not be normalized[1], we define a form of constrained rational polynomials by

$$\tau^{n;q}(A, \mathbf{v}_1, \cdots, \mathbf{v}_n; \mathbf{w}) \equiv \int_A \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle \cdots \langle \mathbf{v}_n, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^q} \, d\sigma(\mathbf{u}) \tag{4.15}$$

for integers $n, q \geq 0$. We call them *rational angular moments*, denoted by $\tau^{n;q}(A, \mathbf{v}_1, \cdots, \mathbf{v}_n; \mathbf{w})$. The axes $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ serve as two purposes: some of them encode the emission distribution of a monomially-varying luminaire, which in combination with $\mathbf{w}$ characterize the radiance distribution $f(\mathbf{u})$ from this emitter; while other axes define a polynomial weighting function for $f(\mathbf{u})$. Similar to $\mathbf{T}^{n;q}$, we also indicate the numerator and denominator orders of its rational polynomial integrand of $\tau^{n;q}$ with the superscript $n; q$. For

[1]The vectors are required to be normalized only when equation (3.39) about the axial moment is used.

consistency, in the argument list, we separate the numerator parameters $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and the denominator parameter $\mathbf{w}$ also by a semicolon. In this section we shall find that these rational polynomials associated with $\tau^{n;q}$ can be integrated exactly over polygonal domains with the aid of generalized irradiance tensors $\mathbf{T}^{n;q}$.

4.3.1 Contour Integral Representations

From equations (4.5), (4.13), (3.32) and (4.14), we may obtain corresponding expressions for rational moments, which reduce the area integrals of $\tau^{n;q}$ to a contour integral representation. This is the key idea underlying our tensor formalism and also of great practical importance since the number of elements of $\mathbf{T}^{n;q}$ increases exponentially with n. Although it is these individual *scalar* moments that are required essentially for image synthesis, the tensor formulation $\mathbf{T}^{n;q}$ provides us a powerful tool to represent a family of rational polynomials parameterized by weighting axes $\mathbf{v}_1, \ldots, \mathbf{v}_n$ by means of tensor composition. It is more advantageous in that the recurrence formula derived for a single tensor $\mathbf{T}^{n;q}$ may lead to a family of expressions for individual rational moments.

We begin by expressing the integrand of equation (4.15) as a composition of tensors:

$$\frac{\langle \mathbf{v}_1, \mathbf{u}\rangle \cdots \langle \mathbf{v}_n, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^q} \;=\; \frac{\mathbf{u}_I^n}{\langle \mathbf{w}, \mathbf{u}\rangle^q} \left(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_n\right)_I ,$$

where the summation convention is implied for the n-index I. It follows from the definition (4.4) that

$$\tau^{n,q} \;=\; \mathbf{T}_I^{n,q} \left(\mathbf{v}_1 \otimes \cdots \mathbf{v}_n\right)_I , \tag{4.16}$$

which associates a rational angular moment with a generalized irradiance tensor of the same order. Consequently, recurrence expressions for $\tau^{n;q}$ follow directly from expanding $\mathbf{T}^{n;q}$ in equation (4.16). To express these new formulas, we introduce a short-hand notation for the parameters of $\tau^{n;q}$. Let $J = (1, 2, \cdots, n)$ be a n-index. We shall define $J\backslash k = (1, \cdots, k-1, k+1, \cdots, n)$ and use $\mathbf{v}_J$ for the axis set $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$. Using these

notations, we get

$$\tau^{n;q}(A, \mathbf{v}_J; \mathbf{w}) = \frac{1}{q-1}\left[\sum_{k=1}^{n} \langle \mathbf{w}, \mathbf{v}_k \rangle \, \tau^{n-1;q-1}(A, \mathbf{v}_{J\setminus k}; \mathbf{w}) - (n-q+3)\tau^{n;q-2}(A, \mathbf{v}_J; \mathbf{w}) - \int_{\partial A} \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle \cdots \langle \mathbf{v}_n, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^{q-1}} \langle \mathbf{w}, \mathbf{n} \rangle \, ds\right], \tag{4.17}$$

$$\tau^{n;1}(A, \mathbf{v}_J; \mathbf{w}) = \langle \mathbf{w}, \mathbf{v}_n \rangle \, \tau^{n-1}(A, \mathbf{v}_{J\setminus n}) + \frac{\mathbf{v}_n{}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})}{n}\left[\sum_{k=1}^{n-1} \mathbf{v}_k \tau^{n-2;1}(A, \mathbf{v}_{J\setminus\{k,n\}}; \mathbf{w}) - \int_{\partial A} \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle \cdots \langle \mathbf{v}_{n-1}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \mathbf{n} \, ds\right], \tag{4.18}$$

$$\tau^{n}(A, \mathbf{v}_J) = \frac{1}{n+1}\left[\sum_{k=1}^{n-1} \langle \mathbf{v}_n, \mathbf{v}_k \rangle \, \tau^{n-2}(A, \mathbf{v}_{J\setminus\{n,k\}}) - \int_{\partial A} \langle \mathbf{v}_1, \mathbf{u} \rangle \cdots \langle \mathbf{v}_{n-1}, \mathbf{u} \rangle \langle \mathbf{v}_n, \mathbf{n} \rangle \, ds\right], \tag{4.19}$$

which correspond to recurrence relations (4.5), (4.13) and (3.32), respectively. Similarly, Equation (4.17) holds for $q \geq 2$. In accordance with the base cases of $\mathbf{T}^{n;q}$, we have $\tau^{-1}(A) = 0$, $\tau^{0}(A) = \sigma(A)$ and $\tau^{0;1}(A; \mathbf{w}) = \mathbf{T}^{0;1}(A, \mathbf{w})$, given by (4.14).

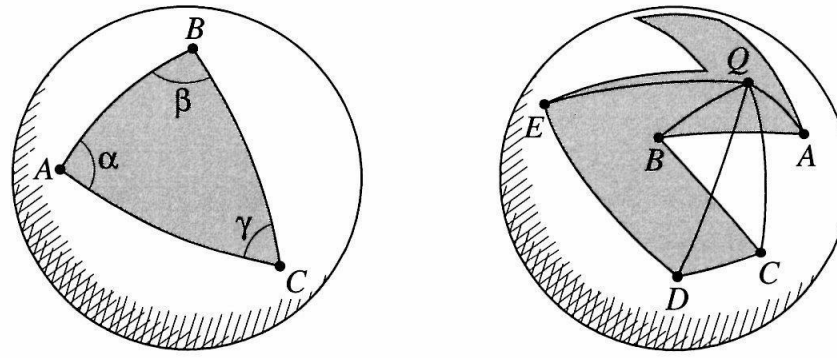
With the contour integral representation now derived for $\tau^{n;q}$, we may proceed to evaluate the resulting boundary integral terms for certain geometries.

4.3.2 Exact Evaluation

Repeated applications of equations (4.17), (4.18), (4.19) and (4.14) reduce the integrals $\tau^{n;q}$ to a sequence of boundary integrals and/or solid angle $\sigma(A)$. In this section, we shall specialize the integration domain A as a spherical polygon and describe how to compute each component exactly (depending on a single special function $\Lambda(\alpha, \beta)$, or essentially, $\mathrm{Cl}_2(x)$), which becomes feasible due to those properties of spherical polygons mentioned in section 3.4.1.

Solid Angle

The solid angle $\sigma(A)$ subtended by a polygon A on a unit sphere is equivalent to the area of A, which can be computed directly using two methods given below.



(a) (b)

Figure 4.2: *(a) The solid angle of a spherical triangle is easily obtained from Girard's formula. (b) One approach for computing the solid angle of non-convex polygons is to break it into triangles. (Courtesy of James Arvo [9]).*

The first approach was proposed by Arvo [9]. Girard's formula for spherical triangles [20, pp.278] states that

$$\sigma(\triangle ABC) \;=\; \alpha + \beta + \gamma - \pi, \tag{4.20}$$

where α, β, γ are the three internal angles of $\triangle ABC$, as shown in Figure 4.2a. The internal angles are the dihedral angles between the planes containing the edges. To compute the solid angle subtended by an arbitrary spherical polygon A shown in Figure 4.2b, we may break it into k triangles, all sharing an arbitrary vertex $Q \in \mathcal{S}^2$. The solid angle is then the sum of the triangle areas signed according to orientation.

However, the identity (3.41) described on Page 72 provides us an alternative to generalize the formula (4.20) to an arbitrary spherical polygon A. To apply equation (3.41), we pick a vector $\mathbf{w} \in \mathcal{S}^2$ such that $-\mathbf{w} \notin A$. Considering A as the spherical projection of a polygon $P \in \mathbb{R}^3$ at the origin, we may choose $\mathbf{w}$ as the unit vector orthogonal to the plane, and thus $-\mathbf{w} \notin A$. Using the property of constant $\mathbf{n}$ over the edge, we may

compute $\sigma(A)$ from (3.41) by

$$\sigma(A) \;=\; -\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{1 + \langle \mathbf{w}, \mathbf{u} \rangle}\, ds \;=\; -\sum_{i=1}^{k} \langle \mathbf{w}, \mathbf{n} \rangle \int_{\zeta} \frac{1}{1 + \langle \mathbf{w}, \mathbf{u} \rangle}\, ds$$

for a k-sided polygon. With the edge parameterization introduced in section 3.4.1, each line integral above can be evaluated within constant time as follows [49, p.180]:

$$\begin{aligned}
\int_{\zeta} \frac{1}{1 + \langle \mathbf{w}, \mathbf{u} \rangle}\, ds \;&=\; \int_{0}^{\Theta} \frac{1}{1 + c_{\mathbf{w}} \cos(\theta - \phi_{\mathbf{w}})}\, d\theta \\
&=\; -\int_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}} \frac{1}{1 + c_{\mathbf{w}} \cos\theta}\, d\theta \\
&=\; \begin{cases}
\left[\frac{2}{\sqrt{1 - c_{\mathbf{w}}^2}} \tan^{-1}\left(\frac{(1 - c_{\mathbf{w}}) \tan(\theta/2)}{\sqrt{1 - c_{\mathbf{w}}^2}} \right) \right]_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}}, & (0 \le c_{\mathbf{w}} < 1) \\
\left[\tan\left(\frac{\theta}{2}\right) \right]_{-\phi_{\mathbf{w}}}^{\Theta - \phi_{\mathbf{w}}}, & (c_{\mathbf{w}} = 1)
\end{cases}
\end{aligned}$$

where $c_{\mathbf{w}}, \phi_{\mathbf{w}}$ are defined by (3.106) and (3.107), and $0 \le c_{\mathbf{w}} \le 1$ for a unit vector $\mathbf{w}$, and Θ is the arc length of ζ.

Base Moment $\tau^{0;1}(A; \mathbf{w})$

Given by a boundary integral in equation (4.14), we evaluate the base moment $\tau^{0;1}$ along each edge, yielding

$$\tau^{0;1}(A; \mathbf{w}) \;=\; \sum_{i} \langle \mathbf{w}, \mathbf{n} \rangle \int_{\zeta} \eta(\mathbf{w}, \mathbf{u})\, ds \;=\; \sum_{i} \langle \mathbf{w}, \mathbf{n} \rangle\, B^*(\zeta), \tag{4.21}$$

where B^*, depending on ζ, is given in (3.112) on Page 95. Section 3.4.2 has discussed the evaluation of B^* in detail, the solution is *closed form* except for one special function Λ that can be expressed in terms of Clausen integrals. The pseudo-code for computing this base moment for a polygon P is given by *BaseMoment*. If we assume that it takes constant time to compute the special function Λ using any one of the approaches (such as direct method, bilinear interpolation, hybrid method) discussed in section 3.4.3, then

the time complexity for evaluating $\tau^{0;1}$ is $\mathrm{O}(k)$ for a k-sided polygon A.

BaseMoment(**Polygon** P, **vector w**)

real $S \leftarrow 0$; *Accumulate the sum*
for each edge **AB** in P **do**
vector s $\leftarrow$ **unit**[**A**];
vector t $\leftarrow$ **unit**[$(\mathbf{I} - \mathbf{ss}^{\mathrm{T}})\mathbf{B}$];
real $a \leftarrow \langle \mathbf{w}, \mathbf{s} \rangle$;
real $b \leftarrow \langle \mathbf{w}, \mathbf{t} \rangle$;
real $c \leftarrow \sqrt{a^2 + b^2}$;
real $\phi \leftarrow$ **sign**[b] $* \cos^{-1}(a/c)$;
real $\Theta \leftarrow \cos^{-1}(\langle$**unit**[**A**], **unit**[**B**]$\rangle)$;
vector n $\leftarrow$ **unit**[**A** $\times$ **B**];
real $\overline{B^*} \leftarrow \Lambda(c, \Theta - \phi) - \Lambda(c, -\phi)$;
$S \leftarrow S + \langle \mathbf{w}, \mathbf{n} \rangle * \overline{B^*}$;
endfor

Boundary Integrals of Rational Polynomials

After moving the factors involving **n** outside the boundary integrals involved in equations (4.17), (4.18) and (4.19), we are required to compute a type of line integral, defined by

$$B^{m,l}(\mathbf{v}_1, \ldots, \mathbf{v}_m) \equiv \int_\zeta \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle \cdots \langle \mathbf{v}_m, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^l} \, ds \tag{4.22}$$

for integers m, l and an arc ζ, where **w** is considered as constant (for a planar luminaire) and $0 \leq m \leq n$, $0 \leq l \leq q$ for a given moment $\tau^{n;q}$ to be evaluated.

For smaller m and l, we can easily derive a direct formula to evaluate $B^{m,l}$ from our

edge parameterization. Without further notice, the varibles c and ϕ in the remainder of this thesis are always defined on an arc ζ by (3.106) and (3.107), and Θ denotes the arc length of ζ. We shall use $\overline{B}$ for the expression to compute B. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be unit vectors, here are some examples of these formulas that are relevant to our work in the next chapter.

$$\begin{aligned}
\overline{B}^1(\mathbf{a}) &= c_{\mathbf{a}} \int_0^{\Theta} \cos(\theta - \phi_{\mathbf{a}})\, d\theta \\
&= c_{\mathbf{a}} \left(\sin(\Theta - \phi_{\mathbf{a}}) + \sin(\phi_{\mathbf{a}})\right), \qquad (4.23) \\
\overline{B}^2(\mathbf{a}, \mathbf{b}) &= c_{\mathbf{a}} c_{\mathbf{b}} \int_0^{\Theta} \cos(\theta - \phi_{\mathbf{a}}) \cos(\theta - \phi_{\mathbf{b}})\, d\theta \\
&= \frac{c_{\mathbf{a}} c_{\mathbf{b}}}{4} \left[2\theta \cos(\phi_{\mathbf{a}} - \phi_{\mathbf{b}}) - \sin(\phi_{\mathbf{a}} + \phi_{\mathbf{b}} - 2\theta)\right]_0^{\Theta} \\
&= \frac{c_{\mathbf{a}} c_{\mathbf{b}}}{4} \Big[2\Theta \cos(\phi_{\mathbf{a}} - \phi_{\mathbf{b}}) - \sin(\phi_{\mathbf{a}} + \phi_{\mathbf{b}} - 2\Theta) + \\
&\qquad \sin(\phi_{\mathbf{a}} + \phi_{\mathbf{b}})\Big] \qquad (4.24)
\end{aligned}$$

for $l = 0$ and

$$\begin{aligned}
\overline{B}^{1,1}(\mathbf{a}) &= \frac{c_{\mathbf{a}}}{c_{\mathbf{w}}} \left[\cos(\phi_{\mathbf{w}} - \phi_{\mathbf{a}})\Theta + \sin(\phi_{\mathbf{w}} - \phi_{\mathbf{a}}) \ln \frac{\cos(\Theta - \phi_{\mathbf{w}})}{\cos \phi_{\mathbf{w}}}\right], \qquad (4.25) \\
\overline{B}^{3,1}(\mathbf{a}, \mathbf{b}, \mathbf{c}) &= \frac{c_{\mathbf{a}} c_{\mathbf{b}} c_{\mathbf{c}}}{c_{\mathbf{w}}} \int_0^{\Theta} \frac{\cos(\theta - \phi_{\mathbf{a}}) \cos(\theta - \phi_{\mathbf{b}}) \cos(\theta - \phi_{\mathbf{c}})}{\cos(\theta - \phi_{\mathbf{w}})}\, d\theta \\
&= \frac{c_{\mathbf{a}} c_{\mathbf{b}} c_{\mathbf{c}}}{4 c_{\mathbf{w}}} \left[a_1 \theta + a_2 \ln \cos(\theta - \phi_{\mathbf{w}}) + \sin(t + 2\theta)\right]_0^{\Theta} \\
&= \frac{c_{\mathbf{a}} c_{\mathbf{b}} c_{\mathbf{c}}}{4 c_{\mathbf{w}}} \left[a_1 \Theta + a_2 \ln \frac{\cos(\Theta - \phi_{\mathbf{w}})}{\cos(\phi_{\mathbf{w}})} + \sin(t + 2\Theta) - \sin t\right] \qquad (4.26)
\end{aligned}$$

for $l = 1$, where a_1, a_2, t are given by

$$\begin{aligned}
t &= \phi_{\mathbf{w}} - \phi_{\mathbf{a}} - \phi_{\mathbf{b}} - \phi_{\mathbf{c}} \\
t_1 &= 3\phi_{\mathbf{w}} - \phi_{\mathbf{a}} - \phi_{\mathbf{b}} - \phi_{\mathbf{c}} \\
t_2 &= \phi_{\mathbf{w}} + \phi_{\mathbf{a}} - \phi_{\mathbf{b}} - \phi_{\mathbf{c}} \\
t_3 &= \phi_{\mathbf{w}} - \phi_{\mathbf{a}} + \phi_{\mathbf{b}} - \phi_{\mathbf{c}}
\end{aligned}$$

$$
\begin{aligned}
t_4 &= \phi_{\mathbf{w}} - \phi_{\mathbf{a}} - \phi_{\mathbf{b}} + \phi_{\mathbf{c}} \\
a_1 &= -\cos t_1 + \cos t_2 + \cos t_3 + \cos t_4 \\
a_2 &= -\sin t_1 + \sin t_2 + \sin t_3 + \sin t_4.
\end{aligned}
$$

Equation (4.25) has appeared in the closed-form solution for the irradiance due to a linearly-varying polygonal luminaire shown in (3.116).

Next we investigate the exact evaluation of $B^{m,l}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ for more general values of m and l. For notational convenience, let $c_i = c_{\mathbf{v}_i}$, $\phi_i = \phi_{\mathbf{v}_i}$ for $1 \le i \le m$ and $c = c_{\mathbf{w}}$, $\phi = \phi_{\mathbf{w}}$. We define

$$
\begin{aligned}
\alpha_i &= \cos(\phi_i - \phi), \\
\beta_i &= \sin(\phi_i - \phi), \\
r_i &= \alpha_i / \beta_i
\end{aligned}
$$

for $i = 1, 2, \ldots, m$. Then by a change of variable, we have

$$
\begin{aligned}
\overline{B}^{m,l} &= \frac{c_1 \cdots c_m}{c^l} \int_0^{\Theta} \frac{\cos(\theta - \phi_1) \cdots \cos(\theta - \phi_m)}{\cos^l(\theta - \phi)} \, d\theta \\
&= \frac{c_1 \cdots c_m}{c^l} \int_{-\phi}^{\Theta - \phi} \frac{\prod_{i=1}^{m} \cos(\theta + \phi - \phi_i)}{\cos^l \theta} \, d\theta \\
&= \frac{c_1 \cdots c_m}{c^l} \int_{-\phi}^{\Theta - \phi} \frac{\prod_{i=1}^{m} (\alpha_i \cos\theta + \beta_i \sin\theta)}{\cos^l \theta} \, d\theta \\
&= \frac{c_1 \cdots c_m \, \beta_1 \cdots \beta_m}{c^l} \int_{-\phi}^{\Theta - \phi} \left[\prod_{i=1}^{m} (r_i + \tan\theta) \right] \cos^{m-l} \theta \, d\theta. \qquad (4.27)
\end{aligned}
$$

Let

$$
P_m(x) \equiv \prod_{i=1}^{m} (x + r_i) = \sum_{i=0}^{m} p_i(r_1, \ldots, r_m) \, x^i \qquad (4.28)
$$

denote a mth-order polynomial of x that has $-r_1, -r_2, \cdots, -r_m$ as its m roots, where the coefficients $p_m = 1$ and p_i $(i = 0, 1, \ldots, m-1)$ are uniquely determined by $r_1, \ldots, r_m$.

Then equation (4.27) becomes

$$\begin{aligned}\overline{B}^{m,l} &= \frac{c_1 \cdots c_m \beta_1 \cdots \beta_m}{c^l} \int_{-\phi}^{\Theta-\phi} P_m(\tan\theta) \cos^{m-l}\theta \, d\theta \\ &= \frac{c_1 \cdots c_m \beta_1 \cdots \beta_m}{c^l} \sum_{i=0}^{m} p_i(r_1, \ldots, r_m) \int_{-\phi}^{\Theta-\phi} \sin^i\theta \cos^{m-l-i}\theta \, d\theta. \quad (4.29)\end{aligned}$$

Therefore, the analytical evaluation of the boundary integral in equation (4.29) is based on integrals of the form

$$G(r, s, x, y) \equiv \int_x^y \sin^r\theta \, \cos^s\theta \, d\theta, \quad (4.30)$$

where the integers $r \geq 0$ and s may be negative. From equation (4.29), we obtain

$$\overline{B}^{m,l} = \frac{c_1 \cdots c_m \beta_1 \cdots \beta_m}{c^l} \sum_{i=0}^{m} p_i(r_1, \ldots, r_m) \, G(i, m-l-i, -\phi, \Theta-\phi). \quad (4.31)$$

The integrals defined in equation (4.30) can be evaluated exactly using the recurrence relations given below (see Appendix A.12)

$$\begin{aligned}G(r, s, x, y) &= \frac{1}{r+s}\left[\sin^{r+1} y \cos^{s-1} y - \sin^{r+1} x \cos^{s-1} x + (s-1)G(r, s-2, x, y)\right] \\ G(r, s, x, y) &= \frac{1}{s+1}\left[\sin^{r+1} x \cos^{s+1} x - \sin^{r+1} y \cos^{s+1} y + (r+s+2)G(r, s+2, x, y)\right] \\ G(r, s, x, y) &= \frac{1}{r+s}\left[\sin^{r-1} x \cos^{s+1} x - \sin^{r-1} y \cos^{s+1} y + (r-1)G(r-2, s, x, y)\right], \quad (4.32)\end{aligned}$$

where the first two formulas apply for $s > 0$ and $s < 0$ respectively and the base cases are comprised of

$$\begin{aligned}G(0, -1, x, y) &= \ln\left(\frac{\tan(\pi/4 + y/2)}{\tan(\pi/4 + x/2)}\right), \\ G(1, -1, x, y) &= \ln\left(\frac{\cos x}{\cos y}\right), \\ G(0, 0, x, y) &= y - x,\end{aligned}$$

$$G(1,0,x,y) \quad = \quad \cos x - \cos y.$$

Consequently, the line integral $B^{m,l}$ defined in (4.22) can be evaluated exactly by equation (4.31).

We now analyze the complexity for evaluating $B^{m,l}$ using equation (4.31), which is comprised of two parts: (1) time used to expand P_m in (4.28) to obtain the coefficients p_i for $i = 1, 2, \ldots, m$; (2) time to compute $G(r, s, x, y)$ for given integers r and s. To obtain all the m coefficients $p_0, p_1, \cdots, p_{m-1}$ for the polynomial $P_m(x)$ defined in (4.28), we expand the product $\prod_{i=1}^{m}(x + r_i)$ incrementally, as shown in the pseudo-code *PolynomialCoeffs* below. In the internal loop, we multiply the current $(k-1)$th order

PolynomialCoeffs(**real** $r[\,]$, **integer** m)

real $p[\,]$ *coefficient array*
Initialize the current polynomial as 1
$p[0] \leftarrow 1$
Incrementally multiply the current $(k-1)$*th order polynomial by* $x + r_k$
for $k \leftarrow 1, 2, \ldots, m$ **do**
  $p[k] \leftarrow p[k-1]$
  for $i \leftarrow k-1, k-2, \ldots, 1$ **do**
    $p[i] \leftarrow p[i] * r[k] + p[i-1]$
  endfor
  $p[0] \leftarrow p[0] * r[k]$
endfor
return p

polynomial by a factor $(x + r_k)$, and update the current coefficients using the following recurrence relation:

$$\left(\sum_{i=0}^{k-1} p_i\, x^i\right)(x + r_k) \quad = \quad \sum_{i=0}^{k-1} p_i\, x^{i+1} + \sum_{i=0}^{k-1} p_i r_k\, x^i$$

$$= \quad p_0 r_k + \sum_{i=1}^{k-1} (p_i r_k + p_{i-1})\, x^i + p_{k-1}\, x^k,$$

taking about $O(k+1)$ for each k. Thus the whole loop takes

$$\sum_{k=1}^{m} (k+1) \;=\; O(m^2).$$

Discussion: We may convert n polynomial roots to coefficients in $O(n \log n)$ time (for an nth degree polynomial), by using a fact that two nth degree polynomials in coefficient form can be multiplied in $O(n \log n)$ time using the FFT and its inverse. Thus we can multiply n linear terms $(x+r_k)$ $(k = 1, 2, \ldots, n)$ via divide-and-conquer as follows: divide the terms into two groups and multiply them recursively, then multiply the two resulting polynomials in $\frac{n}{2} \log \frac{n}{2}$ time. It then follows from the recurrence relation

$$T(n) = 2T(n/2) + \frac{n}{2} \log \frac{n}{2}$$

that the overall complexity is $O(n \log n)$. However, this algorithm is not practical for two reasons: 1)the constant will be large, thus $O(n^2)$ algorithm will probably win until n is very large. 2)all the FFTs and inverse FFTs will cause too much numerical errors.

On the other hand, by means of those recurrence relations given above, we can evaluate the function $G(r, s, x, y)$ for a given r and s within $\lfloor (r + |s|)/2 + 1 \rfloor$ steps. Therefore, with all the p_i's determined, the summation in equation (4.31) will take approximately

$$T \;=\; \sum_{i=0}^{m} \frac{i + |m - l - i|}{2},$$

which can be analyzed by distinguishing the following three cases:

1. $m > l$: that is, $m - l > 0$, we have

$$T \;=\; \sum_{i=0}^{m-l} \frac{m-l}{2} + \sum_{i=m-l+1}^{m} \left(i - \frac{m-l}{2} \right)$$

$$
\begin{aligned}
&= \frac{(m-l)(m-l+1)}{2} + \sum_{j=1}^{l}\left(j + \frac{m-l}{2}\right) \\
&= \frac{(m-l)(m+1)}{2} + \frac{l(l+1)}{2} \\
&= \frac{m^2 + l^2 + m - lm}{2} \\
&< \frac{m(m+1)}{2}.
\end{aligned}
$$

2. $m = l$:

$$T = \sum_{i=0}^{m} i = \frac{m(m+1)}{2}.$$

3. $m < l$: thus $m - l < 0$, we have

$$
\begin{aligned}
T &= \sum_{i=0}^{m}\left(i + \frac{l-m}{2}\right) \\
&= \frac{m(m+1)}{2} + \frac{(l-m)(m+1)}{2} \\
&= \frac{l(m+1)}{2}.
\end{aligned}
$$

Therefore, equation (4.31) for a given l and m takes approximately $\lceil \max(m,l)(m+1)/2 \rceil$ to evaluate exactly. Considering the cost of $O(m^2)$ for determining p_i's, the time complexity for computing a line integral $B^{m,l}$ is $O(\max(m^2, l * m))$.

4.4 Moments Corresponding to Phong Distributions

In the previous section we have defined a general form of rational angular moments in (4.15), which places no restrictions on the parameters $\mathbf{v}_1, \ldots, \mathbf{v}_n$. However, while simulating non-Lambertian rendering effects for environments with Phong-like emission and reflectance distributions [87], we often encounter a special class of rational moments,

where the axes in $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are repeated multiple times. For example,

$$\tau^{n;1}(A, \underbrace{\mathbf{v}, \ldots, \mathbf{v}}_{n}; \mathbf{w}),\ \tau^{n+2;1}(A, \underbrace{\mathbf{v}, \ldots, \mathbf{v}}_{n}, \mathbf{b}, \mathbf{a}; \mathbf{w}),\ \ldots \tag{4.33}$$

To explicitly specify the number of distinct axes appearing in the numerator of such moments, we introduce a special notation $\tilde{\tau}$ for moments with the following form

$$\tilde{\tau}^{p_1,p_2,\cdots;q}(A, \mathbf{v}_1, \mathbf{v}_2, \ldots; \mathbf{w}) \equiv \int_A \frac{\langle \mathbf{v}_1, \mathbf{u}\rangle^{p_1} \langle \mathbf{v}_2, \mathbf{u}\rangle^{p_2} \cdots}{\langle \mathbf{w}, \mathbf{u}\rangle^{q}} d\sigma(\mathbf{u}), \tag{4.34}$$

where the factors involving $\mathbf{v}_1, \mathbf{v}_2, \ldots$ are raised respectively to the orders $p_1, p_2, \ldots$, integers arranged before the semicolon in the superscript, while the factor $\langle \mathbf{w}, \mathbf{u}\rangle$ in the denominator is raised to order q, the single number following the semicolon. Accordingly, the arguments of $\tilde{\tau}^{p_1,p_2,\cdots;q}$ consist of the same number of $\mathbf{v}$ axes as that of integers appearing before the superscript semicolon. As we can see, this special moment is related to the general rational moment $\tau^{n;q}$ by

$$\tilde{\tau}^{p_1,p_2,\cdots;q}(A, \mathbf{v}_1, \mathbf{v}_2, \ldots; \mathbf{w}) = \tau^{p_1+p_2+\cdots;q}(A, \underbrace{\mathbf{v}_1, \ldots, \mathbf{v}_1}_{p_1}, \underbrace{\mathbf{v}_1, \ldots, \mathbf{v}_2}_{p_2}, \ldots; \mathbf{w}).$$

Consequently, using this new notation, the examples shown in (4.33) can be characterized as $\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w})$ and $\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}, \mathbf{b}, \mathbf{a}; \mathbf{w})$, respectively. In this section, we shall show that the occurrence of repeated axes in $\tilde{\tau}$ simplifies the recurrence formulas derived for $\tau^{n;q}$ in section 4.3, leading to more efficient evaluation algorithms. Specifically, we focus on the case of $q = 1$,that is, $\tilde{\tau}^{p_1,p_2,\cdots;1}$. But similar ideas can be extended to the recurrence relations for $q \geq 2$. The result, however, may become a little complicated.

4.4.1 Simplified Recurrence Formulas

For simplicity, we shall only consider the case where only the first numerator axis is raised to the power n, and all the others are of order 1, i.e. the case where $p_1 = n$ and $p_i = 1$ for $i > 1$. We first consider Phong-like moments about one axis, and then extend it to multiple axes.

Given an arbitrary region $A \subset \mathcal{S}^2$ and a vector $\mathbf{v}$, we define the moment about $\mathbf{v}$ by

$$\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w}) \equiv \int_A \frac{\langle \mathbf{v}, \mathbf{u} \rangle^n}{\langle \mathbf{w}, \mathbf{u} \rangle} d\sigma(\mathbf{u}), \tag{4.35}$$

Letting $\mathbf{v}_1 = \mathbf{v}_2 = \cdots = \mathbf{v}_n = \mathbf{v}$ in equation (4.18), we get a simplified recurrence formula for this one-axis moment, given by

$$\begin{aligned} \tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w}) &= \langle \mathbf{w}, \mathbf{v} \rangle \, \bar{\tau}^{n-1}(A, \mathbf{v}) + \frac{\mathbf{v}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})}{n} \times \\ &\quad \left[(n-1) \tilde{\tau}^{n-2;1}(A, \mathbf{v}; \mathbf{w}) \, \mathbf{v} - \int_{\partial A} \mathbf{n} \frac{\langle \mathbf{v}, \mathbf{u} \rangle^{n-1}}{\langle \mathbf{w}, \mathbf{u} \rangle} \, ds \right], \end{aligned} \tag{4.36}$$

where $\tilde{\tau}^{0;1}(A, \mathbf{v}; \mathbf{w}) = \tau^{0;1}(A, \mathbf{v}; \mathbf{w}) = \mathbf{T}^{0;1}(A, \mathbf{w})$ and $\tau^{-1;1}(A, \mathbf{v}; \mathbf{w}) = 0$, and we have used the fact that $\tilde{\tau}^{n-1;0}(A, \mathbf{v}; \mathbf{w}) = \bar{\tau}^{n-1}(A)$, the $(n-1)$th order axial moment about $\mathbf{v}$ described in section 3.2.3. Equation (4.36) is useful as a component of more general moments about multiple axes.

It is straightforward to generalize $\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w})$ to allow for Phong-like moments with respect to multiple axes. Given axis vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, moments about two axes and three axes are defined respectively by

$$\tilde{\tau}^{n,1;1}(A, \mathbf{v}_1, \mathbf{v}_2; \mathbf{w}) \equiv \int_A \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle^n \langle \mathbf{v}_2, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} d\sigma(\mathbf{u}), \tag{4.37}$$

$$\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3; \mathbf{w}) \equiv \int_A \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle^n \langle \mathbf{v}_2, \mathbf{u} \rangle \langle \mathbf{v}_3, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} d\sigma(\mathbf{u}). \tag{4.38}$$

Similarly, by specializing equation (4.18) with copies of $\mathbf{v}_1$ and representing the resulting terms using the notation $\tilde{\tau}$, we may derive the following recurrence relations for $\tilde{\tau}^{n,1;1}$ and $\tilde{\tau}^{n,1,1;1}$, expressed as

$$\begin{aligned} \tilde{\tau}^{n,1;1}(A, \mathbf{v}_1, \mathbf{v}_2; \mathbf{w}) &= \langle \mathbf{w}, \mathbf{v}_2 \rangle \, \bar{\tau}^{n}(A, \mathbf{v}_1) + \frac{\mathbf{v}_2{}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})}{n+1} \times \\ &\quad \left[n \, \tilde{\tau}^{n-1;1}(A, \mathbf{v}_1; \mathbf{w}) \, \mathbf{v}_1 - \int_{\partial A} \mathbf{n} \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle^n}{\langle \mathbf{w}, \mathbf{u} \rangle} \, ds \right], \end{aligned} \tag{4.39}$$

$$\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3; \mathbf{w}) = \langle \mathbf{w}, \mathbf{v}_3 \rangle \, \bar{\tau}^{n,1}(A, \mathbf{v}_1, \mathbf{v}_2) + \frac{\mathbf{v}_3{}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})}{n+2} \times$$

$$\Big[n\tilde{\tau}^{n-1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2;\mathbf{w})\,\mathbf{v}_1+\tilde{\tau}^{n;1}(A,\mathbf{v}_1;\mathbf{w})\,\mathbf{v}_2-$$

$$\int_{\partial A}\mathbf{n}\frac{\langle\mathbf{v}_1,\mathbf{u}\rangle^n\,\langle\mathbf{v}_2,\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds\Big], \tag{4.40}$$

where we replace $\tilde{\tau}^{n,1;0}(A,\mathbf{v}_1,\mathbf{v}_2;\mathbf{w})$ by the double-axis moment $\bar{\bar{\tau}}^{n,1}(A,\mathbf{v}_1,\mathbf{v}_2)$, which can be expressed in terms of axial moments using equation (3.38) shown in section 3.2.3, yielding

$$\begin{aligned}(n+2)\tilde{\tau}^{n,1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3;\mathbf{w}) \;=\; & n\,\langle\mathbf{w},\mathbf{v}_3\rangle\,\langle\mathbf{v}_1,\mathbf{v}_2\rangle\,\bar{\tau}^{n-1}(A,\mathbf{v}_1)-\\ & \langle\mathbf{w},\mathbf{v}_3\rangle\int_{\partial A}\langle\mathbf{v}_1,\mathbf{u}\rangle^n\,\langle\mathbf{v}_2,\mathbf{n}\rangle\,ds+\mathbf{v}_3{}^{\mathrm{T}}(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\times\\ & \Big[n\tilde{\tau}^{n-1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2;\mathbf{w})\,\mathbf{v}_1+\tilde{\tau}^{n;1}(A,\mathbf{v}_1;\mathbf{w})\,\mathbf{v}_2-\\ & \int_{\partial A}\mathbf{n}\frac{\langle\mathbf{v}_1,\mathbf{u}\rangle^n\,\langle\mathbf{v}_2,\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds\Big]\end{aligned} \tag{4.41}$$

4.4.2 Exact Evaluation

Simplified recurrence formulas (4.36), (4.39) and (4.41) reduces the surface integrals $\tilde{\tau}^{n;1}$, $\tilde{\tau}^{n,1;1}$ and $\tilde{\tau}^{n,1,1;1}$ shown in (4.35), (4.37) and (4.38) to boundary integrals of rational polynomials expressed as *steerable lobes*, sums of axial moments and in some case $\tau^{0;1}(A;\mathbf{w})(=\mathbf{T}^{0;1})$. This section demonstrated that each of these components can be integrated exactly over a polygonal domain. In section 4.3.2, we have shown that the base moment $\tau^{0;1}(A;\mathbf{w})$ can be evaluated exactly using one single special function known as the Clausen integral. We shall concentrate on the other two components next.

As for the axial moment $\bar{\tau}$, equation (3.39) reduces it to a sum of integrals, that is,

$$n\bar{\tau}^{n-1}(A,\mathbf{v}) \;=\; \bar{\tau}^{p-1}-\int_{\partial A}\Big[\langle\mathbf{v},\mathbf{u}\rangle^{n-2}+\langle\mathbf{v},\mathbf{u}\rangle^{n-4}+\cdots+\langle\mathbf{v},\mathbf{u}\rangle^{p}\Big]\,\langle\mathbf{v},\mathbf{n}\rangle\,ds, \tag{4.42}$$

where $p=0$ when n is even, and $p=1$ when n is odd, and $\bar{\tau}^{-1}(A)=0$ and $\bar{\tau}^{0}(A)=\sigma(A)$, which has been covered in section 4.3.2. With the parameterization described in section 3.4.1, the analytical evaluation of each term in (4.42) over each edge is based on

a function defined by

$$F(n,x,y) \equiv \int_x^y \cos^n\theta\, d\theta, \tag{4.43}$$

which satisfies the following recurrence relation

$$F(n,x,y) = \frac{1}{n}\left[\cos^{n-1} y \sin y - \cos^{n-1} x \sin x + (n-1)F(n-2,x,y)\right], \tag{4.44}$$

with

$$\begin{aligned} F(1,x,y) &= \sin y - \sin x, \\ F(0,x,y) &= y - x. \end{aligned}$$

By means of the recurrence given in (4.44), the complete integral in (4.42) can be evaluated exactly within $O(nk)$ time for a k-sided polygon by computing the functions $F(n,x,y)$ incrementally.

Finally, we are required to compute two types of line integrals resulting from applying formulas (4.36), (4.39) and (4.41). They are

$$\tilde{B}^n(\mathbf{v}_1) = \int_\zeta \frac{\langle \mathbf{v}_1, \mathbf{u}\rangle^n}{\langle \mathbf{w}, \mathbf{u}\rangle}\, ds, \tag{4.45}$$

$$\tilde{B}^{n,1}(\mathbf{v}_1, \mathbf{v}_2) = \int_\zeta \frac{\langle \mathbf{v}_1, \mathbf{u}\rangle^n \langle \mathbf{v}_2, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle}\, ds, \tag{4.46}$$

which are actually special cases of $B^{m,l}$ defined in Page 123, where the integrands are rational polynomials that are expressed as steerable lobes, with the cosine angle with respect to one axis raised to some power. Let us define

$$F_1(n,a,b,x,y) \equiv \int_x^y \frac{(a\cos\theta + b\sin\theta)^n}{\cos\theta}\, d\theta, \tag{4.47}$$

$$F_2(n,a_1,b_1,a_2,b_2,x,y) \equiv \int_x^y \frac{(a_1\cos\theta + b_1\sin\theta)^n (a_2\cos\theta + b_2\sin\theta)}{\cos\theta}\, d\theta. \tag{4.48}$$

It then follows from the binomial theorem that F_1 and F_2 can be evaluated in terms of

the function $G(r,s,x,y)$ defined in (4.30), that is

$$F_1(n,a,b,x,y) = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k G(k,n-k-1,x,y), \tag{4.49}$$

$$F_2(n,a_1,b_1,a_2,b_2,x,y) = \sum_{k=0}^{n} \binom{n}{k} a_1^{n-k} b_1^k \left[a_2 G(k,n-k,x,y) + b_2 G(k+1,n-k-1,x,y)\right]. \tag{4.50}$$

Accordingly, we may evaluate $\tilde{B}^n$ and $\tilde{B}^{n,1}$ in equations (4.45) and (4.46) exactly as follows, where we use $c_i, \phi_i, \alpha_i, \beta_i$ as described on Page 125:

$$\begin{aligned}
\tilde{B}^n(\mathbf{v}_1) &= \frac{c_1^n}{c} \int_0^{\Theta} \frac{\cos^n(\theta-\phi_1)}{\cos(\theta-\phi)}\, d\theta \\
&= \frac{c_1^n}{c} \int_{-\phi}^{\Theta-\phi} \frac{\cos^n(\theta-(\phi_1-\phi))}{\cos\theta}\, d\theta \\
&= \frac{c_1^n}{c} \int_{-\phi}^{\Theta-\phi} \frac{(\alpha_1\cos\theta+\beta_1\sin\theta)^n}{\cos\theta}\, d\theta \\
&= \frac{c_1^n}{c} F_1(n,\alpha_1,\beta_1,-\phi,\Theta-\phi),
\end{aligned} \tag{4.51}$$

and

$$\begin{aligned}
\tilde{B}^{n,1}(\mathbf{v}_1,\mathbf{v}_2) &= \frac{c_1^n c_2}{c} \int_0^{\Theta} \frac{\cos(\theta-\phi_1)^n \cos(\theta-\phi_2)}{\cos(\theta-\phi)}\, d\theta \\
&= \frac{c_1^n c_2}{c} \int_{-\phi}^{\Theta-\phi} \frac{\cos(\theta-(\phi_1-\phi))^n \cos(\theta-(\phi_2-\phi))}{\cos\theta}\, d\theta \\
&= \frac{c_1^n c_2}{c} \int_{-\phi}^{\Theta-\phi} \frac{(\alpha_1\cos\theta+\beta_1\sin\theta)^n(\alpha_2\cos\theta+\beta_2\sin\theta)}{\cos\theta}\, d\theta \\
&= \frac{c_1^n c_2}{c} F_2(n,\alpha_1,\beta_1,\alpha_2,\beta_2,-\phi,\Theta-\phi).
\end{aligned} \tag{4.52}$$

As shown in section 4.3.2, we may evaluate $G(r,s,x,y)$ within $\lfloor (r+|s|)/2+1 \rfloor$ steps, thus the computation of $\tilde{B}^n$ or $\tilde{B}^{n,1}$ using equation (4.51) or equation (4.52) takes $O(n^2)$ time.

Consequently, let $T(n)$ denote the complexity for computing $\tilde{\tau}^{n;1}$ directly use the

recursive relation (4.36). Then, for a k-sided polygon A, we have

$$T(n) \;=\; T(n-2) + \mathrm{O}(nk) + \mathrm{O}(n^2 k),$$

which leads to a complexity of $\mathrm{O}(n^3 k)$. It can be similarly reasoned that the complexity of evaluating $\tilde{\tau}^{n,1;1}$ and $\tilde{\tau}^{n,1,1;1}$ using equations (4.39) and (4.41) is also cubic.

4.4.3 Efficient Evaluation of $\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w})$

The *cubic* complexity for computing $\tilde{\tau}^{n;1}$ may however be reduced to $\mathrm{O}(n^2 k)$ by reorganizing the terms obtained from the recursive relation (4.36) in an optimal way and using dynamic programming.

Let $d = \mathbf{v}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})\mathbf{v} = 1 - \langle \mathbf{w}, \mathbf{v} \rangle^2$, then we can rewrite equation (4.36) as

$$n\,\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w}) = d(n-1)\tilde{\tau}^{n-2;1} + n\,\langle \mathbf{w}, \mathbf{v} \rangle\, \bar{\tau}^{n-1}\mathbf{v}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}}) \int_{\partial A} \mathbf{n} \frac{\langle \mathbf{v}, \mathbf{u} \rangle^{n-1}}{\langle \mathbf{w}, \mathbf{u} \rangle}\, ds,$$

which leads to

$$n\,\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w}) \;=\; T_1^n + \langle \mathbf{w}, \mathbf{v} \rangle\, T_2^n(\mathbf{v}) - \mathbf{v}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}}) T_3^n(\mathbf{v}), \tag{4.53}$$

where T_1, T_2 and T_3 depend on the order n, which was indicated in the superscript. Defining $k = \lfloor n/2 \rfloor$, the three terms in equation (4.53) are given by

$$T_1^n \;=\; d^{k-q} \left(\frac{n-1}{n-2}\right) \left(\frac{n-3}{n-4}\right) \cdots \left(\frac{q+3}{q+2}\right) (q+1)\tau^{q;1}, \tag{4.54}$$

$$\begin{aligned} T_2^n(\mathbf{v}) \;=\; & n\bar{\tau}^{n-1}(A, \mathbf{v}) + d \left(\frac{n-1}{n-2}\right) (n-2)\bar{\tau}^{n-3}(A, \mathbf{v}) + \\ & \cdots + d^{k-q-1} \left(\frac{n-1}{n-2}\right) \cdots \left(\frac{q+3}{q+2}\right) (q+2)\bar{\tau}^{q+1}(A, \mathbf{v}), \end{aligned} \tag{4.55}$$

$$\begin{aligned} T_3^n(\mathbf{v}) \;=\; & \int_{\partial A} \frac{\mathbf{n}}{\langle \mathbf{w}, \mathbf{u} \rangle} \left[\langle \mathbf{v}, \mathbf{u} \rangle^{n-1} + d \left(\frac{n-1}{n-2}\right) \langle \mathbf{v}, \mathbf{u} \rangle^{n-3} + \right. \\ & \left. \cdots + d^{k-q-1} \left(\frac{n-1}{n-2}\right) \left(\frac{n-3}{n-4}\right) \cdots \left(\frac{q+3}{q+2}\right) \langle \mathbf{v}, \mathbf{u} \rangle^{q+1} \right] ds, \end{aligned} \tag{4.56}$$

where

$$q = \begin{cases} 0 & \text{Even(n)} \\ -1 & \text{Odd(n)} \end{cases} \tag{4.57}$$

It is more advantageous in practice to expand the recurrence relation (4.36) into a big boundary integral, as equation (4.53) allows us to save a great deal of redundant computation by making use of the incremental structures manifested by T_2^n and T_3^n, as described next.

Observing the common coefficients

$$1, \frac{n-1}{n-2}, \frac{(n-1)(n-3)}{(n-2)(n-4)}, \cdots$$

occurring in equations (4.54), (4.55) and (4.56), the first optimization is to cache the series values before evaluating T_1, T_2, T_3, as shown in the pseudo-code below.

CachedFraction(**integer** n)

integer $k \leftarrow$ **if even**(n) $n/2$ **else** $(n-1)/2$;
integer $q \leftarrow$ **if even**(n) 0 **else** -1;
real $\mathbf{a}[k-q]$;
$\mathbf{a}[0] \leftarrow 1$
for $i \leftarrow 1, 2, \ldots, k-q-1$ **do**
    $\mathbf{a}[i] \leftarrow \mathbf{a}[i-1] * (n-(2i-1))/(n-2i)$;
endfor
return a

Using equation (4.42), we may express T_2^n for a k-sided polygon A as

$$T_2^n(\mathbf{v}) = \left[1 + d\left(\frac{n-1}{n-2}\right) + \cdots + d^{k-q-1}\left(\frac{n-1}{n-2}\right)\cdots\left(\frac{q+3}{q+2}\right)\right]\bar{\tau}^{p-1}$$

$$
\begin{aligned}
&- \int_{\partial A} \langle \mathbf{v}, \mathbf{n} \rangle \left(\langle \mathbf{v}, \mathbf{u} \rangle^{n-2} + \left[1 + d \left(\frac{n-1}{n-2} \right) \right] \langle \mathbf{v}, \mathbf{u} \rangle^{n-4} + \cdots \right. \\
&\left. + \left[1 + d \left(\frac{n-1}{n-2} \right) + \cdots + d^{k-1} \left(\frac{n-1}{n-2} \right) \cdots \left(\frac{p+3}{p+2} \right) \right] \langle \mathbf{v}, \mathbf{u} \rangle^{p} \right) ds \\
= \; & \left[1 + d \left(\frac{n-1}{n-2} \right) + \cdots + d^{k-q-1} \left(\frac{n-1}{n-2} \right) \cdots \left(\frac{q+3}{q+2} \right) \right] \bar{\tau}^{p-1} - \\
& \sum_{i=1}^{k} \left[\langle \mathbf{v}, \mathbf{n} \rangle \, S_1(n-2, d, c_{\mathbf{v}}, -\phi_{\mathbf{v}}, \Theta - \phi_{\mathbf{v}}) \right], \qquad (4.58)
\end{aligned}
$$

where $\mathbf{n}$, Θ, c and ϕ all depend on the edge ζ . Here $S_1(n-2, d, c, x, y)$ is a sum of the integrals F of different exponents, given by

$$
\begin{aligned}
& c^{n-2} F(n-2, x, y) + \left[1 + d \left(\frac{n-1}{n-2} \right) \right] c^{n-4} F(n-4, x, y) + \cdots \\
& + \left[1 + d \left(\frac{n-1}{n-2} \right) + \cdots + d^{k-1} \left(\frac{n-1}{n-2} \right) \cdots \left(\frac{p+3}{p+2} \right) \right] c^{p} F(p, x, y), \qquad (4.59)
\end{aligned}
$$

where the integer p satisfies

$$
p = \begin{cases} 0 & \text{Even(n)} \\ 1 & \text{Odd(n)} \end{cases} \qquad (4.60)
$$

Equation (4.59) and thus the term T_2^n can be computed incrementally in linear time using the identity (4.44), the algorithm for computing T_2^n with equations (4.58) and (4.59) is presented in the pseudo-code *AxialMomentSum* shown below, where the parameter **a** is the cached fractional coefficients from *CachedFraction*(n).

AxialMomentSum(**Polygon** P, **vector** $\mathbf{w}, \mathbf{v}$, **integer** n, **real** $d, \mathbf{a}[\,]$)

real $S \leftarrow$ BoundaryIntegral$(P, \mathbf{v}, d, n, \mathbf{a})$;
if odd(n) **then**
  integer $q \leftarrow$ **if even**(n) **then** 0 **else** -1;
  integer $k \leftarrow$ **if even**(n) $n/2$ **else** $(n-1)/2$;
  real $b \leftarrow 1$;
  real $dp \leftarrow d$; *Power of d*
  for $i \leftarrow 1, 2, \ldots, k - q - 1$ **do**
    $b \leftarrow b + dp * \mathbf{a}[i]$;
    $dp \leftarrow dp * d$;
  endfor
  $S \leftarrow S +$ SolidAngle(P);
endif
return S

The procedure *AxialMomentSum* is built upon three procedures: *CosIntegralSum* (computing S_1), *EdgeIntegral* (computing S_1 over an edge) and *BoundaryIntegral* (summing S_1 over the boundary of P), which are presented in pseudo-code as follows, where the operator **unit** normalizes a given vector.

CosIntegralSum(**integer** n, **real** $d, c, x, y, \mathbf{a}[\,]$)

integer $p \leftarrow$ **if even**(n) **then** 0 **else** 1;

Store fractional coefficients of each cosine integral into the array **b**

integer $k \leftarrow$ **if even**(n) $n/2$ **else** $(n-1)/2$; *Number of terms*

$\mathbf{b}[k-1] \leftarrow 1$;

$dp \leftarrow d$; *Power of d*

for $i \leftarrow k-2, k-3, \ldots, 0$ **do**

$\mathbf{b}[i] \leftarrow \mathbf{b}[i+1] + dp * \mathbf{a}[k-1-i]$;

$dp \leftarrow dp * d$;

endfor

Accumulate the final sum

real $F \leftarrow$ **if even**(n) **then** $y - x$ **else** $\sin y - \sin x$; *Cosine integral*

real $S \leftarrow 0$;

real $cp \leftarrow$ **if even**(n) 1 **else** c; *Power of c*

integer $j \leftarrow 0$; *Pointer to b*

while $p \leq n-2$ **do**

$S \leftarrow \mathbf{b}[j] * cp * F$;

$F \leftarrow \left[\cos^{p+1} y \sin y - \cos^{p+1} x \sin x + (p+1)F\right]/(p+2)$;

$cp \leftarrow c * c$;

$p \leftarrow p + 2$;

$j \leftarrow j + 1$;

endwhile

return S

EdgeIntegral(**vector** $\mathbf{A}, \mathbf{B}, \mathbf{v}$, **real** $d, \mathbf{a}[\,]$, **integer** n)

vector $\mathbf{s} \leftarrow$ **unit**$[\,\mathbf{A}\,]$;

vector $\mathbf{t} \leftarrow$ **unit**$[\,(\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{T}})\mathbf{B}\,]$;

real $a \leftarrow \langle \mathbf{v}, \mathbf{s} \rangle$;

real $b \leftarrow \langle \mathbf{v}, \mathbf{t} \rangle$;

real $c \leftarrow \sqrt{a^2 + b^2}$;

real $\Theta \leftarrow \cos^{-1}(\langle$**unit**$[\,\mathbf{A}\,]$, **unit**$[\,\mathbf{B}\,]\rangle)$;

real $\phi \leftarrow$ **sign**$[\,b\,] * \cos^{-1}(a/c)$;

vector $\mathbf{n} \leftarrow$ **unit**$[\,\mathbf{A} \times \mathbf{B}\,]$;

return $\langle \mathbf{v}, \mathbf{n} \rangle *$ CosIntegralSum$(n, d, c, -\phi, \Theta - \phi, \mathbf{a})$;

BoundaryIntegral(**Polygon** P, **vector** $\mathbf{v}$, **real** $d, \mathbf{a}[\,]$, **integer** n)

real $S \leftarrow 0$;

each edge **AB in** P **do**

$S \leftarrow S +$ EdgeIntegral$(\mathbf{A}, \mathbf{B}, \mathbf{v}, d, n, \mathbf{a})$;

endfor

return S

The last term given in equation (4.56) requires us to compute a weighted sum of integrals $\tilde{B}^k$, with k ranging from $q+1$ to $n-1$. It follows from equation (4.51) that this is equivalent to ealuating sums of integrals F_1 with increasing orders. That is,

$$\begin{aligned} T_3^n(\mathbf{v}) &= \sum_{i=1}^{k} \mathbf{n} \left[\tilde{B}^{n-1}(\mathbf{v}) + d\left(\frac{n-1}{n-2}\right) \tilde{B}^{n-3}(\mathbf{v}) + \cdots + \right. \\ &\qquad \left. d^{k-q-1} \left(\frac{n-1}{n-2}\right)\left(\frac{n-3}{n-4}\right) \cdots \left(\frac{q+3}{q+2}\right) \tilde{B}^{q+1}(\mathbf{v}) \right] \\ &= \sum_{i=1}^{k} \frac{\mathbf{n}}{c} \left[c_{\mathbf{v}}^{n-1} F_1(n-1, \alpha, \beta, -\phi, \Theta - \phi) + d\left(\frac{n-1}{n-2}\right) c_{\mathbf{v}}^{n-3} F_1(n-3, \alpha, \beta, -\phi, \Theta - \phi) + \right. \\ &\qquad \left. \cdots + d^{k-q-1} \left(\frac{n-1}{n-2}\right)\left(\frac{n-3}{n-4}\right) \cdots \left(\frac{q+3}{q+2}\right) c_{\mathbf{v}}^{q+1} F_1(q+1, \alpha, \beta, -\phi, \Theta - \phi) \right] \end{aligned} \tag{4.61}$$

for a k-sided polygon, where

$$\begin{aligned} \alpha &= \cos(\phi_{\mathbf{v}} - \phi), \\ \beta &= \sin(\phi_{\mathbf{v}} - \phi). \end{aligned}$$

From equation (4.49) and the recurrence relations about G shown in (4.32), we observe that exact evaluation of these F_1 terms (of different exponents) in equation (4.61) will share a lot of lower-order values of $G(r, s, x, y)$. To avoid redundant computations, we may precompute and cache n^2 values of $G(r, s, x, y)$ for an edge, i.e. $x = -\phi$ and $y = \Theta - \phi$, where $0 \leq r \leq n-1$ and $-1 \leq s \leq n-2$. This common technique known as *dynamic programming* reduces the *cubic* complexity for the exact evaluation of T_3 down to *quadratic*. The creation of the table for G is shown in the pseudo-code below, where optimizations can be done by computing powers of sine and cosine with repeated multiplication. Note that the first column of the resulting table ($\mathbf{G}[r][0]$) represent the values of $G(r, -1, x, y)$ for $r \in [0, n-1]$.

CachedSinCosIntegral(**integer** n, **real** x, y)

real $\mathbf{G}[n][n]$; $n \times n$ *table*
$\mathbf{G}[0][0] \leftarrow \ln\left(\frac{\tan(\pi/4+y/2)}{\tan(\pi/4+x/2)}\right)$; $G(0,-1,x,y)$
if $n > 1$ **then**
Base cases
$\mathbf{G}[0][1] \leftarrow y - x$; $G(0,0,x,y)$
$\mathbf{G}[1][0] \leftarrow \ln\left(\frac{\cos x}{\cos y}\right)$; $G(1,-1,x,y)$
$\mathbf{G}[1][1] \leftarrow \cos x - \cos y$; $G(1,0,x,y)$
Fill the first two columns
for $r \leftarrow 2, 3, \ldots, n-1$ **do**
$\mathbf{G}[r][0] \leftarrow \left[\sin^{r-1} x - \sin^{r-1} y + (r-1)\mathbf{G}[r-2][0]\right]/(r-1)$;
$\mathbf{G}[r][1] \leftarrow \left[\sin^{r-1} x \cos x - \sin^{r-1} y \cos y + (r-1)\mathbf{G}[r-2][1]\right]/r$;
endfor
Fill the remaining terms along each row
for $r \leftarrow 0, 1, \ldots, n-1; s \leftarrow 2, 3, \ldots, n-1$ **do**
$\mathbf{G}[r][s] \leftarrow \left[\sin^{r+1} y \cos^{s-1} y - \sin^{r+1} x \cos^{s-1} x + (s-1)\mathbf{G}[r][s-2]\right]/(r+s)$;
endfor
endif
return G

With this table $\mathbf{G}$ computed for each edge, the term T_3^n can be computed according to equation (4.61), which is embodied in the following procedures: *Fone* that evaluates the function $F_1(n, a, b, x, y)$ using equation (4.49); *FoneSum* that computes the weighted sums of $F_1(i, a, b, x, y)$ in equation (4.61); *EdgeFoneSum* that evaluates *FoneSum* over an edge; *RationalBndIntegralSum* that finally computes the term T_3^n.

Fone(**integer** n, **real** $a, b, x, y, \mathbf{G}[\,][\,]$)

real $S \leftarrow 0$;
for $k \leftarrow 0, 1, \ldots, n$ **do**
    $S \leftarrow S + \text{BinomialCoeff}(n, k) * a^{n-k} * b^k * \mathbf{G}[k][n-k]$;
endfor
return S

FoneSum(**integer** n, **real** $c, d, a, b, x, y, \mathbf{G}[\,][\,]$, **double** $\mathbf{a}[\,]$)

integer $q \leftarrow$ **if even**(n) 0 **else** -1;
integer $k \leftarrow$ **if even**(n) $n/2$ **else** $(n-1)/2$;
real $cp \leftarrow$ **if even**(n) c **else** 1; *Power of c*
real $dp \leftarrow 1$; *Power of d*
real $S \leftarrow 0$;
integer $j = 0$; *Pointer to the cache* **a**
integer $i \leftarrow$ **if even**(n) 1 **else** 0;
while $i \leq n - 1$ **do**
    $S \leftarrow S + \mathbf{a}[k - q - 1 - j] * \text{Fone}(i, a, b, x, y, \mathbf{G})/dp$;
    if $i < n - 1$ **then** $dp \leftarrow dp * p$;
    $j \leftarrow j + 1$;
endfor
return $S * dp$

EdgeFoneSum(**vector** $\mathbf{A}, \mathbf{B}, \mathbf{w}, \mathbf{v}$, **real** $d, \mathbf{a}[\,]$, **integer** n)

vector $\mathbf{s} \leftarrow$ **unit**$[\,\mathbf{A}\,]$;
vector $\mathbf{t} \leftarrow$ **unit**$[\,(\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{T}})\mathbf{B}\,]$;
real $a \leftarrow \langle \mathbf{w}, \mathbf{s} \rangle$;
real $b \leftarrow \langle \mathbf{w}, \mathbf{t} \rangle$;
real $c \leftarrow \sqrt{a^2 + b^2}$;
real $\phi \leftarrow$ **sign**$[\,b\,] * \cos^{-1}(a/c)$;
real $a' \leftarrow \langle \mathbf{v}, \mathbf{s} \rangle$;
real $b' \leftarrow \langle \mathbf{v}, \mathbf{t} \rangle$;
real $c' \leftarrow \sqrt{a'^2 + b'^2}$;
real $\phi' \leftarrow$ **sign**$[\,b'\,] * \cos^{-1}(a'/c')$;
real $\alpha \leftarrow \cos(\phi' - \phi)$;
real $\beta \leftarrow \sin(\phi' - \phi)$;
real $\Theta \leftarrow \cos^{-1}(\langle$**unit**$[\,\mathbf{A}\,]$, **unit**$[\,\mathbf{B}\,]\rangle)$;
vector $\mathbf{n} \leftarrow$ **unit**$[\,\mathbf{A} \times \mathbf{B}\,]$;
real $** \mathbf{G} \leftarrow$ CachedSinCosIntegral$(n, -\phi, \Theta - \phi)$; *Cache $G(r, s, x, y)$ for each edge*
return $\mathbf{n} *$ FoneSum$(n, c', d, \alpha, \beta, -\phi, \Theta - \phi, \mathbf{G}, \mathbf{a})/c$

RationalBndIntegralSum(**Polygon** P, **vector** $\mathbf{w}, \mathbf{v}$, **integer** n, **real** $d, \mathbf{a}[\,]$)

vector $\mathbf{S} \leftarrow 0$;
each edge **AB in** P **do**
    $\mathbf{S} \leftarrow \mathbf{S} +$ EdgeFoneSum$(\mathbf{A}, \mathbf{B}, \mathbf{w}, \mathbf{v}, d, n, \mathbf{a})$;
endfor
return S

Consequently, $\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w})$ may be evaluated analytically using *AxialMomentSum* and *RationalBndIntegralSum* as well as *BaseMoment* shown on Page 122, taking $\mathrm{O}(n^2 k)$ time for a k-sided polygon A. Further savings can be performed by sharing the com-

putations about edge parameterization appearing in *BaseMoment*, *EdgeIntegral* and *EdgeFoneSum*.

OneAxisPhongMoment(**Polygon** P, **vector** $\mathbf{w}, \mathbf{v}$, **integer** n)

if $n = 0$ **then**
  return BaseMoment$(P, \mathbf{w})$;
else
  real $*\mathbf{a} \leftarrow$ CachedFraction(n);
  integer $k \leftarrow$ **if even**(n) $n/2$ **else** $(n-1)/2$;
  integer $q \leftarrow$ **if even**(n) 0 **else** -1;
  real $d \leftarrow 1 - \langle \mathbf{w}, \mathbf{v} \rangle^2$;
  real $s_1 \leftarrow$ **if** $q = -1$ 0 **else** $d^{k-q} * (q+1) * \mathbf{a}[k-q-1] *$ BaseMoment$(P, \mathbf{w})/n$;
  real $s_2 \leftarrow$ AxialMomentSum$(P, \mathbf{w}, \mathbf{v}, n, d, \mathbf{a})$;
  vector $\mathbf{s} \leftarrow$ RationalBndIntegralSum$(P, \mathbf{w}, \mathbf{v}, n, d, \mathbf{a})$;
  vector $\mathbf{t} \leftarrow (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})\mathbf{s}$;
  real $s_3 \leftarrow \langle \mathbf{v}, \mathbf{t} \rangle$;
  return $s_1 + (\langle \mathbf{w}, \mathbf{v} \rangle * s_2 - s_3)/n$;
endif

4.4.4 Efficient Evaluation of $\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3; \mathbf{w})$

It follows from equations (4.39) and (4.40) that we can evaluate $\tilde{\tau}^{n,1;1}$ and $\tilde{\tau}^{n,1,1;1}$ using routines described in the previous section for $\tilde{\tau}^{n;1}$, as well as procedures to compute $\bar{\tau}^n$ or $\bar{\bar{\tau}}^{n,1}$ and an extra boundary integral ($\tilde{B}^n$ for $\tilde{\tau}^{n,1;1}$ and $\tilde{B}^{n,1}$ for $\tilde{\tau}^{n,1,1;1}$). The nth order axial moment and double-axis moment can be evaluated within $O(n)$ time for an edge using equation (4.42) and equation (3.38) [8, pp.94–97]. Consequently, the complexity for computing these two moments analytically is still quadratic. However, observing that the extra boundary integral terms aside from $\tilde{\tau}^{n;1}$ in equations (4.39) and (4.41) are closely related to T_2 and T_3 from the expanded formula of $\tilde{\tau}^{n;1}$ given in (4.53), we may combine these remaining terms into T_2 and T_3 in such a way that

we are able to take advantage of the recurrence relation (4.44) for $F(n,x,y)$ and share the cached table for $G(r,s,x,y)$ while evaluating them. In this section we shall take $\tilde{\tau}^{n,1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3;\mathbf{w})$ as an example and present an efficient approach for its evaluation, which embodies such optimizations and dramatically reduces the constant associated with the quadratic complexity. The moment $\tilde{\tau}^{n,1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3;\mathbf{w})$ about three axes $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ is very important in simulating some scattering effects to be discussed in the next chapter.

Expanding $\tilde{\tau}^{n-1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2;\mathbf{w})$ and $\tilde{\tau}^{n;1}(A,\mathbf{v}_1;\mathbf{w})$ in (4.41) with equation (4.39) and equation (4.36), we obtain

$$\begin{aligned}
&(n+2)\tilde{\tau}^{n,1,1;1}(A,\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3;\mathbf{w}) \\
= \;& r_1\tilde{\tau}^{n-2;1}(A,\mathbf{v}_1;\mathbf{w}) + r_2\bar{\tau}^{n-1}(A,\mathbf{v}_1) + r_3\int_{\partial A}\langle\mathbf{v}_1,\mathbf{u}\rangle^n\langle\mathbf{v}_2,\mathbf{n}\rangle\,ds \\
&+ \int_{\partial A}\langle\mathbf{r}_4,\mathbf{n}\rangle\frac{\langle\mathbf{v}_1,\mathbf{u}\rangle^{n-1}}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds + \int_{\partial A}\langle\mathbf{r}_5,\mathbf{n}\rangle\frac{\langle\mathbf{v}_1,\mathbf{u}\rangle^n\langle\mathbf{v}_2,\mathbf{u}\rangle}{\langle\mathbf{w},\mathbf{u}\rangle}\,ds, \qquad (4.62)
\end{aligned}$$

where the coefficients are given by

$$\begin{aligned}
r_1 &= -(n-1)\mathbf{v}_3{}^{\mathrm{T}}\left[(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\left(\mathbf{v}_1\mathbf{v}_2{}^{\mathrm{T}}+\frac{\mathbf{v}_2\mathbf{v}_1{}^{\mathrm{T}}}{n}\right)(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\right]\mathbf{v}_1 \\
r_2 &= n\langle\mathbf{w},\mathbf{v}_3\rangle\langle\mathbf{v}_1,\mathbf{v}_2\rangle - \mathbf{v}_3{}^{\mathrm{T}}(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\left[n\langle\mathbf{w},\mathbf{v}_2\rangle\mathbf{v}_1+\langle\mathbf{w},\mathbf{v}_1\rangle\mathbf{v}_2\right] \\
r_3 &= -\langle\mathbf{w},\mathbf{v}_3\rangle \\
\mathbf{r}_4 &= \mathbf{v}_3{}^{\mathrm{T}}\left[(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\left(\mathbf{v}_1\mathbf{v}_2{}^{\mathrm{T}}+\frac{\mathbf{v}_2\mathbf{v}_1{}^{\mathrm{T}}}{n}\right)(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\right] \\
\mathbf{r}_5 &= -\mathbf{v}_3{}^{\mathrm{T}}(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}}).
\end{aligned}$$

Using equation (4.53), we replace $\tilde{\tau}^{n-2;1}(A,\mathbf{v}_1;\mathbf{w})$ in equation (4.62) by

$$\frac{1}{n-2}\left[T_1^{n-2}+\langle\mathbf{w},\mathbf{v}_1\rangle T_2^{n-2}(\mathbf{v}_1)-\mathbf{v}_1{}^{\mathrm{T}}(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})T_3^{n-2}(\mathbf{v}_1)\right].$$

Letting

$$\begin{aligned} r_6 &= r_1/(n-2) \\ r_7 &= r_6 \langle \mathbf{w}, \mathbf{v}_1 \rangle \\ \mathbf{r}_8 &= r_6 \mathbf{v}_1{}^{\mathrm{T}}(\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}}), \end{aligned}$$

Then, by regrouping terms of similar structures, we arrive at a formula for $\tilde{\tau}^{n,1,1;1}$ that is similar to equation (4.53), expressed as

$$\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3; \mathbf{w}) = R_1^n + R_2^n + R_3^n, \tag{4.63}$$

where R_1^n, R_2^n and R_3^n are terms depending on the order n. Let $k = \lfloor n/2 \rfloor - 1$, d and q are defined as before. It follows from equation (4.62) that they are given by

$$\begin{aligned} R_1^n &= r_6 T_1^{n-2} \\ &= r_6 d^{k-q} \left(\frac{n-3}{n-4}\right) \cdots \left(\frac{q+3}{q+2}\right) (q+1) \tau^{q;1}, \qquad (4.64) \\ R_2^n &= r_3 \int_{\partial A} \langle \mathbf{v}_1, \mathbf{u} \rangle^n \langle \mathbf{v}_2, \mathbf{n} \rangle \, ds + r_2 \bar{\tau}^{n-1}(A, \mathbf{v}_1) + r_7 T_2^{n-2}(\mathbf{v}_1), \qquad (4.65) \\ R_3^n &= \int_{\partial A} \langle \mathbf{r}_5, \mathbf{n} \rangle \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle^n \langle \mathbf{v}_2, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \, ds + \int_{\partial A} \langle \mathbf{r}_4, \mathbf{n} \rangle \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle^{n-1}}{\langle \mathbf{w}, \mathbf{u} \rangle} \, ds \\ &\qquad - \left\langle \mathbf{r}_8, T_3^{n-2}(\mathbf{v}_1) \right\rangle. \qquad (4.66) \end{aligned}$$

Given A as a k-sided polygon, we may evaluate R_2^n with equations (4.42) and (4.58) as follows:

$$R_2^n = r_9 \bar{\tau}^{p-1} + \sum_{i=1}^{k} S_2(n, d, c_1, -\phi_1, \Theta - \phi_1, r_3, r_2, r_7), \tag{4.67}$$

where the coefficient $r_9 = r_2/n$ for $n \leq 2$ and

$$r_9 = \frac{r_2}{n} + r_7 \left[1 + d\left(\frac{n-3}{n-4}\right) + \cdots + d^{k-q-1}\left(\frac{n-3}{n-4}\right) \cdots \left(\frac{q+3}{q+2}\right)\right].$$

The summation S_2 connects all the terms involving the function F, weighted properly by r_3, r_2, r_7, which is given by

$$\begin{aligned} r_3 \langle \mathbf{v}_2, \mathbf{n} \rangle c_1^n F(n, -\phi_1, \Theta - \phi_1) \quad - \quad & \frac{r_2}{n} \langle \mathbf{v}_1, \mathbf{n} \rangle \Big[c_1^{n-2} F(n-2, -\phi_1, \Theta - \phi_1) + \cdots \\ & + c_1^p F(p, -\phi_1, \Theta - \phi_1) \Big] \\ - \quad & r_7 \langle \mathbf{v}_1, \mathbf{n} \rangle S_1(n-4, d, c_1, -\phi_1, \Theta - \phi_1), \end{aligned} \tag{4.68}$$

which can be efficiently evaluated by modifying the routine *CosIntegralSum* for $S_1(n-4, d, c_1, -\phi_1, \Theta - \phi_1)$ to account for extra terms in equation (4.68). Moreover, it follows from equation (4.61) that

$$\begin{aligned} R_3^n \quad = \quad & \sum_{i=1}^{k} \Bigg(\langle \mathbf{r}_5, \mathbf{n} \rangle \tilde{B}^{n,1}(\mathbf{v}_1, \mathbf{v}_2) + \langle \mathbf{r}_4, \mathbf{n} \rangle \tilde{B}^{n-1}(\mathbf{v}_1) - \\ & \langle \mathbf{r}_8, \mathbf{n} \rangle \left[\tilde{B}^{n-3}(\mathbf{v}_1) + d \left(\frac{n-3}{n-4} \right) \tilde{B}^{n-5}(\mathbf{v}_1) + \cdots + \right. \\ & \left. d^{k-q-1} \left(\frac{n-3}{n-4} \right) \left(\frac{n-5}{n-6} \right) \cdots \left(\frac{q+3}{q+2} \right) \tilde{B}^{q+1}(\mathbf{v}_1) \right] \Bigg). \end{aligned} \tag{4.69}$$

Similarly, we may evaluate equation (4.69) efficiently by allowing the routine *FoneSum* that computes the sum of $\tilde{B}^k$s with $q+1 \leq k \leq n-3$ to also return the values of $\tilde{B}^{n,1}(\mathbf{v}_1, \mathbf{v}_2)$ and $\tilde{B}^{n-1}(\mathbf{v}_1)$, in order to share the cached table for $G(r, s, x, y)$. For this purpose, it follows from equation (4.50) that we are required to cache values of $G(r, s, x, y)$ for $0 \leq r \leq n+1$ and $-1 \leq s \leq n-1$. Notice that we should compute a larger table of G than those needed for T_3^{n-2} to be able to exactly evaluate the two extra terms.

4.4.5 Approximations

The bottleneck to exactly evaluating $\tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w})$ and $\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3; \mathbf{w})$ using equations (4.53) and (4.63) respectively is to compute the third term, namely, T_3^n or R_3^n, which integrates a summation of $\dfrac{\langle \mathbf{v}, \mathbf{u} \rangle^n}{\langle \mathbf{w}, \mathbf{u} \rangle}$ or $\dfrac{\langle \mathbf{v}_1, \mathbf{u} \rangle^n}{\langle \mathbf{w}, \mathbf{u} \rangle}$ of increasing orders over a boundary of a spherical polygon. Due to the cost of computing the $n \times n$ ($(n+2) \times (n+2)$ for $\tilde{\tau}^{n,1,1;1}$) table for $G(r, s, x, y)$, the time complexity for T_3^n or R_3^n is $\mathrm{O}(kn^2)$, where k is

the number of polygon edges.

Another means to speedup the computation is to settle for an approximation for the bottleneck part T_3 or R_3. Numerical quadrature is one option. This is particularly effective as T_3 (or R_3) has the highest computational cost, yet is typically very small in magnitute compared to T_1 and T_2 (or R_1, R_2). In terms of accuracy, this approach is preferable to approximating the original integral using two-dimensional quadrature, as one-dimensional quadrature rules are more robust and more amenable to higher-order methods. By computing the powers of $\langle \mathbf{v}, \mathbf{u} \rangle$ in equation (4.56) incrementally through repeated multiplication for each sampled $\mathbf{u}$, the complete integral (4.56) can be evaluated within $O(l * n)$ time for each edge, where l is the number of samples. Similarly, we may express equation (4.66) as a big boundary integral, that is,

$$R_3^n = \int_{\partial P} \frac{\langle \mathbf{n}, \mathbf{S}(\mathbf{u}) \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \, ds$$

where

$$\begin{aligned} \mathbf{S}(\mathbf{u}) &= \mathbf{r}_5 \langle \mathbf{v}_1, \mathbf{u} \rangle^n \langle \mathbf{v}_2, \mathbf{u} \rangle + \mathbf{r}_4 \langle \mathbf{v}_1, \mathbf{u} \rangle^{n-1} - \mathbf{r}_8 \Bigg[\langle \mathbf{v}_1, \mathbf{u} \rangle^{n-3} + \\ &\quad d \left(\frac{n-3}{n-4} \right) \langle \mathbf{v}_1, \mathbf{u} \rangle^{n-5} + \cdots + d^{k-q-1} \left(\frac{n-3}{n-4} \right) \left(\frac{n-5}{n-6} \right) \cdots \left(\frac{q+3}{q+2} \right) \langle \mathbf{v}_1, \mathbf{u} \rangle^{q+1} \Bigg] \end{aligned}$$

for $k = \lfloor n/2 \rfloor$, the same strategy can then be applied. We have used this numerical approximation for the applications demonstrated in chapter 5.

Looking for better approximation strategies for T_3 is still a future direction we shall pursue. We speculate that research on hypergeometric functions may lead to some efficient algorithms for our problem. To see how the hypergeometric representation arises, let us refomulate the integrand in equation (4.56). By moving the constant $\mathbf{n}$ outside the integral, we are left with the following summation:

$$\begin{aligned} S_n(\mathbf{v}, \mathbf{w}, \mathbf{u}) &\equiv \frac{1}{\langle \mathbf{w}, \mathbf{u} \rangle} \Bigg[\langle \mathbf{v}, \mathbf{u} \rangle^{n-1} + d \left(\frac{n-1}{n-2} \right) \langle \mathbf{v}, \mathbf{u} \rangle^{n-3} + \\ &\quad \cdots + d^{k-q-1} \left(\frac{n-1}{n-2} \right) \left(\frac{n-3}{n-4} \right) \cdots \left(\frac{q+3}{q+2} \right) \langle \mathbf{v}, \mathbf{u} \rangle^{q+1} \Bigg], \end{aligned}$$

where q is defined in (4.57) and $k = \lfloor n/2 \rfloor$. For odd and even n, we have respectively

$$S_{2k+1} = \frac{1}{\langle \mathbf{w}, \mathbf{u} \rangle} \left[\langle \mathbf{v}, \mathbf{u} \rangle^{2k} + d \left(\frac{2k}{2k-1} \right) \langle \mathbf{v}, \mathbf{u} \rangle^{2(k-1)} + \cdots + d^k \left(\frac{(2k)!!}{(2k-1)!!} \right) \langle \mathbf{v}, \mathbf{u} \rangle^0 \right], \tag{4.70}$$

$$S_{2k} = \frac{1}{\langle \mathbf{w}, \mathbf{u} \rangle} \left[\langle \mathbf{v}, \mathbf{u} \rangle^{2k-1} + d \left(\frac{2k-1}{2k-2} \right) \langle \mathbf{v}, \mathbf{u} \rangle^{2k-3} + \cdots + d^{k-1} \left(\frac{(2k-1)!!}{(2k-2)!!} \right) \langle \mathbf{v}, \mathbf{u} \rangle \right], \tag{4.71}$$

which may be expressed as hypergeometric functions, that is,

$$S_{2k+1} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle^{2k}}{\langle \mathbf{w}, \mathbf{u} \rangle} \, {}_2F_1 \left[\begin{matrix} -k, 1 \\ -k + \frac{1}{2} \end{matrix} ; \frac{d}{\langle \mathbf{v}, \mathbf{u} \rangle^2} \right], \tag{4.72}$$

$$S_{2k} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle^{2k-1}}{\langle \mathbf{w}, \mathbf{u} \rangle} \, {}_2F_1 \left[\begin{matrix} -k + \frac{1}{2}, 1 \\ -k + 1 \end{matrix} ; \frac{d}{\langle \mathbf{v}, \mathbf{u} \rangle^2} \right]. \tag{4.73}$$

Moreover, we may combine equations (4.72) and (4.73) into a single hypergeometric representation for S_n, expressed as

$$S_n = \frac{\langle \mathbf{v}, \mathbf{u} \rangle^{n-1}}{\langle \mathbf{w}, \mathbf{u} \rangle} \, {}_2F_1 \left[\begin{matrix} -\frac{n-1}{2}, 1 \\ -\frac{n}{2} + 1 \end{matrix} ; \frac{d}{\langle \mathbf{v}, \mathbf{u} \rangle^2} \right]. \tag{4.74}$$

Consequently, the evaluation of T_3 is equivalent to integrating a hypergeometric function:

$$\int_\zeta \frac{\langle \mathbf{v}, \mathbf{u} \rangle^{n-1}}{\langle \mathbf{w}, \mathbf{u} \rangle} \, {}_2F_1 \left[\begin{matrix} -\frac{n-1}{2}, 1 \\ -\frac{n}{2} + 1 \end{matrix} ; \frac{d}{\langle \mathbf{v}, \mathbf{u} \rangle^2} \right] ds,$$

which suggests that studies in the area of hypergeometric functions may shed some light on discovering efficient evaluation strategies for these costly terms in the rational moments.

Chapter 5

Applications of Generalized Irradiance Tensors

Deterministic rendering algorithms are often quite limited in the optical effects they simulate; for the most part they are limited to diffuse and pure specular effects. A common assumption is that of a uniform luminaire with constant radiant exitance in all directions and positions, for which a wide assortment of closed-form expressions exist for computing the radiative exchange [58, 18, 97] and some Phong-like scattering effects [9]. Unfortunately, these formulas rarely apply to non-uniform luminaires, especially *inhomogeneous* or *spatially-varying* luminaires. To a large extent, this limitation stems from the difficulty of computing the surface integrals associated with spatially-varying luminaires. Unlike uniform luminaires, they generally cannot be expressed as a polynomial integrated over regions of the sphere. In chapter 3 we have derived a closed-form solution for the irradiance due to a polygonal linearly-varying luminaire using tedious Taylor expansion, which becomes impossible to be extended to higher order polynomials. Although DiLaura et al. [39] have addressed the problem of computing the irradiance at a point from polynomially-varying luminaires, their results hardly lead to closed-form solutions. Moreover, as far as we know, no further efforts have been made towards simulating other lighting effects involving spatially-varying luminaires and non-diffuse surfaces.

Generalized irradiance tensors and their associated rational moments introduced in

the previous chapter are well suited to quantify emission and scattering features due to polynomially-varying luminaires, especially for polygonal environments with Phong-like reflection or transmission distributions, which makes them quite useful in the simulation of some rendering effects involving this type of emitters. The formulas and evaluation methods derived for these tensors and moments in chapter 4 provide us new direct methods for computing some illuminating and scattering effects involving polynomially-varying luminaires and a wide range of surfaces from diffuse to highly specular. Applications include the computation of irradiance due to polynomially-varying luminaires or directional luminaires with Phong-like distribution, view-dependent glossy reflections and transmissions from such emitters. Our results greatly extend the repertoire of non-Lambertian phenomena that can be handled deterministically. In this chapter we shall present these applications of generalized irradiance tensors in image synthesis, and verify our results with the Monte Carlo and Finite element methods.

5.1 Irradiance due to Polynomially-Varying Luminaires

The irradiance at a point due to a polynomially-varying luminaire L is defined as a weighted integral of the radiance function $f(\mathbf{u})$ caused by the emitter. That is,

$$\Phi(A) \;=\; \int_A f(\mathbf{u})\,\langle \mathbf{b},\mathbf{u}\rangle \; d\sigma(\mathbf{u}), \tag{5.1}$$

where A is the spherical projection of L at the origin, and $\mathbf{b}$ is the normal there. By default, we shall assume that the receiver point is located at the origin. As we discussed in section 4.1, the radiance field due to any polynomially-varying luminaire can be expressed as combinations of equation (4.3), the radiance distribution due to a monomially-varying luminaire. Consequently, the problem of computing the irradiance at the origin due to polynomially-varying luminaires is simplified to computing an integral of the form

$$\Phi_n(A) \;=\; \frac{h^n}{\pi}\int_A \frac{\langle \mathbf{e}_1,\mathbf{u}\rangle^p\,\langle \mathbf{e}_2,\mathbf{u}\rangle^q\,\langle \mathbf{e}_3,\mathbf{u}\rangle^r\,\langle \mathbf{b},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle^n}\, d\sigma(\mathbf{u}), \tag{5.2}$$

where $\mathbf{w}$ is the unit vector perpendicular to the luminaire and h is the distance from the origin to the plane, and $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is an orthonormal coordinate system at the origin. When $p + q + r = n$, equation (5.2) corresponds to the irradiance at the origin due to a luminaire whose radiant exitance varies according to a nth order monomial function $\phi(\mathbf{x}) = x_1^p x_2^q x_3^r$. Comparing equation (5.2) with the definition (4.15), we realize that the concept of rational angular moments is closely related to the irradiance due to polynomially-varying luminaires. Therefore, the expressions derived for rational moments and generalized irradiance tensors associated with them may be used to derive new expressions for the irradiance due to polynomially-varying luminaires, extending the result in chapter 3 to higher order polynomials.

In this section we first examine linearly-varying luminaires again and re-derive the closed-form solution found in chapter 3 using the machinery of generalized irradiance tensors. Then, by studying quadratically-varying luminaires, we relate the special function Λ to the irradiance due to an emitter with a special quadratically-varying radiant exitance. The solutions for linear and quadratic cases exhibit a general structure subsuming Lambert's formula, which motivates us to generalize them to higher-order polynomials with a recurrence formula that relates the solution to those for lower-order polynomials.

5.1.1 Linearly-Varying Luminaires

Generalized irradiance tensors provide us a more elegant machinery to re-derive the closed-form solution for the irradiance due to a linearly-varying luminaire identified in chapter 3. As we have seen, the original approach based on Taylor expansion and a formula derived for *triple-axis moments* is quite complicated and hard to follow.

As analyzed in section 3.1, the irradiance at the origin due to a planar diffuse emitter with linearly-varying radiant exitance determined by three non-collinear points can be formulated by equation (3.7). Corresponding to the definitions (4.15) and (4.4), we may express the irradiance integral in (3.7) with our tensor notation as follows:

$$\Phi_1(A) \;=\; \frac{1}{\pi}\tau^{2;1}(A, \mathbf{a}, \mathbf{b}; \mathbf{w}) \;=\; \frac{1}{\pi}\mathbf{T}_{ij}^{2;1}(A, \mathbf{w})\, \mathbf{a}_i \mathbf{b}_j, \tag{5.3}$$

Applying the recursive formula (4.13), we have

$$\mathbf{T}_{ij}^{2,1} = \mathbf{w}_j \mathbf{T}_i^1 + \frac{1}{2}\left(\delta_{jm} - \mathbf{w}_j \mathbf{w}_m\right)\left[\delta_{im}\mathbf{T}^{0,1} - \int_{\partial A} \frac{\mathbf{u}_i \mathbf{n}_m}{\langle \mathbf{w}, \mathbf{u} \rangle}\, ds\right],$$

which reduces to a boundary integral by using equations (3.32) and (4.14), that is

$$\begin{aligned}
\mathbf{T}_{ij}^{2,1} &= \frac{1}{2}\int_{\partial A}\left[-\mathbf{w}_j \mathbf{n}_i + \left(\delta_{jm} - \mathbf{w}_j \mathbf{w}_m\right)\left(\delta_{im}\langle \mathbf{w}, \mathbf{n} \rangle \frac{\ln \langle \mathbf{w}, \mathbf{u} \rangle}{1 - \langle \mathbf{w}, \mathbf{u} \rangle^2} - \frac{\mathbf{u}_i \mathbf{n}_m}{\langle \mathbf{w}, \mathbf{u} \rangle}\right)\right] ds \\
&= -\frac{1}{2}\int_{\partial A}\left[\delta_{ik}\mathbf{w}_j - \left(\delta_{jm} - \mathbf{w}_j \mathbf{w}_m\right)\left(\delta_{im}\mathbf{w}_k\, \eta - \frac{\delta_{km}\mathbf{u}_i}{\langle \mathbf{w}, \mathbf{u} \rangle}\right)\right] \mathbf{n}_k\, ds, \qquad (5.4)
\end{aligned}$$

where η is given by equation (3.82). Combining equations (5.3) and (5.4), we arrive at exactly the same general boundary integral for the irradiance at the origin, given by

$$\boxed{\Phi_1(A, \mathbf{w}, \mathbf{a}, \mathbf{b}) = -\frac{1}{2\pi}\int_{\partial A} \mathbf{M}_{ijk}^3 \mathbf{a}_i \mathbf{b}_j \mathbf{n}_k\, ds,} \qquad (5.5)$$

where the 3-tensor $\mathbf{M}^3$ is defined in (3.84). Here and in the remainder of this chapter we will indicate the order of polynomial by subscript of Φ and the order of tensor $\mathbf{M}$ by superscript. Using the same technique demonstrated in section 3.4, we obtain a closed-form solution given in (3.116) for a polygonal linearly-varying luminaire.

The pseudo-code for computing the irradiance at the origin $\mathbf{o}$ due to a linearly-varying luminaire P is shown in *LinearIrradiance*. Here, $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ are any three non-collinear points on P with radiant exitance w_1, w_2 and w_3, respectively. We assume that the polygon vertices of P are ordered *counterclockwise* viewed from the receiver point $\mathbf{o}$, which guarantees that $\mathbf{n}$ computed above is outward pointing and the irradiance computed is positive. For vertices that are oriented clockwise, the resulting irradiance will be negative. Preserving the sign is useful for integrating over polygons with holes. Note that the ordering of the points $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ are immaterial, as the $\mathbf{w}$ vector is always adjusted to point toward the polygon. $\overline{B}^{1,1}$ and $\overline{B^*}$ in *LinearIrradiance* are evaluated using equations (3.113) and (3.114), respectively. The special function $\Lambda(\alpha, \beta)$

LinearIrradiance(**Polygon** P, **vector** $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, **real** w_1, w_2, w_3)

vector $\mathbf{w} \leftarrow$ **unit**$[\,(\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_2 - \mathbf{p}_3)\,]$

if $\langle \mathbf{w}, \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 \rangle < 0$

  $\mathbf{w} \leftarrow -\mathbf{w}$

endif

vector $\mathbf{b} \leftarrow$ *surface normal at the origin* $\mathbf{o}$

vector $\mathbf{a} \leftarrow \langle \mathbf{p}_1, \mathbf{w} \rangle\, [w_1\ w_2\ w_3]\, [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]^{-1}$

real $s \leftarrow 0$

for each edge **AB** in P **do**

  vector $\mathbf{s} \leftarrow$ **unit**$[\,\mathbf{A}\,]$

  vector $\mathbf{t} \leftarrow$ **unit**$[\,(\mathbf{I} - \mathbf{s}\mathbf{s}^{\mathrm{T}})\mathbf{B}\,]$

  real $a_1 \leftarrow \langle \mathbf{w}, \mathbf{s} \rangle$; $\quad a_2 \leftarrow \langle \mathbf{a}, \mathbf{s} \rangle$

  real $b_1 \leftarrow \langle \mathbf{w}, \mathbf{t} \rangle$; $\quad b_2 \leftarrow \langle \mathbf{a}, \mathbf{t} \rangle$

  real $c_1 \leftarrow \sqrt{a_1^2 + b_1^2}$; $\quad c_2 \leftarrow \sqrt{a_2^2 + b_2^2}$

  real $\phi_1 \leftarrow$ **sign**$[\,b_1\,] * \cos^{-1}(a_1/c_1)$

  real $\phi_2 \leftarrow$ **sign**$[\,b_2\,] * \cos^{-1}(a_2/c_2)$

  real $\Theta \leftarrow \cos^{-1}(\langle$**unit**$[\,\mathbf{A}\,],$ **unit**$[\,\mathbf{B}\,]\rangle)$

  vector $\mathbf{n} \leftarrow$ **unit**$[\,\mathbf{A} \times \mathbf{B}\,]$

  vector $\mathbf{v} \leftarrow (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}})\,\mathbf{b}$

  real $s_0 \leftarrow \langle \mathbf{a}, \mathbf{n} \rangle * \langle \mathbf{b}, \mathbf{w} \rangle * \Theta$

  real $s_1 \leftarrow \overline{B}^{1,1}(c_1, c_2, \Theta, \phi_1, \phi_2) * \langle \mathbf{v}, \mathbf{n} \rangle$

  real $s_2 \leftarrow \overline{B^*}(c_1, \Theta, \phi_1) * \langle \mathbf{w}, \mathbf{n} \rangle * \langle \mathbf{v}, \mathbf{a} \rangle$

  $s \leftarrow s + s_0 + s_1 - s_2$

endfor

return $-s\,/\,2\pi$

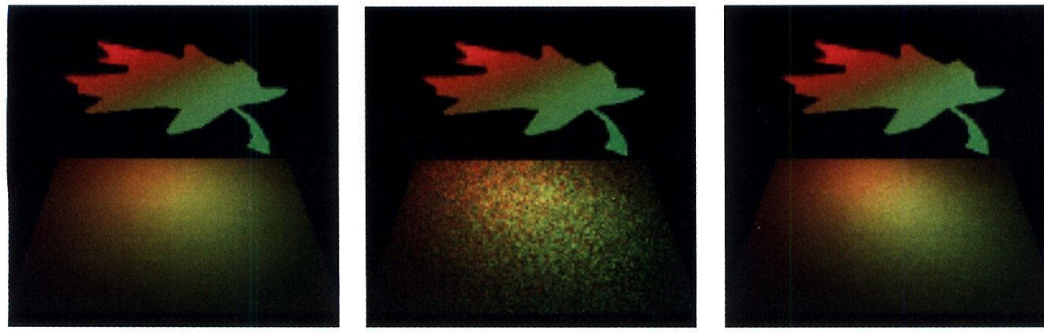
end

Analytic (1.7m) Monte Carlo (2.6m) Monte Carlo (42m)

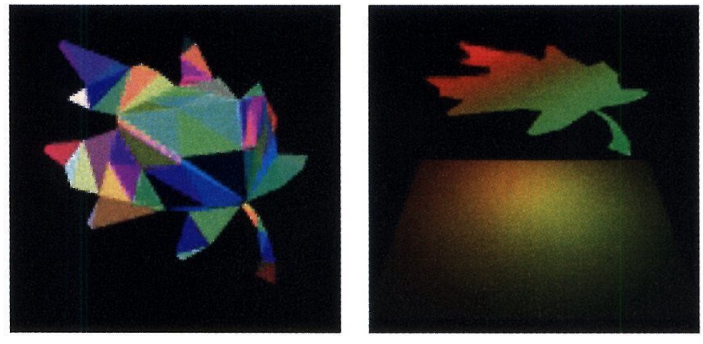
Figure 5.1: *Images of a polygonal environment generated using the analytic solution and Monte Carlo. The numbers beneath each image indicate the computation time. (Left) The analytic solution is applied by clipping the luminaire against all blockers with respect to each point on the receiving surface and computing the contributions from the remaining polygons. (Middle) Monte Carlo solution using four stratified samples per pixel, which requires a comparable amount of time. (Right) Monte Carlo solution using 100 stratified samples per pixel, which produces an image comparable to the analytic solution.*

involved in $\overline{B^*}$ can be computed directly from equation (3.126), where the evaluation of the Clausen integral has been shown in pseudo-code on Page 105.

Figures 5.1 and 5.2 show two simple scenes, each illuminated by an area light source with two linearly-varying superimposed colors. This effect is simulated by integrating two linearly-varying scalar values over the luminaire and weighting the corresponding colors by them. Both scenes were rendered using the new technique and by Monte Carlo for comparison. In implementing our analytical method, several optimizations were employed: 1) computing the vector $\mathbf{a}$ at each receiver point incrementally using equation (3.12) rather than inverting a matrix at each point, 2) computing an irradiance vector of the form given in equation (3.117) to save the computation cost for two color variations and 3) bilinear interpolation for Λ evaluation. Stratified sampling was used for the Monte Carlo solutions. All the images have a resolution of 300×300, and were generated on a SGI Indigo2 working at 175 Mhz $R10000$. The numbers beneath each picture indicate the computation time in minutes.



Analytic (3.2m) Monte Carlo (4.2m) Monte Carlo (80m)



Triangulation (68 patches) Semi-Analytic (12.5m)

Figure 5.2: *Non-convex polygonal luminaires with linearly-varying colors can be efficiently handled with our proposed algorithm. (Top Row) Comparison similar to that shown in Figure 5.1. (Bottom Row) Image generated by subdividing the luminaire into small regions that are treated as constant. Although each pixel of the resulting image is within 2% of the exact solution, this approach is slower than the analytic solution. Moreover, the accuracy depends on an appropriate level of subdivision.*

Figure 5.1 depicts a box blocker in front of a luminaire. Polygonal occlusions are handled by clipping the luminaire against all blockers and computing the contribution from each visible portion using our closed-form solution. Our noise-free result clearly matches the Monte Carlo image generated by using 100 stratified samples per pixel, yet is over 20 times faster. Moreover, our method results in much higher quality than the simple Monte Carlo solution, using nearly the same amount of time.

For the leaf-like polygonal luminaire in Figure 5.2, our proposed algorithm shows

a significant advantage over Monte Carlo in that it can handle non-convex polygons directly, and produce a noise-free image in far less time, as shown in the top row. The Monte Carlo images are rendered by stratified sampling of the bounding rectangle of the leaf-shaped polygon. On the bottom, we show results of rendering the same scene using a semi-analytical method for further comparison. This approach subdivides the luminaire into small pieces, and computes the contribution from each piece using Lambert's formula by assuming that they are uniform. Although each pixel of the resulting image by breaking the leaf into 68 triangles shown in the left is within 2% of the exact solution from our new method, this finite element approach is much slower than the analytic solution. Moreover, the accuracy depends on an appropriate level of subdivision, which is hard to choose.

These experiments provide numerical evidence of the correctness of our formulas (3.116) and (3.117), and also demonstrate its application to computing direct illumination. The approach is also applicable to indirect illumination, however, as it provides a means of computing form-factors involving non-constant surface elements [72].

5.1.2 Quadratically-Varying Luminaires

As a first step to extend equation (5.5) to higher order polynomials, we study luminaires with *quadratically-varying* radiant exitance in this section.

Before we consider a general quadratic radiant exitance distribution, we divert our attention to a special quadratically-varying luminaire whose irradiance problem essentially leads to the special function Λ arising in the closed-form solution in the linear case. While expressing Λ in terms of Clausen integrals in chapter 3, a key step is to relate $\Lambda(\alpha, \beta)$ to another special function $\Upsilon(\mu, \nu)$ by equation (3.123), which was derived there by a change of variable that is not intuitive at all. The alternative derivation presented here is to make some connections between the function Λ and the irradiance due to a two-parameter family of triangular luminaires with quadratically-varying radiant exitance. This is actually how we found relation (3.123) in the first place.

Figure 5.3 shows a triangular luminaire in the $z = 1$ plane with two parameters: the angle μ and the edge length ν. Suppose that the radiant exitance distribution of this

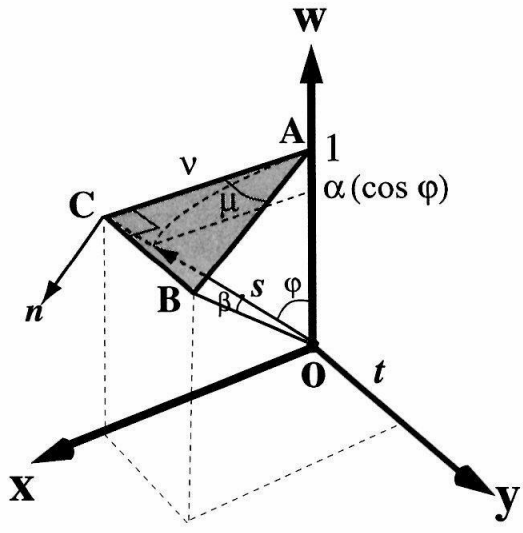


Figure 5.3: *The irradiance at the origin due to a triangular luminaire whose radiant exitance varies according to the quadratic function* $\phi(\mathbf{x}) = x^2 + y^2 + z^2$ *can be represented in terms of either* $\Upsilon(\mu, \nu)$ *or* $\Lambda(\alpha, \beta)$.

luminaire is given by

$$\phi(\mathbf{x}) \equiv x^2 + y^2 + z^2 = \langle \mathbf{x}, \mathbf{x} \rangle \tag{5.6}$$

for the point $\mathbf{x} = (x, y, z)$ on the luminaire plane. Thus the radiant exitance varies quadratically with position, but is independent of direction. It then follows from (4.2) that the radiance distribution $f(\mathbf{u})$ at the origin due to this luminaire can be expressed as

$$\begin{aligned} f(\mathbf{u}) &= \frac{1}{\pi} \phi\left(h \frac{\mathbf{u}}{\langle \mathbf{w}, \mathbf{u} \rangle}\right) \\ &= \left\langle \frac{\mathbf{u}}{\langle \mathbf{w}, \mathbf{u} \rangle}, \frac{\mathbf{u}}{\langle \mathbf{w}, \mathbf{u} \rangle} \right\rangle \\ &= \frac{1}{\langle \mathbf{w}, \mathbf{u} \rangle^2}, \end{aligned} \tag{5.7}$$

where the distance $h = 1$ for this particular configuration, and the perpendicular vector $\mathbf{w}$ is coincident with the $\mathbf{z}$ axis, which is also the normal $\mathbf{b}$ at the origin. Let P denote the spherical projection of this right triangle at the origin $\mathbf{o}$. Combining equation (5.1) and equation (5.7), the irradiance at the origin due to this luminaire is given by

$$\begin{aligned}\Phi(\mu,\nu) &= \int_P \frac{1}{\langle \mathbf{w},\mathbf{u}\rangle^2}\,\langle \mathbf{w},\mathbf{u}\rangle\; d\sigma(\mathbf{u}) \\ &= \int_P \frac{1}{\langle \mathbf{w},\mathbf{u}\rangle}\; d\sigma(\mathbf{u}) \\ &= \tau^{0;1}(P;\mathbf{w}). \end{aligned} \tag{5.8}$$

We denote the irradiance by $\Phi(\mu,\nu)$ to indicate its dependency on the geometry of the lumnaire. Using equation (4.21), we may evaluate equation (5.8) along three edges of the triangle by

$$\Phi(\mu,\nu) = \sum_{i=1}^{3} \langle \mathbf{w},\mathbf{n}_i\rangle\, B^*(\zeta_i). \tag{5.9}$$

where $\zeta_1, \zeta_2, \zeta_3$ are the great arcs corresponding to the three edges AB, BC and CA, respectively, and $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ are their outgoing normals, and B^* is given in equation (3.112). Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ denote the position vectors for the three triangle vertices, then

$$\begin{aligned}\mathbf{n}_1 &= \mathbf{unit}[\,\mathbf{A}\times\mathbf{B}\,] \\ \mathbf{n}_2 &= \mathbf{unit}[\,\mathbf{B}\times\mathbf{C}\,] \\ \mathbf{n}_3 &= \mathbf{unit}[\,\mathbf{C}\times\mathbf{A}\,].\end{aligned}$$

Notice that $\mathbf{w}$ is perpendicular to both $\mathbf{n}_1$ and $\mathbf{n}_3$, we have $\langle \mathbf{w},\mathbf{n}_1\rangle = \langle \mathbf{w},\mathbf{n}_3\rangle = 0$. Consequently, equation (5.9) simplifies to

$$\Phi(\mu,\nu) = \langle \mathbf{w},\mathbf{n}\rangle\, B^*(\zeta), \tag{5.10}$$

where $\mathbf{n}$ and ζ refer to the edge BC. If we apply the parameterization described in section 3.4.1 to BC, then the two orthonormal vectors $\mathbf{s}$, $\mathbf{t}$ are respectively $\mathbf{unit}[\,\mathbf{C}\,]$ and

y axis. It thus follows from equations (3.106) and (3.107) that

$$\begin{aligned} \langle \mathbf{w}, \mathbf{s} \rangle &= \cos\varphi, \quad \langle \mathbf{w}, \mathbf{t} \rangle = 0, \\ c_{\mathbf{w}} &= \cos\varphi, \quad \phi_{\mathbf{w}} = 0, \end{aligned}$$

where φ is the angle between $\mathbf{w}$ and $\mathbf{s}$, as shown in Figure 5.3. Let $\alpha = \cos\varphi$ and $\beta = \Theta$, arc length of ζ, we may evaluate $B^*(\zeta)$ using equation (3.114) by

$$\begin{aligned} \overline{B^*} &= \Lambda(c_{\mathbf{w}}, \Theta - \phi_{\mathbf{w}}) - \Lambda(c_{\mathbf{w}}, -\phi_{\mathbf{w}}) \\ &= \Lambda(\alpha, \beta), \end{aligned}$$

which allows us to express equation (5.10) as

$$\Phi(\mu, \nu) = \langle \mathbf{w}, \mathbf{n} \rangle \, \Lambda(\alpha, \beta). \tag{5.11}$$

In this special configuration, $\mathbf{n}$ is perpendicular to the triangle OBC, and thus orthogonal to the **y** axis. Consequently, $\mathbf{n}$ lies on the XZ plane and points in the outwards direction, as demonstrated in Figure 5.3. Then

$$\langle \mathbf{w}, \mathbf{n} \rangle = \cos\left(\frac{\pi}{2} + \varphi\right) = -\sin\varphi = -\sqrt{1 - \alpha^2}.$$

From equation (5.11), we may express the irradiance in terms of Λ by

$$\Phi(\mu, \nu) = -\sqrt{1 - \alpha^2} \, \Lambda(\alpha, \beta). \tag{5.12}$$

On the other hand, expressing the 2-form $d\omega$ by

$$d\omega = \frac{\cos\theta}{r^2} dA = \frac{\langle \mathbf{w}, \mathbf{u} \rangle}{r^2} dA$$

where r is the distance from the origin to the point $\mathbf{x}$ in the differential area dA of the luminaire, and $\mathbf{u}$ is the unit vector from $\mathbf{o}$ to $\mathbf{x}$, the irradiance in equation (5.8) may also

be computed by an integral over the area of the triangle. That is,

$$\begin{aligned}\Phi(\mu,\nu) &= \int\int_{\triangle} \frac{1}{r^2}\, dA \\ &= \int\int_{\triangle} \frac{1}{x^2+y^2+1}\, dxdy. \end{aligned} \tag{5.13}$$

By changing variables to polar coordinates, the integral in (5.13) reduces to $\Upsilon(\mu,\nu)$ after simplification. We have

$$\begin{aligned}\Phi(\mu,\nu) &= \frac{1}{2}\int_0^{\mu}\int_0^{\nu\sec\theta} \frac{2}{1+\rho^2}\,\rho\, d\rho d\theta \\ &= \frac{1}{2}\int_0^{\mu} \ln(1+\rho^2)\Big|_0^{\nu\sec\theta}\, d\theta \\ &= \frac{1}{2}\int_0^{\mu} \ln(1+\nu^2\sec^2\theta)\, d\theta \\ &= \frac{1}{2}\Upsilon(\mu,\nu). \end{aligned} \tag{5.14}$$

Incorporating equation (5.12) and equation (5.14), we may relate these two special functions by

$$\Lambda(\alpha,\beta) = -\frac{\Upsilon(\mu,\nu)}{2\sqrt{1-\alpha^2}} \tag{5.15}$$

for $\alpha \neq 1$, where (μ,ν) and (α,β) are dependent on each other. As seen from the geometric configuration in Figure 5.3, we have

$$\begin{aligned}\alpha &= \cos\varphi = \frac{1}{\sqrt{1+\nu^2}}, \\ \tan\beta &= \frac{\nu\tan\mu}{\sqrt{1+\nu^2}}. \end{aligned} \tag{5.16}$$

Solving for μ and ν, we get

$$\begin{aligned}\mu(\alpha,\beta) &= \tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right), \\ \nu(\alpha) &= \frac{\sqrt{1-\alpha^2}}{\alpha}. \end{aligned} \tag{5.17}$$

When the parameter (α, β) ranges over the required domain, $(0,1) \times [0, \pi/2]$, it is easy to check from equation (5.17) that the corresponding parameter (μ, ν) varies over $[0, \pi/2) \times [0, \infty)$, defining a family of triangular luminaires shown in Figure 5.3. Finally, the relation (3.123) on Page 99 follows directly from equation (5.15) and equation (5.17). An important consequence of equation (5.12) is that any method for computing the irradiance from quadratically-varying polygonal luminaires must also be capable of evaluating Λ over its entire domain. In other words, Λ becomes an inescapable component of such irradiance computations, thus it is not surprising that it appears in the problem we have addressed in chapter 3.

Next we shall turn our analysis to a planar luminaire with an arbitrary quadratic radiant exitance distribution function $\phi(\mathbf{x})$. As mentioned at the beginning of section 5.1, we are only required to compute the irradiance due to a quadratic monomial distribution $\phi(\mathbf{x}) = x_i x_j$, given by

$$\Phi_2(A) \;=\; \frac{h^2}{\pi} \int_A \frac{\langle \mathbf{e}_i, \mathbf{u} \rangle \langle \mathbf{e}_j, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2} \, d\sigma(\mathbf{u}) \tag{5.18}$$

for $i, j = 1, 2, 3$, where h, $\mathbf{w}$ and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are defined as before. Let $\mathbf{a} = h\,\mathbf{e}_i$ and $\mathbf{c} = h\,\mathbf{e}_j$, equation (5.18) is equivalent to

$$\Phi_2(A) \;=\; \frac{1}{\pi} \int_A \frac{\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle \langle \mathbf{c}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2} \, d\sigma(\mathbf{u}). \tag{5.19}$$

Similarly, we can represent the integral in equation (5.19) in terms of generalized irradiance tensors or rational moments as

$$\Phi_2(A) \;=\; \frac{1}{\pi} \tau^{3;2}(A, \mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{w}) \;=\; \frac{1}{\pi} \mathbf{T}^{3;2}_{ijk} \, \mathbf{a}_i \mathbf{b}_j \mathbf{c}_k. \tag{5.20}$$

In order to obtain a boundary integral formula for Φ_2, we proceed by replacing $\mathbf{T}^{3;2}$ with equation (4.5), attaining

$$\mathbf{T}^{3,2}_{ijk} \;=\; \mathbf{w}_i \mathbf{T}^{2,1}_{jk} + \mathbf{w}_j \mathbf{T}^{2,1}_{ik} + \mathbf{w}_k \mathbf{T}^{2,1}_{ij} - 4\mathbf{T}^{3}_{ijk} - \int_{\partial A} \frac{\mathbf{u}^3_{ijk} \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \, ds. \tag{5.21}$$

Using equation (5.4) and the 3-tensor $\mathbf{M}^3$, we may express the tensor $\mathbf{T}^{2;1}$ above as a

boundary integral by

$$\mathbf{T}_{ij}^{2;1} = -\frac{1}{2}\int_{\partial A} \mathbf{M}_{ijl}^{3}\,\mathbf{n}_l\,ds. \tag{5.22}$$

Moreover, it follows from equation (3.32) that

$$\begin{aligned} 4\mathbf{T}_{ijk}^{3} &= \delta_{ki}\mathbf{T}_{j}^{1} + \delta_{kj}\mathbf{T}_{i}^{1} - \int_{\partial A} \mathbf{u}_{ij}^{2}\mathbf{n}_k\,ds \\ &= -\frac{1}{2}\int_{\partial A} \left(\delta_{ki}\delta_{jl} + \delta_{kj}\delta_{il} + 2\,\delta_{kl}\,\mathbf{u}_i\mathbf{u}_j\right)\mathbf{n}_l\,ds. \end{aligned} \tag{5.23}$$

Substituting equation (5.22) and equation (5.23) into equation (5.21), we get

$$\mathbf{T}_{ijk}^{3,2} = -\frac{1}{2}\int_{\partial A} \mathbf{M}_{ijkl}^{4}(\mathbf{w},\mathbf{u})\,\mathbf{n}_l\,ds, \tag{5.24}$$

where the 4-tensor $\mathbf{M}^4$ depending on $\mathbf{w}$ and $\mathbf{u}$ is defined in terms of the lower-order tensor $\mathbf{M}^3$ by

$$\mathbf{M}_{ijkl}^{4} \equiv \mathbf{w}_i\mathbf{M}_{jkl}^{3} + \mathbf{w}_j\mathbf{M}_{ikl}^{3} + \mathbf{w}_k\mathbf{M}_{ijl}^{3} + 2\,\mathbf{w}_l\,\frac{\mathbf{u}_i\mathbf{u}_j\mathbf{u}_k}{\langle\mathbf{w},\mathbf{u}\rangle} - \delta_{ki}\delta_{jl} - \delta_{kj}\delta_{il} - 2\,\delta_{kl}\,\mathbf{u}_i\mathbf{u}_j \tag{5.25}$$

Combining equation (5.20) and equation (5.24), we are able to reduce the irradiance integral in (5.19) to a one-dimensional integral, given by

$$\boxed{\Phi_2(A,\mathbf{w},\mathbf{a},\mathbf{c},\mathbf{b}) = -\frac{1}{2\pi}\int_{\partial A} \mathbf{M}_{ijkl}^{4}\,\mathbf{a}_i\mathbf{b}_j\mathbf{c}_k\mathbf{n}_l\,ds.} \tag{5.26}$$

Equation (5.26) and equation (5.25) suggest that the solution for the irradiance due to a quadratically-varying luminaire is based on that for linearly-varying luminaires. Let $\Phi_1(A,\mathbf{w},\mathbf{a},\mathbf{b})$ denote the irradiance given in equation (5.5), we may express $\Phi_2(A)$ in equation (5.26) by

$$\Phi_2(A,\mathbf{w},\mathbf{a},\mathbf{b},\mathbf{c}) = \langle\mathbf{w},\mathbf{a}\rangle\,\Phi_1(A,\mathbf{w},\mathbf{b},\mathbf{c}) + \langle\mathbf{w},\mathbf{b}\rangle\,\Phi_1(A,\mathbf{w},\mathbf{a},\mathbf{c}) + \langle\mathbf{w},\mathbf{c}\rangle\,\Phi_1(A,\mathbf{w},\mathbf{a},\mathbf{b})$$

$$-\frac{1}{2\pi}\left[2\int_{\partial A}\left(\langle \mathbf{w},\mathbf{n}\rangle\,\frac{\langle \mathbf{a},\mathbf{u}\rangle\,\langle \mathbf{b},\mathbf{u}\rangle\,\langle \mathbf{c},\mathbf{u}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle}-\langle \mathbf{c},\mathbf{n}\rangle\,\langle \mathbf{a},\mathbf{u}\rangle\,\langle \mathbf{b},\mathbf{u}\rangle\right)ds\right.$$
$$\left.-\langle \mathbf{a},\mathbf{c}\rangle\int_{\partial A}\langle \mathbf{b},\mathbf{n}\rangle\,ds-\langle \mathbf{b},\mathbf{c}\rangle\int_{\partial A}\langle \mathbf{a},\mathbf{n}\rangle\,ds\right], \tag{5.27}$$

Furthermore, by specializing equation (5.27) to a polygonal luminaire and using assorted exact formulas for $B^{m,l}$ that was listed on Page 124, we may derive a closed-form solution for the irradiance at the origin due to a luminaire whose radiant exitance varies according to a quadratic monomial function. Let us define

$$H(\zeta,\mathbf{v}_1,\mathbf{v}_2)\;\equiv\;\langle \mathbf{v}_1,\mathbf{n}\rangle\,\langle \mathbf{v}_2,\mathbf{w}\rangle\,\Theta+\mathbf{v}_2{}^{\mathrm{T}}\,(\mathbf{I}-\mathbf{w}\mathbf{w}^{\mathrm{T}})\left(\mathbf{n}\,\overline{B}^{1,1}(\mathbf{v}_1)-\langle \mathbf{w},\mathbf{n}\rangle\,\mathbf{v}_1\,\overline{B^*}\right), \tag{5.28}$$

where Θ and $\mathbf{n}$ depends on ζ. It follows from equation (3.116) that H represents the contribution from each edge to the irradiance due to a linearly-varying luminaire. Thus we have

$$\Phi_1(P,\mathbf{w},\mathbf{a},\mathbf{b})\;=\;-\frac{1}{2\pi}\sum_{i=1}^{k}H(\zeta_i,\mathbf{a},\mathbf{b})$$

for a k-sided polygon P. Consequently, we can evaluate equation (5.27) for P by

$$\begin{aligned}\Phi_2(P,\mathbf{w},\mathbf{a},\mathbf{b},\mathbf{c})\;=\;-\frac{1}{2\pi}\sum_{i=1}^{k}\Big\{&\langle \mathbf{w},\mathbf{a}\rangle\,H(\zeta_i,\mathbf{b},\mathbf{c})+\langle \mathbf{w},\mathbf{b}\rangle\,H(\zeta_i,\mathbf{a},\mathbf{c})+\langle \mathbf{w},\mathbf{c}\rangle\,H(\zeta_i,\mathbf{a},\mathbf{b})\\&+2\,\langle \mathbf{w},\mathbf{n}\rangle\,\overline{B}^{3,1}(\mathbf{a},\mathbf{b},\mathbf{c})-2\,\langle \mathbf{c},\mathbf{n}\rangle\,\overline{B}^{2}(\mathbf{a},\mathbf{b})\\&-\left[\langle \mathbf{a},\mathbf{c}\rangle\,\langle \mathbf{b},\mathbf{n}\rangle+\langle \mathbf{b},\mathbf{c}\rangle\,\langle \mathbf{a},\mathbf{n}\rangle\right]\Theta\Big\}.\end{aligned} \tag{5.29}$$

In equation (5.28) and equation (5.29), $\overline{B}^{1,1}, \overline{B^*}, \overline{B}^{3,1}$ and $\overline{B}^{2}$ are given respectively by equations (4.25), (3.114), (4.26) and (4.24).

Finally, we end this section with a discussion about the more general case of a receiver point $\mathbf{q}$ other than the origin. Consider a luminaire whose radiant exitance varies according to a quadratic monomial function $\phi(\mathbf{x})$, which is defined in terms of the position vector $\mathbf{x}=(x_1,x_2,x_3)$ of a luminaire point $\mathbf{p}$ with respect to the coordinate system $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ at the origin. We shall demonstrate that the irradiance at $\mathbf{q}$ due to this lu-

minaire is completely determined by the formulas derived for Φ_0, Φ_1, Φ_2 at the origin, where Φ_0 denotes Lambert's formula for uniform luminaires. The idea is to express $\phi(\mathbf{x})$ as a function $\hat{\phi}$ in terms of the coordinates of $\mathbf{x}-\mathbf{q}$, the position vector of the same point $\mathbf{p}$ with respect to the coordinate system $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ translated to $\mathbf{q}$. It turns out that $\hat{\phi}$ becomes a combination of monomials of orders 0, 1 and 2, and thus our previous formulas can be applied as if the receiver point is at the origin. In detail, given a quadratic monomial function $\phi(\mathbf{x}) = \langle \mathbf{e}_i, \mathbf{x} \rangle \langle \mathbf{e}_j, \mathbf{x} \rangle$, we may derive $\hat{\phi}$ as follows:

$$\begin{aligned}
\hat{\phi}(\mathbf{x}-\mathbf{q}) &= \phi(\mathbf{x}-\mathbf{q}+\mathbf{q}) \\
&= \langle \mathbf{x}-\mathbf{q}+\mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{x}-\mathbf{q}+\mathbf{q}, \mathbf{e}_j \rangle \\
&= [\langle \mathbf{x}-\mathbf{q}, \mathbf{e}_i \rangle + \langle \mathbf{q}, \mathbf{e}_i \rangle] [\langle \mathbf{x}-\mathbf{q}, \mathbf{e}_j \rangle + \langle \mathbf{q}, \mathbf{e}_j \rangle] \\
&= \langle \mathbf{x}-\mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{x}-\mathbf{q}, \mathbf{e}_j \rangle + \langle \mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{x}-\mathbf{q}, \mathbf{e}_j \rangle + \\
&\quad \langle \mathbf{q}, \mathbf{e}_j \rangle \langle \mathbf{x}-\mathbf{q}, \mathbf{e}_i \rangle + \langle \mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{q}, \mathbf{e}_j \rangle . \qquad (5.30)
\end{aligned}$$

Therefore, the radiant exitance function of the luminaire may be expressed in terms of the position vector $\mathbf{r}$ with respect to the receiver point by

$$\hat{\phi}(\mathbf{r}) = \langle \mathbf{r}, \mathbf{e}_i \rangle \langle \mathbf{r}, \mathbf{e}_j \rangle + \langle \mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{r}, \mathbf{e}_j \rangle + \langle \mathbf{q}, \mathbf{e}_j \rangle \langle \mathbf{r}, \mathbf{e}_i \rangle + \langle \mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{q}, \mathbf{e}_j \rangle ,$$

which leads to a radiance distribution $f(\mathbf{u})$ at $\mathbf{q}$ given by

$$\frac{1}{\pi} \left(h^2(\mathbf{q}) \frac{\langle \mathbf{e}_i, \mathbf{u} \rangle \langle \mathbf{e}_j, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2} + h(\mathbf{q}) \left[\langle \mathbf{q}, \mathbf{e}_i \rangle \frac{\langle \mathbf{e}_j, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} + \langle \mathbf{q}, \mathbf{e}_j \rangle \frac{\langle \mathbf{e}_i, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \right] + \langle \mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{q}, \mathbf{e}_j \rangle \right) , \qquad (5.31)$$

where we have used the identity $\mathbf{r} = h(\mathbf{q})\mathbf{u}/\langle \mathbf{w}, \mathbf{u} \rangle$, with $h(\mathbf{q})$ as the distance from $\mathbf{q}$ to the luminaire plane. Consequently, the irradiance at $\mathbf{q}$ can be computed by integrating $f(\mathbf{u})$ over the spherical projection A of the luminaire onto the hemisphere about $\mathbf{q}$, yielding

$$\begin{aligned}
\Phi_2(A, \mathbf{q}) &= h^2(\mathbf{q}) \Phi_2(A, \mathbf{e}_i, \mathbf{e}_j, \mathbf{b}) + \langle \mathbf{q}, \mathbf{e}_i \rangle \langle \mathbf{q}, \mathbf{e}_j \rangle \Phi_0(A, \mathbf{b}) + \\
&\quad h(\mathbf{q}) [\langle \mathbf{q}, \mathbf{e}_i \rangle \Phi_1(A, \mathbf{w}, \mathbf{e}_j, \mathbf{b}) + \langle \mathbf{q}, \mathbf{e}_j \rangle \Phi_1(A, \mathbf{w}, \mathbf{e}_i, \mathbf{b})] , \qquad (5.32)
\end{aligned}$$

where Φ_1, Φ_2 are given by formulas (5.5) and (5.26) respectively, and $\Phi_0(A, \mathbf{b})$ denotes the irradiance due to a uniform luminaire with a spherical projection A at the origin (with normal $\mathbf{b}$), given by Lambert's formula as

$$\Phi_0(A, \mathbf{b}) = \frac{1}{\pi} \int_A \langle \mathbf{b}, \mathbf{u} \rangle \, d\sigma(\mathbf{u}) = -\frac{1}{2\pi} \int_{\partial A} \langle \mathbf{b}, \mathbf{n} \rangle \, ds. \tag{5.33}$$

It can be seen from equation (5.32) that the irradiance $\Phi_2(A, \mathbf{q})$ at an arbitrary receiver point $\mathbf{q}$ is comprised of contributions from radiant exitance variations of orders 0, 1 and 2.

5.1.3 Higher Order Polynomials

In this section we shall generalize the closed-form solutions for the irradiance due to linearly-varying and quadratically-varying luminaires derived in previous sections to higher order polynomials. For simplicity, we only consider the luminaire with a radiant exitance distribution given by a nth-order ($n = 1, 2, \ldots$) monomial of x_1, x_2, x_3, directional cosines of the luminaire point $\mathbf{x}$ with respect to an orthogonal coordinate system $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ at the origin, as these monomials serve as basis functions of the space P_n consisting of all the polynomials of at most degree n.

As discussed before, the irradiance at the origin due to such an emitter is equivalent to computing the integral given in equation (5.2). Let

$$\begin{aligned} \mathbf{r} &= h\,\mathbf{e}_1, \\ \mathbf{s} &= h\,\mathbf{e}_2, \\ \mathbf{t} &= h\,\mathbf{e}_3, \end{aligned}$$

then the required integral has the following form

$$\Phi_n(A, \mathbf{w}, \mathbf{V}, \mathbf{b}) = \frac{1}{\pi} \int_A \frac{\langle \mathbf{r}, \mathbf{u} \rangle^p \langle \mathbf{s}, \mathbf{u} \rangle^q \langle \mathbf{t}, \mathbf{u} \rangle^r \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^n} \, d\sigma(\mathbf{u}), \tag{5.34}$$

where $\mathbf{b}$, $\mathbf{w}$ are defined as before, and we define $\mathbf{V}$ as the sequence of n vectors, comprised

of p occurrence of $\mathbf{r}$ followed by q occurrence of $\mathbf{s}$ and r occurrence of $\mathbf{t}$, that is

$$\mathbf{V} \equiv (\underbrace{\mathbf{r}, \cdots, \mathbf{r}}_{p}, \underbrace{\mathbf{s}, \cdots, \mathbf{s}}_{q}, \underbrace{\mathbf{t}, \cdots, \mathbf{t}}_{r}) \tag{5.35}$$

for $p + q + r = n$. Obviously, equation (5.34) subsumes the irradiance integrals (5.33), (3.7) and (5.19) as its special cases. That is, $\mathbf{V} = \emptyset$ for uniform luminaires, $\mathbf{V} = \mathbf{a}$ for the linear case, and $\mathbf{V} = (\mathbf{a}, \mathbf{c})$ for the second order monomial. Using this formulation, we now state and prove the central theorem of this section.

Formulas (5.33), (5.5) and (5.26) exhibit a general structure that is easy to be extended to higher order cases. Recall that

$$\begin{aligned}
\Phi_0(A, \mathbf{b}) &= -\frac{1}{2\pi} \int_{\partial A} \delta_{jk} \mathbf{b}_j \mathbf{n}_k \, ds, \\
\Phi_1(A, \mathbf{w}, \mathbf{a}, \mathbf{b}) &= -\frac{1}{2\pi} \int_{\partial A} \mathbf{M}^3_{ijk}(\mathbf{w}, \mathbf{u}) \, \mathbf{a}_i \mathbf{b}_j \mathbf{n}_k \, ds, \\
\Phi_2(A, \mathbf{w}, \mathbf{a}, \mathbf{c}, \mathbf{b}) &= -\frac{1}{2\pi} \int_{\partial A} \mathbf{M}^4_{iljk}(\mathbf{w}, \mathbf{u}) \, \mathbf{a}_i \mathbf{c}_l \mathbf{b}_j \mathbf{n}_k \, ds.
\end{aligned}$$

Following the common pattern shown above, we speculate that $\Phi_n(A, \mathbf{V}, \mathbf{b}, \mathbf{w})$ may be expressed in a similar way by introducing a $(n+2)$-tensor $\mathbf{M}^{n+2}$, yielding

$$\boxed{\Phi_n(A, \mathbf{w}, \mathbf{V}, \mathbf{b}) = -\frac{1}{2\pi} \int_{\partial A} \mathbf{M}^{n+2}_{Ijk}(\mathbf{w}, \mathbf{u}) \, \mathbf{V}_I \mathbf{b}_j \mathbf{n}_k \, ds,} \tag{5.36}$$

where I is a n-index $(i_1, \ldots, i_n)$. With $\mathbf{V}$ given in equation (5.35), the notation $\mathbf{V}_I$ denotes

$$\mathbf{V}_I = (\mathbf{r})_{i_1} \cdots (\mathbf{r})_{i_p} (\mathbf{s})_{i_{p+1}} \cdots (\mathbf{s})_{i_{p+q}} (\mathbf{t})_{i_{p+q+1}} \cdots (\mathbf{t})_{i_n}. \tag{5.37}$$

To prove this speculation and also derive the expression for the unknown tensor $\mathbf{M}^{n+2}$ in equation (5.36), we express Φ_n given in equation (5.34) in terms of generalized

irradiance tensors, attaining

$$\Phi_n(A, \mathbf{w}, \mathbf{V}, \mathbf{b}) \;=\; \frac{1}{\pi}\tau^{n+1;n}(A, \mathbf{V}, \mathbf{b}; \mathbf{w}) \;=\; \frac{1}{\pi}\mathbf{T}^{n+1;n}_{Ij}(A, \mathbf{w})\, \mathbf{V}_I \mathbf{b}_j, \tag{5.38}$$

where $\mathbf{V}_I$ is interpreted by equation (5.37). Equation (5.38) suggests that the irradiance due to an emitter with its radiant exitance varying according to a nth-order monomial is associated with a special generalized irradiance tensor, $\mathbf{T}^{n+1;n}$, where the orders of the numerator and the denominator differ by one. Actually, it follows from equations (5.33), (5.3) and (5.20) that this claim holds for the special cases of uniform, linearly-varying and quadratically emitters, which correspond to $n = 0, 1, 2$, respectively. As a consequence, Equation (5.36) can be easily derived by proving a recurrence relation about this special tensor $\mathbf{T}^{n+1;n}$, which is phrased as a theorem.

Theorem 9 *Given an integer $n \geq 0$, the generalized irradiance tensor $\mathbf{T}^{n+1;n}$ is given by a boundary integral*

$$\mathbf{T}^{n+1;n}_{Ij}(A, \mathbf{w}) \;=\; -\frac{1}{2}\int_{\partial A} \mathbf{M}^{n+2}_{Ijk}(\mathbf{w}, \mathbf{u})\, \mathbf{n}_k \, ds, \tag{5.39}$$

where I is a n-index, and the $(n+2)$-tensor $\mathbf{M}$ depending on unit vectors $\mathbf{w}$ and $\mathbf{u}$ satisfies a recurrence relation

$$\begin{aligned}\mathbf{M}^{n+2}_{Ijk} \;=\; \frac{1}{n-1}\Bigg[&\sum_{i=1}^{n}\left(\mathbf{w}_{I_i}\mathbf{M}^{n+1}_{(I\setminus i)jk} - \delta_{jI_i}\mathbf{M}^{n}_{(I\setminus i)k}\right) + (n-1)\,\mathbf{w}_j\mathbf{M}^{n+1}_{Ik} \\ &+2\mathbf{w}_k\frac{\mathbf{u}^{n+1}_{Ij}}{\langle \mathbf{w}, \mathbf{u}\rangle^{n-1}} - 2\delta_{jk}\frac{\mathbf{u}^{n}_{I}}{\langle \mathbf{w}, \mathbf{u}\rangle^{n-2}}\Bigg],\end{aligned} \tag{5.40}$$

which is completed by the base cases:

$$\mathbf{M}^{3}_{ijk} \;=\; \delta_{ik}\mathbf{w}_j + (\delta_{jm} - \mathbf{w}_j\mathbf{w}_m)\left(\frac{\delta_{km}\mathbf{u}_i}{\langle \mathbf{w}, \mathbf{u}\rangle} - \delta_{im}\mathbf{w}_k\,\eta\right), \tag{5.41}$$

$$\mathbf{M}^{2}_{jk} \;=\; \delta_{jk}. \tag{5.42}$$

Proof: The proof is performed by induction on n. First, from equation (3.33) we have

$$\mathbf{T}_j^{1;0} \;=\; -\frac{1}{2}\int_{\partial A} \mathbf{n}_j \, ds \;=\; -\frac{1}{2}\int_{\partial A} \delta_{jk}\mathbf{n}_k \, ds.$$

Thus equation (5.39) holds for $n = 0$. The validity of equation (5.39) for $n = 1$ has been verified in equation (5.4) of section 5.1.1. Now we proceed with induction. Suppose that

$$\mathbf{T}_{Lj}^{k+1;k} \;=\; -\frac{1}{2}\int_{\partial A} \mathbf{M}_{Ljk}^{k+2}\, \mathbf{n}_k \, ds \tag{5.43}$$

is true for all $2 \leq k < n$, where L is a k-index. When $n \geq 2$, we may apply the recursive formula (4.5), obtaining

$$\mathbf{T}_{Ij}^{n+1;n} \;=\; \frac{1}{n-1}\left[\sum_{i=1}^{n} \mathbf{w}_{I_i}\mathbf{T}_{(I\setminus i)j}^{n;n-1} + \mathbf{w}_j\mathbf{T}_I^{n;n-1} - 4\mathbf{T}_{Ij}^{n+1;n-2} - \int_{\partial A} \frac{\mathbf{u}_{Ij}^{n+1}\,\langle \mathbf{w},\mathbf{n}\rangle}{\langle \mathbf{w},\mathbf{u}\rangle^{n-1}}\, ds\right] \tag{5.44}$$

According to Theorem 6, we may replace $\mathbf{T}^{n+1;n-2}$ above by

$$4\mathbf{T}_{Ij}^{n+1;n-2} \;=\; \sum_{i=1}^{n} \delta_{jI_i}\mathbf{T}_{I\setminus i}^{n-1;n-2} - (n-2)\,\mathbf{w}_j\mathbf{T}_I^{n;n-1} - \int_{\partial A} \frac{\mathbf{u}_I^n\mathbf{n}_j}{\langle \mathbf{w},\mathbf{u}\rangle^{n-2}}\, ds.$$

Thus equation (5.44) becomes

$$\begin{aligned}\mathbf{T}_{Ij}^{n+1;n} \;=\; \frac{1}{n-1}\Bigg[&\sum_{i=1}^{n} \mathbf{w}_{I_i}\mathbf{T}_{(I\setminus i)j}^{n;n-1} + (n-1)\mathbf{w}_j\mathbf{T}_I^{n;n-1} - \sum_{i=1}^{n} \delta_{jI_i}\mathbf{T}_{I\setminus i}^{n-1;n-2} \\ &+ \int_{\partial A}\left(\delta_{jk}\frac{\mathbf{u}_I^n}{\langle \mathbf{w},\mathbf{u}\rangle^{n-2}} - \mathbf{w}_k\frac{\mathbf{u}_{Ij}^{n+1}}{\langle \mathbf{w},\mathbf{u}\rangle^{n-1}}\right)\mathbf{n}_k\, ds.\Bigg]\end{aligned} \tag{5.45}$$

Expanding the tensors $\mathbf{T}^{n;n-1}$ and $\mathbf{T}^{n-1;n-2}$ in equation (5.45) with the induction hypothesis (5.44), and then extracting a common factor $\mathbf{n}_k$, we obtain

$$\begin{aligned}\mathbf{T}_{Ij}^{n+1;n} \;=\; -\frac{1}{2}\int_{\partial A}\frac{1}{n-1}\Bigg[&\sum_{i=1}^{n} \mathbf{w}_{I_i}\mathbf{M}_{(I\setminus i)jk}^{n+1} + (n-1)\,\mathbf{w}_j\mathbf{M}_{Ik}^{n+1} - \sum_{i=1}^{n} \delta_{jI_i}\mathbf{M}_{(I\setminus i)k}^{n} \\ &+2\mathbf{w}_k\frac{\mathbf{u}_{Ij}^{n+1}}{\langle \mathbf{w},\mathbf{u}\rangle^{n-1}} - 2\delta_{jk}\frac{\mathbf{u}_I^n}{\langle \mathbf{w},\mathbf{u}\rangle^{n-2}}\Bigg]\mathbf{n}_k\, ds.\end{aligned} \tag{5.46}$$

Comparing equation (5.46) with equations (5.39) and (5.40), we have proved equation (5.39) for all $n \geq 2$ under the assumption of (5.43). It then follows from mathematical induction that Theorem 9 applies for all $n \geq 0$. □□

Consequently, equation (5.36) follows directly from equation (5.39) and equation (5.38). The recurrence formula (5.40) about the tensor $\mathbf{M}$ relates the irradiance solution for higher-order polynomials to those for lower-order cases, as demonstrated in equation (5.27). In addition, since equation (5.36) reduces to Lambert's formula when $n = 0$, it provides a natural generalization of Lambert's formula for homogeneous emitters to non-homogeneous emitters with polynomially-varying radiant exitance.

Similarly, when A is restricted to be a spherical polygon with m edges, the irradiance Φ_n given in equation (5.36) can be evaluated by summing the contribution from each edge, resulting in

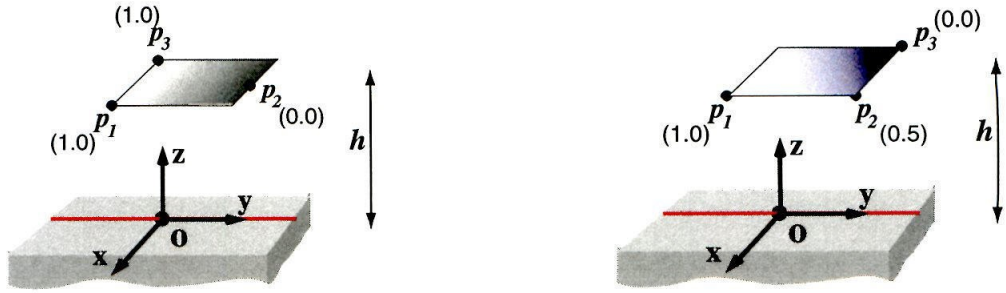
$$\Phi_n(A, \mathbf{w}, \mathbf{V}, \mathbf{b}) \;=\; -\frac{1}{2\pi}\left[\sum_{i=1}^{m}\left(\int_{\zeta} \mathbf{M}^{n+2}_{Ijk}\,\mathbf{V}_I \mathbf{b}_j \; ds\right)\mathbf{n}_k(\zeta)\right]. \tag{5.47}$$

It follows from equation (5.40) that the exact evaluation of each line integral in equation (5.47) finally relies on the analytical computations of base moment $\tau^{0;1}(A;\mathbf{w})$ and boundary integrals $B^{m,l}$, which have been discussed in section 4.3.2.

While it is not practical to evaluate Φ_n analytically for higher order polynomial due to its complexity, the new boundary integral representation given in equation (5.36) is primarily of theoretical interest as it proves that there exist closed-form solutions for such irradiance problems. Moreover, this new formulation may also suggests better numerical approximations.

5.1.4 Numerical Verification

Currently, few tools exist for computing the irradiance from inhomogeneous emitters aside from two commonly used numerical means: *Monte Carlo ray tracing* and *finite element method.* In Monte Carlo simulation, rays emitted from the origin to random jitterly-sampled points on the emitter are used to approximate the irradiance integral (3.1); while



(a) (b)

Figure 5.4: *A floor is illuminated by a square emitter with a linear radiant exitance distribution specified by three points* $\mathbf{p}_1$, $\mathbf{p}_2$ *and* $\mathbf{p}_3$ *and their associate radiant exitance values shown beside.* h *is the distance between the floor and the light source. (a) distribution* $\phi(\mathbf{x}) = y + \frac{1}{2}z$. *(b) distribution* $\phi(\mathbf{x}) = \frac{1}{2}(x - y + z)$.

finite element approach subdivides the emitter into several small patches (considered as homogeneous) and then compute the contribution of each patch using Lambert's formula. For the purpose of verifying the present analysis, we compare our analytical solutions with these two numerical approximations.

The study is carried out for a simple configuration: a floor is directly illuminated by a square emitter of size 1×1, which is located at a distance of $h = 1$ above the floor, as shown in Figure 5.4. The radiant exitance varies spatially over parts of the emitter surface. An orthogonal coordinate system $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ is defined at the center of the floor. We are computing the irradiance values along the red horizontal center line, where the y coordinates range from -1 to 1, and x, z coordinates remain zero.

Our attention is first directed to the linear case. Two linear distributions of radiant exitance are specified by positioning three non-collinear points $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ as marked in Figure 5.4a and Figure 5.4b and assigning the numbers nearby as their corresponding radiant exitance values. According to equation (3.4), the settings in Figure 5.4a and Figure 5.4b define the distribution of radiant exitance over the emitter as $\phi(\mathbf{x}) = y + \frac{1}{2}z$ and $\phi(\mathbf{x}) = \frac{1}{2}(x - y + z)$ respectively, where $(x, y, z) \equiv \mathbf{x}$. These two distributions approximately correspond to linear variations along horizontal and diagonal directions. Each of figures 5.5 through 5.8 shows the variation of the irradiance along the center

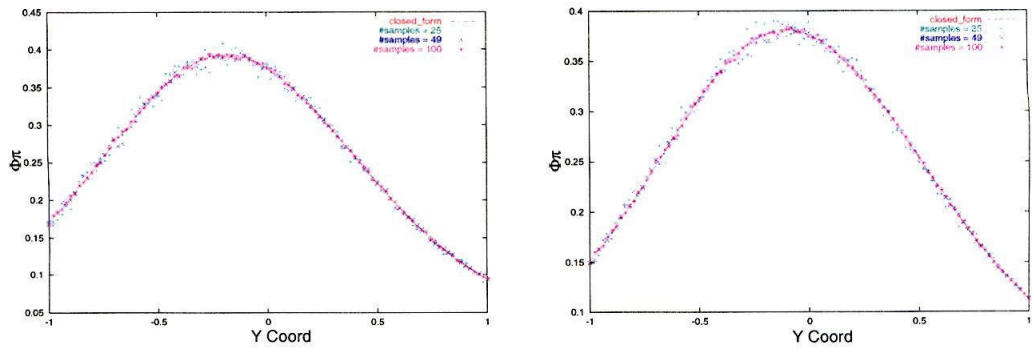


Figure 5.5: *Comparing our closed-form solutions with monte carlo simulations by varying the number of stratified samples. (a)* $\phi(\mathbf{x}) = y + \frac{1}{2}z$. *(b)* $\phi(\mathbf{x}) = \frac{1}{2}(x - y + z)$.

line, where the abscissa of y coordinate ranges from -1 to 1 and the irradiance value Φ scaled by π appears as the ordinate variable. Figure 5.5 has compared our exact solutions computed from equation (3.116) with Monte Carlo simulations from various numbers of stratified samples, that is, 25, 49 and 100. Figure 5.6 shows comparisons made with finite element results. There are plots for different levels of subdivision that correspond to 2, 8 and 50 triangle patches obtained from regular grids. Both comparisons show good agreement between our analytical results and numerical approximations when the number of samples or patches increases.

To test the formula (5.36) derived for a general nth order monomial radiant exitance distribution, we chose a quadratic function given by $\phi(\mathbf{x}) = y^2$, that is, $\mathbf{V} = (\mathbf{y}, \mathbf{y})$ in equation (5.36), and we performed the similar comparisons in Figure 5.7, where the close-form solution is computed using equation (5.29). As a final demonstration, we also show the numerical comparisons for a special quadratic radiant exitance distribution given in equation (5.6) in Figure 5.8. Since the light source and the floor are parallel, the normal to the floor $\mathbf{b}$ and the vector $\mathbf{w}$ orthogonal to the light source are coincident, thus our exact results can be computed from equations (5.8) and (3.114), in terms of Clausen integrals. All these plots show that our closed-form solutions match very well with both Monte Carlo simulations and finite element approximations in the limit.

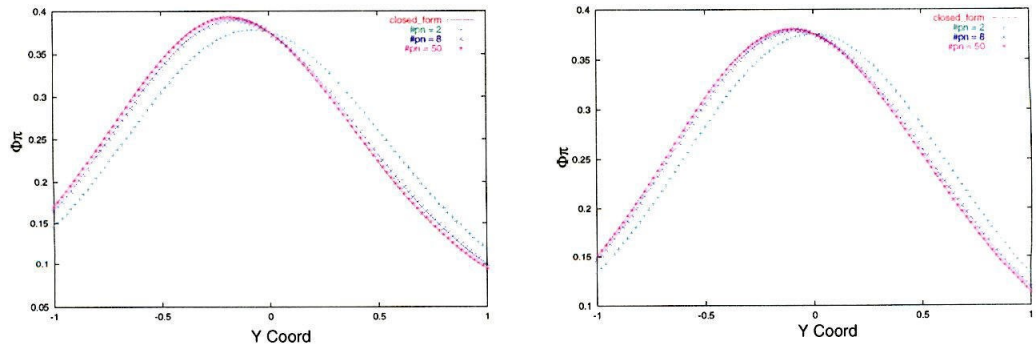


Figure 5.6: *Comparing our closed-form solutions with finite element approximations by varying subdivision levels. (a)* $\phi(\mathbf{x}) = y + \frac{1}{2}z$. *(b)* $\phi(\mathbf{x}) = \frac{1}{2}(x - y + z)$.

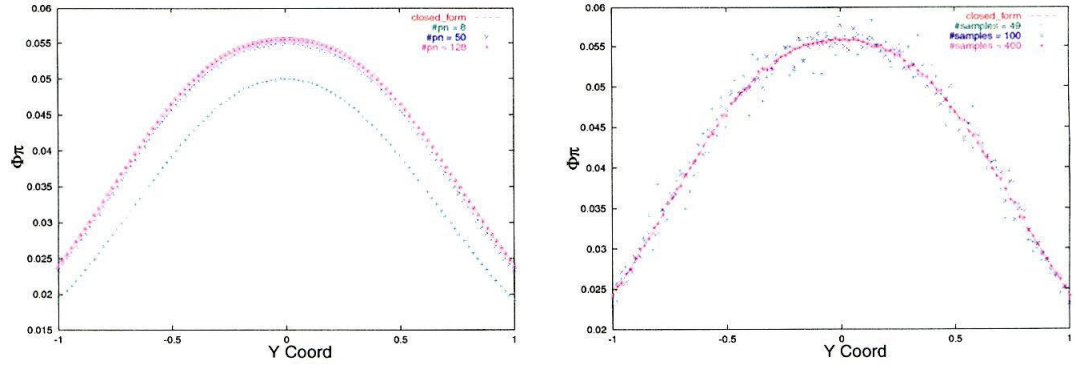
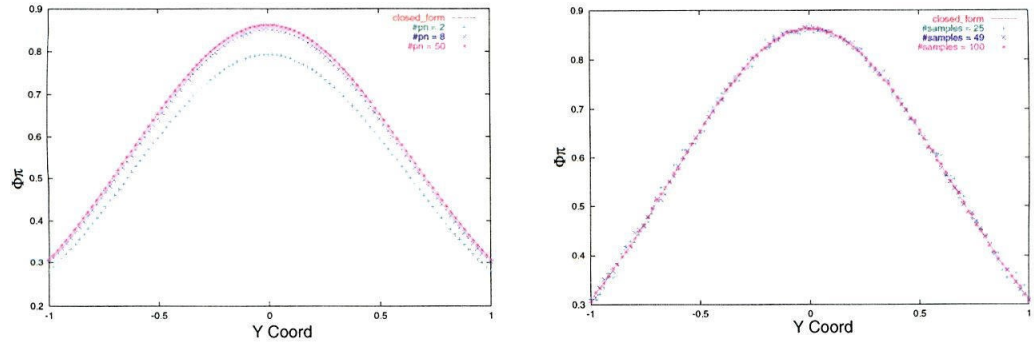


Figure 5.7: *Quadratic radiant exitance distribution* $\phi(\mathbf{x}) = y^2$. *(a) Comparison between our closed-form solutions with finite element approximation. (b) Comparison between our closed-form solutions with Monte Carlo simulations.*

5.2 Non-Lambertian Scattering Simulations

For ideal diffuse surfaces, irradiance is sufficient to compute the reflected radiance. However, the situation is quite different for non-diffuse surfaces, where the reflected or transmitted radiance also depends on the scattering features of the surfaces, which are in general specified by BRDF (*Bidirectional Reflectance Distribution Function*) for a reflecting surface, or BTDF (*Bidirectional Transmission Distribution Function*) for a transparent surface. For environments with reflection or transmission distributions defined in



(a) (b)

Figure 5.8: *Quadratic radiant exitance distribution* $\phi(\mathbf{x}) = x^2+y^2+z^2$. *(a) Comparison between our closed-form solutions with finite element approximation. (b) Comparison between our closed-form solutions with Monte Carlo simulations.*

terms of Phong exponents, the mathematical tools derived for generalized irradiance tensors $\mathbf{T}^{n;q}$ and associated rational moments $\tau^{n;q}$ may be applied to simulations of non-Lambertian scattering effects involving spatially-varying luminaires and non-diffuse surfaces. The similar approach employing irradiance tensors $\mathbf{T}^n$ has been used to simulate non-Lambertian scattering effects involving only uniform luminaires [9], while $\mathbf{T}^{n;q}$ as a generalization of $\mathbf{T}^n$ allows us to extend such simulations to involve spatially-varying luminaires as well. See Figure 5.9.

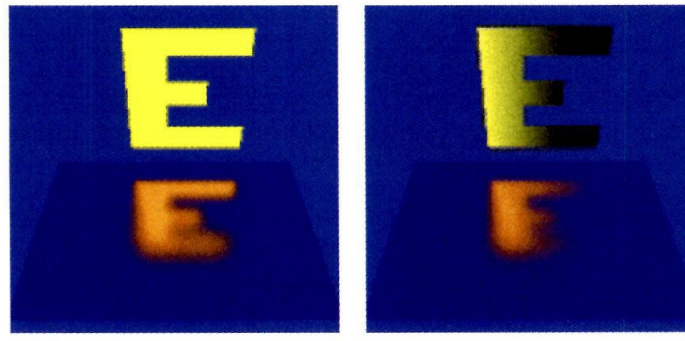


Figure 5.9: *Analytical glossy reflection of an E-shaped luminaire whose radiant exitance is uniform (left) and linearly-varying (right) with respect to position.*

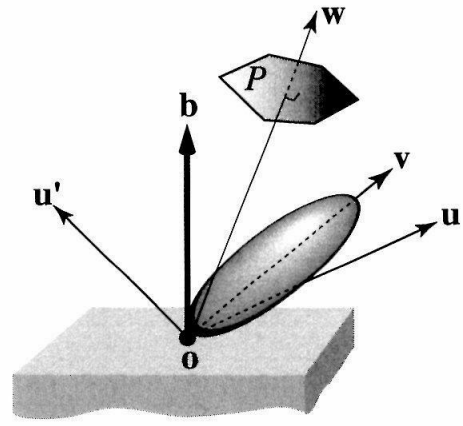


Figure 5.10: *Given a glossy surface with a simple BRDF defined in terms of a Phong exponent around the mirror reflection* $\mathbf{v}$*, the reflected radiance along the view direction* $\mathbf{u}'$ *at* $\mathbf{o}$ *due to a monomially-varying luminaire can be formulated as a rational moment* $\tilde{\tau}$*.*

5.2.1 Phong-Like Glossy Reflection and Transmission

The problem of computing view-dependent glossy reflection involving a monomially-varying luminaire and a non-diffuse surface with Phong-like BRDF can be formulated in terms of those special rational moments $\tilde{\tau}$ defined in equation (4.34) of section 4.4.

Let $\mathbf{o}$ be a point on a reflective surface and $\mathbf{b}$ be the normal at $\mathbf{o}$, see Figure 5.10. If $A = \Pi(P)$ denotes the spherical projection of the luminaire P at $\mathbf{o}$, the reflected radiance at $\mathbf{o}$ in the direction $\mathbf{u}'$ due to P is given by

$$f(\mathbf{u}') \;=\; \int_A \rho(\mathbf{u} \to \mathbf{u}') f(\mathbf{u}) \, \langle \mathbf{b}, \mathbf{u} \rangle \; d\sigma(\mathbf{u}), \tag{5.48}$$

where ρ is the BRDF of the reflecting surface, and $f(\mathbf{u})$ is the radiance distribution at $\mathbf{o}$ due to the luminaire P. Suppose that the radiant exitance of the emitter P varies with respect to position according to a monomial function, then $f(\mathbf{u})$ can be characterized by a constrained rational polynomial given in equation (4.3). Now consider a simple BRDF defined in terms of a Phong exponent, that is, a cosine raised to a power m [87],

$$\rho(\mathbf{u} \to \mathbf{u}') \;\equiv\; -c \left[\mathbf{u}'^{\mathrm{T}} \left(\mathbf{I} - 2\mathbf{b}\mathbf{b}^{\mathrm{T}} \right) \mathbf{u} \right]^m \;=\; c \, \langle \mathbf{v}, \mathbf{u} \rangle^m , \tag{5.49}$$

where $\mathbf{v} = -(\mathbf{I} - 2\mathbf{b}\mathbf{b}^{\mathrm{T}})\mathbf{u}'$ and the Householder matrix $\mathbf{I} - 2\mathbf{b}\mathbf{b}^{\mathrm{T}}$ performs a reflection about the tangent plane at $\mathbf{o}$. This BRDF defines a cosine lobe about an axis $\mathbf{v}$ in the direction of mirror reflection of $\mathbf{u}'$, as shown in Figure 5.10. Using equation (5.49) and equation (4.3), it follows from equation (5.48) that the reflected radiance in the direction $\mathbf{u}'$ at $\mathbf{o}$ on this reflective surface due to a monomially-varying luminaire P of nth order may be expressed as a type of rational angular moment introduced in section 4.4. That is,

$$\begin{aligned} f(\mathbf{u}') &= \int_A \frac{\langle \mathbf{v}, \mathbf{u} \rangle^m \langle \mathbf{e}_1, \mathbf{u} \rangle^p \langle \mathbf{e}_2, \mathbf{u} \rangle^q \langle \mathbf{e}_3, \mathbf{u} \rangle^r \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^n} \, d\sigma(\mathbf{u}) \\ &= \tilde{\tau}^{m,p,q,r,1;n}(A, \mathbf{v}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{b}; \mathbf{w}) \end{aligned} \tag{5.50}$$

up to some scalar, where $\mathbf{w}$ is the unit vector from $\mathbf{o}$ that is orthogonal to P. See Figure 5.10.

5.2.2 Glossy Scattering from Linearly-Varying Luminaires

When the radiant exitance of the luminaire P involved in the previous section varies linearly with position, the moment given in equation (5.50) is simplified by $n = 1$ and the fact that only one of p, q, r is nonzero and 1, resulting in a kind of moment described in section 4.4. Therefore, the algorithms described there for their efficient evaluation may be applied to directly simulate view-dependent reflections from or transmissions through glossy surfaces due to a linearly-varying luminaire.

Specifically, letting $\mathbf{a}$ given in equation (3.6) denote the vector encoding the radiant exitance variation across a linearly-varying luminaire P, then the reflected radiance expressed in (5.48) reduces to $\tilde{\tau}^{m,1,1;1}(A, \mathbf{v}, \mathbf{a}, \mathbf{b}; \mathbf{w})$ $(= \tilde{\tau}^{m,1,1;1}(A, \mathbf{v}, \mathbf{b}, \mathbf{a}; \mathbf{w}))$, or equivalently

$$f(\mathbf{u}') = \int_A \frac{\langle \mathbf{v}, \mathbf{u} \rangle^m \langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \, d\sigma(\mathbf{u}). \tag{5.51}$$

The procedures described in section 4.4.4 can thus be directly used to simulate the glossy reflection effect involving linearly-varying emitters and a reflective surface with Phong-

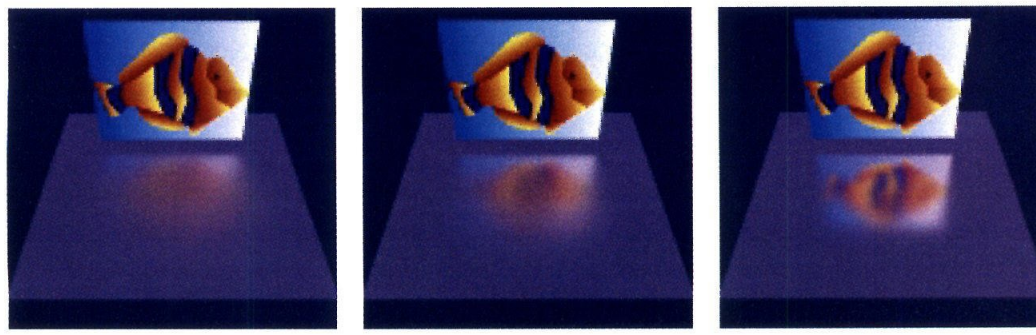


Figure 5.11: *Glossy reflection of a stained glass window with linearly-varying colors, where the Phone exponents are* 15, 50 *and* 300 *from left to right.*

like reflection distribution. The technique is demonstrated in Figure 5.11 using a variety of exponents to simulate surfaces with varying finishes. From left to right, the floor becomes from relatively rough to near specular. In order to efficiently handle two color variations superimposed on the patches of the fish-shaped luminaire shown in Figure 5.11, we separate the vector $\mathbf{a}$ depending on the color variation from equation (4.40) and rewrite $\tilde{\tau}^{n,1,1;1}$ as an inner product of $\mathbf{a}$ and a vector-valued function $\mathbf{t}(A, \mathbf{v}, \mathbf{b}, \mathbf{w})$, that is

$$\tilde{\tau}^{n,1,1;1}(A, \mathbf{v}, \mathbf{a}, \mathbf{b}; \mathbf{w}) \;=\; \langle \mathbf{a}, \mathbf{t}(A, \mathbf{v}, \mathbf{b}, \mathbf{w}) \rangle \,, \tag{5.52}$$

where the vector-valued $\mathbf{t}$ can be derived from (4.40), given by

$$\begin{aligned} \mathbf{t}(A, \mathbf{v}, \mathbf{b}, \mathbf{w}) \;=\; & \mathbf{w}\bar{\tilde{\tau}}^{n,1}(A, \mathbf{v}, \mathbf{b}) + \frac{\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathrm{T}}}{n+2} \left[n\tilde{\tau}^{n-1,1;1}(A, \mathbf{v}, \mathbf{b}; \mathbf{w})\, \mathbf{v} + \tilde{\tau}^{n;1}(A, \mathbf{v}; \mathbf{w})\, \mathbf{b} \right. \\ & \left. - \int_{\partial A} \mathbf{n} \frac{\langle \mathbf{v}, \mathbf{u} \rangle^n \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle}\, ds \right], \end{aligned} \tag{5.53}$$

which may be evaluated efficiently using the similar technique described in section 4.4.4. The procedures exactly parallel to those presented in section 4.4.4. Similar optimization of amortizing the computation cost among several variations was also used in simulating the direct illumination from linearly-varying luminaires in section 5.1.1, where we used (3.117) to compute an irradiance vector $\mathbf{\Phi}(P, \mathbf{w}, \mathbf{b})$ instead of a scalar irradiance value.

In Figure 5.12, our analytical glossy reflection is compared with a Monte Carlo solution based on 64 samples per pixel, where the samples are stratified and distributed according to a cosine (a form of importance sampling) lobe about $\mathbf{v}$. The analytical

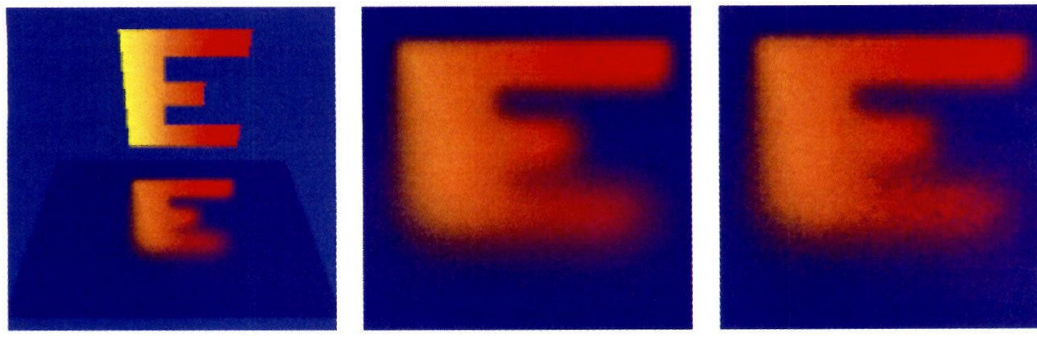


Figure 5.12: *A simple test scene of a glossy surface of order* 200*, the closeup images of the reflection on the floor were computed analytically (middle) and by Monte Carlo using* 64 *samples per pixel (right).*

solution matches very well with Monte Carlo estimate with the advantage of eliminating statistical noise. Moreover, its running time (2 min) in SGI Onyx2 is also slightly faster than Monte Carlo method (2.6 min) for this example.

Nearly the same strategy can be used to compute glossy transmission through a glossy surface from linearly-varying luminaires. By choosing $\mathbf{v}$ in equation (5.53) as the reversed view direction $-\mathbf{u}'$, which is now located on the other side of the transparent material, $\rho(\mathbf{u} \rightarrow \mathbf{u}')$ defined in equation (5.49) can be interpreted as a phong-like BTDF. Consequently, the transmitted radiance in the direction $\mathbf{u}'$ at a surface point due to a polygonal luminaire P is represented by the same integral shown in equation (5.51). Figure 5.13 shows three images depicting a frosted glass fish tank, with different finishes corresponding to different Phong exponents. This effect due to Lambertian luminaires was demonstrated by Wallace et al. [116] and Arvo [9].

5.2.3 More Complex BRDF

In this section we shall discuss the simulation of scattering effects from non-diffuse surfaces with more complex BRDFs. The mathematical formulation using tensors and mo-

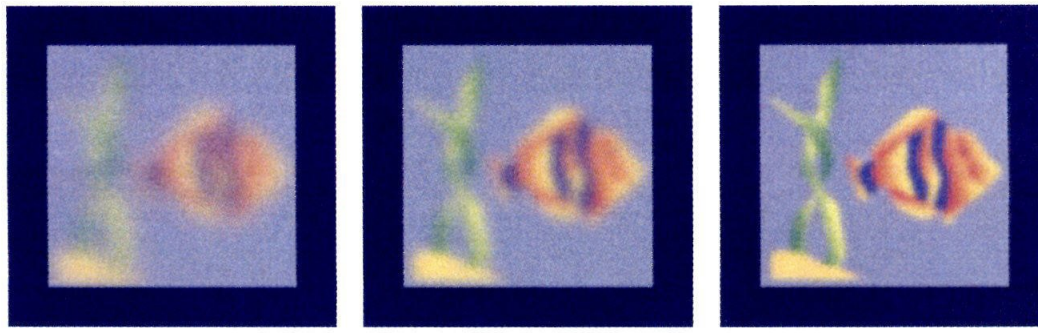


Figure 5.13: *Glossy transmission through a frosted glass fish tank, where the tropical fish and sea weeds are superimposed with linearly-varying colors. From left to right, the Phong exponents are* 5, 15, *and* 65, *respectively.*

ments discussed in section 5.2.1 still applies, as long as the BRDF may be approximated by polynomials. Examples include BRDFs obtained from superposing cosine lobes along different axes and of different orders, and *Lafortune's Model* [66], etc.

Lafortune's model is a generalization of the cosine lobe model shown in equation (5.49), which is defined by

$$\begin{aligned} \rho(\mathbf{u} \to \mathbf{u}') &\equiv c\left[\mathbf{u}^{\mathrm{T}} M \mathbf{u}'\right]^n \\ &= c\left[\mathbf{u}^{\mathrm{T}} Q^{\mathrm{T}} D Q \mathbf{u}'\right]^n, \end{aligned} \tag{5.54}$$

where M is a symmetric 3×3 matrix and $M = Q^{\mathrm{T}}DQ$ is a singular value decomposition of M. The vectors of Q specify a new coordinate system on the surface, which is aligned with the surface and principal directions of anisotropy. Then, by expressing the coordinates of $\mathbf{u}$ and $\mathbf{u}'$ in this new coordinate system, we get the general form of $\rho(\mathbf{u} \to \mathbf{u}')$, given by

$$\rho(\mathbf{u} \to \mathbf{u}') = c[C_x u_x u'_x + C_y u_y u'_y + C_z u_z u'_z]^n,$$

which provides a non-linear approximation to ρ with respect to the coefficients C_x, C_y, C_z, determined from measured data using nonlinear optimization techniques. Anisotropic reflection and specular reflection near grazing can be simulated by allowing c and n in the

definition (5.54) of ρ to vary with the incident direction $\mathbf{u}$. It follows from equation (5.54) that we may still express such a nonlinear BRDF as a cosine raised to a power, that is

$$\rho(\mathbf{u} \to \mathbf{u}') \;=\; c\,\langle \mathbf{v}, \mathbf{u} \rangle^n ,$$

where $\mathbf{v} = M\mathbf{u}'$. Consequently, we may use the same procedures in the previous sections to simulate some scattering effects involving surfaces with such a BRDF and spatially-varying luminaires.

In fact, a more interesting topic may be to model more complex and realistic BRDFs using the class of constrained rational polynomials that can be represented by our generalized irradiance tensors. With such a BRDF, the reflected radiance from non-diffuse surfaces due to a Lambertian luminaire can still be formulated in terms of rational moments, which may be amenable to closed-form solutions. In approximating these BRDFs, it would be interesting that they manifest some of the reflection properties listed below:

1. Non-Lambertian diffuse reflection: the diffuse component fades out near grazing angles.

2. Increased specular reflection near grazing

3. Off-specular reflection

4. Anisotropic reflection.

In the meanwhile, the modeled function should observe the inherent properties of BRDF, that is

- energy preserving
- reciprocity

The ideal case is that the functions can fit into our rational moment definitions and are amenable to closed-form solutions. However, a big challenge arising here is that we only allow one factor to appear in the denominator of approximating rational polynomials.

5.3 Directional Luminaires

The idea of using generalized irradiance tensors and rational moments for non-Lambertian simulations of spatially-varying luminaires may be similarly applied to the general case of non-uniform luminaires with emission distributions that vary polynomially in positions as well as in directions.

Let P be a polygonal luminaire whose emission distribution is spatially uniform but varies directionally according to a Phong distribution, i.e. as a cosine to a power l, where the direction of maximum radiance $\mathbf{w}$ may be normal to the planar luminaire, falling off rapidly in other directions, as shown in Figure 1.1a. The irradiance at the origin $\mathbf{o}$ with the normal $\mathbf{b}$ can be computed by

$$\begin{aligned} \Phi &= \int_P \langle \mathbf{w}, \mathbf{u} \rangle^l \langle \mathbf{b}, \mathbf{u} \rangle \, d\sigma(\mathbf{u}) \\ &= \tilde{\tau}^{l,1;0}(P, \mathbf{w}, \mathbf{b}) \\ &= \bar{\tilde{\tau}}^{n,1}(P, \mathbf{w}, \mathbf{v}), \end{aligned} \tag{5.55}$$

where the corresponding rational moment is reduced to a double-axis moment, as was already demonstrated by Arvo [9].

Furthermore, if the radiant exitance of P also varies in position according to a nth order monomial function, then the radiance distribution at $\mathbf{o}$ due to this luminaire becomes

$$f(\mathbf{u}) = \frac{\langle \mathbf{e}_1, \mathbf{u} \rangle^p \langle \mathbf{e}_2, \mathbf{u} \rangle^q \langle \mathbf{e}_3, \mathbf{u} \rangle^r \langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^n} \langle \mathbf{w}, \mathbf{u} \rangle^l .$$

Consequently, the reflected radiance at $\mathbf{o}$ on a glossy surface with a simple BRDF defined in equation (5.49) due to P is still given by a moment that follows from equation (5.48). In particular, depending on the relation between the monomial order n and Phong exponent

l regarding the directional variation, we have

$$f(\mathbf{u}') = \begin{cases} \tilde{\tau}^{m,p,q,r,1;n-l}(A, \mathbf{v}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{b}; \mathbf{w}) & n > l \\ \tilde{\tau}^{m,p,q,r,1;0}(A, \mathbf{v}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{b}) & n = l \\ \tilde{\tau}^{m,p,q,r,l-n,1;0}(A, \mathbf{v}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{w}, \mathbf{b}) & n < l \end{cases} \tag{5.56}$$

Chapter 6

Conclusion and Future Work

6.1 Summary

Chapter 2 presented an analytical perturbation formula for perturbing a general ray-traced path involving both specular reflections and refractions. The formula is based on closed-form expressions for path Jacobians and path Hessians. In addition, as a continuation of our previous work [27], we simplified the original formula for path Jacobians, generalized the path perturbation theory to parametric surfaces, and made some insightful comparisons with two similar ideas. In addition, the virtual object algorithm is extended to refraction for lens simulation. Some potential application areas of our path perturbation framework were discussed in the end.

Chapter 3 presented a closed-form solution for the irradiance at a point due to a linearly-varying luminaire. The expression was derived using Taylor expansion and a generalization of double-axis moments and then verified by Stokes' theorem. Although the whole procedure is quite tedious and out-of-dated, the work serves as a motivation to the theory developed in the chapter 4.

Chapter 4 presented new mathematical tools for computing direct illumination and scattering involving polynomially-varying luminaires. The building block of this chapter is a generalization of irradiance tensors called *generalized irradiance tensors*, which were shown to satisfy several recurrence relations. These tensors depict the emission features

of polynomially-varying luminaires and are well suited to represent a class of constrained rational polynomials integrated over the sphere, called *rational angular moments.* Particularly, we derived a number of closed-form expressions for integrating a subclass of these rational polynomials that correspond to Phong distributions and presented complete algorithms for their efficient evaluation. These special moments have direct applications in image synthesis, and the algorithms were verified by comparison with Monte Carlo simulation and Finite element method in chapter 5.

Chapter 5 demonstrated several applications of new expressions derived for generalized irradiance tensors and rational moments. First of all, we presented some theoretical results about the irradiance at a point due to a polynomially-varying luminaire. With a new formulation using the concept of generalized irradiance tensors, we not only rederived the same closed-form solution for the irradiance due to a linearly-varying luminaire in a more elegant way than that described in chapter 3, but also found a general recursive formula for the irradiance at a point due to a luminaire whose radiant exitance varies with respect to the position according to a monomially function of any order. The recurrence relation relates the solutions for higher-order polynomials to those for lower-order cases, subsuming Lambert's formula as its base case. Furthermore, we have reasoned that there exist a closed-form solution for the irradiance at a point due to a Lambertian polygonal luminaire with polynomially-varying radiant exitance, depending on a single well-behaved special function known as the *Clausen integral.* We then illustrated other applications of generalized irradiance tensors, including glossy reflections and transmissions on non-diffuse surfaces due to spatially-varying luminaires as well as directionally-varying luminaires.

6.2 Future Work

The algorithm for simulating lens effects presented in chapter 2 should be enhanced with a better mechanism to identify and split multiple refraction images of a single scene point. This problem also exists in the application of interactive specular reflection, but becomes more severe in the case of refraction. Interval analysis may provide us a

general technique to find all the reflection or refraction points corresponding to an object point. Also, path perturbation theory may find various potential applications, such as stereoscopic ray tracing, image-based rendering of non-diffuse environments, caustics acceleration, variance reduction in Metropolis light transport, as well as shape recovery in computer vision. The details have been presented in section 2.7.

There are a number of extensions to the work presented in chapter 4. The most important one is to make those new expressions more practical. According to section 4.3.2, all the rational moments admit closed-form expressions, except for a single special function known as the Clausen integral. However, it is still necessary to look for practical algorithms to efficiently evaluate a general rational moment, or at least those special moments corresponding to Phong distributions with a denominator order greater than one.

Motivated by the problem of computing the irradiance due to a polynomially-varying luminaire, we extracted a form of irradiance tensor given in (4.4) to accommodate a class of constrained rational polynomials over the sphere arising in the irradiance problem. Another important extension is to look into other possible tensor formulations to account for more general rational polynomial functions, or even other forms of basis functions over the sphere, which may lead to other non-Lambertian simulations.

The analysis of the irradiance problem in chapter 5 may be extended in several ways. Given the expression for area-to-point form factor, we may integrate over an area to compute area-to-area form factor involving a spatially-varying polygon; we should also explore the occluded irradiance from spatially-varying emitters. Taking this a bit further, we may consider computing indirect illumination scattered from one or multiple reflectors. Since the irradiance across a Lambertian reflector can be represented as a polyhedral spline [107], then the problem of computing the irradiance on a diffuse surface scattered from this reflector reduces to the computation of direct illumination from a luminaire with spline-like spatial variations. Unfortunately, since polyhedral splines are not polynomials, our machinery cannot be directly applied. We may even include some kinds of BRDF at intermediate surfaces. All these studies will boil down to expressing them in terms of more general rational polynomials, where the denominators may become

more complex at some point (i.e. more than one factor). This remains to be seen.

Aside from these directions, we may extend it from the perspective of differentiation. As learned from path perturbation, we have seen that derivatives always show important information of the function we are studying. With the irradiance formula derived for the irradiance due to inhomogeneous emitters, we may consider differentiating it with respect to the receiver point or other parameters, as the work of *irradiance Jacobian* done by Arvo [7], or that of *irradiance gradient* by Ward and Heckbert [117]. Consequently, we may approach the problem of computing direct illumination due to an inhomogeneous emitter from a different view, that is, formulating it as an ODE instead of an integration problem. Moreover, the irradiance derivatives may be of use in such applications as interpolation, mesh construction, and so on.

Appendix A

Additional Proofs

A.1 Solving the Linear System (2.50)

In this appendix, we show how to compute N satisfying equation (2.50) in section 2.3. Obviously, equation (2.50) is equivalent to the following linear system:

$$\begin{aligned}
\hat{\Gamma}N + \mathbf{h}\mathbf{v}^{\mathrm{T}} &= 1 \quad (1)\\
\hat{\Gamma}\mathbf{v} + d\mathbf{h} &= 0 \quad (2)\\
\mathbf{h}^{\mathrm{T}}N &= 0 \quad (3)\\
\mathbf{h}^{\mathrm{T}}\mathbf{v} &= 1 \quad (4)
\end{aligned}$$

Multiplying both sides of (1) by $\hat{\Gamma}$, we have

$$\hat{\Gamma}N\hat{\Gamma} + \mathbf{h}\mathbf{v}^{\mathrm{T}}\hat{\Gamma} = \hat{\Gamma}.$$

Using the fact that $\hat{\Gamma}$ is symmetric, we get

$$\hat{\Gamma}N\hat{\Gamma} + \mathbf{h}(\hat{\Gamma}\mathbf{v})^{\mathrm{T}} = \hat{\Gamma}. \tag{A.1}$$

Replacing $\hat{\Gamma}\mathbf{v} = -d\mathbf{h}$ from (2), equation (A.1) can be written as

$$\hat{\Gamma}N\hat{\Gamma} - d\mathbf{h}\mathbf{h}^{\mathrm{T}} = \hat{\Gamma}.$$

If $\hat{\Gamma}$ is non-singular, we can solve for N, attaining

$$N = \hat{\Gamma}^{-1}\left(\mathbf{I} + d\mathbf{h}\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}\right), \tag{A.2}$$

where d above can be computed from (2) and (4). In detail, multiplying $\hat{\Gamma}^{-1}$ and then $\mathbf{h}^{\mathrm{T}}$ on both sides of (2), we get

$$\mathbf{h}^{\mathrm{T}}\mathbf{v} + d\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}\mathbf{h} = 0.$$

Using (4), we can solve d as

$$d = -\frac{1}{\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}\mathbf{h}}.$$

Plugging into equation (A.2), we get

$$N = \hat{\Gamma}^{-1}\left(\mathbf{I} - \frac{\mathbf{h}\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}}{\mathbf{h}^{\mathrm{T}}\hat{\Gamma}^{-1}\mathbf{h}}\right). \tag{A.3}$$

It can be easily verified that the expression (A.3) for N satisfies (3).

A.2 Proof of Identity (3.41)

In this appendix we supply the proof of the identity (3.41).

Proof: The proof is done by using Stokes' Theorem, which can be stated in three steps:

Step 1: By using the fact (3.90), we may rewrite the boundary integral on the left of equation (3.41) in terms of the position vector $\mathbf{r}$ and its derivatives, obtaining

$$\begin{aligned}\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{1 + \langle \mathbf{w}, \mathbf{u} \rangle}\, ds &= \int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{r} \times d\mathbf{r} \rangle}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)\, r} \\ &= \int_{\partial A} \frac{\varepsilon_{jpl} \mathbf{w}_{\mathrm{j}} \mathbf{r}_{\mathrm{p}}\, d\mathbf{r}_{\mathrm{l}}}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)\, r} \\ &= \int_{\partial A} B_{\mathrm{l}}\, d\mathbf{r}_{\mathrm{l}}, \end{aligned} \tag{A.4}$$

where

$$B_{\mathrm{l}} = \frac{\varepsilon_{jpl} \mathbf{w}_{\mathrm{j}} \mathbf{r}_{\mathrm{p}}}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)\, r}.$$

Step 2: To convert the resulting boundary integral into a surface integral to compare with the right hand side of equation (3.41), we need to compute the partial derivatives of B_{l}. By using the fact that

$$\frac{\partial}{\partial \mathbf{r}_{\mathrm{m}}} \left(\frac{1}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)\, r} \right) = -\frac{2r\mathbf{r}_{\mathrm{m}} + r^2 \mathbf{w}_{\mathrm{m}} + \langle \mathbf{w}, \mathbf{r} \rangle\, \mathbf{r}_{\mathrm{m}}}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)^2\, r^3},$$

the partial derivative $B_{l,m}$ can be computed as

$$\begin{aligned} B_{l,m} &= \varepsilon_{jpl} \mathbf{w}_{\mathrm{j}} \left(\frac{\delta_{pm}}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)\, r} - \frac{2r\mathbf{r}_{\mathrm{p}}\mathbf{r}_{\mathrm{m}} + r^2 \mathbf{r}_{\mathrm{p}} \mathbf{w}_{\mathrm{m}} + \langle \mathbf{w}, \mathbf{r} \rangle\, \mathbf{r}_{\mathrm{p}}\mathbf{r}_{\mathrm{m}}}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)^2\, r^3} \right) \\ &= \varepsilon_{jpl} \mathbf{w}_{\mathrm{j}} \left(\frac{\delta_{pm}\, (r + \langle \mathbf{w}, \mathbf{r} \rangle)\, r^2 - 2r\mathbf{r}_{\mathrm{p}}\mathbf{r}_{\mathrm{m}} - r^2 \mathbf{r}_{\mathrm{p}} \mathbf{w}_{\mathrm{m}} - \langle \mathbf{w}, \mathbf{r} \rangle\, \mathbf{r}_{\mathrm{p}}\mathbf{r}_{\mathrm{m}}}{(r + \langle \mathbf{w}, \mathbf{r} \rangle)^2\, r^3} \right) \end{aligned}$$

We then apply Stokes' theorem to equation (A.4), yielding

$$\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{1 + \langle \mathbf{w}, \mathbf{u} \rangle}\, ds = \int_A B_{l,m}\, d\mathbf{r}_{\mathrm{m}} \wedge d\mathbf{r}_{\mathrm{l}}$$

$$= \int_A \varepsilon_{kml} B_{l,m} \left[\frac{\varepsilon_{kst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2} \right], \tag{A.5}$$

where the last step follows from the transformation (3.26).

Step 3: Using the definition (3.27), we may express the surface integral (A.5) in terms of solid angle. First, we simplify $\varepsilon_{kml} B_{l,m}$ as follows:

$$\begin{aligned}
\varepsilon_{kml} B_{l,m} &= \varepsilon_{kml}\varepsilon_{jpl}\mathbf{w}_j \left(\frac{\delta_{pm}\left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right) r^2 - 2r\mathbf{r}_p\mathbf{r}_m - r^2\mathbf{r}_p\mathbf{w}_m - \langle \mathbf{w}, \mathbf{r} \rangle \mathbf{r}_p\mathbf{r}_m}{\left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right)^2 r^3} \right) \\
&= \frac{\delta_{pm}\delta_{jk} - \delta_{pk}\delta_{jm}}{\left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right)^2 r^3} \mathbf{w}_j \left(\delta_{pm}\left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right) r^2 - 2r\mathbf{r}_p\mathbf{r}_m - r^2\mathbf{r}_p\mathbf{w}_m - \langle \mathbf{w}, \mathbf{r} \rangle \mathbf{r}_p\mathbf{r}_m \right) \\
&= \frac{1}{\left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right)^2 r^3} \Big[3\mathbf{w}_k \left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right) r^2 - 2r^3\mathbf{w}_k - r^2 \langle \mathbf{w}, \mathbf{r} \rangle \mathbf{w}_k - r^2 \langle \mathbf{w}, \mathbf{r} \rangle \mathbf{w}_k \\
&\quad - \mathbf{w}_k \left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right) r^2 + 2r \langle \mathbf{w}, \mathbf{r} \rangle \mathbf{r}_k + r^2 \langle \mathbf{w}, \mathbf{w} \rangle \mathbf{r}_k + \langle \mathbf{w}, \mathbf{r} \rangle^2 \mathbf{r}_k \Big] \\
&= \frac{\mathbf{r}_k}{\left(r + \langle \mathbf{w}, \mathbf{r} \rangle\right)^2 r^3} \left[2r \langle \mathbf{w}, \mathbf{r} \rangle + r^2 + \langle \mathbf{w}, \mathbf{r} \rangle^2 \right] \\
&= \frac{\mathbf{r}_k}{r^3}.
\end{aligned}$$

Consequently, equation (A.5) becomes

$$\begin{aligned}
-\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{1 + \langle \mathbf{w}, \mathbf{u} \rangle}\, ds &= -\int_A \frac{\varepsilon_{kst}\mathbf{r}_k\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2r^3} \\
&= \int_A d\omega \\
&= \sigma(A),
\end{aligned}$$

which proves our identity (3.41). □□

A.3 Proof of Binomial Identities (3.63) to (3.66)

In this appendix we prove the four binomial identities (3.63), (3.64), (3.65) and (3.66) shown in section 3.3.3, which all involve binomial coefficients $\binom{n}{k}$. Our approach is to use *generating function.*

First, we prove identities (3.63) and (3.65), where each summation term has a denominator of $(k+1)$.

Proof: Integrating both sides of

$$\sum_{k=0}^{n}(-1)^k \binom{n}{k} x^k = (1-x)^n$$

with respect to x, we have

$$\sum_{k=0}^{n}(-1)^k \binom{n}{k} \frac{x^{k+1}}{k+1} = c(n) - \frac{(1-x)^{n+1}}{n+1}.$$

where $c(n)$ is a function only dependent on n. Letting $x = 0$, we get

$$c(n) = \frac{1}{n+1}.$$

Thus, we get equation (3.63), that is

$$\sum_{k=0}^{n}(-1)^k \binom{n}{k} \frac{x^k}{k+1} = \frac{1-(1-x)^{n+1}}{(n+1)x}.$$

By letting $x = 1$ above, we get another identity (3.65). □□

A similar strategy applies to identities (3.64) and (3.66) whose summation terms have a denominator of $(k+3)$.

Proof: Integrating both sides of

$$\sum_{k=0}^{n}(-1)^k \binom{n}{k} x^{k+2} = (1-x)^n x^2$$

with respect to the variable x, we get

$$\begin{aligned}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^{k+3}}{k+3} &= \int(1-x)^n x^2 dx + c(n)\\ &= \int(1-x)^n(1-x-1)^2 dx + c(n)\\ &= \int\left[(1-x)^{n+2} - 2(1-x)^{n+1} + (1-x)^n\right]dx + c(n)\\ &= -\frac{(1-x)^{n+3}}{n+3} + \frac{2(1-x)^{n+2}}{n+2} - \frac{(1-x)^{n+1}}{n+1} + c(n).\end{aligned}$$

where $c(n)$ is a function only dependent on n. Letting $x = 0$, we get

$$c(n) = \frac{1}{n+3} - \frac{2}{n+2} + \frac{1}{n+1} = \frac{2}{(n+1)(n+2)(n+3)}.$$

Thus, we have

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^k}{k+3} = \frac{1}{x^3}\left(\frac{2}{(n+1)(n+2)(n+3)} - \frac{(1-x)^{n+3}}{n+3} + \frac{2(1-x)^{n+2}}{n+2} - \frac{(1-x)^{n+1}}{n+1}\right),$$

which is identity (3.64). By letting $x = 1$, we can get another identity (3.66). □□

A.4 Proof of Identity (3.75)

In this appendix we prove the identity (3.75) used in simplifying T_2 in section 3.3.4.

Proof: Since both sides of equation (3.75) are boundary integrals, we prove this identity by converting them into surface integrals and comparing the resulting integrands. Specifically, by applying Stokes' theorem, we can change the left hand side as follows:

$$
\begin{aligned}
\int_{\partial A} \frac{\langle \mathbf{v}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2}\, ds &= \int_{\partial A} \frac{\langle \mathbf{v}, \mathbf{r} \times d\mathbf{r} \rangle}{\langle \mathbf{w}, \mathbf{r} \rangle^2} \quad \text{(Eq.(3.90))} \\
&= \int_{\partial A} \frac{\varepsilon_{jpl} \mathbf{v}_{\mathrm{j}} \mathbf{r}_{\mathrm{p}}\, d\mathbf{r}_{\mathrm{l}}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} \\
&= \int_A \varepsilon_{jpl} \mathbf{v}_{\mathrm{j}} \left(\frac{\partial}{\partial \mathbf{r}_{\mathrm{m}}} \left[\frac{\mathbf{r}_{\mathrm{p}}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} \right] \right) d\mathbf{r}_{\mathrm{m}} \wedge d\mathbf{r}_{\mathrm{l}} \quad (Stokes' theorem) \\
&= \int_A \varepsilon_{jpl} \mathbf{v}_{\mathrm{j}} \left[\frac{\delta_{pm}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} - \frac{2\mathbf{w}_{\mathrm{m}} \mathbf{r}_{\mathrm{p}}}{\langle \mathbf{w}, \mathbf{r} \rangle^3} \right] d\mathbf{r}_{\mathrm{m}} \wedge d\mathbf{r}_{\mathrm{l}} \\
&= \int_A \varepsilon_{kml} \varepsilon_{jpl} \mathbf{v}_{\mathrm{j}} \left[\frac{\delta_{pm}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} - \frac{2\mathbf{w}_{\mathrm{m}} \mathbf{r}_{\mathrm{p}}}{\langle \mathbf{w}, \mathbf{r} \rangle^3} \right] \left[\frac{\varepsilon_{kst}\, d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}}}{2} \right] \quad \text{(Eq.(3.26))} \\
&= \int_A \left(\delta_{pm} \delta_{jk} - \delta_{pk} \delta_{jm} \right) \mathbf{v}_{\mathrm{j}} \left[\frac{\delta_{pm}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} - \frac{2\mathbf{w}_{\mathrm{m}} \mathbf{r}_{\mathrm{p}}}{\langle \mathbf{w}, \mathbf{r} \rangle^3} \right] \left[\frac{\varepsilon_{kst}\, d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}}}{2} \right] \quad \text{(Eq.(3.18))} \\
&= \int_A \left[\frac{3\mathbf{v}_{\mathrm{k}}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} - \frac{2 \langle \mathbf{w}, \mathbf{r} \rangle \mathbf{v}_{\mathrm{k}}}{\langle \mathbf{w}, \mathbf{r} \rangle^3} - \frac{\mathbf{v}_{\mathrm{k}}}{\langle \mathbf{w}, \mathbf{r} \rangle^2} + \frac{2\mathbf{r}_{\mathrm{k}} \langle \mathbf{w}, \mathbf{v} \rangle}{\langle \mathbf{w}, \mathbf{r} \rangle^3} \right] \left[\frac{\varepsilon_{kst}\, d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}}}{2} \right] \\
&= \langle \mathbf{w}, \mathbf{v} \rangle \int_A \frac{2\mathbf{r}_{\mathrm{k}}}{\langle \mathbf{w}, \mathbf{r} \rangle^3} \left[\frac{\varepsilon_{kst}\, d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}}}{2} \right], \qquad \text{(A.6)}
\end{aligned}
$$

where we have labeled the crucial steps with the transformations used. Since equation (A.6) holds for all vectors $\mathbf{w}$ and $\mathbf{v}$. Letting $\mathbf{v} = \mathbf{w}$ in (A.6) and using $\langle \mathbf{w}, \mathbf{w} \rangle = 1$ for unit vector $\mathbf{w}$, we get

$$
\int_{\partial A} \frac{\langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^2}\, ds = \int_A \frac{2\mathbf{r}_{\mathrm{k}}}{\langle \mathbf{w}, \mathbf{r} \rangle^3} \left[\frac{\varepsilon_{kst}\, d\mathbf{r}_{\mathrm{s}} \wedge d\mathbf{r}_{\mathrm{t}}}{2} \right]. \qquad \text{(A.7)}
$$

Multiplying equation (A.7) by $\langle \mathbf{w}, \mathbf{v} \rangle$ and comparing the result with the right hand side of equation (A.6), we have proved the identity (3.75). □□

A.5 Proof of Identity (3.79)

In this appendix we supply the proof for another identity (3.79) used in section 3.3.4. We first state and prove a useful lemma.

Lemma 1 *Suppose that* $\mathbf{u}$ *is the unit vector in the same direction as the position vector* $\mathbf{r}$ *of a point with respect to the origin and* $r = ||\,\mathbf{r}\,||$. *Let*

$$\xi_n(\mathbf{u}) \equiv \frac{\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^n}, \tag{A.8}$$

then the partial derivative of ξ_n *with respect to the* m*th coordinate* $\mathbf{r}_m$, *denoted by* $\xi_{n,m}$, *is given by*

$$\xi_{n,m} = \frac{1 - n\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^{n+1} r^2}\left[r\mathbf{w}_m - \langle \mathbf{w}, \mathbf{u}\rangle \mathbf{r}_m\right] \tag{A.9}$$

for any integer $n \geq 0$.

Proof: The proof is done by induction. When $n = 0$, $\xi_0 = \ln\langle \mathbf{w}, \mathbf{u}\rangle$, we have

$$\begin{aligned}\frac{\partial \ln\langle \mathbf{w}, \mathbf{u}\rangle}{\partial \mathbf{r}_m} &= \frac{\mathbf{w}_m}{\langle \mathbf{w}, \mathbf{r}\rangle} - \frac{\mathbf{r}_m}{r^2} \\ &= \frac{1}{\langle \mathbf{w}, \mathbf{u}\rangle r^2}\left[r\mathbf{w}_m - \langle \mathbf{w}, \mathbf{u}\rangle \mathbf{r}_m\right].\end{aligned}$$

Thus, equation (A.9) is true for $n = 0$. Suppose that equation (A.9) holds for $k-1$, that is,

$$\xi_{k-1,m} = \frac{1 - (k-1)\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^{k} r^2}\left[r\mathbf{w}_m - \langle \mathbf{w}, \mathbf{u}\rangle \mathbf{r}_m\right]. \tag{A.10}$$

Then the partial derivative of ξ_k can be computed as:

$$\begin{aligned}\xi_{k,m} &= \frac{\partial}{\partial \mathbf{r}_m}\left[\xi_{k-1}\frac{1}{\langle \mathbf{w}, \mathbf{u}\rangle}\right] \\ &= \xi_{k-1,m}\frac{1}{\langle \mathbf{w}, \mathbf{u}\rangle} + \xi_{k-1}\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{r}{\langle \mathbf{w}, \mathbf{r}\rangle}\right] \\ &= \frac{1 - (k-1)\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^{k+1} r^2}\left[r\mathbf{w}_m - \langle \mathbf{w}, \mathbf{u}\rangle \mathbf{r}_m\right] + \frac{\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^{k-1}} \cdot \frac{\mathbf{r}_m\langle \mathbf{w}, \mathbf{u}\rangle - r\mathbf{w}_m}{\langle \mathbf{w}, \mathbf{u}\rangle^2 r^2}\end{aligned}$$

$$= \frac{1 - k \ln \langle \mathbf{w}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle^{k+1} r^2} \left[r \mathbf{w}_m - \langle \mathbf{w}, \mathbf{u} \rangle \mathbf{r}_m \right],$$

which has shown that equation (A.9) also holds for $n = k$ if it is true for $n = k - 1$. By induction, we can claim that equation (A.9) is true for all integer $n \geq 0$. □□

Now we are ready to prove the identity (3.79), where we will use lemma 1.

Proof: To prove an identity about two boundary integrals, we convert them into their corresponding surface integrals, as we did in Appendix A.4. Respectively,

Step 1: (The left hand side)

First, we break it into two parts:

$$\begin{aligned} \int_{\partial A} \left[\langle \mathbf{a}, \mathbf{n} \rangle - \frac{\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} \right] ds &= \int_{\partial A} \langle \mathbf{a}, \mathbf{n} \rangle \, ds - \int_{\partial A} \frac{\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} ds \\ &= A_1 - A_2, \end{aligned} \tag{A.11}$$

where

$$A_1 = \int_{\partial A} \langle \mathbf{a}, \mathbf{n} \rangle \, ds, \qquad A_2 = \int_{\partial A} \frac{\langle \mathbf{a}, \mathbf{u} \rangle \langle \mathbf{w}, \mathbf{n} \rangle}{\langle \mathbf{w}, \mathbf{u} \rangle} ds.$$

Next, we apply Stokes' theorem to the two terms A_1 and A_2, and express the resulting surface integrals in terms of $d\omega$ using equation (3.27). For this purpose, the transformation (3.26) is used to complete the 2-form corresponding to $d\omega$, and the tensor identity (3.18) is used for simplification. This step is demonstrated below:

$$\begin{aligned} A_1 &= \int_A \varepsilon_{kpl} \mathbf{a}_k \left[\frac{\mathbf{r}_p}{r^2} \right] d\mathbf{r}_l \\ &= \int_A \varepsilon_{qml} \varepsilon_{kpl} \mathbf{a}_k \frac{\partial}{\partial \mathbf{r}_m} \left[\frac{\mathbf{r}_p}{r^2} \right] \left[\frac{\varepsilon_{qst} \, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2} \right] \\ &= \int_A \left[\mathbf{a}_q \delta_{pm} - \mathbf{a}_m \delta_{pq} \right] \left[\frac{\delta_{pm}}{r^2} - \frac{2 \mathbf{r}_p \mathbf{r}_m}{r^4} \right] \left[\frac{\varepsilon_{qst} \, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2} \right] \\ &= \int_A \left[-2 \langle \mathbf{a}, \mathbf{u} \rangle \right] \left[-\frac{\mathbf{r}_q}{r^3} \right] \left[\frac{\varepsilon_{qst} \, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2} \right] \end{aligned} \tag{A.12}$$

$$= -\int_A 2\langle \mathbf{a}, \mathbf{u}\rangle \, d\omega, \tag{A.13}$$

$$\begin{aligned} A_2 &= \int_{\partial A} \varepsilon_{kpl}\mathbf{w}_k \left[\frac{\langle \mathbf{a}, \mathbf{r}\rangle}{\langle \mathbf{w}, \mathbf{r}\rangle}\right]\left[\frac{\mathbf{r}_p}{r^2}\right] d\mathbf{r}_l \\ &= \int_A \varepsilon_{qml}\varepsilon_{kpl}\mathbf{w}_k \frac{\partial}{\partial \mathbf{r}_m}\left(\frac{\langle \mathbf{a}, \mathbf{r}\rangle}{\langle \mathbf{w}, \mathbf{r}\rangle}\left[\frac{\mathbf{r}_p}{r^2}\right]\right)\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right] \\ &= \int_A \left[\mathbf{w}_q\delta_{pm} - \mathbf{w}_m\delta_{pq}\right]\left[\left(\frac{\mathbf{a}_m}{\langle \mathbf{w}, \mathbf{r}\rangle} - \frac{\mathbf{w}_m\langle \mathbf{a}, \mathbf{r}\rangle}{\langle \mathbf{w}, \mathbf{r}\rangle^2}\right)\frac{\mathbf{r}_p}{r^2} + \frac{\langle \mathbf{a}, \mathbf{r}\rangle}{\langle \mathbf{w}, \mathbf{r}\rangle}\left(\frac{\delta_{pm}}{r^2} - \frac{2\mathbf{r}_p\mathbf{r}_m}{r^4}\right)\right] \times \\ &\qquad \left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right] \\ &= \int_A \left[\frac{\langle \mathbf{a}, \mathbf{w}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle} - \frac{\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2} - 2\langle \mathbf{a}, \mathbf{u}\rangle\right]\left[-\frac{\mathbf{r}_q}{r^3}\right]\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right] \\ &= \int_A \left[\frac{\langle \mathbf{a}, \mathbf{w}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle} - \frac{\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2} - 2\langle \mathbf{a}, \mathbf{u}\rangle\right] d\omega. \end{aligned} \tag{A.14}$$

Combining equations (A.11) and (A.13), (A.14), we have

$$\int_{\partial A}\left[\langle \mathbf{a}, \mathbf{n}\rangle - \frac{\langle \mathbf{a}, \mathbf{u}\rangle\langle \mathbf{w}, \mathbf{n}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle}\right] ds = \int_A \left[\frac{\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2} - \frac{\langle \mathbf{a}, \mathbf{w}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle}\right] d\omega. \tag{A.15}$$

Step 2: (The right hand side)

In terms of ξ_n defined in (A.8), we rewrite the integral on the right as

$$\begin{aligned} \int_{\partial A} \xi_2\left[\langle \mathbf{a}, \mathbf{w}\rangle\langle \mathbf{w}, \mathbf{n}\rangle - \langle \mathbf{a}, \mathbf{n}\rangle\right] ds &= \langle \mathbf{a}, \mathbf{w}\rangle \int_{\partial A} \xi_2\langle \mathbf{w}, \mathbf{n}\rangle \, ds - \int_{\partial A} \xi_2\langle \mathbf{a}, \mathbf{n}\rangle \, ds \\ &= \langle \mathbf{a}, \mathbf{w}\rangle B_1 - B_2, \end{aligned} \tag{A.16}$$

where

$$B_1 = \int_{\partial A} \xi_2\langle \mathbf{w}, \mathbf{n}\rangle \, ds, \qquad B_2 = \int_{\partial A} \xi_2\langle \mathbf{a}, \mathbf{n}\rangle \, ds.$$

Notice that B_1 is just a special case of B_2 with $\mathbf{a} = \mathbf{w}$, we only need to derive the surface integral formula for B_2 . It follows from lemma 1 that

$$\xi_{2,m} = \frac{1 - 2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^3 r^2}\left[r\mathbf{w}_m - \langle \mathbf{w}, \mathbf{u}\rangle \mathbf{r}_m\right].$$

Therefore, parallel to the steps used in equations (A.13) and (A.14), we have

$$\begin{aligned}
B_2 &= \int_{\partial A} \xi_2 \langle \mathbf{a}, \mathbf{n}\rangle \, ds \\
&= \int_{\partial A} \varepsilon_{kpl}\mathbf{a}_k \left[\xi_2 \frac{\mathbf{r}_p}{r^2}\right] d\mathbf{r}_l \\
&= \int_A \varepsilon_{qml}\varepsilon_{kpl}\mathbf{a}_k \left[\frac{1 - 2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^3 r^2}\left(r\mathbf{w}_m - \mathbf{r}_m\langle \mathbf{w}, \mathbf{u}\rangle\right)\frac{\mathbf{r}_p}{r^2}\right. \\
&\qquad \left. + \frac{\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2}\left(\frac{\delta_{pm}}{r^2} - \frac{2\mathbf{r}_p\mathbf{r}_m}{r^4}\right)\right]\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right] \\
&= \int_A \left[\mathbf{a}_q\delta_{pm} - \mathbf{a}_m\delta_{pq}\right]\left[\frac{1 - 2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^3 r^3}\mathbf{w}_m\mathbf{r}_p - \frac{\mathbf{r}_p\mathbf{r}_m}{\langle \mathbf{w}, \mathbf{u}\rangle^2 r^4} + \frac{\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2 r^2}\delta_{pm}\right] \times \\
&\qquad \left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right] \\
&= \int_A \left[\frac{1 - 2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^3}\langle \mathbf{a}, \mathbf{w}\rangle - \frac{\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2}\right]\left[-\frac{\mathbf{r}_q}{r^3}\right]\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right] \\
&= \int_A \left[\frac{1 - 2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^3}\langle \mathbf{a}, \mathbf{w}\rangle - \frac{\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2}\right] d\omega. \qquad \text{(A.17)}
\end{aligned}$$

Letting $\mathbf{a} = \mathbf{w}$ above, we get

$$B_1 = \int_A \left[\frac{1 - 2\ln\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^3} - \frac{1}{\langle \mathbf{w}, \mathbf{u}\rangle}\right] d\omega. \qquad \text{(A.18)}$$

It then follows from equations (A.16) and (A.18), (A.17) that

$$\int_{\partial A} \xi_2 \left[\langle \mathbf{a}, \mathbf{w}\rangle \langle \mathbf{w}, \mathbf{n}\rangle - \langle \mathbf{a}, \mathbf{n}\rangle\right] ds = \int_A \left[\frac{\langle \mathbf{a}, \mathbf{u}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle^2} - \frac{\langle \mathbf{a}, \mathbf{w}\rangle}{\langle \mathbf{w}, \mathbf{u}\rangle}\right] d\omega. \qquad \text{(A.19)}$$

Comparing equation (A.15) and equation (A.19), we get the identity (3.79).

A.6 Proving Equation (3.123) by Change of Variable

In this appendix we prove the relationship between $\Lambda(\alpha, \beta)$ and $\Upsilon(\mu, \nu)$ that is expressed in equation (3.123) using change of variables.

Proof: Let

$$\tan\theta = \sqrt{1-\alpha^2}\tan t,$$

we can easily derive from trigonometric identities that

$$\begin{aligned} t &= \tan^{-1}\left(\frac{\tan\theta}{\sqrt{1-\alpha^2}}\right) \\ \cos^2\theta &= \frac{\cos^2 t}{1-\alpha^2\sin^2 t} \\ d\theta &= \sqrt{1-\alpha^2}\,\frac{\cos^2\theta}{\cos^2 t}\,dt \\ &= \frac{\sqrt{1-\alpha^2}}{1-\alpha^2\sin^2 t}\,dt. \end{aligned}$$

It then follows that

$$\alpha^2\cos^2\theta = \frac{\alpha^2\cos^2 t}{1-\alpha^2\sin^2 t}.$$

Consequently, we have

$$\begin{aligned} \Lambda(\alpha,\beta) &= \int_0^{\beta}\frac{\ln(\alpha\cos\theta)}{1-\alpha^2\cos^2\theta}\,d\theta \\ &= \frac{1}{2}\int_0^{\beta}\frac{\ln(\alpha^2\cos^2\theta)}{1-\alpha^2\cos^2\theta}\,d\theta \\ &= \frac{1}{2}\int_0^{\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right)}\ln\left(\frac{\alpha^2\cos^2 t}{1-\alpha^2\sin^2 t}\right)\left[\frac{1-\alpha^2\sin^2 t}{1-\alpha^2}\right]\left[\frac{\sqrt{1-\alpha^2}}{1-\alpha^2\sin^2 t}\right]dt \\ &= \frac{1}{2\sqrt{1-\alpha^2}}\int_0^{\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right)}\ln\left(\frac{\alpha^2\cos^2 t}{1-\alpha^2\sin^2 t}\right)dt \end{aligned}$$

$$= -\frac{1}{2\sqrt{1-\alpha^2}} \int_0^{\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right)} \ln\left(\frac{1-\alpha^2\sin^2 t}{\alpha^2\cos^2 t}\right)\, dt$$

$$= -\frac{1}{2\sqrt{1-\alpha^2}} \int_0^{\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right)} \ln\left(1+\frac{1-\alpha^2}{\alpha^2}\sec^2 t\right)\, dt$$

$$= -\frac{1}{2\sqrt{1-\alpha^2}}\, \Upsilon\left(\tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right),\ \frac{\sqrt{1-\alpha^2}}{\alpha}\right).$$

□□

A.7 Reduction of $\Upsilon(\mu, \nu)$ to Clausen Integrals

In this appendix, we show a complete derivation for representing $\Upsilon(\mu, \nu)$ given in (3.122) in terms of the Clausen integral. First, we state and prove two lemmas regarding two intermediate functions used in this reduction.

Lemma 2 *Lobachevsky's function $L(x)$ defined by equation (3.137) satisfies [79, page 88]*

$$L(x) = \frac{1}{2}\,\mathrm{Cl}_2(\pi - 2x) - x \ln 2. \tag{A.20}$$

Proof: It follows from the series representation of Clausen integral given in equation (3.120) that $\mathrm{Cl}_2(\pi) = 0$. Using the integral definition (3.120), we have

$$\begin{aligned}
\mathrm{Cl}_2(\pi - 2x) &= -\int_0^{\pi-2x} \ln\left(2\sin\frac{\theta}{2}\right)\, d\theta \\
&= -\mathrm{Cl}_2(\pi) - \int_\pi^{\pi-2x} \ln\left(2\sin\frac{\theta}{2}\right)\, d\theta \\
&= 2x\ln 2 - \int_\pi^{\pi-2x} \ln\left(\sin\frac{\theta}{2}\right)\, d\theta \\
&= 2x\ln 2 + 2\int_0^x \ln(\cos t)\, dt \\
&= 2x\ln 2 + 2\,L(x),
\end{aligned}$$

where we have used a change of variable $\theta = \pi - 2t$. Solve for $L(x)$, we get equation (A.20). □□

Lemma 3 *The two-parameter function $F(\alpha, x)$ defined as*

$$F(\alpha, x) \equiv \int_0^x \ln(1 + \sin\alpha\cos\theta)\, d\theta \tag{A.21}$$

satisfies an identity given by

$$F(\alpha, x) = x \ln\left(\sin^2 \frac{\alpha}{2}\right) - \eta \ln\left(\tan^2 \frac{\alpha}{2}\right) - \mathrm{Cl}_2(2x) + \mathrm{Cl}_2(2x - 2\eta) + \mathrm{Cl}_2(2\eta), \quad \text{(A.22)}$$

where

$$\tan \eta \equiv \frac{\sin x}{\tan \frac{\alpha}{2} + \cos x}. \quad \text{(A.23)}$$

Proof: The identity (A.22) was first identified by Newman [79, pp. 89] back in 1892. Our derivation shown here closely follows Newman's, where trigonometric identities play an important role. Let $m = \tan \frac{\alpha}{2}$ and

$$\tan \gamma = \frac{\sin \theta}{m + \cos \theta}, \quad \text{(A.24)}$$

Solving for m, we get

$$m = \frac{\sin(\theta - \gamma)}{\sin \gamma}. \quad \text{(A.25)}$$

It follows that

$$\begin{aligned} \sin \alpha &= \frac{2 \tan \frac{\alpha}{2}}{1 + \tan^2 \frac{\alpha}{2}} = \frac{2m}{1 + m^2}, \\ m + \cos \theta &= \frac{\sin(\theta - \gamma)}{\sin \gamma} + \cos \theta = \frac{\sin \theta \cos \gamma}{\sin \gamma}, \\ 1 + m^2 &= \sec^2 \frac{\alpha}{2}, \end{aligned}$$

where we have used the following identities regarding trigonometry functions:

$$\begin{aligned} \sin \alpha = 2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} &= \frac{2 \tan \frac{\alpha}{2}}{\sec^2 \frac{\alpha}{2}} = \frac{2 \tan \frac{\alpha}{2}}{1 + \tan^2 \frac{\alpha}{2}}, \\ \sin(\theta - \gamma) &= \sin \theta \cos \gamma - \cos \theta \sin \gamma. \end{aligned}$$

Thus, we can express $1 + \sin\alpha\cos\theta$ as a product by

$$\begin{aligned} 1 + \sin\alpha\cos\theta &= \frac{1 + 2m\cos\theta + m^2}{1 + m^2} \\ &= \frac{\sin^2\theta + (\cos\theta + m)^2}{1 + m^2} \\ &= \frac{\sin^2\theta}{\sin^2\gamma \sec^2\frac{\alpha}{2}} \\ &= \frac{\sin^2\theta}{\sin^2\gamma}\cos^2\frac{\alpha}{2}. \end{aligned}$$

According to the definition (3.120), we get

$$\begin{aligned} F(\alpha, x) &= \int_0^x \left(\ln(\sin^2\theta) - \ln(\sin^2\gamma) + \ln\left(\cos^2\frac{\alpha}{2}\right)\right) d\theta \\ &= \int_0^x \left(\ln\left(4\sin^2\theta\right) - \ln\left(4\sin^2\gamma\right) + \ln\left(\cos^2\frac{\alpha}{2}\right)\right) d\theta \\ &= x\ln\left(\cos^2\frac{\alpha}{2}\right) - \mathrm{Cl}_2(2x) - \int_0^x \ln\left(4\sin^2\gamma\right) d\theta. \end{aligned} \tag{A.26}$$

Considering $\gamma(\theta)$ as a C^1 function of θ implicitly defined by equation (A.24) and rewriting $d\theta$ as

$$d\theta = d\gamma(\theta) + d(\theta - \gamma(\theta)), \tag{A.27}$$

we can express the last integral in equation (A.26) in terms of Clausen integrals as follows:

$$\begin{aligned} \int_0^x \ln(4\sin^2\gamma)\, d\theta &= \int_0^x \ln[\,4\sin^2\gamma(\theta)\,]\, d\gamma(\theta) + \int_0^x \ln[\,4\sin^2\gamma(\theta)\,]\, d(\theta - \gamma(\theta)) \\ &= \int_0^{\gamma(x)} \ln[\,4\sin^2 t\,]\, dt + \int_0^x \ln[\,4\sin^2\gamma(\theta)\,]\, d(\theta - \gamma(\theta) \\ &= -\mathrm{Cl}_2(2\gamma(x)) + 2\int_0^x \ln\left[\frac{4\sin^2(\theta - \gamma(\theta))}{m^2}\right] d(\theta - \gamma(\theta)) \\ &= -\mathrm{Cl}_2(2\gamma(x)) + \int_0^{x-\gamma(x)} \ln\left[4\sin^2 t\right] dt - [x - \gamma(x)]\ln m^2 \\ &= -\mathrm{Cl}_2(2\gamma(x)) - \mathrm{Cl}_2(2x - 2\gamma(x)) - \end{aligned}$$

$$(x - \gamma(x)) \ln\left(\tan^2 \frac{\alpha}{2}\right), \tag{A.28}$$

where we have used equation (A.25) and $\gamma(0) = 0$ in changing variables. Let $\eta = \gamma(x)$ and thus

$$\tan\eta \;=\; \frac{\sin x}{\tan\frac{\alpha}{2} + \cos x}.$$

Combining equations (A.26) and (A.28) yields

$$\begin{aligned} F(\alpha, x) &= x \ln\left(\cos^2 \frac{\alpha}{2}\right) + (x - \eta) \ln\left(\tan^2 \frac{\alpha}{2}\right) - \mathrm{Cl}_2(2x) + \mathrm{Cl}_2(2\eta) + \mathrm{Cl}_2(2x - 2\eta) \\ &= x \ln\left(\sin^2 \frac{\alpha}{2}\right) - \eta \ln\left(\tan^2 \frac{\alpha}{2}\right) - \mathrm{Cl}_2(2x) + \mathrm{Cl}_2(2\eta) + \mathrm{Cl}_2(2x - 2\eta), \end{aligned}$$

which establishes equation (A.22). □□

Discussion: In lemma 3, the function η, which depends on α and x, is only implicitly defined by equation (A.23). To uniquely determine η from $\tan\eta$, we need two restrictions on $\eta = \gamma(x)$ that were enforced by our proof shown above. That is,

- $\gamma(x)$ must be a C^1 function in x to make equation (A.27) hold.
- $\gamma(0) = 0$.

Specifically, when $0 \le \alpha \le \pi/2$ and $0 \le x \le \pi$, we have

$$0 \le m \;=\; \tan\frac{\alpha}{2} \le 1.$$

Then,

$$\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{m + \cos\theta}\right) \;=\; \frac{m\cos\theta + 1}{(m + \cos\theta)^2} \;\ge\; 0$$

for all $0 \le \theta \le \pi$ and $\theta \ne \theta_0 = \cos^{-1}(-m)$, that is, the denominator $m + \cos\theta \ne 0$. Therefore, $\tan\gamma$ is non-negative when $\theta \in [0, \theta_0)$ and is non-positive when $\theta \in (\theta_0, \pi]$. It

follows from the two conditions of γ that γ should be in $[0, \pi]$ instead of $[-\pi/2, \pi/2]$, the range of the function $\tan^{-1}$. Therefore, we obtain an explicit formula for η below:

$$\eta = \begin{cases} \tan^{-1}\left(\dfrac{\sin x}{\tan(\alpha/2) + \cos x}\right) & \tan\eta \geq 0 \\ \\ \pi + \tan^{-1}\left(\dfrac{\sin x}{\tan(\alpha/2) + \cos x}\right) & \tan\eta < 0 \end{cases} \tag{A.29}$$

Now we are ready to prove the identity (3.124).

Proof: We first represent $\Upsilon(\mu, \nu)$ in terms of the two intermediate functions $L(x)$ and $F(\alpha, x)$. That is,

$$\begin{aligned} \Upsilon(\mu,\nu) &= \int_0^{\mu} \ln\left(1 + \frac{\nu^2}{\cos^2\theta}\right)\, d\theta \\ &= \int_0^{\mu} \ln\left(\nu^2 + \cos^2\theta\right)\, d\theta - 2\int_0^{\mu} \ln\left(\cos\theta\right)\, d\theta \\ &= \frac{1}{2}\int_0^{2\mu} \ln\left[\frac{\cos\theta + 2\nu^2 + 1}{2}\right]\, d\theta - 2\,L(\mu) \\ &= \frac{1}{2}\int_0^{2\mu} \ln\left(1 + \frac{\cos\theta}{2\nu^2+1}\right)\, d\theta + \mu\,\ln\left(\frac{2\nu^2+1}{2}\right) - 2\,L(\mu) \\ &= \frac{1}{2}F\left(\sin^{-1}\frac{1}{2\nu^2+1}, 2\mu\right) + \mu\,\ln\left(\frac{2\nu^2+1}{2}\right) - 2\,L(\mu). \end{aligned}$$

Then, by replacing F and L above by identities (A.22) and (A.20), we get

$$\begin{aligned} \Upsilon(\mu,\nu) &= \mu\,\ln\left(\sin^2\frac{a}{2}\right) - \frac{1}{2}\eta\,\ln\left(\tan^2\frac{a}{2}\right) + \mu\,\ln\left(4\nu^2+2\right) \\ &\quad -\frac{1}{2}\Big[\mathrm{Cl}_2(4\mu) - \mathrm{Cl}_2(4\mu - 2\eta) - \mathrm{Cl}_2(2\eta) + 2\mathrm{Cl}_2(\pi - 2\mu)\Big] \\ &= -\frac{1}{2}\Big[\mathrm{Cl}_2(4\mu) - \mathrm{Cl}_2(4\mu - 2\eta) - \mathrm{Cl}_2(2\eta) + 2\mathrm{Cl}_2(\pi - 2\mu) \\ &\quad -2\mu\,\ln\left(\left(4\nu^2+2\right)\sin^2\frac{a}{2}\right) + \eta\,\ln\left(\tan^2\frac{a}{2}\right)\Big], \end{aligned} \tag{A.30}$$

where

$$a = \sin^{-1}\frac{1}{2\nu^2+1} \tag{A.31}$$

$$\tan\eta = \frac{\sin(2\mu)}{\tan\frac{a}{2}+\cos(2\mu)}. \tag{A.32}$$

Obviously, we have $0 \leq a \leq \pi/2$ from its definition (A.31). Moreover, it follows from equation (3.123) that the parameter μ corresponding to a given $\Lambda(\alpha,\beta)$ is given by

$$\mu(\alpha,\beta) = \tan^{-1}\left(\frac{\tan\beta}{\sqrt{1-\alpha^2}}\right).$$

Therefore, $0 \leq \mu \leq \pi/2$ and thus η given in (A.32) can be determined using equation (A.29).

Using the following identity for Clausen integrals [67, pp. 94]

$$\frac{1}{2}\mathrm{Cl}_2(2\theta) = \mathrm{Cl}_2(\theta) - \mathrm{Cl}_2(\pi-\theta), \tag{A.33}$$

equation (A.30) simplifies to

$$\begin{aligned}\Upsilon(\mu,\nu) = &-\frac{1}{2}\Big[2\mathrm{Cl}_2(2\mu) - \mathrm{Cl}_2(4\mu-2\eta) - \mathrm{Cl}_2(2\eta)\\ &-2\mu\ln\left(\left(4\nu^2+2\right)\sin^2\frac{a}{2}\right) + \eta\ln\left(\tan^2\frac{a}{2}\right)\Big]\end{aligned} \tag{A.34}$$

Furthermore, it follows from equation (A.31) that

$$\sin a = \frac{1}{2\nu^2+1},$$

$$\cos a = \frac{2\nu\sqrt{\nu^2+1}}{2\nu^2+1},$$

which reduces $\tan(a/2)$ to

$$\begin{aligned}\tan\frac{a}{2} &= \frac{1-\cos a}{\sin a}\\ &= (\nu-\sqrt{\nu^2+1})^2.\end{aligned}$$

Let $\gamma = (\nu - \sqrt{\nu^2+1})^2$, then

$$\begin{aligned} \eta &= \tan^{-1}\left(\frac{\sin(2\mu)}{\gamma + \cos(2\mu)}\right) \\ \sin^2\frac{a}{2} &= \frac{\tan^2(a/2)}{\sec^2(a/2)} = \frac{\gamma^2}{1+\gamma^2}, \end{aligned}$$

where $\tan^{-1}$ for η should be interpreted by equation (A.29), i.e. $\eta \in [0, \pi]$. Notice that

$$\gamma = \frac{\sqrt{\nu^2+1} - \nu}{\sqrt{\nu^2+1} + \nu},$$

the above formula for $\sin^2\frac{a}{2}$ can be further simplified as

$$\sin^2\frac{a}{2} = \frac{\gamma^2}{1+\gamma^2} = \frac{\gamma}{4\nu^2+2}. \tag{A.35}$$

It then follows from equation (A.34) that

$$\Upsilon(\mu,\nu) = -\frac{1}{2}[2\mathrm{Cl}_2(2\mu) - \mathrm{Cl}_2(4\mu - 2\eta) - \mathrm{Cl}_2(2\eta) + 2(\eta - \mu)\ln\gamma]. \tag{A.36}$$

□□

Alternatively, the application of various trigonometric identities provides us another two simpler derivations for equation (A.35), that is

1. $\sin^2\frac{a}{2} = \frac{1}{2}\tan\frac{a}{2}\sin a = \frac{\gamma}{4\nu^2+2}$.
2. It follows from the identity $\cos a = 1 - 2\sin^2\frac{a}{2}$ that

$$\sin^2\frac{a}{2} = \frac{1-\cos a}{2} = \frac{\gamma}{4\nu^2+2}.$$

A.8 Proof of Equation (3.136)

In this appendix we derive the exact expression (3.136) for a boundary curve of Λ', given by

$$\Lambda'(\alpha, \pi/2) \;=\; \int_0^{\pi/2} \frac{\ln\cos\theta}{1-(\alpha\cos\theta)^2}\, d\theta.$$

We first prove the following identity [8, pp. 113]

$$\Upsilon(\frac{\pi}{2}, \beta) \;=\; \pi \ln\left(\beta + \sqrt{1+\beta^2}\right), \tag{A.37}$$

where Υ is given in equation (3.122).

Proof: Equation (A.37) follows from the identity

$$\int_0^{\pi/2} \ln\left(a^2\cos^2\theta + b^2\sin^2\theta\right)\, d\theta \;=\; \pi \ln\left(\frac{a+b}{2}\right), \tag{A.38}$$

due to Carlson [24]. That is

$$\begin{aligned}
\Upsilon(\frac{\pi}{2}, \beta) &= \int_0^{\pi/2} \ln(1+\beta\sec^2\theta)\, d\theta \\
&= \int_0^{\pi/2} \ln(1+\beta^2+\beta^2\tan^2\theta)\, d\theta \\
&= \int_0^{\pi/2} \ln((1+\beta^2)\cos^2\theta + \beta^2\sin^2\theta)\, d\theta - \int_0^{\pi/2} \ln(\cos^2\theta)\, d\theta \\
&= \pi \ln\left(\frac{\beta+\sqrt{1+\beta^2}}{2}\right) - \pi\ln\left(\frac{1}{2}\right) \\
&= \pi \ln\left(\beta + \sqrt{1+\beta^2}\right).
\end{aligned}$$

□□

Using the identity (A.37) and equation (3.123), we have

$$\Lambda(\alpha, \frac{\pi}{2}) \;=\; -\frac{1}{2\sqrt{1-\alpha^2}}\, \Upsilon(\frac{\pi}{2}, \frac{\sqrt{1-\alpha^2}}{\alpha})$$

$$= -\frac{\pi}{2\sqrt{1-\alpha^2}} \ln\left(\frac{1+\sqrt{1-\alpha^2}}{\alpha}\right)$$

$$= \frac{\pi}{2\sqrt{1-\alpha^2}} \ln\alpha - \frac{\pi}{2\sqrt{1-\alpha^2}} \ln\left(1+\sqrt{1-\alpha^2}\right), \tag{A.39}$$

where we used the fact that

$$\tan^{-1}\left(\frac{\tan(\pi/2)}{\sqrt{1-\alpha^2}}\right) = \frac{\pi}{2},$$

On the other hand, it follows from equation (3.131) that

$$\Lambda(\alpha, \frac{\pi}{2}) = \lambda(\alpha, \frac{\pi}{2}) + \Lambda'(\alpha, \frac{\pi}{2})$$

$$= \frac{\pi}{2\sqrt{1-\alpha^2}} \ln\alpha + \Lambda'(\alpha, \frac{\pi}{2}). \tag{A.40}$$

Combining equations (A.39) and (A.40), we get equation (3.136).

A.9 Proof of Theorem 6

In this appendix we supply the proof of theorem 6, the steps are identical to those for theorem 5.

Proof: Corresponding to equation (4.7), we have

$$\int_{\partial A} \frac{\mathbf{u}_I^{n-1}\mathbf{n}_j}{\langle \mathbf{w},\mathbf{u}\rangle^q}\, ds \;=\; \int_{\partial A} B_{Ijl}\, d\mathbf{r}_l \;=\; \int_A \varepsilon_{qml} B_{Ijl,m}\left[\frac{\varepsilon_{qst}\; d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right], \tag{A.41}$$

where the $(n+1)$-order tensor B is given by

$$B_{Ijl} \;=\; \varepsilon_{jpl}\left[\frac{\mathbf{r}_p \mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^q\, r^{n-q+1}}\right].$$

and its derivative $B_{Ijl,m}$ is computed from the chain rule by

$$B_{Ijl,m} \;=\; \varepsilon_{jpl}\left(\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{1}{\langle \mathbf{w},\mathbf{r}\rangle^q}\right]\left[\frac{\mathbf{r}_p \mathbf{r}_I^{n-1}}{r^{n-q+1}}\right] + \left[\frac{1}{\langle \mathbf{w},\mathbf{r}\rangle^q}\right]\frac{\partial}{\partial \mathbf{r}_m}\left[\frac{\mathbf{r}_p \mathbf{r}_I^{n-1}}{r^{n-q+1}}\right]\right) \;=\; \varepsilon_{jpl}\,(A_1 + A_2)\,.$$

Here A_1 and A_2 are given by

$$\begin{aligned} A_1 &= -q\,\frac{\mathbf{w}_m \mathbf{r}_p \mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^{q+1}\, r^{n-q+1}}, \\ A_2 &= \frac{1}{\langle \mathbf{w},\mathbf{r}\rangle^q}\left[\frac{\delta_{pm}\mathbf{r}_I^{n-1} + \mathbf{r}_p \mathbf{r}_{I,m}^{n-1}}{r^{n-q+1}} - (n-q+1)\frac{\mathbf{r}_m \mathbf{r}_p \mathbf{r}_I^{n-1}}{r^{n-q+3}}\right]. \end{aligned}$$

After simplification, we have

$$\begin{aligned} \varepsilon_{qml}\varepsilon_{jpl} A_1 &= -\frac{q}{\langle \mathbf{w},\mathbf{r}\rangle^{q+1}\, r^{n-q+1}}\left[\delta_{qj}\delta_{pm} - \delta_{pq}\delta_{jm}\right]\mathbf{w}_m \mathbf{r}_p \mathbf{r}_I^{n-1} \\ &= \frac{q\,\mathbf{w}_j \mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^{q+1}\, r^{n-q-2}}\left(\frac{\mathbf{r}_q}{r^3}\right) - \frac{q\,\delta_{qj}\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^q\, r^{n-q+1}}, \\ \varepsilon_{qml}\varepsilon_{jpl} A_2 &= \frac{\delta_{qj}\delta_{pm} - \delta_{pq}\delta_{jm}}{\langle \mathbf{w},\mathbf{r}\rangle^q}\left[\frac{\delta_{pm}\mathbf{r}_I^{n-1} + \mathbf{r}_p \mathbf{r}_{I,m}^{n-1}}{r^{n-q+1}} - (n-q+1)\frac{\mathbf{r}_m \mathbf{r}_p \mathbf{r}_I^{n-1}}{r^{n-q+3}}\right] \\ &= \frac{q\,\delta_{qj}\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^q} + \left[(n-q+1)\frac{\mathbf{r}_j \mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^q\, r^{n-q}} - \frac{\mathbf{r}_{I,j}^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^q\, r^{n-q-2}}\right]\left(\frac{\mathbf{r}_q}{r^3}\right). \end{aligned}$$

Consequently, by expanding $\mathbf{r}_{I,j}^{n-1}$ using equation (4.9) and representing the result in terms of $d\omega$ using equation (3.27), equation (A.41) simplifies to

$$\begin{aligned}\int_{\partial A} B_{Ijl}\ d\mathbf{r}_l &= \int_A \left[-\frac{q\,\mathbf{w}_j \mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^{q+1}\, r^{n-q-2}} - (n-q+1)\frac{\mathbf{r}_I^{n-1}\mathbf{r}_j}{\langle \mathbf{w},\mathbf{r}\rangle^{q}\, r^{n-q}} + \frac{\sum_{k=1}^{n-1}\delta_{jI_k}\mathbf{r}_{I\backslash k}^{n-2}}{\langle \mathbf{w},\mathbf{r}\rangle^{q}\, r^{n-q-2}} \right] d\omega \\ &= -q\,\mathbf{w}_j \mathbf{T}_I^{n-1;q+1} - (n-q+1)\mathbf{T}_{Ij}^{n;q} + \sum_{k=1}^{n-1}\delta_{jI_k}\mathbf{T}_{I\backslash k}^{n-2;q}. \qquad \text{(A.42)}\end{aligned}$$

When $n \neq q-1$, $\mathbf{T}_{Ij}^{n;q}$ can be solved from equation (A.42), yielding equation (4.12). □□

A.10 Proof of Theorem 7

The steps are almost identical to those shown in proving Theorem 5 and Appendix A.9.

Proof: Corresponding to equations (4.7), we have

$$\int_{\partial A} \frac{\mathbf{u}_I^{n-1}\mathbf{n}_m}{\langle \mathbf{w},\mathbf{u}\rangle}\, ds \;=\; \int_{\partial A} B_{Iml}\, d\mathbf{r}_l \;=\; \int_A \varepsilon_{qsl} B_{Iml,s}\left[\frac{\varepsilon_{qht}\, d\mathbf{r}_h \wedge d\mathbf{r}_t}{2}\right], \tag{A.43}$$

where the $(n+1)$-order tensor B is given by

$$B_{Iml} \;=\; \frac{\varepsilon_{mpl}\mathbf{r}_I^{n-1}\mathbf{r}_p}{\langle \mathbf{w},\mathbf{r}\rangle\, r^n},$$

and its derivative $B_{Iml,s}$ is computed from the chain rule by

$$B_{Iml,s} = \varepsilon_{mpl}\left(\frac{\partial}{\partial \mathbf{r}_s}\left[\frac{1}{\langle \mathbf{w},\mathbf{r}\rangle}\right]\left[\frac{\mathbf{r}_p\mathbf{r}_I^{n-1}}{r^n}\right] + \left[\frac{1}{\langle \mathbf{w},\mathbf{r}\rangle}\right]\frac{\partial}{\partial \mathbf{r}_s}\left[\frac{\mathbf{r}_p\mathbf{r}_I^{n-1}}{r^n}\right]\right) = \varepsilon_{mpl}\left[A_1 + A_2\right].$$

Here, A_1, A_2 are given by

$$A_1 \;=\; -\frac{\mathbf{w}_s\mathbf{r}_p\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^2\, r^n}, \quad A_2 \;=\; \frac{\delta_{ps}\mathbf{r}_I^{n-1} + \mathbf{r}_p\mathbf{r}_{I,s}^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle\, r^n} - n\frac{\mathbf{r}_s\mathbf{r}_p\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle\, r^{n+2}}.$$

After simplification, we get

$$\begin{aligned}
\varepsilon_{qsl}\varepsilon_{mpl}A_1 &= \frac{\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^2\, r^n}\left[\delta_{qm}\langle \mathbf{w},\mathbf{r}\rangle - \mathbf{w}_m\mathbf{r}_q\right], \\
\varepsilon_{qsl}\varepsilon_{mpl}A_2 &= \frac{1}{\langle \mathbf{w},\mathbf{r}\rangle}\left[\frac{\delta_{qm}\mathbf{r}_I^{n-1}}{r^n} - \frac{\mathbf{r}_q\mathbf{r}_{I,m}^{n-1}}{r^n} + n\frac{\mathbf{r}_q\mathbf{r}_m\mathbf{r}_I^{n-1}}{r^{n+2}}\right].
\end{aligned}$$

Consequently, in terms of $d\omega$ defined in equation (3.27) and using equation (4.9), equation (A.43) simplifies to

$$\int_{\partial A} B_{Iml}\, d\mathbf{r}_l \;=\; -\int_A \left[\frac{\mathbf{w}_m\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle^2\, r^{n-3}} - \frac{\sum_{k=1}^{n-1}\delta_{mI_k}\mathbf{r}_{I\setminus k}^{n-2}}{\langle \mathbf{w},\mathbf{r}\rangle\, r^{n-3}} + n\frac{\mathbf{r}_m\mathbf{r}_I^{n-1}}{\langle \mathbf{w},\mathbf{r}\rangle\, r^{n-1}}\right] d\omega. \tag{A.44}$$

Multiplying both sides of equation (A.44) by $(\delta_{jm} - \mathbf{w}_j\mathbf{w}_m)/n$ and expressing the resulting terms in terms of $\mathbf{T}^{n;1}$ or $\mathbf{T}^n$, we have proved equation (4.13). □□

A.11 Proof of Theorem 8

In this appendix we provide the proof of Theorem 8 using generalized Stokes' theorem.

Proof: In accordance with equation (4.7), we get the following equations:

$$\int_{\partial A} \frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\langle \mathbf{w},\mathbf{n}\rangle\, ds \;=\; \int_{\partial A} B_l\, d\mathbf{r}_l \;=\; \int_A \varepsilon_{qml} B_{l,m}\left[\frac{\varepsilon_{qst}\, d\mathbf{r}_s \wedge d\mathbf{r}_t}{2}\right], \qquad \text{(A.45)}$$

where the vector B is given by

$$B_l \;=\; \varepsilon_{kpl}\mathbf{w}_k \frac{\ln\langle \mathbf{w},\mathbf{u}\rangle}{1-\langle \mathbf{w},\mathbf{u}\rangle^2}\cdot\frac{\mathbf{r}_p}{r^2} \;=\; \varepsilon_{kpl}\mathbf{w}_k\eta\frac{\mathbf{r}_p}{r^2},$$

and its derivative $B_{l,m}$ is computed from the chain rule by

$$B_{l,m} \;=\; \varepsilon_{kpl}\mathbf{w}_k\left[\eta\frac{\partial}{\partial \mathbf{r}_m}\left(\frac{\mathbf{r}_p}{r^2}\right)+\eta_m\frac{\mathbf{r}_p}{r^2}\right].$$

Here, we have

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{r}_m}\left(\frac{\mathbf{r}_p}{r^2}\right) &= \frac{\delta_{pm}r^2-2\mathbf{r}_p\mathbf{r}_m}{r^4},\\
\eta_m\left(=\frac{\partial \eta}{\partial \mathbf{r}_m}\right) &= \frac{r^2-\langle \mathbf{w},\mathbf{r}\rangle^2+2\langle \mathbf{w},\mathbf{r}\rangle^2\ln\langle \mathbf{w},\mathbf{u}\rangle}{\left(r^2-\langle \mathbf{w},\mathbf{r}\rangle^2\right)^2}\left[\frac{r^2\mathbf{w}_m}{\langle \mathbf{w},\mathbf{r}\rangle}-\mathbf{r}_m\right].
\end{aligned}$$

After simplification, we attain

$$\begin{aligned}
\varepsilon_{qml}\varepsilon_{kpl}\mathbf{w}_k\eta\frac{\partial}{\partial \mathbf{r}_m}\left(\frac{\mathbf{r}_p}{r^2}\right) &= \frac{\mathbf{r}_q}{r^3}\left[2\eta\langle \mathbf{w},\mathbf{u}\rangle\right],\\
\varepsilon_{qml}\varepsilon_{kpl}\mathbf{w}_k\eta_m\frac{\mathbf{r}_p}{r^2} &= -\frac{\mathbf{r}_q}{r^3}\left[\frac{1}{\langle \mathbf{w},\mathbf{u}\rangle}+2\eta\langle \mathbf{w},\mathbf{u}\rangle\right].
\end{aligned}$$

Consequently, equation (4.14) follows directly from indentifying the 2-form $d\omega$ appearing on the right hand side of equation (A.45). □□

A.12 Proof of Equation (4.32)

In this appendix we prove the recurrence formulas shown in (4.32) for $G(r,s,x,y)$, using integration by parts.

$$\begin{aligned} G(r,s,x,y) &= \int_x^y \sin^r\theta\cos^s\theta\,d\theta \\ &= \frac{1}{r+s}\int_x^y \cot^{s-1}\theta\,d\sin^{r+s}\theta \\ &= \frac{1}{r+s}\left[\sin^{r+1}\theta\cos^{s-1}\theta\,\Big|_x^y + (s-1)\int_x^y \sin^r\theta\cos^{s-2}\theta\,d\theta\right] \\ &= \frac{1}{r+s}\left[\sin^{r+1}y\cos^{s-1}y - \sin^{r+1}x\cos^{s-1}x + (s-1)G(r,s-2,x,y)\right]. \end{aligned}$$

Similarly, we can prove the other two recurrences as follows:

$$\begin{aligned} G(r,s,x,y) &= \int_x^y \sin^r\theta\cos^s\theta\,d\theta \\ &= -\frac{1}{s+1}\int_x^y \sin^{r+s+2}\theta\,d\cot^{s+1}\theta \\ &= -\frac{1}{s+1}\left[\sin^{r+1}\theta\cos^{s+1}\theta\,\Big|_x^y - (r+s+2)\int_x^y \sin^r\theta\cos^{s+2}\theta\,d\theta\right] \\ &= \frac{1}{s+1}\left[\sin^{r+1}x\cos^{s+1}x - \sin^{r+1}y\cos^{s+1}y + (r+s+2)G(r,s+2,x,y)\right]. \end{aligned}$$

$$\begin{aligned} G(r,s,x,y) &= \int_x^y \sin^r\theta\cos^s\theta\,d\theta \\ &= -\frac{1}{r+s}\int_x^y \tan^{r-1}\theta\,d\cos^{r+s}\theta \\ &= -\frac{1}{r+s}\left[\sin^{r-1}\theta\cos^{s+1}\theta\,\Big|_x^y - (r-1)\int_x^y \sin^{r-2}\theta\cos^s\theta\,d\theta\right] \\ &= \frac{1}{r+s}\left[\sin^{r-1}x\cos^{s+1}x - \sin^{r-1}y\cos^{s+1}y + (r-1)G(r-2,s,x,y)\right]. \end{aligned}$$

Bibliography

[1] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions.* Dover Publications, New York, 1965.

[2] S. J. Adelson and L. F.Hodges. Generating exact ray-traced animation frames by reprojection. *IEEE Computer Graphics and Applications*, 15:43–52, 1995.

[3] S. J. Adelson and L.F. Hodges. Stereoscopic ray tracing. *The Visual Computer*, 10(3):127–144, 1993.

[4] George E. Andrews, Richard Askey, and Ranjan Roy. *Special Functions.* Cambridge University Press, New York, 1999.

[5] Tom M. Apostol. *Calculus II: Multi-Variable Calculus and Linear Algebra, with Applications to Differential Equations and Probability.* John Wiley & Sons, New York, 1969.

[6] A. Appel. Some techniques for shading machine renderings of solids. In *Proceedings of the Spring Joint Computer Conference*, pages 37–45, 1968.

[7] James Arvo. The irradiance Jacobian for partially occluded polyhedral sources. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 343–350, July 1994.

[8] James Arvo. *Analytic Methods for Simulated Light Transport.* Ph.D. thesis, Yale University, December 1995.

[9] James Arvo. Applications of irradiance tensors to the simulation of non-Lambertian phenomena. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 335–342, August 1995.

[10] James Arvo and David Kirk. A survey of ray tracing acceleration techniques. In Andrew S. Glassner, editor, *An Introduction to Ray Tracing*, chapter 6, pages 201–262. Academic Press, New York, 1989.

[11] James Arvo and Kevin Novins. Iso-contour volume rendering. In *1994 ACM/IEEE Symposium on Volume Visualization*, Washington DC, October 1994.

[12] A. Ashour and A. Sabri. Tabulation of the function $\psi(\theta) = \sum_{n=1}^{\infty} \sin n\theta/n^2$. *Mathematical Tables and other Aids to Computation*, 10(54):57–65, April 1956.

[13] S. Badt, Jr. Two algorithms taking advantage of temporal coherence in ray tracing. *The Visual Computer*, 4(3):123–132, 1988.

[14] Hujun Bao and Qunsheng Peng. Shading models for linear and area light sources. *Computers and Graphics*, 17(2):137–145, 1993.

[15] Hujun Bao, Jianguo Ying, and Qunsheng Peng. Shading with curved light source. In *Proceedings of Eurographics 95*, pages 217–227, 1995.

[16] Albert A. Bartlett and Rodger Lucero. Note on a common virtual image. *American Journal of Physics*, 52(7):640–643, 1984.

[17] Rui M. Bastos, António A. de Sousa, and Fernando N. Ferreira. Reconstruction of illumination functions using hermite bicubic interpolation. In *Proceedings of the Fourth Eurographics Workshop on Rendering*, Paris, France, June 1993.

[18] Daniel R. Baum, Holly E. Rushmeier, and James M. Winget. Improving radiosity solutions through the use of analytically determined form-factors. *Computer Graphics*, 23(3):325–334, July 1989.

[19] Philippe Bekaert, Philip Dutré, and Yves D. Willems. Final radiosity gather step using a monte carlo technique with importance sampling. Technical Report Report CW275, Department of Computer Science, K.U.Leuven, 1998.

[20] Marcel Berger. *Geometry*, volume 2. Springer-Verlag, New York, 1987. Translated by M. Cole and S. Levy.

[21] T. Binford. Inferring surfaces from images. *Artificial Intelligence*, 17, 1981.

[22] Max Born and Emil Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light.* Pergamon Press, New York, third edition, 1965.

[23] N. Briere and P. Poulin. Hierarchical view-dependent structures for interactive scene manipulation. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 83–90, August 1996.

[24] B. C. Carlson. Invariance of an integral average of a logarithm. *The American Mathematical Monthly*, 82(4):379–382, April 1975.

[25] J. Chapman, T. W. Calvert, and J. Dill. Exploiting temporal coherence in ray tracing. In *Proceedings of Graphics Interface'90*, pages 196–204, May 1990.

[26] J. Chapman, T. W. Calvert, and J. Dill. Spatio-temporal coherence in ray tracing. In *Proceedings of Graphics Interface'91*, pages 101–108, June 1991.

[27] Min Chen. Perturbation methods for image synthesis. Master's thesis, California Institute of Technology, May 1999. Technical report CS-TR-99-05, ftp://ftp.cs.caltech.edu/tr/cs-tr-99-05.ps.Z.

[28] Min Chen and James Arvo. Perturbation methods for interactive specular reflections. *IEEE Transactions on Visualization and Computer Graphics*, 6(3):253–264, July-September 2000.

[29] Min Chen and James Arvo. Theory and application of specular path perturbation. *ACM Transactions on Graphics*, 19(4):246–278, October 2000.

[30] R. Cipolla and P. Giblin. *Visual motion of curves and surfaces.* Cambridge University Press, New York, 2000.

[31] C.W. Clenshaw. *Chebyshev Series for Mathematical Functions, Mathematical Tables, Vol.5.* National Physics Laboratory, Her Majesty's Stationery Office, London, 1962.

[32] Michael F. Cohen and Donald P. Greenberg. The hemi-cube: A radiosity solution for complex environments. *Computer Graphics*, 19(3):75–84, July 1985.

[33] Michael F. Cohen and John R. Wallace. *Radiosity and Realistic Image Synthesis.* Academic Press, New York, 1993.

[34] Robert L. Cook. Shade trees. *Computer Graphics*, 18(3):137–145, July 1984.

[35] Robert L. Cook. Stochastic sampling in computer graphics. *ACM Transactions on Graphics*, 5(1):51–72, 1986.

[36] Robert L. Cook, Thomas Porter, and Loren Carpenter. Distributed ray tracing. *Computer Graphics*, 18(3):137–145, July 1984.

[37] Mathieu Desbrun, Mark Meyer, Peter Schröder, and Alan Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 317–324, August 1999.

[38] David L. DiLaura. Non-Diffuse Radiative Transfer II: Planar Area Sources and Receivers. In *1996 Illuminating Engineering Society Annual Conference Technical Papers*, pages 106–125, 1995.

[39] David L. DiLaura. Non-diffuse radiative transfer 3: Inhomogeneous planar area sources and point receivers. *Journal of the Illuminating Engineering Society*, 26(1):182–187, 1997.

[40] David L. DiLaura and Jeffrey Quinlan. Non-Diffuse Radiative Transfer I: Planar Area Sources and Point Receivers. *Journal of the Illuminating Engineering Society*, 24(2):102–112, 1995.

[41] Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces.* Prentice-Hall, New Jersey, 1976.

[42] Yoshinori Dobashi, Kazufumi Kaneda, Hideki Nakatani, and Hideo Yamashita. A quick rendering method using basis functions for interactive lighting design. In *Proceedings of Eurographics 95*, pages 229–240, 1995.

[43] Julie O'B. Dorsey, James Arvo, and Donald P. Greenberg. Interactive design of complex time-dependent lighting. *IEEE Computer Graphics and Applications*, 15(2):26–36, March 1995.

[44] Julie O'B. Dorsey, Franccois X. Sillion, and Donald P. Greenberg. Design and simulation of opera lighting and projection effects. *Computer Graphics*, 25(4):41–50, August 1991.

[45] George Drettakis and Eugene L. Fiume. Concrete computation of global illumination using structured sampling. In *Proceedings of the Third Eurographics Workshop on Rendering,* Bristol, United Kingdom, 1992.

[46] A. F. Emery, O. Johansson, M. Lobo, and A. Abrous. A comparative study of methods for computing the diffuse radiation viewfactors for complex structures. *ASME Journal of Heat Transfer*, 113(2):413–422, May 1991.

[47] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modeling the interaction of light between diffuse surfaces. *Computer Graphics*, 18(3):213–222, July 1984.

[48] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 43–54, August 1996.

[49] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, New York, fifth edition, 1994.

[50] C. C. Grosjean. Formulae concerning the computation of the Clausen integral $cl_2(\theta)$. *Journal of Computational and Applied Mathematics*, 11:331–342, 1984.

[51] M. Halsead, A. Barsky, S. Klein, and R. Mandell. Reconstructing curved surfaces from specular reflection patterns using spline surface fitting normals. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 335–345, August 1996.

[52] Pat Hanrahan and Wolfgang Krueger. Reflection from layered surfaces due to subsurface scattering. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 165–174, August 1993.

[53] Pat Hanrahan, David Salzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. *Computer Graphics*, 25(4):197–206, July 1991.

[54] David Hart, Philip Dutré, and Donald P. Greenberg. Direct illumination with lazy visibility evaluation. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 147–154, August 1999.

[55] G. Healey and T. Binford. Local shape from specularity. *Computer Vision, Graphics, and Image Processing*, 42(1):62–86, 1988.

[56] Paul S. Heckbert. *Simulating Global Illumination Using Adaptive Meshing.* Ph.D. thesis, University of California, Berkeley, June 1991.

[57] Nicholas Holzschuch and Franccois X. Sillion. Accurate computation of the radiosity gradient for constant and linear emitters. In *Proceedings of the Sixth Eurographics Workshop on Rendering*, Dublin, Eire, June 1995.

[58] John R. Howell. *A Catalog of Radiation Configuration Factors.* McGraw-Hill, New York, 1982.

[59] Homan Igehy. Tracing ray differentials. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 179–186, August 1999.

[60] Henrik Wann Jensen, Steve Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages ??–??, August 2001.

[61] David A. Jevans. Object space temporal coherence for ray tracing. In *Proceedings of Graphics Interface '92*, pages 176–183, May 1992.

[62] James T. Kajiya. The rendering equation. *Computer Graphics*, 20(4):143–150, August 1986.

[63] K. S. Kölbig. Chebyshev coefficients for the Clausen function $cl_2(\theta)$. *Journal of Computational and Applied Mathematics*, 64:295–297, 1995.

[64] Erwin Kreyszig. *Differential Geometry.* Dover Publications, New York, 1991.

[65] Wolfgang Krueger and B. Froehlich. The responsive workbench: virtual work environment. *IEEE Computer Graphics and Applications*, 14(3):12–15, May 1994.

[66] Eric P. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 117–126, August 1997.

[67] Leonard Lewin. *Dilogarithms and associated functions.* Macdonald, London, 1958.

[68] Dani Lischinski and Ari Rappoport. Image-based rendering for non-diffuse synthetic scenes. In *Proceedings of the Ninth Eurographics Workshop on Rendering*, Vienna, Austria, June 1998.

[69] Dani Lischinski, Filippo Tampieri, and Donald P. Greenberg. Discontinuity meshing for accurate radiosity. *IEEE Computer Graphics and Applications*, 12(6):25–39, November 1992.

[70] David G. Luenberger. *Optimization by Vector Space Methods.* John Wiley & Sons, New York, 1969.

[71] Jerrold E. Marsden and Michael J. Hoffman. *Elementary Classical Analysis.* W. H. Freeman, New York, 1993.

[72] Nelson Max and Michael Allison. Linear radiosity approximation using vertex-to-vertex form factors. In David Kirk, editor, *Graphics Gems III.* Academic Press, New York, 1992.

[73] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, and Augusta H. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[74] Don Mitchell and Pat Hanrahan. Illumination from curved reflectors. *Computer Graphics*, 26(2):283–291, July 1992.

[75] P. Moon. *The Scientific Basis of Illuminating Engineering.* Dover Publications, New York, 1961.

[76] Claus Müller. *Spherical Harmonics.* Springer-Verlag, New York, 1966.

[77] K. Murakami and K. Hirota. Incremental ray tracing. In *Eurographics Workshop on Photosimulation, Realism and Physics in Computer Graphics, Conference Proceedings*, pages 15–29, Rennes, France, June 1990.

[78] Peter M. Neumann. Reflections on reflection in a spherical mirror. *The American Mathematical Monthly*, 105(6):523–528, June–July 1998.

[79] Francis W. Newman. *The Higher Trigonometry and Superrationals of Second Order.* Macmillan and Bowes, 1892.

[80] T. Nishita and E. Nakamae. A shading model of parallel cylinderical light sources. In *Proceedings of CG International'92*, pages 429–445, 1992.

[81] T. Nishita and E. Nakamae. A new radiosity approach using area sampling for parametric patches. In *Proceedings of Eurographics 93*, pages 385–398, 1993.

[82] T. Nishita, I. Okamura, and E. Nakamae. Shading models for point and linear sources. *ACM Transactions on Graphics*, 4(2):124–146, 1985.

[83] Tomoyuki Nishita and Eihachiro Nakamae. Half-tone representation of 3-D objects illuminated by area sources or polyhedron sources. In *Proceedings of the IEEE Computer Software and Applications Conference*, pages 237–242, Chicago, November 1983.

[84] Tomoyuki Nishita and Eihachiro Nakamae. Continuous tone representation of 3-D objects taking account of shadows and interreflection. *Computer Graphics*, 19(3):23–30, July 1985.

[85] Eyal Ofek and Ari Rappoport. Interactive reflections on curved objects. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 333–342, July 1998.

[86] Matt Pharr and Pat Hanrahan. Monte carlo evaluation of non-linear scattering equations for subsurface reflection. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 75–84, August 2000.

[87] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, June 1975.

[88] K. P. Picott. Extensions of the linear and area lighting models. *IEEE Computer Graphics and Applications*, 12:31–38, 1992.

[89] G. C. Pomraning. *The Equations of Radiation Hydrodynamics.* Pergamon Press, New York, 1973.

[90] Pierre Poulin and John Amanatides. Shading and shadowing with linear light sources. *Computers and Graphics*, 15(2):259–265, 1991.

[91] E. O. Powell. An integral related to the radiation integrals. *Philosophical Magazine*, 34(236):600–607, September 1943.

[92] Rudolph W. Preisendorfer. *Radiative Transfer on Discrete Spaces.* Pergamon Press, New York, 1965.

[93] Mark C. Reichert. A two-pass radiosity method driven by lights and viewer position. Master's thesis, Cornell University, January 1992.

[94] David Salesin, Dani Lischninski, and Tony DeRose. Reconstructing illumination functions with selected discontinuities. In *Proceedings of the Third Eurographics Workshop on Rendering,* Bristol, United Kingdom, pages 99–112, May 1992.

[95] Silvio Savarese and Pietro Perona. Local analysis for 3d reconstruction of specular surfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

[96] M. Schreiber. *Differential Forms: A Heuristic Introduction.* Springer-Verlag, New York, 1984.

[97] Peter Schröder and Pat Hanrahan. On the form factor between two polygons. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 163–164, August 1993.

[98] Lee A. Segel and G. H. Handelman. *Mathematics Applied to Continuum Mechanics.* Macmillan Publishing Company, New York, 1977.

[99] Carlo H. Séquin and Eliot K. Smyrl. Parameterized ray tracing. *Computer Graphics*, 23(3):307–314, July 1989.

[100] Mikio Shinya, Tokiichiro Takahashi, and Seiichiro Naito. Principles and applications of pencil tracing. *Computer Graphics*, 21(4):45–54, July 1987.

[101] Peter Shirley. A ray tracing algorithm for global illumination. *Proceedings of Graphics Interface '90*, pages 205–212, May 1990.

[102] Franccois Sillion, James Arvo, Stephen Westin, and Donald P. Greenberg. A global illumination solution for general reflectance distributions. *Computer Graphics*, 25(4):187–196, July 1991.

[103] Brian Smits, James Arvo, and David Salesin. An importance-driven radiosity algorithm. *Computer Graphics*, 26(2):273–282, July 1992.

[104] Martin Avery Snyder. *Chebyshev Methods in Numerical Approximation.* Dover Publications, New York, 1995. Good introductory book on Chebyshev approximation.

[105] E. M. Sparrow. A new and simpler formulation for radiative angle factors. *ASME Journal of Heat Transfer*, 85(2):81–88, May 1963.

[106] Michael Spivak. *Calculus on Manifolds.* Benjamin/Cummings, Reading, Massachusetts, 1965.

[107] Michael M. Stark, Elaine Cohen, Tom Lynche, and Richard F. Riesenfeld. Computing exact shadow irradiance using splines. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 155–164, August 1999.

[108] Dirk J. Struik. *Lectures on Classical Differential Geometry.* Dover Publications, New York, second edition, 1961.

[109] Frank Suykens and Yves Willems. Path differentials and applications. In *The 12th Eurographics Workshop on Rendering*, pages 254–265, London, 2001.

[110] Takayuki Tanaka and T. Takahashi. Shading with area light sources. In *Proceedings of Eurographics 91*, pages 235–246, 1991.

[111] Roy Troutman and Nelson L. Max. Radiosity algorithms using higher-order finite element methods. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 209–212, August 1993.

[112] Eric Veach and Leonidas J. Guibas. Metropolis light transport. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 65–76, August 1997.

[113] Christophe Vedel. Improved storage and reconstruction of light intensities on surfaces. In *Proceedings of the Third Eurographics Workshop on Rendering,* Bristol, United Kingdom, pages 113–121, May 1992.

[114] Christophe Vedel. Computing illumination from area light sources by approximate contour integration. In *Proceedings of Graphics Interface '93*, pages 237–244, 1993.

[115] C. P. Verbeck and D. P. Greenberg. A comprehensive light-source description for computer graphics. *IEEE Computer Graphics and Applications*, 4(7):66–75, July 1984.

[116] John Wallace, M. F. Cohen, and D. P. Greenberg. A two-pass solution to the rendering equation: A synthesis of ray tracing and radiosity methods. *Computer Graphics*, 21(3):311–320, July 1987.

[117] Gregory J. Ward. Measuring and modeling anisotropic reflection. *Computer Graphics*, 26(2):265–272, July 1992.

[118] Gregory J. Ward and Paul S. Heckbert. Irradiance gradients. In *Proceedings of the Third Eurographics Workshop on Rendering,* Bristol, United Kingdom, pages 85–98, May 1992.

[119] Turner Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 32(6):343–349, June 1980.

[120] Harold Zatz. Galerkin radiosity: A higher order solution method for global illumination. Master's thesis, Cornell University, August 1992.