Area-Preserving Parameterizations for Spherical Ellipses

Ibón Guillén[1,2]   Carlos Ureña[3]   Alan King[4]   Marcos Fajardo[4]   Iliyan Georgiev[4]   Jorge López-Moreno[2]   Adrian Jarabo[1]

[1]Universidad de Zaragoza, I3A   [2]Universidad Rey Juan Carlos   [3]Universidad de Granada   [4]Solid Angle

Abstract

We present new methods for uniformly sampling the solid angle subtended by a disk. To achieve this, we introduce two novel area-preserving mappings from the unit square $[0,1]^2$ to a spherical ellipse (i.e. the projection of the disk onto the unit sphere). These mappings allow for low-variance stratified sampling of direct illumination from disk-shaped light sources. We discuss how to efficiently incorporate our methods into a production renderer and demonstrate the quality of our maps, showing significantly lower variance than previous work.

CCS Concepts
•*Computing methodologies* → *Rendering; Ray tracing; Visibility;*

1. Introduction

Illumination from area light sources is among the most important lighting effects in realistic rendering, due to the ubiquity of such sources in real-world scenes. Monte Carlo integration is the standard method for computing the illumination from such luminaires [SWZ96]. This method is general and robust, supports arbitrary reflectance models and geometry, and predictively converges to the actual solution as the number of samples increases. Accurately sampling the illumination from area light sources is crucial for minimizing the amount of noise in rendered images.

Estimating the direct illumination at a point requires sampling the radiance contribution from directions inside the solid angle subtended by the given luminaire. A sensible strategy is to distribute those directions uniformly. This, however, is hard to achieve for an arbitrary-shaped luminaire, as it involves first computing and then uniformly sampling its subtended solid angle. Specialized methods have been proposed for spherical [Wan92], triangular [Arv95, Ure00], rectangular [UFK13], and polygonal lights [Arv01]. These elaborate solid angle sampling techniques are more computationally expensive than naïve methods that uniformly sample the surface area of the luminaire. However, in most nontrivial scenes, where the sample contribution evaluation is orders of magnitude more costly than the sample generation, their lower variance improves overall efficiency.

Few papers have focused on sampling oriented disk-shaped light sources. Disk lights are important in practice, both for their artistic expressiveness and their use in a number of real-world scenarios, generally including man-made light sources such as in architectural lighting, film and photography. Moreover, disk lights form the base for some approximate global illumination algorithms [HKWB09,

SHD15]. So far, the only practical method for uniformly sampling the solid angle of disk lights is the work by Gamito [Gam16], who proposed a rejection sampling approach that generates candidates using spherical quad sampling [UFK13]. Unfortunately, achieving good sample stratification with this method requires special care.

In this paper we present a set of methods for uniformly sampling the solid angle subtended by an oriented disk. We exploit the fact that a disk, as seen from a point, is bounded by an elliptical cone [Ebe99] and thus its solid angle defines a spherical ellipse whose properties have been analyzed in depth [Boo44]. This allows us to define two different exact area-preserving mappings that can be used to transform stratified unit-square sample patterns to stratified directions on the subtended spherical ellipse. We describe how to efficiently implement these mappings in practice and demonstrate the lower variance they achieve compared to previous work.

2. Problem Statement and Previous Work

Our goal is to compute the radiance L_s scattered at a point $\mathbf{o}$ in direction $\widehat{\omega}_o$ due to irradiance from a disk-shaped luminaire D. This can be written as an integral over the solid angle Ω_D subtended by the luminaire:

$$L_s(\mathbf{o}, \widehat{\omega}_o) = \int_{\Omega_D} f(\mathbf{o}, \mathbf{x}_{\mathbf{o}\widehat{\omega}}, \widehat{\omega}_o, \widehat{\omega}) \, d\mu(\widehat{\omega}), \qquad (1)$$

where $\mathbf{x}_{\mathbf{o}\widehat{\omega}}$ is the first visible point from $\mathbf{o}$ in direction $\widehat{\omega}$, μ is the solid angle measure, and the contribution function f is

$$f(\mathbf{o}, \mathbf{x}, \widehat{\omega}_o, \widehat{\omega}) = \begin{cases} L_e(\mathbf{x}, -\widehat{\omega}) f_s(\mathbf{o}, \widehat{\omega}_o, \widehat{\omega}) |\widehat{\omega} \cdot \hat{n}_\mathbf{o}|, & \text{if } \mathbf{o} \text{ is on a surface,} \\ L_e(\mathbf{x}, -\widehat{\omega}) \rho(\mathbf{o}, \widehat{\omega}_o, \widehat{\omega}) T(\mathbf{o}, \mathbf{x}), & \text{if } \mathbf{o} \text{ is in a medium,} \end{cases}$$

with f_s, $\hat{n}_\mathbf{o}$, and ρ being respectively the BSDF, surface normal, and medium phase function (times the scattering coefficient) at $\mathbf{o}$.
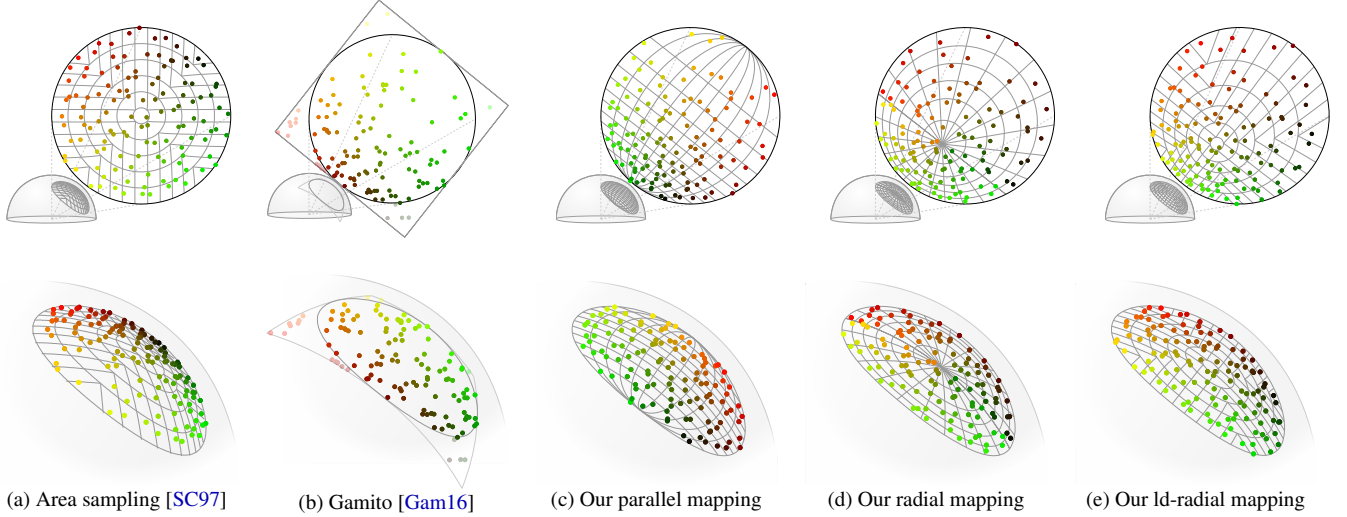
Figure 1: *A stratified unit-square sample pattern transformed onto the surface of a disk using existing techniques and our proposed maps (with solid angle projections on the bottom row). The points are colored according to their canonical $[0,1]^2$ coordinates to illustrate the continuity of the maps. Gamito's rejection sampling does not allow for direct stratification, so we show the candidate low-discrepancy pattern for that case.*

$L_e(\mathbf{x}, -\widehat{\omega})$ is the luminaire emission radiance at $\mathbf{x}$ in direction $-\widehat{\omega}$ and $T(\mathbf{o}, \mathbf{x})$ is the medium transmitance between $\mathbf{o}$ and $\mathbf{x}$.

Solid angle sampling. Monte Carlo estimation of Equation (1) using N randomly sampled directions $\widehat{\omega}_i$ has the following form:

$$L_s(\mathbf{o}, \widehat{\omega}_o) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{f(\mathbf{o}, \mathbf{x}_{\mathbf{o}\widehat{\omega}_i}, \widehat{\omega}_o, \widehat{\omega}_i)}{p(\widehat{\omega}_i)}, \quad (2)$$

where $p(\widehat{\omega})$ is the pdf for sampling $\widehat{\omega}$. The choice of sampling density p is important, since a lower variation of f/p makes the estimator more efficient [SWZ96]. For disk lights the traditional choice is uniform density over the luminaire surface D. This *area sampling* technique is easy to implement and its resulting solid angle pdf is $p(\widehat{\omega}) = \|\mathbf{o} - \mathbf{x}_{\mathbf{o}\widehat{\omega}}\| / (A(D) |\widehat{\omega} \cdot \widehat{n}_{\mathbf{x}_{\mathbf{o}\widehat{\omega}}}|)$, where $A(D)$ is the area of D. This pdf can lead to very high variance in the radiance estimator (2), especially when the point $\mathbf{o}$ is close to the luminaire. Our goal in this paper is to devise uniform *solid angle sampling* techniques that generate directions $\widehat{\omega}$ with constant density $p(\widehat{\omega}) = 1/|\Omega_D|$, yielding estimators with significantly lower variance than uniform area sampling.

Area-preserving mapping. Sample stratification can greatly improve the efficiency of Monte Carlo estimators [Shi91, SK13, PSC*15]. Most existing stratification techniques generate samples in the canonical unit square $[0,1]^2$, however our goal is to sample directions inside the solid angle Ω_D. Therefore, in order to take advantage of these techniques, we need to find a mapping M from $[0,1]^2$ to Ω_D such that for any two regions $R_1, R_2 \subseteq [0,1]^2$:

$$\frac{A(R_1)}{A(R_2)} = \frac{\mu(M(R_1))}{\mu(M(R_2))},$$

where A is the area measure, and μ is the solid angle measure as in Equation (1). We call such maps *area-preserving maps*. This key property makes it possible to generate stratified samples in Ω_D, because stratification is far more easily achieved in $[0,1]^2$.

Area-preserving solid angle maps have been developed for triangles [Arv95] and rectangles [UFK13]. For sampling the solid angles of disks, Gamito [Gam16] proposed to use a rectangle map [UFK13] followed by rejection sampling. This technique cannot be used with fixed-size canonical point sets, and needs a low-discrepancy sequence capable of progressively generating stratified sample candidates. The rejection sampling also makes it very difficult to achieve good high-dimensional stratification in the presence of other distributed effects, e.g. volumetric scattering, where the coordination of the sample patterns of different effects is desired. In this paper we focus on area-preserving maps for disks that do not require rejection sampling and work with any canonical sample pattern. Figure 1 compares our proposed maps against existing techniques.

For surface scattering points $\mathbf{o}$, an even better strategy is to importance sample the term $|\widehat{\omega} \cdot \widehat{n}_\mathbf{o}|$ in the contribution f. Such uniform sampling of the *projected solid angle* has been described by Arvo [Arv01] for polygonal lights. Extending our approach to projected solid angle sampling is an interesting avenue for future work.

3. Solid Angle Sampling of an Oriented Disk

We base our sampling techniques on the key observation that the projected area of any ellipse, including a disk, forms a spherical ellipse on the unit sphere around the shading point (Figure 2). Thus, in order to sample the solid angle subtended at point $\mathbf{o}$ by an oriented disk with center $\mathbf{c}$, normal $\widehat{n}$, and radius r, we will uniformly sample a point $\mathbf{q}$ on the spherical ellipse, and then backproject it to the disk.

Spherical ellipse. To compute the projected spherical ellipse, we first define a local reference frame for the disk $\mathcal{R}_d = (\widehat{x}_d, \widehat{y}_d, \widehat{z}_d)$:

$$\widehat{z}_d = -\widehat{n}, \qquad \widehat{y}_d = \widehat{z}_d \times \frac{\mathbf{c} - \mathbf{o}}{\|\mathbf{c} - \mathbf{o}\|}, \qquad \widehat{x}_d = \widehat{y}_d \times \widehat{z}_d. \quad (3)$$

We then take the boundary disk coordinates y_0 and y_1 w.r.t. the $\widehat{y}_d$ axis and project them onto the sphere (Figure 2, left). From the
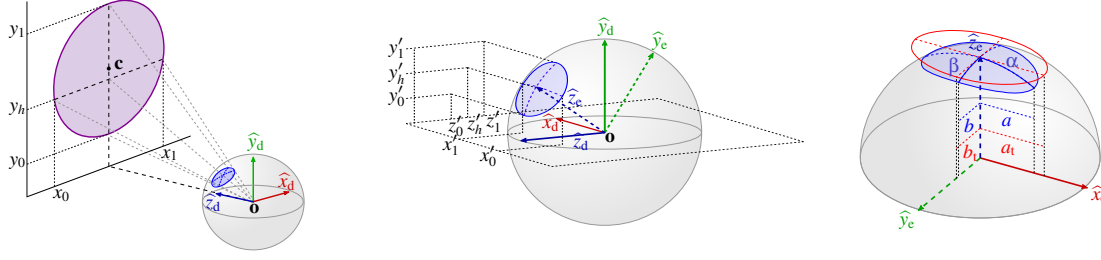
Figure 2: *Left:* The disk's local reference system $\mathcal{R}_d = (\widehat{x}_d, \widehat{y}_d, \widehat{z}_d)$ and the local coordinates required to characterize its solid angle projection. *Center:* Projections of the relevant coordinates onto the unit sphere, defining the spherical ellipse and its local reference system $\mathcal{R}_e = (\widehat{x}_e, \widehat{y}_e, \widehat{z}_e)$, where $\widehat{x}_e \equiv \widehat{x}_d$. *Right:* The spherical ellipse is defined by either its semi-arcs α and β or its semi-axes a and b in $\mathcal{R}_e$. The tangent ellipse (in red), which lies on a plane tangent to the sphere at the ellipse center, is defined by its semi-axes a_t and b_t.

coordinates y'_0, y'_1, z'_0, z'_1 of these projections (Figure 2, middle) we can compute the spherical ellipse center: it is the result $\widehat{z}_e$ of normalizing the vector $(0, y'_h, z'_h)$, where $y'_h = (y'_0 + y'_1)/2$ and $z'_h = (z'_0 + z'_1)/2$. (Note that $\widehat{z}_e$ in general does not coincide with the spherical projection of the disk center $\mathbf{c}$.) Reprojecting $\widehat{z}_e$ onto the disk, the obtained y_h coordinate defines a chord $\overline{x_0 x_1}$ parallel to $\widehat{x}_d$. The chord endpoint projections onto the sphere, with coordinates x'_0 and x'_1, allow us to compute the lengths of the ellipse's semi-axes, a and b, and semi-arcs, α and β (Figure 2, right):

$$a = x'_1, \qquad b = \frac{1}{2}\sqrt{(y'_1 - y'_0)^2 + (z'_1 - z'_0)^2}, \qquad (4)$$

$$\alpha = \sin^{-1} a, \qquad \beta = \sin^{-1} b. \qquad (5)$$

Finally, from α and β we can compute the semi-axes $a_t = \tan \alpha$ and $b_t = \tan \beta$ of the ellipse tangent to the sphere at $\widehat{z}_e$ (Figure 2, right).

In the following, we use both the spherical and the tangent ellipses to derive two different mappings for uniformly sampling points $\mathbf{q}$ on the spherical ellipse which we then map to the surface of the disk. These mappings operate in a coordinate system $\mathcal{R}_e = (\widehat{x}_e, \widehat{y}_e, \widehat{z}_e)$, where $\widehat{x}_e \equiv \widehat{x}_d$ and $\widehat{y}_e = \widehat{z}_e \times \widehat{x}_e$, shown in Figures 2 and 3.
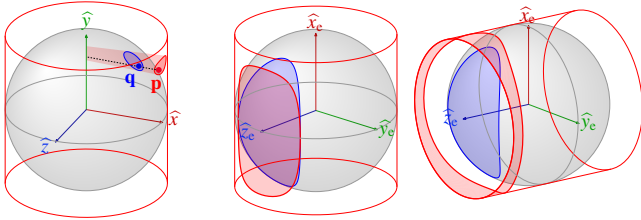

Figure 3: *Left*: Any region (blue) on the unit sphere can be radially projected to another region (red) onto the cylinder (aligned with any axis, here $\widehat{y}$). Any point $\mathbf{q}$ on the sphere can be expressed in cylindrical coordinates as (ϕ, h, r) (angle about axis, distance along axis, distance from center). This point can be mapped to a point $\mathbf{p}$ on the unit cylinder with coordinates (ϕ, h). This mapping preserves the areas of both finite and differential regions. Thus, to obtain a point in the spherical region, we can sample in the cylindrical region and project back onto the sphere. *Center*: A spherical ellipse (blue, with center on $\widehat{z}_e$ axis) projected onto a $\widehat{y}_e$-axis aligned cylinder. *Right*: The same spherical ellipse projected onto a $\widehat{z}_e$-axis aligned cylinder. In this case, the projected region (red) has a ring-like shape.

3.1. Area-preserving mappings

Our new mappings are based on a generalization of the so-called *Archimedes Hat-Box theorem*, which states that the area of a region on a sphere between two parallels is equal to the area of that region's cylindrical projection. This area-preserving property also holds for any arbitrarily shaped region on the sphere (Figure 3, left). The latter property was used by Tobler et al. [TNSP98] to define compact metallic BRDFs (they provide a demonstration). It was also used (although not explicitly stated) by Arvo [Arv95] and Ureña et al. [UFK13] to define area-preserving parameterizations for spherical triangles and rectangles, respectively.

We use this property to derive our area-preserving mappings. We use that the center of the spherical ellipse is on the $\widehat{z}_e$ axis. The ellipse can be radially projected onto a cylinder, obtaining a *cylindrical ellipse*. Two different unit-radius cylinders can be used. One is aligned with the $\widehat{y}_e$ axis (Figure 3, center), which we call a *parallel map*. The second one is aligned with the $\widehat{z}_e$ axis (Figure 3, right), which we call a *radial map*. We also propose a variant of the radial map that uses Shirley's low-distortion map [SC97], which we call a *low-distortion radial map*, or *ld-radial map*.

Maps overview. The basic idea behind our maps is to first select a point $\mathbf{p}$ on the cylindrical ellipse, as a function of $(\varepsilon_1, \varepsilon_2) \in [0,1]^2$, and then project it back onto the sphere (perpendicularly to the cylinder axis; see Figure 3, left) to get the point $\mathbf{q}$. Let (ϕ, h) be the cylindrical coordinates of $\mathbf{p}$. First, the angle ϕ about the cylinder axis is obtained by finding the slice of the cylindrical ellipse whose solid angle is $\varepsilon_1 \Omega_D$ (Figure 4). With ϕ fixed, the altitude h is computed as a simple linear interpolation using ε_2 along the lateral line segment that is the intersection between the lateral plane at angle ϕ and the cylindrical ellipse (green segment in Figure 4). The sampling of ϕ involves numerical inversion of incomplete elliptic integrals, as we show next.

3.2. Parallel Mapping

Our parallel mapping, whose cylinder axis is aligned with $\widehat{x}_e$, operates by considering a portion (sector) of the cylindrical ellipse – the red-shaded region in Figure 4, left. This region is determined by the green line segment, whose endpoints have cylindrical coordinates (ϕ_p, h_p) and $(\phi_p, -h_p)$. The angle ϕ_p goes from $-\beta$ to β, as the spherical ellipse is centered on the $\widehat{z}_e$ axis.

Due to the Hat-Box theorem, the differential solid angle covered by the green segment is equal to its length, $2h_p$. This length is in fact a function of ϕ_p. Thus, the solid angle subtended by the red region onto the spherical ellipse (the blue region in Figure 4, left) can be written as the integral of the segment length:

$$\Omega_p(\phi_p) = \int_{-\beta}^{\phi_p} 2h_p(\phi_p') \, d\phi_p', \qquad (6)$$

where the full solid angle of the spherical ellipse is $\Omega_D = \Omega_p(\beta)$. Due to symmetry, for any angle $\phi_p \in [0, \beta]$ it holds $h_p(-\phi_p) = h_p(\phi_p)$. We use this to express $\Omega_p(\phi_p)$ as a sum of integrals $\Omega_p^+(\phi_p)$ over positive angles:

$$\Omega_p(\phi_p) = \begin{cases} \Omega_p^+(\beta) + \Omega_p^+(\phi_p) & : \phi_p \geq 0 \\ \Omega_p^+(\beta) - \Omega_p^+(|\phi_p|) & : \phi_p < 0 \end{cases}, \qquad (7)$$

where

$$\Omega_p^+(\phi_p) = \int_0^{\phi_p} 2h_p(\phi_p') \, d\phi_p'. \qquad (8)$$

In Appendix A we derive an expression for $h_p(\phi_p)$:

$$h_p(\phi_p) = c_t \sqrt{\frac{1 - (p+1)\sin^2 \phi_p}{1 - (mp+1)\sin^2 \phi_p}}, \qquad (9)$$

where

$$p = \frac{1}{b_t^2}, \qquad m = \frac{a_t^2 - b_t^2}{a_t^2 + 1}, \qquad c_t = \frac{a_t}{\sqrt{1 + a_t^2}}. \qquad (10)$$

Substituting Equation (9) into Equation (8) and simplifying, we get:

$$\Omega_p^+(\phi_p) = 2b_t c_t \left[p \, F(\varphi_p|m) - (p+1) \Pi\left(-b_t^2; \varphi_p|m\right) \right], \qquad (11)$$

where $F(\varphi|m)$ and $\Pi(n; \varphi|m)$ are Legendre incomplete elliptic integrals of respectively the first and third kind, $m \in [0, 1)$, and

$$\varphi_p = -\sin^{-1}\left(\frac{\tan \phi_p}{b_t}\right). \qquad (12)$$

Unfortunately, no closed-form expressions are known for $F(\varphi|m)$ and $\Pi(n; \varphi|m)$, so Equation (11) must be evaluated numerically.

Sampling. With the fractional spherical ellipse area Ω_p characterized, we can map a point on the unit square $(\varepsilon_1, \varepsilon_2) \in [0, 1]^2$ to a point on the spherical ellipse **q**. We first need to find the angle ϕ_p that satisfies $\Omega_p(\phi_p) = \varepsilon_1 \Omega_D$, for which we need to evaluate the inverse function Ω_p^{-1}. This function has no analytical form, so we resort to numerically finding the roots of the equation

$$\Omega_p(\phi_p) - \varepsilon_1 \Omega_D = 0. \qquad (13)$$

Having sampled ϕ_p, we obtain the point **p** on the cylindrical ellipse by computing $h_p(\phi_p)$ using Equation (9) and linearly interpolating the altitude coordinate between $-h_p(\phi_p)$ and $h_p(\phi_p)$ using ε_2:

$$\mathbf{p} = (\phi_p, (2\varepsilon_2 - 1)h_p(\phi_p)) = (\phi_p, h). \qquad (14)$$

Finally, the corresponding point **q** on the ellipse is obtained by radially projecting **p** onto the sphere (see Figure 3, left):

$$\mathbf{q} = \left(h, \sqrt{1 - h^2} \sin \phi_p, \sqrt{1 - h^2} \cos \phi_p\right). \qquad (15)$$
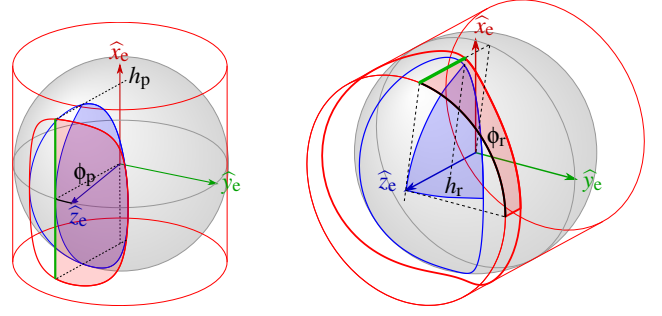
Figure 1c shows the resulting map.

Figure 4: Illustration of our parallel (*left*) and radial (*right*) maps. Given a canonical sample $(\varepsilon_1, \varepsilon_2) \in [0; 1]^2$, we first find the angle ϕ_p, respectively ϕ_r, that cuts a region of the cylindrical ellipse, via the green line segment, with area $\varepsilon_1 \Omega_D$ (in red). A sample point on the cylinder is then obtained by linearly interpolating between the green segment endpoints using ε_2. For the parallel map (*left*), the endpoint cylindrical coordinates are $(\phi_p, -h_p)$ and (ϕ_p, h_p), with $\phi_p \in [-\beta, \beta]$. For the radial map (*right*) these coordinates are (ϕ_r, h_r) and $(\phi_r, 1)$, with $\phi_r \in [0, 2\pi]$ (we use $\phi_r \in [0, \pi/2]$ in each quadrant).

3.3. Radial Mapping

The parallel mapping presented in Section 3.2 involves two elliptic integrals and introduces noticeable distortions (high variations in the angles), which can increase discrepancy and ruin any blue noise properties of the unit-square samples. In this section we present an alternative radial mapping that uses a single elliptical integral and also exhibits less distortion. It is based on the analysis of the spherical ellipse topology by Booth [Boo44].

We will exploit the fact that the four quadrants in the spherical ellipse are radially symmetric (see Figure 2, right), so the total area can be expressed as $\Omega_D = 4\Omega_r$, with Ω_r being the area of each quadrant. Within a quadrant, the azimuth angle is $\phi_r \in [0, \pi/2]$.

We now consider a bounding cylinder aligned with the $\widehat{z}_e$ axis (Figure 4, right). Specifically, we are interested in the lateral region (in red in Figure 5) that is the radial projection of a fraction of the spherical quadrant. This region is determined by the position of the green line segment whose endpoints have cylindrical coordinates (ϕ_r, h_r) and $(\phi_r, 1)$. The segment length is $1 - h_r$, which is a function of ϕ_r. Similarly to Equation (6), we use the Hat-Box theorem to express the fractional quadrant area as the integral of this length:

$$\Omega_r(\phi_r) = \int_0^{\phi_r} \left[1 - h_r(\phi_r')\right] d\phi_r' = \phi_r - \int_0^{\phi_r} h_r(\phi_r') \, d\phi_r'. \qquad (16)$$

Using the Pythagorean theorem, we express $h_r(\phi_r)$ as (see Figure 5)

$$h_r(\phi_r) = \sqrt{1 - r^2(\phi_r)}, \qquad (17)$$

where $r(\phi_r)$ is the (planar) elliptical radius of the spherical ellipse:

$$r(\phi_r) = \frac{ab}{\sqrt{a^2 \sin^2 \phi_r + b^2 \cos^2 \phi_r}}, \qquad (18)$$

which we derive in Appendix B.

Plugging Equations (17) and (18) back into Equation (16), and using Booth's derivations [Boo44], we can now express the fractional quadrant area $\Omega_r(\phi_r)$ using Legendre's incomplete elliptic

integral of the third kind $\Pi(n;\varphi|m)$, so it becomes

$$\Omega_r(\phi_r) = \phi_r - \frac{b(1-a^2)}{a\sqrt{1-b^2}}\,\Pi(n;\varphi_r|m), \qquad (19)$$

where

$$\varphi_r = \tan^{-1}\left(\frac{a_t}{b_t}\tan\phi_r\right), \quad n = \frac{a^2-b^2}{a^2(1-b^2)}, \quad m = \frac{a^2-b^2}{1-b^2}. \quad (20)$$

Above, φ_r is the parametric angle of the tangent ellipse, and n and m are the elliptic characteristic and module that characterize the elliptic integral [Boo52].

Unfortunately, as with Equation (11), no closed-form expression is known for an incomplete third-kind elliptic integral in the general case, so we need to evaluate Equation (19) numerically.

Direct radial mapping. Having an expression for the fractional spherical ellipse Ω_r, we can map a unit-square point $(\varepsilon_1,\varepsilon_2) \in [0,1]^2$ to a point $\mathbf{q}$ on the spherical ellipse. Below we only consider sampling the first ellipse quadrant (shown in Figure 5); the entire ellipse can be covered by flipping the $\widehat{x}_e$- and $\widehat{y}_e$-coordinates of $\mathbf{q}$.

First, we need to find the azimuth angle $\phi_r \in [0,\pi/2]$ satisfying

$$\Omega_r(\phi_r) - \varepsilon_1\,\Omega_r(\pi/2) = 0. \qquad (21)$$

Since we do not have a method to analytically invert $\Omega_r(\phi_r)$, we compute ϕ_r by numerically finding the roots of the above equation.

Having sampled ϕ_r, we obtain point $\mathbf{p}$ on the cylindrical ellipse by computing $h_r(\phi_r)$ using Equation (17) and linearly interpolating the altitude coordinate between $h_r(\phi_r)$ and 1 using ε_2:

$$\mathbf{p} = \big(\phi_r, (1-\varepsilon_2)h_r(\phi_r)+\varepsilon_2\big) = (\phi_r, h). \qquad (22)$$

We find the corresponding point $\mathbf{q}$ on the ellipse by projecting $\mathbf{p} = (\phi_r, h)$ using Equation (15) with swapped $\widehat{x}_e$- and $\widehat{z}_e$-coordinates. Figure 1d shows the resulting map.

Low-distortion radial mapping. The direct mapping from above resembles the classical 2D polar mapping from Cartesian coordinates, as seen in Figure 1d. As such, it exhibits the same distortion – the lines converging at the ellipse center, which does not preserve relative distances between samples and damages their stratification. In the planar case, the mapping of Shirley and Chiu [SC97] alleviates this distortion by instead warping concentric squares into concentric disks. To achieve the analogous mapping on the spherical ellipse, we first warp our input unit-square samples $(\varepsilon_1,\varepsilon_2)$ to the unit disk using Shirley and Chiu's concentric mapping. We then move back to the unit square using the *inverse 2D polar mapping*:

$$u = \begin{cases} 2\dfrac{\theta}{\pi} & : \theta \in [0,\frac{\pi}{2}) \\[2mm] 1 - 2\dfrac{\theta-\pi/2}{\pi} & : \theta \in [\frac{\pi}{2},\pi) \\[2mm] 2\dfrac{\theta-\pi}{\pi} & : \theta \in [\pi,\frac{3\pi}{2}) \\[2mm] 1 - 2\dfrac{\theta-3/2\pi}{\pi} & : \theta \in [\frac{3\pi}{2},2\pi) \end{cases}, \qquad v = r^2. \quad (23)$$

The result of this detour is a unit-square point set that, when warped using the classical 2D polar mapping, exhibits the well-stratified disk distribution of Shirley and Chiu. We, instead, feed this point set to the direct radial mapping to produce a distortion-free distribution on the spherical ellipse. Figure 1e shows this distribution.
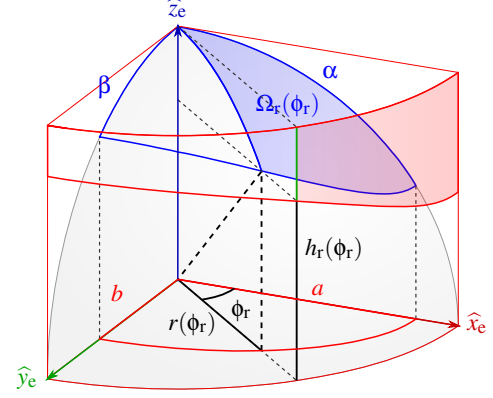
Figure 5: Illustration of our polar mapping. Projected perpendicularly to its axis onto the $\widehat{x}_e\widehat{y}_e$ plane, the spherical ellipse forms a planar ellipse with semi-axes $a = \sin\alpha$ and $b = \sin\beta$. We use the Pythagorean theorem to express the altitude $h_r(\phi_r)$ of the cylindrical projection of the spherical ellipse's curve in terms of a, b, and ϕ_r.

4. Implementation

We have implemented our maps as custom sampling procedures for disk lights in two different systems: the Mitsuba renderer [Jak10] and the Arnold production renderer [Faj10].

In order to sample from each mapping, we need to find the roots of Equations (13) and (21) respectively. Since the elliptic integrals they contain prevent analytical inversion, we resort to numerical root finding using an iterative Newton-Raphson method. However, this method can become very expensive, since for each iteration we need to numerically evaluate two and one incomplete elliptic integrals (for the parallel and radial mappings, respectively).

4.1. Tabulation

In order to reduce the significant cost of Newton-Raphson over area sampling (up to $10\times$ in simple scenes; see Figure 6) and avoid the expensive numerical inversion, we approximate our radial mapping by tabulating Equation (19). We choose to tabulate this mapping as it introduces less distortion in the output sample distribution than the parallel one, as shown in Figure 1.

We can write the fractional solid angle $\Omega_r' = \Omega_r(\phi_r)/\Omega_r(\pi/2)$ as a function of $\alpha \in [0,\pi/2]$, $\beta \in [0,\alpha]$ and $\phi_r \in [0,\pi/2]$. We can tabulate this function by discretizing each of the three parameters, producing a 3D array of values. A quick binary search based on ϕ_r (combined with interpolation) then allows us to get approximate values of Ω_r' with good accuracy. However, storing such a table would require a large amount of memory. To address this, we reparametrize Ω_r' in terms of α, the ratio $\beta' = \beta/\alpha$ (which is in $[0,1]$), and ϕ_r. This version of Ω_r' has very low variation w.r.t. to α, so we can remove this parameter altogether, reducing the tabulation to a 2D array of Ω_r' values for a set of β' and ϕ_r values. Each entry in this array corresponds to a spherical triangle defined by the fraction of ϕ_r covered by the given entry and the value of $\theta(\phi_r)$ at the start of the entry's interval, which can be easily sampled [Arv95]. This approximation causes some generated samples to lie outside the

spherical ellipse, which we reject. Note that this rejection ensures unbiasedness; however, for practical reasons our production renderer implementation simply assigns zero weight to such invalid samples, resulting in a slight understimation of the illumination. We have found the rejection ratio to be negligible, the storage requirement low, and the accuracy satisfactory for realistic rendering. In our implementation we use a 2D table with resolution 1024^2, which we found to be accurate enough to provide an insignificant difference in variance compared to the analytic solution. Finally, note that in order to compute the samples' pdf $p(\widehat{\omega}) = 1/|\Omega_D|$, we still need to compute Ω_D numerically. This computation is amortized among all samples for a given shading point.

4.2. Efficiency

Similarly to existing solid angle sampling techniques (e.g. for spherical triangles [Arv95] and rectangles [UFK13]), the cost of drawing a sample with our technique is higher than that of uniform area sampling (though some of it is amortized over multiple samples). This overhead pays off when the luminaire is close to the shading point (i.e. the subtended solid angle is relatively large).

As an optimization, our Arnold implementation (Figure 9) employs a simple heuristic to switch to uniform area sampling when the luminaire is far away. In order to provide a fair comparison against existing techniques (in Figures 7 and 8), our Mitsuba implementation does not take advantage of this optimization.

Even with the above tabulation scheme, we still need to compute the solid angle of the spherical ellipse Ω_D for the sampling pdf, using either Equation (11) or (19). The elliptical integrals involved can be computed using Carlson's fast numerical algorithms [Car95].

5. Results

Figures 7 and 8 show a comparison between traditional area sampling [SC97], Gamito's rejection-based solid angle sampling [Gam16] and our techniques (Mitsuba implementation), without and with the presence of participating media. In both figures only direct illumination (single scattering) is computed, using 16 samples/pixel. Inside a medium, we first sample a point along the ray via equiangular sampling [KF12] w.r.t. the disk light center, and then use the corresponding disk sampling technique to generate a point on the light. In the case of uniform area sampling, a better strategy is to first sample the light surface and then perform equiangular sampling w.r.t. that chosen point. We therefore include this strategy in Figure 8 (called "Area sampling (first)"), which is incompatible with the solid angle mappings. The results show that our sampling methods outperform Gamito's method on surfaces, and perform at least on par in participating media, where variance due to medium sampling dominates when using solid angle sampling. In all cases, area sampling yields much higher variance. In the supplemental document we provide global illumination comparisons between our tabulated radial sampling and Mitsuba's built-in disk area sampling.

Figure 6 shows a comparison between the convergence and the cost of the different techniques from Figures 7 and 8. For the same number of samples, our mappings produce lower error than Gamito on surfaces and perform virtually identically in media. In terms of
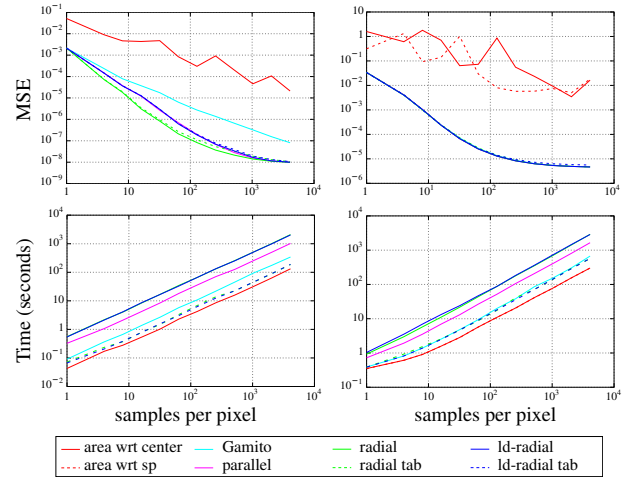

Figure 6: Error (*top*) and cost (*bottom*) for the results shown in Figure 7 (*left*) and Figure 8 (*right*), w.r.t. sample count.

cost, our tabulated version is almost as fast as area sampling, and the fully numerical implementation can be up to $10\times$ slower. Note that the performance of the iterative numerical inversion depends on the geometrical configuration: the starting point for the inversion affects the number of iterations required for convergence. The parallel and radial mappings take respectively 1-3 and 1-4 Newton-Raphson iterations in our tests. Also note that in scenes with higher geometric and shading complexity, the relative cost of the different methods has less impact on the overall rendering performance.

Finally, Figure 9 shows a scene rendered in Arnold, comparing our tabulated radial map implementation to the renderer's built-in uniform area sampling. The scene features many production features, including high-resolution texture maps, fur, displacement, subsurface scattering, indirect surface and volume-to-surface light transport. In such cases the higher cost of our technique has a negligible impact on the total rendering performance. With 256 samples/pixel our tabulated radial map yields a noise-free image, while area sampling suffers from a substantial amount of noise.

6. Conclusions

We have presented two new area-preserving mappings that enable the uniform solid angle sampling of oriented disk light sources. Following the key observation that this solid angle is a spherical ellipse, we make use of the Hat-Box theorem to transform canonical unit-square sample points onto the spherical ellipse in a way that preserves their stratification. To avoid costly numerical inversion, we develop a practical mapping tabulation that introduces little overhead over traditional uniform area sampling [SC97] while significantly reducing the variance of the illumination estimate. Our mappings are also competitive to existing disk solid angle sampling techniques [Gam16], without imposing restrictions on the sample generator.

As a by-product of our work, we have proposed two new expressions for the subtended solid angle of a disk, which in addition to graphics is important in other fields such as particle transport. In this context, most previous analytic formulations [Pax59,TPK*07,Con10] have included at least two incomplete elliptic integrals that need to be computed numerically. In contrast, our radial formulation, based on Booth's spherical topology analysis [Boo44], involves only one elliptic integral, making it more simple and practical than previous work.

While this work only considers circular disks, our approach could be extrapolated to other shapes whose subtended solid angle is also an ellipse, such as elliptical disks and ellipsoids [Hei17]. Including these geometries would only require finding the spherical ellipses subtended by them. Moreover, our mappings could reduce variance of other shapes such as cylinders, following Gamito [Gam16].

The main limitation of our method is the lack of analytical inversion of the proposed mappings, which requires using either costly numerical inversion or tabulation. Unfortunately, it seems impossible to find a spherical ellipse mapping that does not involve incomplete elliptic integrals, whose inversion is unknown. The presented mappings also only consider the solid angle, but not the other contribution terms in Equation (1), e.g. the BRDF or the foreshortening term. Developing methods for including at least some of these other terms is an interesting direction for future work. Furthermore, while our mappings are nearly optimal for uniformly emitting disk luminaires, it would be interesting to take into account spatially-varying emission profiles, in the spirit of the work of Bitterli et al. [BNJ15].

Acknowledgements

References

[Arv95] ARVO J.: Stratified sampling of spherical triangles. In *Proceedings of SIGGRAPH '95* (1995), ACM, pp. 437–438. 1, 2, 3, 5, 6

[Arv01] ARVO J.: Stratified sampling of 2-manifolds. *SIGGRAPH 2001 Course Notes 29*, 2 (2001). 1, 2

[BNJ15] BITTERLI B., NOVÁK J., JAROSZ W.: Portal-masked environment map sampling. *Computer Graphics Forum 34*, 4 (2015), 13–19. 7

[Boo44] BOOTH J.: On the rectification and quadrature of the spherical ellipse. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 25*, 163 (1844), 18–38. 1, 4, 7

[Boo52] BOOTH J.: Researches on the geometrical properties of elliptic integrals. *Philosophical Transactions of the Royal Society of London 142* (1852), 311–416. 5

[Car95] CARLSON B. C.: Numerical computation of real or complex elliptic integrals. *Numerical Algorithms 10*, 1 (1995), 13–26. 6

[Con10] CONWAY J. T.: Analytical solution for the solid angle subtended at any point by an ellipse via a point source radiation vector potential. *Nuclear Instruments and Methods in Physics Research 614*, 1 (2010), 17–27. 7

[Ebe99] EBERLY D.: Perspective projection of an ellipse, 1999. 1

[Faj10] FAJARDO M.: Ray tracing solution for film production rendering. *ACM SIGGRAPH Talks: Global Illumination Across Industries* (2010). 5

[Gam16] GAMITO M. N.: Solid angle sampling of disk and cylinder lights. *Computer Graphics Forum 34*, 4 (2016). 1, 2, 6, 7, 8

[Hei17] HEITZ E.: Analytical calculation of the solid angle subtended by an arbitrarily positioned ellipsoid to a point source. *Nuclear Instruments and Methods in Physics Research 852* (2017). 7

[HKWB09] HAŠAN M., KŘIVÁNEK J., WALTER B., BALA K.: Virtual spherical lights for many-light rendering of glossy scenes. *ACM Trans. Graph. 28*, 5 (dec 2009), 143:1–143:6. 1

[Jak10] JAKOB W.: Mitsuba renderer, 2010. http://www.mitsuba-renderer.org. 5

[KF12] KULLA C., FAJARDO M.: Importance sampling techniques for path tracing in participating media. *Computer Graphics Forum 31*, 4 (2012), 1519–1528. 6

[Pax59] PAXTON F.: Solid angle calculation for a circular disk. *Review of Scientific Instruments 30*, 4 (1959), 254–258. 7

[PSC*15] PILLEBOUE A., SINGH G., COEURJOLLY D., KAZHDAN M., OSTROMOUKHOV V.: Variance analysis for Monte Carlo integration. *ACM Trans. Graph. 34*, 4 (2015), 124. 2

[SC97] SHIRLEY P., CHIU K.: A low distortion map between disk and square. *Journal of graphics tools 2*, 3 (1997), 45–52. 2, 3, 5, 6, 8

[SHD15] SIMON F., HANIKA J., DACHSBACHER C.: Rich-VPLs for improving the versatility of many-light methods. *Computer Graphics Forum 34*, 2 (2015), 575–584. 1

[Shi91] SHIRLEY P.: Discrepancy as a quality measure for sample distributions. In *Proceedings of Eurographics '91* (1991), vol. 91, pp. 183–194. 2

[SK13] SUBR K., KAUTZ J.: Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Trans. Graph. 32*, 4 (jul 2013), 128:1–128:12. 2

[SWZ96] SHIRLEY P., WANG C., ZIMMERMAN K.: Monte Carlo techniques for direct lighting calculations. *ACM Trans. Graph. 15*, 1 (jan 1996), 1–36. 1, 2

[TNSP98] TOBLER R. F., NEUMANN L., SBERT M., PURGATHOFER W.: A new form factor analogy and its application to stochastic global illumination algorithms. In *Proceedings of EGSR '98* (1998), pp. 35–44. 3

[TPK*07] TIMUS D., PRATA M., KALLA S., ABBAS M., ONER F., GALIANO E.: Some further analytical results on the solid angle subtended at a point by a circular disk using elliptic integrals. *Nuclear Instruments and Methods in Physics Research 580*, 1 (2007), 149–152. 7

[UFK13] UREÑA C., FAJARDO M., KING A.: An area-preserving parametrization for spherical rectangles. *Computer Graphics Forum 32*, 4 (2013), 59–66. 1, 2, 3, 6

[Ure00] UREÑA C.: Computation of irradiance from triangles by adaptive sampling. *Computer Graphics Forum 19*, 2 (2000), 165–171. 1

[Wan92] WANG C.: Physically correct direct lighting for distribution ray tracing. In *Graphics Gems III* (1992), Academic Press Professional, Inc., pp. 307–313. 1

(a) Area sampling [SC97] (b) Gamito [Gam16] (c) Our parallel mapping (d) Our radial mapping (e) Our ld-radial mapping
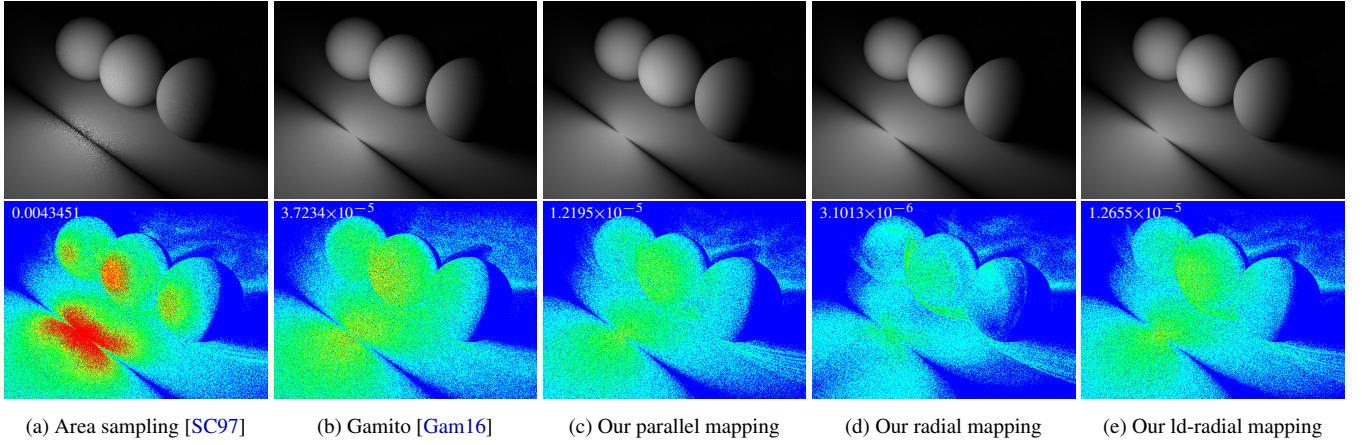
Figure 7: *Top*: Scene illuminated by a double-sided disk light, rendered with 16 samples/pixel. The light is perpendicular to the ground and is invisible to camera rays. *Bottom*: False-color differences and MSE w.r.t. to a reference image computed with 32K samples/pixel.

(a) Area sampling [SC97] (b) Area sampling (first) (c) Gamito [Gam16] (d) Our parallel mapping (e) Our radial mapping (f) Our ld-radial mapping
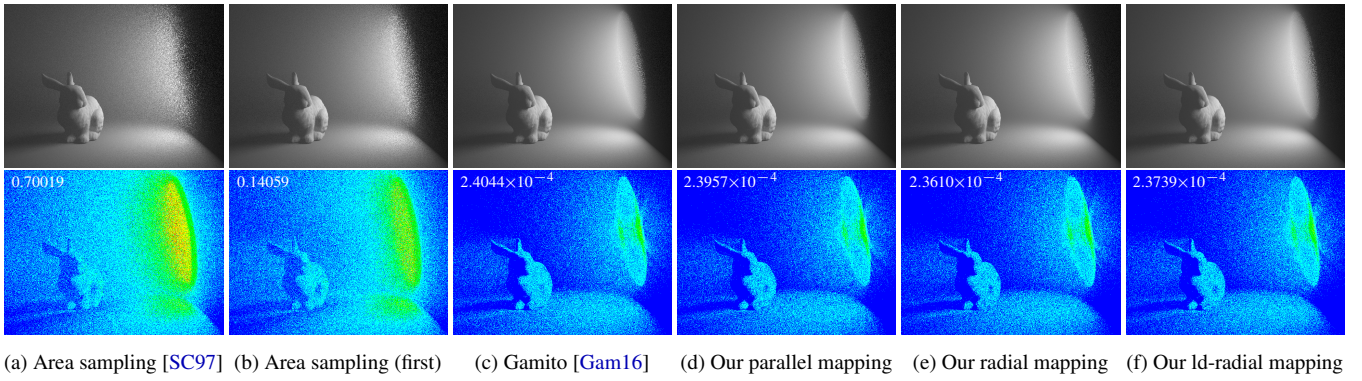
Figure 8: *Top*: A scene with a participating medium illuminated by a single-sided disk light (invisible), rendered with 16 samples/pixel. Please refer to Section 5 for details. *Bottom*: False-color differences and MSE w.r.t. to a reference image computed with 64K samples/pixel.
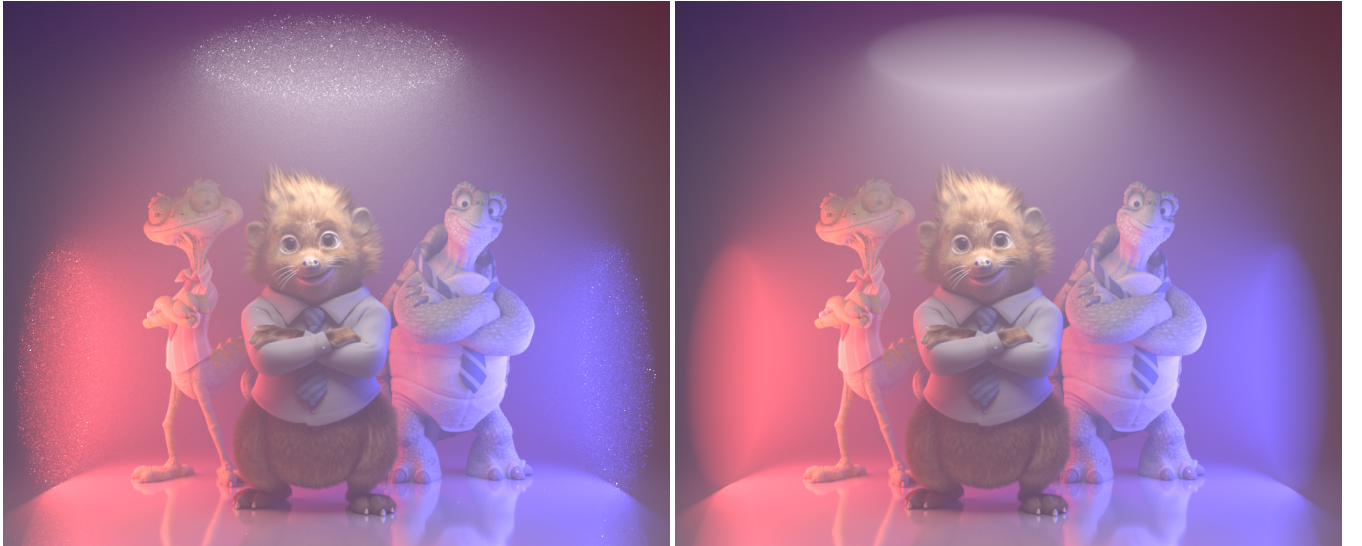

Figure 9: Colored disk lights (invisible) rendered in Arnold using area sampling (*left*) and our tabulated radial map (*right*) with 256 samples/pixel. Due to the use of complex surface and hair shaders, the higher cost of our technique has a negligible impact on performance.

Appendix A: Derivation of Equation (9)

Here we derive the expression for $h_p(\phi_p)$ in Equation (9), whose integral we then express as a combination of incomplete elliptic integral functions in Equation (11). We use the tangent ellipse, shown in Figure 10 and introduced in Section 3 and Figure 2, right. The ellipse semi-axes $a_t \geq b_t$ are aligned with $\widehat{x}_e$ and $\widehat{y}_e$, respectively.

For any angle $\phi_p \in [-\beta, \beta]$, we first obtain a coordinate $y = \tan \phi_p$ along the $\widehat{y}_e$ axis. (We only consider $\phi_p > 0$, thus $y > 0$, and convert negative ϕ_p to positive using symmetry, as described in Section 3.2.) Using the ellipse equation $(x/a_t)^2 + (y/b_t)^2 = 1$, we can get the corresponding $x \geq 0$ coordinate along $\widehat{x}_e$ as a function of y:

$$x = a_t \sqrt{1 - (y/b_t)^2}. \tag{24}$$

We then consider the point $\mathbf{t} = (x, y, 1)$ on the tangent ellipse and its spherical projection $\mathbf{s} = \mathbf{t}/\|\mathbf{t}\|$. The $\widehat{x}_e$-coordinate of $\mathbf{s}$, and also of its cylindrical projection (see Figure 4, left), is

$$h_p = \frac{x}{\sqrt{x^2 + y^2 + 1}}. \tag{25}$$

Substituting Equation (24) into (25):

$$h_p = a_t \frac{\sqrt{1 - py^2}}{\sqrt{y^2 + 1 + a_t^2 (1 - py^2)}} = c_t \sqrt{\frac{1 - py^2}{1 - mpy^2}} \tag{26}$$

where p, m and c_t are as in Equation (10).

Using $y = \tan \phi_p$ and $0 \leq \phi_p \leq \beta \leq \pi/2$, in Equation (26) we can substitute y^2 by $(\sin^2 \phi_p)/(1 - \sin^2 \phi_p)$. With this we can finally write h_p explicitly as a function of ϕ_p:

$$h_p(\phi_p) = c_t \sqrt{\frac{1 - (p+1)\sin^2 \phi_p}{1 - (mp+1)\sin^2 \phi_p}} \tag{27}$$

which is exactly Equation (9).

Appendix B: Derivation of Equation (18)

Here we derive the expression for $r(\phi_r)$ in Equation (18), which is used in the radial mapping (Section 3.3). We consider the planar ellipse resulting from the parallel projection of the spherical ellipse onto the $\widehat{x}_e \widehat{y}_e$ plane. This ellipse's semi-major and semi-minor axes are a and b, respectively (see Figure 5).

Consider a point $\mathbf{r}$ whose polar coordinates $(\phi_r, r(\phi_r))$ and Cartesian coordinates (x, y) are related as

$$x = r(\phi_r) \cos \phi_r, \qquad y = r(\phi_r) \sin \phi_r. \tag{28}$$

We want to define $r(\phi_r)$ in such a way that $\mathbf{r}$ is on the planar ellipse curve. Thus, x and y must obey the ellipse equation, i.e.

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1. \tag{29}$$

We can substitute x and y from Equation (28) into the ellipse equation, resulting in

$$r(\phi_r)^2 \left(\frac{b^2 \cos^2 \phi_r}{b^2 a^2} + \frac{a^2 \sin^2 \phi_r}{a^2 b^2}\right) = 1. \tag{30}$$

We can thus write

$$r(\phi_r) = \frac{ab}{\sqrt{b^2 \cos^2 \phi_r + a^2 \sin^2 \phi_r}}, \tag{31}$$
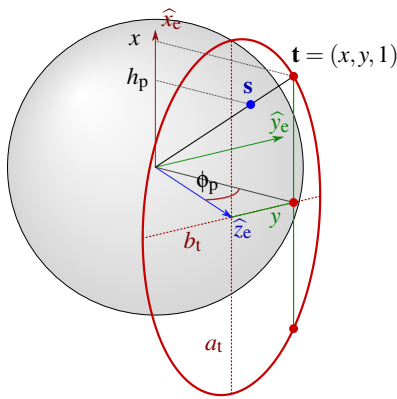
which is Equation (18).

Figure 10: A view of the tangent ellipse (red). For a given angle ϕ_p, we first obtain the coordinate $y = \tan \phi_p$ along axis $\widehat{y}_e$. Using the ellipse equation, we then find the corresponding coordinate x along $\widehat{x}_e$. This gives point $\mathbf{t}$ on the tangent ellipse, whose spherical projection $\mathbf{s}$ has $\widehat{x}_e$-coordinate h_p – the quantity we are interested in.