APPENDIX E: HUYGHENS-FRAUNHOFER-KIRCHHOFF APPROXIMATION

We shall use the WKB (eikonal) approximation up to the exit surface of the lens, but construct a solution of the wave equation which is better than the WKB expression in the space beyond the lens. This requires input of just the WKB values for U and ∇U values at the exit surface of the lens. The solution beyond the lens is provided by Green's theorem:

$$U(\mathbf{x}) = \frac{1}{4\pi}\int_{\Sigma_0} d\mathbf{\Sigma_0}\cdot[U(\mathbf{x_0})\nabla_0 G(\mathbf{x},\mathbf{x}_0) - G(\mathbf{x},\mathbf{x}_0)\nabla_0 U(\mathbf{x_0})]. \tag{E1}$$

In what follows, it is helpful to use symbols illustrated in Fig. (19), although the details associated with this particular lens are not needed. In Eq. (E1), $\mathbf{x} = \mathbf{r}$ is the observation point beyond the lens. $\mathbf{x}_0 = -L\hat{\mathbf{k}} + \mathbf{r}_1$ represents a point on Σ_0, the exit surface of the lens. (It also includes the screen, but on it U and ∇U are taken to vanish). $d\mathbf{\Sigma}_0 = d\Sigma_0\hat{n}$ is the surface element of the lens, whose normal $\hat{n}$ points radially outward from it. $G(\mathbf{x},\mathbf{x}_0)$ is the Green's function for the vacuum, given by Eq. (C13) with $r = |\mathbf{x}-\mathbf{x}_0| \equiv D$. It satisfies the wave equation with a point source (Eq. (C1), with $n = 1$ and with the argument of the delta function changed to $\mathbf{D}$). Thus, Eq. (E1) describes U as a continuous sum (integral) of solutions of the wave equation so, of course, it is a solution of the wave equation.

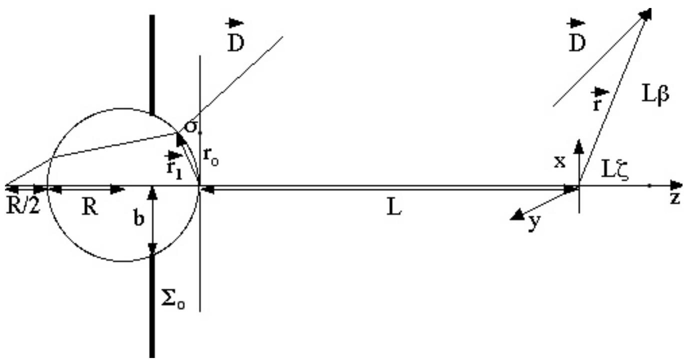

FIG. 19: Ray geometry for a ball lens

Eq. (E1) can be simplified. From (C13),

$$\nabla_0 G = -G\hat{\mathbf{D}}[ik - D^{-1}] \approx -G\hat{\mathbf{D}}ik,$$

where the approximation is valid for $D >> \lambda$. From (C12),

$$\nabla_0 U(\mathbf{x}_0) = ikG(\mathbf{x}_0)\nabla_0\Phi(\mathbf{x}_0) \approx ikG(\mathbf{x}_0)\hat{\mathbf{v}}_0,$$

where the approximation replaces Φ by Φ_0 (since Φ_1 is quite constant over the lens exit surface) and uses (C4).

Thus, (E1) becomes:

$$U(\mathbf{x}) = \frac{-ik}{4\pi}\int_{\Sigma_0} d\Sigma_0 U(\mathbf{x_0})\frac{1}{D}e^{ikD(\mathbf{x},\mathbf{x}_0)}\hat{n}\cdot[\hat{\mathbf{v}}_0 + \hat{\mathbf{D}}].$$

We are interested in the solution for large L, on the image plane far from the lens. There, D^{-1} varies slowly, and may be taken out of the integral.

As shown at the end of section (B 7), the outgoing ray from the lens surface satisfies is almost parallel to the z-axis (the optic axis), i.e., $\hat{\mathbf{v}}_0 \approx \hat{\mathbf{k}}$. (For a perfect lens, $\hat{\mathbf{v}}_0 = \hat{\mathbf{k}}$ since then the source point is imaged at ∞.) Similarly, $\hat{\mathbf{D}} \approx \hat{\mathbf{k}}$ since the intensity at $\mathbf{x}$ we wish to explore is not very much off-axis. The normal to the exit lens surface is not parallel to $\hat{\mathbf{k}}$, but $d\mathbf{\Sigma_0}\cdot\hat{\mathbf{k}} = d\Sigma_0\hat{n}\cdot\hat{\mathbf{k}} = dA_0$, where dA_0 is the surface element of S_0, the plane tangent to the exit surface of the lens at the point where it intersects the optic axis and perpendicular to the optic axis (the "tangent plane"). Therefore, the surface integral can be converted from being over the exit surface of the lens to being over the tangent plane. With $U(\mathbf{x_0}) \sim \exp ik\Phi_0(\mathbf{x_0})$ given by the WKB approximation, the approximate solution to be evaluated is

$$U(\mathbf{x}) \sim \int_{S_0} dA_0 e^{ik[\Phi_0(\mathbf{x_0})+|\mathbf{x}-\mathbf{x}_0|]}. \tag{E2}$$

Eq. (E2) is what we need hereafter. Since we are only interested in relative values of $|U(\mathbf{x})|^2$, constant factors may be dropped or chosen at pleasure.

It is worth re-emphasis, that $\Phi_0(\mathbf{x_0})$ in Eq. (E2) is the optical path length (C9), from the source to the exit surface of the lens, at height r_0 above the optic axis. It is *not* the optical path length from the source to the tangent plane whose surface area element is integrated over in Eq. (E2).

APPENDIX F: POINT SPREAD FUNCTION AND CONSEQUENCES

1. The Diffraction Integral

To integrate (E2), we need $D = |\mathbf{x}-\mathbf{x}_0|$. Again, refer to Fig.(19). The origin of the coordinate system is on the optic axis, a large distance L away from the exit surface of the lens. $\mathbf{D}$ makes a small angle β with respect to the optic axis, and its horizontal component extends a small distance ζL beyond the origin. Therefore, the observation point is

$$\mathbf{x} = \mathbf{r} = \hat{\mathbf{i}}L\beta + \hat{\mathbf{k}}L\zeta.$$

The point on the surface of the lens is

$$\mathbf{x}_0 = \hat{\mathbf{i}}r_0\cos\phi + \hat{\mathbf{j}}r_0\sin\phi - \hat{\mathbf{k}}[L+\sigma]$$

where ϕ is the azimuthal angle in the tangent plane and

$$\sigma = R - \sqrt{R^2 - r_0^2} \approx \frac{r_0^2}{2R^2} + \frac{r_0^4}{8R^3}$$

is the "sagitta," the horizontal distance between the surface of the lens and the tangent plane, at height r_0 above the optic axis. With $D = [(\mathbf{x} - \mathbf{x}_0)^2]^{1/2}$, dropping terms of order L^{-1}, we obtain:

$$e^{ikD} = e^{ikL\{1+(1/2)[\beta^2-\zeta+\frac{1}{2}\zeta^2]\}}e^{ik[-\beta r_0\cos\phi+(-\zeta+1)\sigma]},$$

and (E2) becomes

$$U(\beta) \sim \int_0^b r_0 dr_0 \int_0^{2\pi} d\phi e^{ik[\Phi_0(\mathbf{x_0})-\beta r_0\cos\phi+(1-\zeta)\sigma]}, \tag{F1}$$

where b is the radius of the exit pupil.

The purpose of section F 4 is to show that the optical path length from the source point to the exit surface of the lens at distance r_0 from the optic axis is

$$\Phi_0 = 3.5R - \frac{r_0^2}{2R} - \frac{37R}{216}\left[\frac{r_0}{R}\right]^4 = 3.5R - \sigma - \frac{R}{21.6}\left[\frac{r_0}{R}\right]^4. \tag{F2}$$

Suppose a ray exits the lens at at distance r_0 from the optic axis in a direction almost parallel to $\hat{\mathbf{k}}$, and travels the distance σ further to the tangent plane. As Eq. (F2) shows, it still has to travel a bit further than that to achieve the same optical path from source to tangent plane as the axial ray ($r_0 = 0$), whose optical path is $(R/2) + n2R = 3.5R$. Thus, the wavefront is slightly converging.

With change of variable to $\rho \equiv r_0/b$ and

$$\bar{b} \equiv b/R, \quad \bar{\sigma}(\rho) \equiv \frac{(\bar{b}\rho)^2}{2} + \frac{(\bar{b}\rho)^4}{8},$$

Eq. (F1) becomes

$$U(\beta) \sim \int_0^1 \rho d\rho \int_0^{2\pi} d\phi e^{-ikR\left[\rho^4\frac{\bar{b}^4}{21.6}+\rho\beta\bar{b}\cos\phi+\zeta\bar{\sigma}(\rho)\right]}. \tag{F3}$$

The integral over ϕ is readily performed:

$$U(\beta) \sim \int_0^1 \rho d\rho e^{-ikR\left[\rho^4\frac{\bar{b}^4}{21.6}+\zeta\bar{\sigma}(\rho)\right]} J_0(kR\rho\beta\bar{b}), \tag{F4}$$

where J_0 is the Bessel function.

We note that if we choose the observation plane to be $\zeta = 0$ and neglect the exponent (a good approximation for sufficiently small exit pupil radius b), the result can be integrated using the identity $d[xJ_1(cx)]/dx = cxJ_0(x)$, with resulting intensity

$$I_{\text{A}}(kb\beta) \sim |U(\beta)|^2 \sim \left[\frac{2J_1(kb\beta)}{kb\beta}\right]^2. \tag{F5}$$

$I_{\text{A}}(kb\beta)$ is the well known and important Airy point spread function of a circular perfect lens or aperture, discussed in Section III H and illustrated in Fig.11.

The exponent in (F4) is responsible for spherical aberration. In geometrical optics, this is caused by the rays at the outer edge of a lens coming to a focus on the optic axis closer to the lens than the paraxial ray focus. In our calculation, this is represented by the converging wavefront. As a result, as one moves a plane along the optic axis, one sees a circle of light of varying radius. One tries to choose the best effective focal plane, the plane where there is the "circle of least confusion" or the plane where the on-axis intensity is largest. The value of the present discussion is that it gives the intensity of the combined diffraction and spherical aberration, something not given by geometrical optics.

In order to slickly choose the best plane of focus, one needs to introduce a complication, We express the exponent in (F4) in terms of orthogonal Zernike polynomials (designed just for this purpose!), $R_{2n}(\rho) \equiv P_n(2\rho^2 - 1)$, where the P_n are the Legendre polynomials. The first three, which we shall need, are $R_0 = 1$, $R_2 = 2\rho^2 - 1$, $R_4 = 6\rho^4 - 6\rho^2 + 1$. They obey the orthogonality relations $\int_0^1 \rho d\rho R_{2n}R_{2m} = \delta_{nm}(4n+2)^{-1}$. They also obey the neat relation $\int_0^1 \rho d\rho R_{2n}(\rho)J_0(c\rho) = (-1)^n J_{2n+1}(c)/c$.

In terms of these polynomials, (F4) becomes

$$U(\beta) = \sim \int_0^1 \rho d\rho e^{-ikR\left[R_4\bar{b}^4\left(\frac{1}{130}+\frac{\zeta}{48}\right)+R_2\bar{b}^2\frac{1}{4}(\zeta+\frac{6\bar{b}^2}{65})\right]} \times J_0(kR\rho\beta\bar{b}). \tag{F6}$$

(In obtaining (F6), we have set $\zeta[1 + \bar{b}^2/4] \approx \zeta$, a few percent error).

It is apparent that one can choose ζ so that the R_2 term vanishes. Moreover, this is essentially the plane of largest intensity on the optic axis ($\beta = 0$). Upon setting $J_0(0) = 1$ in Eq.(F6), expanding the exponential to second order, and using the orthogonality relations, the result is

$$|U(0)|^2 = \left|2\int_0^1 \rho d\rho e^{-i[pR_4+qR_2]}\right|^2 = 1 - \frac{p^2}{5} - \frac{q^2}{3}.$$

Since $dp^2/d\zeta << dq^2/d\zeta$ the intensity is maximized to a high degree of accuracy by setting $\zeta = -6\bar{b}^2/65$ and thus making the q^2 term vanish. This best focus plane is closer to the lens, consistent with the spherical aberration effect. When this value is put into the first term, it becomes $R_4\bar{b}^4[1 + \bar{b}^2/4]/130 \approx R_4\bar{b}^4/130$, using the same approximation already made.

2. The Point Spread Function

Defining

$$\bar{\beta} \equiv kR\beta\bar{b}, \quad \bar{B} \equiv kR\bar{b}^4/130, \tag{F7}$$

we arrive finally at the amplitude which combines diffraction and spherical aberration,

$$U(\bar{\beta}) \sim \sqrt{2}\int_0^1 \rho d\rho e^{i\bar{B}R_4(\rho)} J_0(\rho\bar{\beta}). \tag{F8}$$

Fig. (20) contains plots of $I(r)$, the intensity of light due to a point source imaged by the lens on the image plane, i.e., the point spread function,

$$I(r) = \sqrt{\bar{B}}|U(kbr/f)|^2, \qquad \text{(F9)}$$

for various exit pupil radii b, for our 1mm diameter lens. r is the distance from the optic axis. (F7) has been used (with $\beta = r/f$, as discussed at the end of the introduction to this Appendix), with $\lambda = .55\mu$m, and $f = (3/2)R = .75$mm.

The reason for the factor $\sqrt{2}$ in Eq. (F8) and the factor $\sqrt{\bar{B}}$ in Eq. (F9) are as follows.

Using the Fourier-Bessel relation

$$\int_0^\infty \bar{\beta} d\bar{\beta} J_0(\rho\bar{\beta}) J_0(\rho'\bar{\beta}) = \rho^{-1}\delta(\rho - \rho'),$$

because of the factor $\sqrt{2}$ in (F8), an integral proportional to the total energy emerging from the lens (proportional to the intensity integrated over the area of the image plane) is conveniently normalized to 1:

$$\int_0^\infty \bar{\beta} d\bar{\beta} |U(\bar{\beta})|^2 = 2\int_0^1 \rho d\rho = 1$$

(the maximum $\bar{\beta}$ allowed by our limitation $\sin\beta \approx \beta$, is large enough to allow the integral to be extended to ∞ with good accuracy).

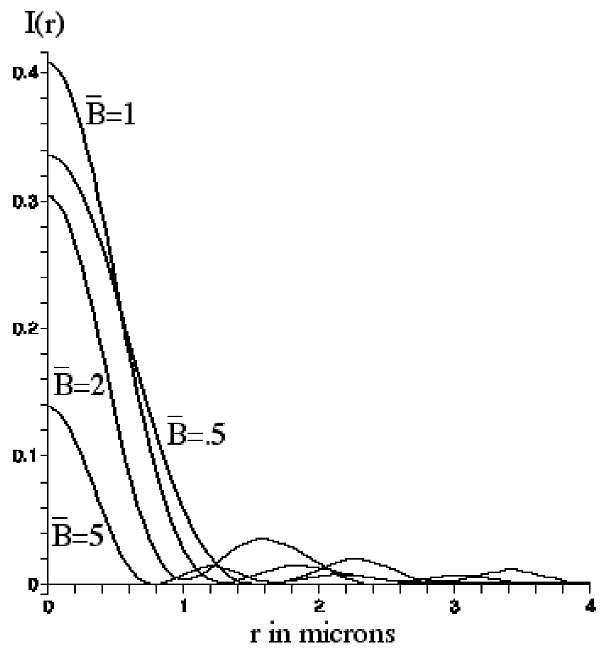

FIG. 20: Point spread functions $I(r)$ for $b = .16\mu$m ($\bar{B} = .5$), $b = .19\mu$m ($\bar{B} = 1$), $b = .23\mu$m ($\bar{B} = 2$), $b = .29\mu$m ($\bar{B} = 5$).

However, the total energy emerging from the lens should be proportional to the exit pupil area, $\sim b^2$. Since $\sqrt{\bar{B}} \sim b^2$ (see (F7)), this factor is included in the definition (F9) of the point spread function.

It is useful to have an approximate analytic expression for the point spread function, at least for small values of $\bar{B}$. Expansion of (F8) to second order in $\bar{B}$ gives

$$U(\bar{\beta}) \sim \sqrt{2}\int_0^1 \rho d\rho [1 - i\bar{B}R_4 - \frac{1}{2}(\bar{B}R_4)^2] J_0(\rho\bar{\beta}).$$

A little algebra shows that $R_4^2 = (18/35)R_8 + (2/7)R_4 + (1/5)$ so, dropping the R_8 term,

$$U(\bar{\beta}) \sim \sqrt{2}\left\{\left[1 - \frac{\bar{B}^2}{10}\right]\frac{J_1(\bar{\beta})}{\bar{\beta}} - \frac{\bar{B}^2}{7}\frac{J_5(\bar{\beta})}{\bar{\beta}} - i\bar{B}\frac{J_5(\bar{\beta})}{\bar{\beta}}\right\}.$$

Thus, to second order in $\bar{B}$ (neglecting the small $\sim J_1 J_5$) term), the point spread function is, approximately,

$$I(r)/\sqrt{\bar{B}} \sim \frac{1}{2}\left(\frac{2J_1(\bar{\beta})}{\bar{\beta}}\right)^2\left(1 - \frac{\bar{B}^2}{5}\right) + \frac{\bar{B}^2}{2}\left(\frac{2J_5(\bar{\beta})}{\bar{\beta}}\right)^2 \qquad \text{(F10)}$$

The normalized total energy is still 1 in this approximation:

$$\int_0^\infty \bar{\beta} d\bar{\beta} I(r)/\sqrt{\bar{B}} = \frac{1}{2}\left\{2\left(1 - \frac{\bar{B}^2}{5}\right) + \bar{B}^2\frac{2}{5}\right\} = 1.$$

Eq. (F10) shows that, for small $\bar{B}$ (small b), the point spread function is essentially the Airy function. As $\bar{B}$ grows, the amplitude of the Airy function decreases, with concomitant growth of a fifth order Bessel function contribution which vanishes for $r = 0$, and whose oscillations are displaced to larger r values than the oscillations of the Airy function.

Eq. (F10) has good accuracy for $\bar{B} = 1$: it is $\approx 2.5\%$ low at $\bar{\beta} = 0$, improving to negligible inaccuracy at $\bar{\beta} = 1$ and beyond. (For $\bar{B} = 1.3$ and 1.5, these percentages are 6% and 12% at $\bar{\beta} = 0$, with negligible inaccuracy beyond $\bar{\beta} = 2.5$, 3 respectively.)

3. The Exit Pupil

These results may be used to choose the optimum exit pupil radius b for our lens.

As can be seen from Fig. (20), as b is increased from a small value, the intensity on the optic axis $r = 0$ initially grows, because the exit pupil is allowing more light to exit the lens. For small values of b ($\bar{B} < 1$), the intensity distribution (F10) is essentially $\sim I_A(\bar{\beta})$, the Airy function, variously given in (F5) or Eq. (3), and illustrated in Fig.11 and the $\bar{B} = .5$ curve in Fig. (20). The Airy radius for our lens is, from Eq. (4),

$$r_A = .61\frac{\lambda}{b/f} = \frac{.50}{\bar{b}}\mu\text{m}. \qquad \text{(F11)}$$

b/f is called the lens "numerical aperture."

For the Airy function, ≈84% of the light energy lies within the Airy disc. But, as b increases further, spherical aberration kicks in, the intensity on the optic axis starts to diminish and a greater percentage of light energy appears beyond r_A. For the values b =.16, .19, .23 and .29μm, used in Fig. (20), Eq. (F11) gives r_A =1.56, 1.3, 1.1 and .86μm respectively. The first two curves appear to reach 0 at these values of r_A, whereas the last two curves deviate somewhat.

One wants b to be as large as possible, to decrease r_A and thus increase resolution, and to let as much light as possible exit the lens. However, as b grows, spherical aberration grows, as seen in Fig. (20): $I(r)$ decreases for $r < r_A$ and more light appears for $r > r_A$, so resolution decreases. A rule of thumb, called the Strehl criterion, suggests increasing b until the maximum intensity, the intensity on the optic axis $I(0)$, is reduced to 80% of the maximum intensity on the optic axis without any spherical aberration. Then, the image is considered still *diffraction limited*, i.e., the image is still essentially the Airy disc. From (F9), we see $I(0) \approx 1/2[1-(\bar{B})^2/5]$. Thus, the Strehl criterion implies $\bar{B} = 1$. It does seem from Fig. (20) that this is an optimal choice.

For $\bar{B} = 1$, the wavefront (the surface of constant phase), for a ray exiting the lens a distance $\equiv \rho b$ above the optic axis, goes beyond the tangent plane by the distance $R\rho^4\bar{b}^4/21.6 = \rho^4 6\bar{B}/k \approx \lambda\rho^4$, according to Eq.(F4). Thus, the wavefront at the edge of the exit pupil, $\rho = 1$, is about a wavelength in front of the tangent plane. For $\bar{B} > 1$, images are available[85] showing appreciable spherical aberration, for path differences from $1.4\lambda\rho^4$ to $17.5\lambda\rho^4$.

4. Optical Path Calculation

The unfinished business remains of showing that the optical path length, of a ray emerging from the source at the lens focal length, passing through the lens and up to its exit surface, is given by Eq. (F2). For the following discussion, refer to Fig. (21). The focal length of the lens, according to Eq. (4), is $f = nR/2(n-1) = 1.5R$ for $n = 1.5$. Thus, the point source at a is at a distance $R/2$ to the left of the lens surface. We shall follow a ray which leaves the source at angle α to the optic axis.

A simplifying feature, which occurs only for $n = 1.5$, is that the angle of refraction cde also happens to be α. That can be seen as follows. The angle of incidence θ and the angle of refraction θ' are related by Snell's law, $\sin\theta = n\sin\theta'$. By the law of sines applied to the triangle adc, $\sin\alpha/R = \sin(\pi-\theta)/f$, or $\sin\theta = (f/R)\sin\alpha$. This is the same as Snell's law provided $f = nR$, which is only true for n=1.5.

The axial ray, $\alpha = 0$, obviously has optical path length $(R/2) + n2R = 3.5R$. For arbitrary α, the optical path length Φ_0 is ad+$n$de or, as can readily be seen from

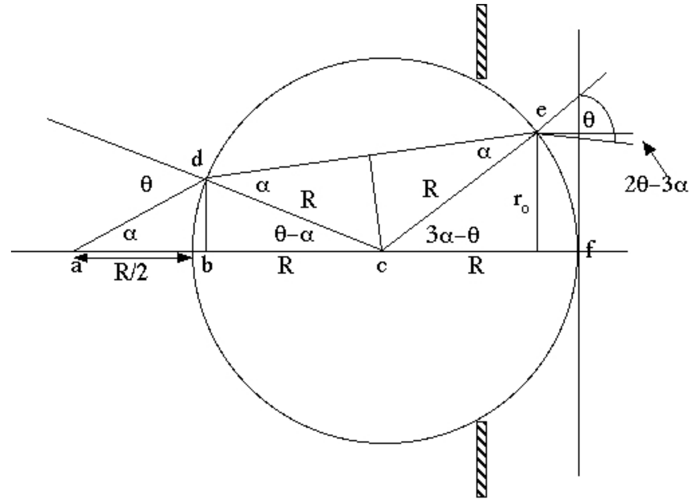

FIG. 21: Optical path length geometry

Fig.(21),

$$\begin{aligned}
\Phi_0 &= \frac{R\sin(\theta-\alpha)}{\sin(\alpha)} + (1.5)2R\cos\alpha \\
&= 4.5R\cos\alpha - R\cos\theta \\
&= 4.5R\cos\alpha - R\sqrt{1-(1.5\sin\alpha)^2} \\
&\approx R[3.5 - (9/8)\alpha^2 + (57/128)\alpha^4 + ...] \quad \text{(F12)}
\end{aligned}$$

The approximation (F12) is good to about 1% at $\alpha = .6 \approx 34°$. We want to express Φ_0 in terms of the distance r_0 between the optic axis and the exit point e. In terms of α, r_0 is

$$\begin{aligned}
r_0 &= R\sin(3\alpha-\theta) \\
&= R\sin 3\alpha\sqrt{1-(1.5\sin\alpha)^2} - R1.5\cos 3\alpha\sin\alpha \\
&\approx R[1.5\alpha - (7/8)\alpha^3 + ...].
\end{aligned}$$

This equation can be inverted,

$$\alpha = R^{-1}[(2/3)r_0 + (14/81)r_0^3 + ...]$$

and inserted into Eq. (F12), with the result (F2).

It was mentioned earlier that the angle γ the exiting ray makes with the horizontal is quite small. Here is the argument. From Fig. (21), $\gamma = 2\theta - 3\alpha$. From Snell's law, $\theta - \theta^3/6 \approx (3/2)[\alpha - \alpha^3/6]$, or $\theta \approx (3/2)\alpha + (5/16)\alpha^3$, so $\gamma \approx (5/8)\alpha^3$. Thus, the horizontal distance σ from e to the tangent plane differs from the actual distance $\sigma[1+(1/2)\gamma^2]$ by a negligibly small amount.

APPENDIX G: EXTENDED OBJECT

Having treated the image of a point source, we shall now consider the image of a uniformly illuminated hole of radius a. The hole models a transparent object such as a spherosome or a polystyrene sphere. We shall suppose