IMAGE ILLUMINATION
($4F^2$ OR $4F^2+1$?)

BACKGROUND

Publications abound with two differing expressions for calculating image illumination, the amount of radiation that transfers from an object through an optical system to an image[1,2]. The discrepancy frequently appears in equations defining the response of an optical or infrared detector to an emitting object in a viewed scene, the two competing forms being

$$\mathcal{R} \propto \frac{1}{4F^2} \qquad \text{and} \qquad \mathcal{R} \propto \frac{1}{4F^2+1}, \tag{1}$$

Where $\mathcal{R}$ is the response and F is the optical f/number. The f/number is defined by

$$F = \frac{f}{D}, \tag{2}$$

where f is the effective focal length of the optical system and D is the diameter of the entrance pupil. The popular understanding is that the first form is an approximation of the second. The purpose of this paper is to demonstrate that the first is rigorously correct for a well-designed optical system, and the second results from a paraxial approximation.

BASIC OPTICS

Figure 2 shows an optical system imaging an object located in object space between the first focal plane F_1 and infinity, the image occurring in image space between the second focal plane F_2 and infinity. There are several points and surfaces of interest. The image-space focal plane F_2 is the location of the image of an object located at infinity in object space. That is, all parallel rays incident on the optical system in object space will converge to a single point in image space, and that point will be in the plane F_2. Similarly, all rays emanating from a point in the plane F_1 will emerge as a parallel bundle in image; *i.e.*, they will create image of the point at infinity in image space.

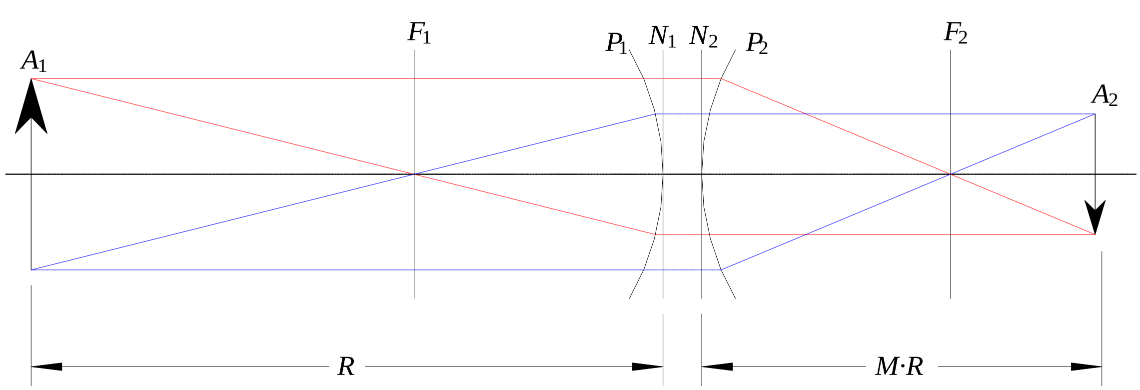

FIGURE 1 – KEY FEATURES OF AN OPTICAL SYSTEM.

The planes N_1 and N_2 in Figure 2 are, respectively, the first and second nodal planes, which are the planes normal to the optical axis and passing through the nodal points. The nodal points are the on-axis points for which a ray passing through one of them at a small angle emerges from the other at the same angle. The effective focal length (efl), or simply "focal length", of an optical system is the distance between the focal plane and the corresponding nodal plane.

The third set of surfaces of interest are the principal surfaces, often call the principal planes because they can be approximated by planes in the paraxial approximation. The principal surface P_2 in image space is defined by the intersection of rays parallel to the optical axis in object space with their corresponding converging rays in image space. The principal surface P_1 in object space is similarly defined.

A pair of planes defined such that an object in one of the planes creates an image in the other is known a set of conjugate planes. Thus, the plane conjugate to F_1 is at infinity in image space, and the plane conjugate to F_2 is at infinity in object space. In Figure 2, the planes A_1 and A_2 are conjugate. A well-known theorem of optics, the Abbe sine condition[3], states that for rays in conjugate planes

$$n_2 y_2 \sin\theta_2 = n_1 y_1 \sin\theta_1 , \tag{3}$$

where n is the index of refraction, y is the ray height in the object or image plane, θ is the angle between the ray and the optical axis, and the subscripts 1 and 2 designate object space and image space, respectively. This condition must be valid if the optical system is well-corrected for coma. Distances are defined as positive above the optical axis and positive to the right of the nodal point in each space. Note that for a converging optic y_1 and y_2 will generally have opposite signs, as will θ_1 and θ_2. If we assume the object and the image both exist in the same medium (usually air), then the condition is

$$y_2 \sin\theta_2 = y_1 \sin\theta_1 . \tag{4}$$

ANALYSIS

GENERAL CASE

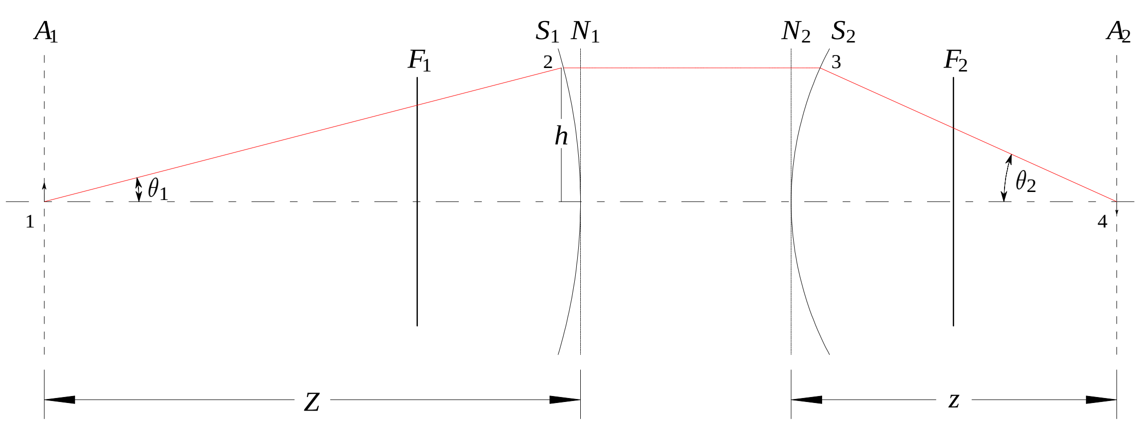

FIGURE 2 – MAPPING OF RAYS FROM CONJUGATE PLANES OF AN OPTICAL SYSTEM.

Plane A_1 in Figure 2 is located a distance z_1 (< 0) to the left of the first nodal plane N_1, and plane A_2 is located a distance z_2 (> 0) to the right of the second nodal plane N_2. The surface S_1 is a sphere of radius

$|z_1|$ centered at point 1, and surface S_2 is a sphere of radius $|z_2|$ centered at point 4, the axial point in the image plane. Consider an object-space ray that emanates from a source at point 1, and that strikes surface S_1 at a height h above the axis. We have

$$\sin\theta_1 = -\frac{h}{z_1}, \tag{5}$$

and therefore, from Eq. (3),

$$\sin\theta_2 = -\frac{y_1}{y_2}\frac{h}{z_1}. \tag{6}$$

Now the magnification of the instrument is

$$M = -\frac{y_2}{y_1} = -\frac{z_2}{z_1}, \tag{7}$$

even though y_1 may be infinitesimally small, so

$$\sin\theta_2 = -\frac{h}{z_2}. \tag{8}$$

This means that a ray emitted from point 1 emerges from spherical surface S_2 at the same height that it strikes spherical surface S_1.

Now, if we allow $z_1 \to \infty$, then $\theta_1 \to 0$, S_1 becomes a plane coincident with the first nodal plane, $z_2 \to f$ (the effective focal length), and S_2 becomes a sphere centered at the intersection of F_2 and the optical axis. If we then consider a marginal ray (*i.e.*, a ray that intersects S_1 at the maximum aperture), then

$$\sin\theta_2 = \frac{D}{2f} = \frac{1}{2F}, \tag{9}$$

where F is the f/number. In words, this says that a ray parallel to the optical axis in object space produces a ray that passes through the image-space focal point, and the two rays intersect on a sphere of radius f centered at the focal point in image space. That is, the principal surface in image space is a sphere[4]. A similar argument with the spaces reversed would show that the same holds true in object space. These spherical surfaces are represented in Figure 2 by P_1 and P_2.

The power received from infinity by a small, Lambertian detector at the focal plane of a lossless optical system is

$$P_{image} = 2\pi B_0 A_{det} \int_0^{\theta_{max}} \cos\theta \sin\theta \, d\theta = \pi B_0 A_{det} \sin^2\theta_{max}, \tag{10}$$

where B_0 is the brightness of the source (W/cm^2·sr), the $\cos\theta$ in the integral reflects the fact that the effective area of the detector for the off-axis contributions is the projected area, and θ_{max} is the cone half-angle of the ray bundle as viewed from the detector.

$$P_{image} = \frac{\pi B_0 A_{det}}{4F^2}, \tag{11}$$

PARAXIAL APPROXIMATION

Note that if we worked in the paraxial approximation where the principal surface in image space is a plane, then we would have

$$\tan\theta_2 = \frac{1}{2F}, \tag{12}$$

and so

$$\sin\theta_2 = \frac{1}{\sqrt{4F^2+1}}. \tag{13}$$

In this case we would have

$$P_{image} = \frac{\pi B_0 A_{det}}{4F^2+1}. \tag{14}$$

The form Eq. (14) is therefore only valid in the paraxial approximation (*i.e.*, large *f*/numbers), Eq. (11) being the generally valid form.

A SIMPLE ALTERNATIVE DERIVATION OF THE CORRECT FORM

Another way to look at the problem is to consider the source side; *i.e.*, to consider the portion of radiation emitted by the source, collected by the optics, and incident on the detector. Consider an on-axis detector pixel of small $x_{det}\times y_{det}$ size, and consider the rays emitted from the periphery of the pixel and passing through the image-space nodal point, as illustrated in Figure 3. These rays will emerge from the object-space nodal point and form a rectangle of size $x_{tgt}\times y_{tgt}$ in object space. This defines the area of object space that is exactly imaged onto the axial pixel; *i.e.*, all the radiation that is emitted from that target area and collected by the lens will impinge on the pixel, but no other scene radiation will. The area of the target is simply

$$A_{tgt} = x_{det}y_{det} = \left(\frac{R}{f}x_{det}\right)\left(\frac{R}{f}y_{det}\right) = \frac{R^2}{f^2}A_{det}. \tag{15}$$

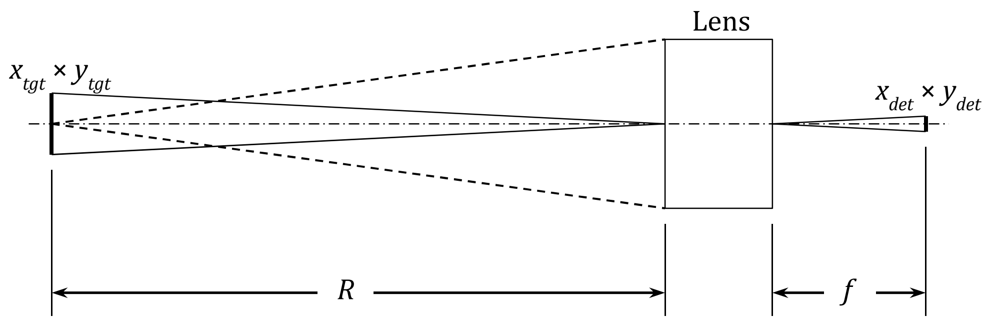

FIGURE 3 - ENERGY COLLECTED BY A SINGLE PIXEL.

The radiation collected by the optics from that area of the target is given by

$$P_{image} = \pi B_0 A_{tgt}\sin^2\theta_{opt}, \tag{16}$$

where θ_{opt} is the half-angle subtended by the optical aperture as viewed from the target. Now,

$$\sin\theta_{opt} = \frac{D/2}{\sqrt{R^2 + (D/2)^2}}, \tag{17}$$

so, from Eqs. (15), (16) and (17),

$$P_{image} = \pi B_0 \left(\frac{R^2}{f^2} A_{det} \right) \frac{D^2/4}{R^2 + D^2/4} = \frac{\pi B_0 A_{det}}{4F^2} \frac{R^2}{R^2 + D^2/4} \xrightarrow{R\to\infty} \frac{\pi B_0 A_{det}}{4F^2}. \tag{18}$$

OTHER CONSIDERATIONS

CAVEATS

It should be noted that image illumination is not necessarily the same as the radiation sensed by the detector, and for more reasons that just absorption efficiency. If the only issue were the absorption efficiency we could simply introduce a multiplier, which could be wavelength dependent, and this is the usual approximation. That approach assumes the detector to be Lambertian; *i.e.*, it assumes that the absorption efficiency is independent of the angle of incidence. A real detector does not generally behave that way, for a couple of reasons. First, when an electromagnetic wave is incident upon a surface, the reflection and transmission depend on both the angle of incidence and the polarization, according to the Fresnel equations. In addition to this, most highly-efficient absorbers make use of anti-reflection coatings or other such structures whose behavior depends in some resonant manner on the wavelength of the radiation. Changing the angle of incidence changes the effective optical thickness of each layer of these structures, thereby changing overall efficiency.

Changing the *f*/number changes the cone angle of the radiation incident on the detector. A small change in *f*/number will change the cone by a thin cone-shaped shell, and the absorption of that energy will, in general, differ on average from the bulk of the cone. Depending on the design of the absorption structure and on the spectral composition of the incident energy, this can either improving or degrading the average absorption efficiency, thereby distorting the apparent dependence on *f*/number of Eq. (11).

ILLUMINATION VS. RANGE

Note that Eq. (18) (before the limit $R \to \infty$) also predicts a change in image illumination vs. distance to the target. For many sensors this is irrelevant because $R >> D$, but there are instances where this is not valid. An important example is the indoor use of a radiometric thermal imaging sensor, where the objects viewed can be very close. Physically, two things happen as the viewed object moves closer to the optical aperture. Assuming the position of the focal is adjusted to keep the object in focus, the magnification increases as the object come closer. That means a smaller portion of the image cover on a single pixel, thereby reducing the illumination. At the same time the solid angle subtended by the optical aperture (relative to the object) increases. This has the effect of increasing the illumination. The first effect “wins” by a slight margin, resulting in a small net loss of illumination. It is also important to remember that the distance to the viewed object should be measured from the first nodal point, not from the first surface of the lens or the lens housing.

This effect can be important, but it can also be overwhelmed by a change in the effective optical aperture as the object moves closer. If the entrance pupil is physical stop in front of the lens, a close object can emit rays that pass through the stop, but their angle is sufficiently steep that they strike another stop behind the entrance pupil. If the stop is internal to the lens the effect will be different still.

[1] Niclaus, F., Decharat, A., Jansson, C. and Stemme, G., "Performance model for uncooled infrared bolometer arrays and performance predictions of bolometers operating at atmospheric pressure," *Infrared Physics & Technology*, Vol. 51, No. 3, Jan 2008, p.168.
[2] Richwine, R., Balcerak, R., Rapach, C., Freyvogel, K. and Sood, A., "A Comprehensive Model for Bolometer Element and Uncooled Array Design and Imaging Sensor Performance Prediction," *Proc. SPIE*, Vol. 6294, 2008.
[3] Born, M. and Wolf, E., *Principles of Optics*, 4th Ed., Pergamon Press, 1970, p. 168.
[4] Smith, Warren J., *Modern Optical Engineering, 4th Ed.*, McGraw-Hill, 2008, p. 23.