The F-word in Optics

F-number, was first found for fixing photographic exposure. But F/number is a far more flexible phenomenon for finding facts about optics. Douglas Goodman finds fertile field for fixes for frequent confusion for F/#.

The F-word of optics is the F-number,* also known as F/number, F/#, etc. The F-number is not a natural physical quantity, it is not defined and used consistently, and it is often used in ways that are both wrong and confusing. Moreover, there is no need to use the F-number, since what it purports to describe can be specified rigorously and unambiguously with direction cosines. The F-number is usually defined as follows:

$$\text{F-number}=\frac{\text{focal length}}{\text{entrance pupil diameter}} \qquad [1]$$

A further identification is usually made with ray angle. Referring to Figure 1, a ray from an object at infinity that is parallel to the axis in object space intersects the front principal plane at height Y and presumably leaves the rear principal plane at the same height. If *f* is the rear focal length, then the angle of the ray in image space θ′′ is given by

$$\tan\,\theta'=\frac{\mathrm{y}}{\mathrm{f}} \qquad [2]$$

In this equation, as in the others here, a sign convention for angles heights and magnifications is not rigorously enforced. Combining these equations gives:

$$F-number=\frac{1}{2\tan\theta'} \qquad [3]$$

For an object not at infinity, by analogy, the F-number is often taken to be the half of the inverse of the tangent of the marginal ray angle.

However, Eq. 2 is wrong. For lenses corrected for coma, the correct and exact equation is given by the sine condition:

$$\sin\theta'\frac{\mathrm{Y}}{\mathrm{f}} \text{ or } \mathrm{n'}\sin\theta'=\mathrm{NA'}=\mathrm{Y}\phi \qquad [4]$$

where ϕ is the power of the lens, *n*' is the image space refractive index, and NA' is the image space numerical aperture. Even if there is some coma, this equation holds very closely. The tangent equation, Eq. 2, arises from the description of what is often taken to be ideal lens behavior, for which every point in object space is stigmatically imaged. If the refractive indices in both spaces are constant, such imaging is described mathematically by the collinear transformation. However, such imaging is fundamentally impossible, except in the special case of an afocal system with magnification n/n'. This postulated behavior is self-contradictory, since the assumption of perfect imaging at every point implies that the path lengths involved in imaging individual points are not equal.

* It would be of interest to learn the history of this term. Is this known by any readers?

The F-number suffers from still more problems:

- The literature also contains the definition F-number = (focal length) / *(exit* pupil diameter). By itself, this ratio is useless, since outgoing ray angles are not determined without also specifying the axial position of the exit pupil.
- The literature contains confusing discussions of the variation of F-number with conjugates. A common equation is: $F_m = (1 + m)\, F_\infty$ where F_m is the F-number for magnification m and F_∞ is that for an object at infinity. This is a relationship between tangents of ray angles, which presumes both the validity of the collinear transformation (which never holds), and the assumption that the pupils are at the principal planes (which may not be the case). These problems cannot be cured by using numerical apertures, since if numerical apertures are known for one magnification, then from this information alone they cannot be found for other magnifications. A further complication is that different structures of a lens can act as the stop for different conjugates, complicating the way that the numerical apertures vary with object position.
- For some inexplicable reason, the F-number for an object at infinity is often given for lenses intended to be used at finite conjugates. (Perhaps this is done since, say, F/1 sounds more impressive than F/2.) From the F-number for infinite conjugates alone, that for finite conjugates cannot be found without additional information about the pupil locations. And again, the numerical aperture analog also fails; if the numerical aperture is given for an object at infinity, then from this information alone, neither numerical aperture can be found for other conjugates.
- What is the relationship between object space and image space F-numbers? The expressions of collinear transformation give $m = \tan\theta / \tan\theta'$, but the sine condition gives $m = n \sin\theta / n' \sin\theta'$. As mentioned above, the sine condition is exact for a lens with no coma, and holds very well in most cases. The tangent relationship comes from the collinear transformation. Note that the sine condition contains refractive indices, whereas the tangent relationship does not, and this is related to the failure of the collinear transformation to account for optical path lengths.
- What is the F-number of an afocal system? When used at finite conjugates, afocal lenses certainly have numerical apertures.

THE RIGHT WAY TO DO THINGS

For lenses that are rotationally symmetric with circular apertures, and for axial object points and object planes perpendicular to the axis, use the sines of the marginal ray angles. In the more general case of arbitrary systems, including arbitrarily shaped apertures and pupils, and tilted or non-flat object and image planes, use direction cosines (Fig. 2). At each field point, construct a coordinate system with an axis perpendicular to the object surface. Determine the extreme rays that pass through the lens, and find their intersection with a unit sphere about the field point. Project the intersection with the unit circle in the plane tangent to the image surface at the field point to obtain the extent of the pupil in direction cosines. The sin θs of the numerical aperture are just direction cosines in disguise for the special case of rotational symmetry. (Optics is filled with direction cosines disguised as sines, for instance, Snell's law and the grating equation.)

When things are done in this way, all is harmonious. The equations are general and exact, and ray optics, wave optics, and radiometry are as one. And by using ray angles and not heights, afocal and non-afocal systems are treated identically.

As an example of the simplification that occurs when this approach is taken, consider the "cosine-to-the-fourth" type variation of image irradiance with field position. Let α and β be the direction cosines used for pupil directions. If an element of the object surface emits according to a Lambertian distribution, then the power collected by the lens is proportional to $\int d\alpha d\beta$, where the integral is over the pupil directions. The common expression $\sin^2\theta$ is just the special case of this integral for a circular pupil, an object plane perpendicular to the axis, and an axial object point. The "cosine-to-the-fourth" law is nothing more than an approximation of the ratio of the integral $\int d\alpha d\beta$ for field points to that on axis in the case of rotationally symmetric situations with planar objects.

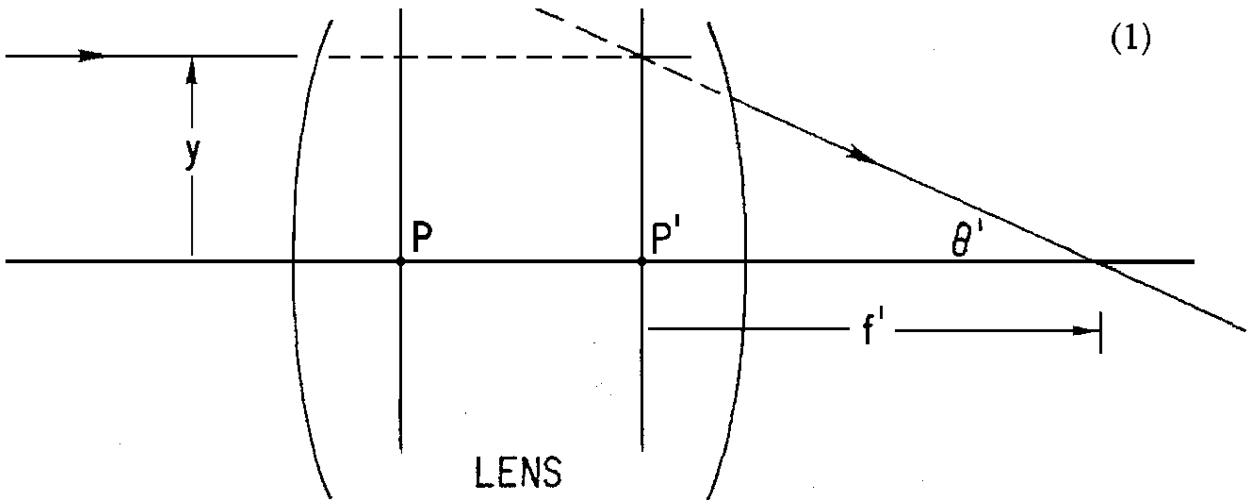
FURTHER READING

M.R. Hatch and D.F. Stoltzman, "The f-stops here," Optical Spectra, June, 1980, pp. 88-91.

R. Kingslake, "The effective aperture of a photographic objective," JOSA **35,** 518-520, (1945)

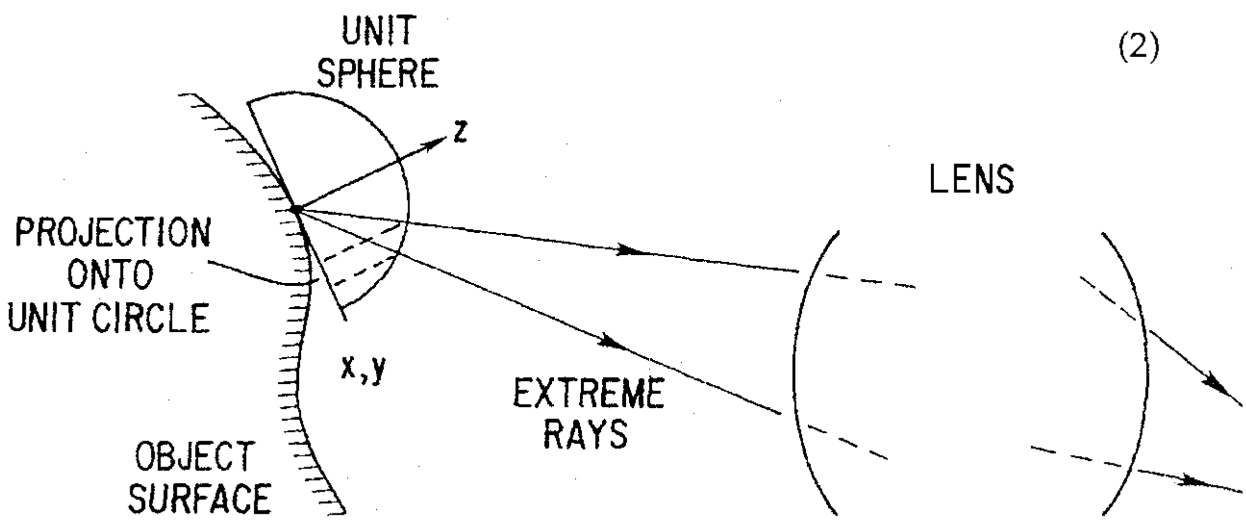
T. Smith, "The optical cosine law," Trans. Opt. Soc. London, **24,** , 31-39. (1922-23)

DOUGLAS S. GOODMAN *was with the IBM T.]. Watson Research Center, Yorktown Heights, N.Y. when this article was penned. He earned his PhD in optics from The Optical Sciences Center, University of Arizona in 1980. This article was scanned and minor corrections made by J.M. Palmer.*



y
P
P'
θ'
f'
LENS

(1)



UNIT
SPHERE
z
LENS
PROJECTION
ONTO
UNIT CIRCLE
x,y
EXTREME
RAYS
OBJECT
SURFACE

(2)