Maximum angular accuracy with an instrument limited by diffraction and photon statistics

by L. Lindegren 77-03-10

This note is a study of the interaction between the Fraunhofer diffraction pattern of an ideal telescope and an ideal periodic grid. The Fraunhofer diffraction of a rectangular aperture with rectangular central obscuration is convolved with the grid pattern to yield the intensity modulation. Assuming Poisson distribution of counted photons we can then get an estimate of the limiting angular accuracy as a function of the integration time, incident flux density and the parameters describing the aperture and grid. Although highly simplified (e.g., monocromatic light assumed) the model may give valuable hints as to the optimum design of the aperture and the grid.

1. The limiting accuracy of the timing of a modulated intensity

Let \( I(t) \) be the true intensity as a function of time, i.e. the probability of detecting a photon during the interval \([t, t+dt]\) is \( I(t)dt \). The actual observation is most completely given by the step-function \( N(t) \) which gives the number of photons counted from time \( t = 0 \) up to time \( t \). Thus, the arriving times of all photons are known. The observation extends from \( t = 0 \) to \( t = T \); outside of this interval we may assume \( I(t) = 0 \).

If \( I(t) \) is parametrised by \( X \), we can express the probability of obtaining the observed process \( N(t) \) by means of the function *

\[
p\left[ \left\{ N(t), 0 \leq t \leq T \right\} \bigg| X \right] = \exp \left[ - \int_0^T I(t; X) dt + \int_0^T \ln I(t; X) dN(t) \right]
\]

(1)

derived from the Poisson distribution function. The maximum likelihood estimate \( \hat{X} \) of \( X \) is found by maximising the likelihood function

\[
L(X) = \ln p\left[ \left\{ N(t), 0 \leq t \leq T \right\} \bigg| X \right] = - \int_0^T I(t; X) dt + \int_0^T \ln I(t; X) dN(t).
\]

(2)

Let us choose as the only parameter the time \( \theta \) such that \( I(t; X) = I(t - \theta) \). Then

\[
L(\theta) = - \int_0^T I(t - \theta) dt + \int_0^T \ln I(t - \theta) dN(t).
\]

(3)

* see, e.g., D.L. Snyder, Random Point Processes, John Wiley and Sons, New York 1975
Provided that the interval \([0, T]\) is sufficiently great to include all expected photons within the possible range of \(\theta\), the first integral is constant, and

\[
\frac{dI}{d\theta} = - \int_0^T \frac{I'(t - \theta)}{I(t - \theta)} dN(t).
\]  

(4)

The maximum likelihood estimate \(\hat{\theta}\) of \(\theta\) is thus the solution of the equation

\[
\int_0^T \frac{I'(t - \hat{\theta})}{I(t - \hat{\theta})} dN(t) = 0.
\]  

(5)

Note that (5) is of the general form \(w f(\hat{\theta}) = 0\), where \(f\) is the observed photon counts and \(w\) is a weight function, investigated by me in relation to the problem of deriving transit times from photoelectric meridian observations of planets. If \(I(t)\) is the expected intensity and photon noise dominates, then \(w = I''/I\) is the optimum weight function.

The accuracy of the estimate \(\hat{\theta}\) is \(\varepsilon\), where \(\varepsilon^2 = E[(\hat{\theta} - \theta)^2|\theta]\).

From the Cramér-Rao Lower Bound Theorem it follows that

\[
\varepsilon^2 \geq \left[ \int_0^T \frac{1}{I(t - \hat{\theta})} \left( \frac{d}{d\theta} \left( \frac{I(t - \theta)}{I(t)} \right) \right)^2 dt \right]^{-1} = \left[ \int_0^T \left( \frac{I'(t)}{I(t)} \right)^2 dt \right]^{-1}
\]

(6)

Equality holds only for an unbiased and efficient estimate. The maximum likelihood estimate is asymptotically both unbiased and efficient under relatively weak conditions.

We now have an estimate of the minimum mean square error of \(\theta\) which can be achieved with any method of estimation, utilising all information available in the recorded process. This lower bound is useful in two respects in the optimisation process:

1. The absolute merit of any proposed (practically useful) method of estimation can be evaluated by comparing its performance with the lower bound estimate.

2. In optimising e.g. the grid pattern, the lower bound estimate is useful because we need not specify which algorithm will actually be used.
2. Sinusoidal intensity modulation

We shall evaluate the lower bound estimate $\xi_p$ for the precision of an angle $p$ measured from the intensity function $I(t)$,

$$I(t) = b_o + b_1 \cos(2\pi vt/s),$$

obtained by passing an image over a grid with period $s$ (radians) at the angular velocity $v$ (rad s$^{-1}$). The total integration time is taken to be $m$ periods: $T = ms/v$. Then

$$\xi_t^{-2} = \int_0^T \frac{4 \pi^2 v^2 s^2 \sin^2(2\pi vt/s)}{b_o + b_1 \cos(2\pi vt/s)} \ dt = \frac{4 \pi^2 v^2 s^2}{b_o T} \left(1 - \sqrt{1 - \left(\frac{b_1}{b_o}\right)^2}\right)$$

(8)

$b_oT$ is the total number of photons counted, and $b_1/b_o = (I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}}) = M$, the modulation factor. Furthermore, $\xi_p = v \xi_t$. Hence

$$\xi_p = \frac{s/2\pi}{N} \left[\sum(1 - \sqrt{1 - M^2})\right]^{-1/2}$$

(9)

3. Fraunhofer diffraction of a rectangular aperture with central obscuration

![Diagram of a rectangular aperture with diffraction pattern](image)

**Fig. 1**
Aperture A

**Fig. 2**
Diffraction pattern

Let $F$ be the incident monochromatic flux (photons s$^{-1}$ m$^{-2}$) on the aperture $A$ with dimensions (in m) as in Fig. 1. The Fraunhofer diffraction has intensity distribution $I(p, q)$ (in photons s$^{-1}$ sr$^{-1}$), where $p$ and $q$ are in radians. With $k = 2\pi/\lambda$, we have approximately

$$I(p, q) = \frac{F k^2}{4 \pi^2} \left| \iint_A \exp[-ik(p\xi + q\eta)] d\xi d\eta \right|^2$$

\[
\begin{align*}
\psi_F & = \frac{4F}{\pi^2 k^2} \left[ \sin^2 kpa \sin^2 (kqb/2) + \sin^2 kpa \sin^2 (kq\beta b/2) \right. \\
& \quad \left. - \frac{\sin kpa \sin kpa}{p^2} \left( \frac{\sin^2 (kqb/2)}{q^2} - \frac{\sin^2 (kq(1-\beta)b/2)}{q^2} + \frac{\sin^2 (kq\beta b/2)}{q^2} \right) \right] \\
& = \frac{4F}{\pi^2 k^2} \left[ \frac{\sin^2 kpa}{p^2} - \frac{\sin^2 kpa \sin kpa}{p^2} \right] \quad (10)
\end{align*}
\]

The slits are parallel to the q axis. We are interested only in the function

\[
J(p) = \int I(p,q) dq = \frac{2Fb}{\pi k} \left[ \frac{\sin^2 kpa}{p^2} - \frac{2\beta}{p^2} \sin kpa \sin kpa \right. \\
& \quad \left. + \frac{\sin^2 kpa}{p^2} \right] \quad (\text{photons s}^{-1} \text{ rad}^{-1}) \quad (11)
\]

The total flux is \( \int J(p) dp = 2ab(1 - \delta\beta)F \).

4. Modulation by the grid

We assume that the transmission of the grid varies periodically with period \( s \) in the \( p \) coordinate, according to Fig. 3.

![Fig. 3. Transmission of grid](image)

The transmission function \( R(p) \) can be expanded in a Fourier series:

\[
R(p) = \sum_{n=0}^{\infty} a_n \cos \left( \frac{2\pi np}{s} \right); \quad a_0 = \frac{r}{s}; \quad a_n = \frac{2r}{s} \frac{\sin(n\pi r/s)}{n\pi r/s}, \quad n > 0. \quad (12)
\]

When the diffraction pattern \( J(p) \) is superposed on the grid and the central peak is at coordinate \( vt \), the transmitted intensity is

\[
I(t) = \int R(p) J(p - vt) dp \quad (\text{photons s}^{-1}). \quad (13)
\]
If \( I(t) \) is expanded in the series

\[
I(t) = \sum_{n=0}^{\infty} b_n \cos(2\pi n vt/s),
\]

we obtain the coefficients

\[
b_n = a_n \int_{-\infty}^{\infty} J(p) \cos(2\pi np/s) \, dp = 2abFk(w)a_n,
\]

where the attenuation function \( k(w) \) is \( w = n\lambda/(2as) \)

\[
k(w) = \begin{cases} 
(1 - \alpha \beta) - (1 + \beta)w & \text{for } 0 \leq w < \alpha \\
(1 - 2\alpha \beta) - w & \text{for } \alpha \leq w < (1-\alpha)/2 \\
(1 - \beta - \alpha \beta) - (1 - 2\beta)w & \text{for } (1-\alpha)/2 \leq w < (1+\alpha)/2 \\
1 - w & \text{for } (1+\alpha)/2 \leq w < 1 \\
0 & \text{for } 1 \leq w
\end{cases}
\]

if \( \alpha \leq 1/3 \), and

\[
k(w) = \begin{cases} 
(1 - \alpha \beta) - (1 + \beta)w & \text{for } 0 \leq w < (1-\alpha)/2 \\
(1 - \beta) - (1 - \beta)w & \text{for } (1-\alpha)/2 \leq w < \alpha \\
(1 - \beta - \alpha \beta) - (1 - 2\beta)w & \text{for } \alpha \leq w < (1+\alpha)/2 \\
1 - w & \text{for } (1+\alpha)/2 \leq w < 1 \\
0 & \text{for } 1 \leq w
\end{cases}
\]

if \( \alpha > 1/3 \).

We observe that the cosine series for \( I(t) \) is finite, since \( b_n = 0 \) for \( n > 2as/\lambda \). When \( s \leq \lambda/2a \) we get \( n_{\text{max}} = 0 \) and no modulation of the light at all; when \( \lambda/2a < s \leq \lambda/a \) we have \( n_{\text{max}} = 1 \) and a purely sinusoidal modulation, in which case the formula derived in Section 2 is valid. It turns out that the optimum value of \( s \) generally lies in this interval, which makes the analysis fairly simple.
5. The weight of an angular observation

We shall use the following notations:

\( \delta, \beta \) defines the obscuration as in Fig. 1

\( \gamma = 2\omega / \lambda \) defines the grid period

\( \delta = r / s \) is the average transmission of the grid

\( N_o = 2abPF \) is the number of incident photons from the star in question

\( c_o = F_o / (2abF) \), where \( F_o \) = dark counts (assumed to be Poisson distributed)

\( c_1 = F_1 / F \), where \( F_1 \) is the incident flux density from background stars within the cathod spot

\( Q = \frac{\varepsilon_p^{-2}}{N_o} \) is the weight of the observation per photon

We then obtain

\[
b_o = 2abF \left[ c_o + (1 + c_1)(1 - \alpha \beta) \delta \right], \quad (17)
\]

\[
b_1 = 2abF \cdot 2 \delta \, k(\frac{1}{\gamma}) \sin \frac{\pi \delta}{\pi \delta}, \quad (18)
\]

and

\[
Q = (4\pi a / \lambda)^2 \gamma^{-2} \delta \left[ c_o + (1 + c_1)(1 - \alpha \beta) \right] \left\{ 1 - \left[ 1 - 4 \delta^2 \frac{1}{\gamma^2} \sin^2 \frac{\pi \delta}{\pi \delta} \right]^{-1/2} \right\} \quad (19)
\]

For a bright star we can neglect \( c_o \) and \( c_1 \) and obtain

\[
Q = (4\pi a / \lambda)^2 \gamma^{-2} (1 - \alpha \beta) \delta \left[ 1 - \sqrt{1 - 4k^2(\frac{1}{\gamma}) \sin^2 \frac{\pi \delta}{(\pi \delta)^2} (1 - \alpha \beta)^{-2}} \right] \quad (20)
\]

(19) and (20) are valid for \( 1 \leq \gamma \leq 2 \).

The quantity \( Q(4\pi a / \lambda)^{-2} \) is plotted in Figs. 1-5 as a function of \( \gamma \) for various values of the parameters \( \alpha, \beta, \delta, c_o, \) and \( c_1 \). In the region \( 1 \leq \gamma \leq 2 \) the formula (19) has been used; for \( \gamma > 2 \) the Fourier coefficients of \( I(t) \) has been evaluated by means of (15) and (16) and the integral in (6) through numerical integration. It is clear that the maximum value of \( Q \) for a given aperture is generally obtained for \( \gamma = 2 / (1 + \alpha) \). Since this is always in the interval \([1, 2]\), we can use (19) or (20) in the optimisation of the other parameters.
With \( \gamma = 2/(1 + \alpha) \) we have

\[
Q = (4\pi a/\lambda)^2 \frac{1}{4} (1 + \alpha)^2 \left\{ c_0 + (1 + c_1)(1 - \alpha \beta) \delta - \sqrt{\left[ c_0 + (1 + c_1)(1 - \alpha \beta) \delta \right]^2 - \pi^{-2} (1 - \alpha)^2 \sin^2(\pi \delta)} \right\}^{+1/2} \tag{21}
\]

With \( c_0 = c_1 = 0 \) we find that the optimum value of \( \delta \) satisfies the transcendental equation

\[
\tan \pi \delta = 2\pi \delta - (1 - \alpha)^2 (1 - \alpha \beta)^{-2} \sin \pi \delta \cos \pi \delta \tag{22}
\]

The solution for \( \delta \) is always in the range \( \delta = 0.314291 \) to \( \delta = 0.371010 \).

As indicated by Fig. 5 the optimum value of \( \delta \) is still near 0.4 when \( c_1 = 1 \), i.e. when the background current equals the current from the star.

Fig. 6 shows \( Q(4\pi a/\lambda)^{-2} \) versus \( \delta \) for \( \alpha = \beta = c_0 = c_1 = 0 \) and \( \gamma = 2 \).

The shape of the curve is independent of \( \alpha \) if \( \beta = 1 \), and so is the optimum value of \( \delta \), viz. \( \delta = 0.314291 \). Fig. 7 shows the variation of \( Q(4\pi a/\lambda)^{-2} \) with \( \alpha \) for this combination. This curve has a maximum at \( \alpha = 1/3 \). Decreasing \( \beta \) below 1 decreases \( Q \). Hence it seems that the overall optimum choice for the parameters is:

\[
\begin{align*}
\alpha &= 1/3 \\
\beta &= 1 \\
\gamma &= 3/2 \\
\delta &= 0.314291
\end{align*}
\Rightarrow Q(4\pi a/\lambda)^{-2} = 0.0433615 \tag{23}
\]

\( M = 0.845 \)

Fig. 8 gives \( Q(4\pi a/\lambda)^{-2} \) versus \( \gamma \) for \( \alpha = 1/3, \beta = 1, \delta = 0.314291, c_0 = 0 \), and for \( c_1 = 0, 0.1, 0.3, 1, \) and 3.

In practice, when we have a finite bandwidth and not monochromatic light, it is of course not possible to match the grid period and the diffraction pattern period as closely as has been assumed here. From the width of the peaks in Fig. 8 it is perhaps possible to estimate the degradation caused by cromatic dispersion. It may be that the optimum values of \( \alpha \) and \( \beta \) become radically different in polychromatic light, but the optimum value of \( \delta \) should be almost the same.
$Q \left( \frac{4\pi a}{\lambda} \right)^{-2}$

$\alpha = 0.3$
$\beta = 0.3$
$c_0 = 0.1$
$c_1 = 1$
\[ Q \left( \frac{4\pi a}{\lambda} \right)^{-2} \]

\[ \alpha = \frac{1}{3} \]
\[ \beta = 1 \]
\[ c_0 = 0 \]
Addendum (77-04-04)

Calculations of efficiency in case of polychromatic light

Let us assume that the incident photon flux density in the wavelength interval \( \lambda, \lambda + d\lambda \) is

\[
F(\lambda) \, d\lambda,
\]

where, as before, \( F \) is the integrated flux density and \( f(\lambda) \) is a spectral distribution function for which \( \int f(\lambda) \, d\lambda = 1 \). By means of the quantities

\[
\lambda_{\text{eff}} = \int \lambda \, f(\lambda) \, d\lambda \quad \text{and} \quad \gamma_{\text{eff}} = 2a s / \lambda_{\text{eff}}
\]

we see that the equations (14) and (15) are still valid if we replace \( k(n) = k(n/y) \) by

\[
\bar{k}(n/y_{\text{eff}}) = \int k \left( \frac{n}{y_{\text{eff}}} \right) f(\lambda) \, d\lambda
\]

We can then go on calculating the efficiency factor \( Q(4\pi a / \lambda_{\text{eff}})^2 \) for polychromatic light.

Figs. 9 - 11 show some results for the spectral distribution

\[
f(\lambda) = 0.25 \, \delta(\lambda - 350) + 0.25 \, \delta(\lambda - 400) + 0.25 \, \delta(\lambda - 450) + 0.25 \, \delta(\lambda - 500)
\]

(\( \lambda \) in nm). \( \delta(x) \) is the Dirac delta function. \( \lambda_{\text{eff}} = 425 \) nm.

Comparing Figs. 9 - 11 with the corresponding Figs. 1 - 3 for monochromatic light the following conclusions may be drawn:

1. The maximum efficiency for any \( (\alpha, \beta) \) is always lower in the polychromatic case as compared to the monochromatic case. The degradation is (as was expected) severe especially when the monochromatic efficiency is very peaked in the grid period/efficiency diagramme. In particular, it is no longer advantageous to increase \( \beta \) beyond the value set by the size of the secondary mirror (we recall that it is possible, in the monochromatic case, to gain efficiency by throwing away some of the light!).
2. The optimum grid period $\chi_{\text{eff}}$ is still approximately $2/(1 + \alpha)$, or 5 - 6% larger, for moderate values of $\alpha$, $\beta$.

3. The optimum value of $\delta$ (the average transmittance of the ideal grid) is in the range 0.3 - 0.4. Perhaps 0.35 is now slightly better than 0.314, which is optimal for monochromatic light.

It should be noted that for $\chi_{\text{eff}} = 1.65$ (optimal for $\alpha = \beta = 0.3$) we still have an almost sinusoidal modulation of the transmitted light (this goes for all values of $\delta$).
\[ Q \left( \frac{4\pi a}{\lambda_{eff}} \right)^{-2} \]

\[ \alpha = 0.3 \]

\[ \beta = 0.3 \]

\[ c_0 = c_1 = 0 \]
\( Q \left( \frac{4\pi a}{\lambda_{\text{eff}}} \right)^{-2} \)

\( \alpha = 0.3 \)

\( \beta = 1 \)

\( C_0 = C_1 = 0 \)

\( \delta = 0.3 \)

\( \delta = 0.4 \)
Maximum angular accuracy in polychromatic light

Lindgren 77-04-13/1

The table below gives numerical results of computations according to the note 77-03-10 (addendum) also for the cases when the secondary mirror is larger than $\alpha = 0.3$. It is worth noticing that the efficiency $Q_L = Q(4\pi a/\lambda_{\text{eff}})^2$ decreases much faster with the relative size $\alpha$ of the secondary mirror than the relative light-gathering area $1 - \alpha^2$; it rather goes like $(1 - \alpha^2)^2 \approx 2.5$.

$\Delta \lambda$ is the FWHM or some equivalent measure of the wavelength range. $\delta = 0.35$ throughout, which is always nearly optimal.

$c_0 = c_1 = 0$, i.e. no background counts (or a bright star).

For each value of $\alpha$, the table gives the optimum choice for the grid period and the corresponding efficiencies $Q_L$ for two different bandwidths $M_0 = \epsilon_0/\epsilon_4$.

<table>
<thead>
<tr>
<th>$\alpha = \beta$</th>
<th>$\gamma_{\text{eff}} = 2\alpha s/\lambda_{\text{eff}}$</th>
<th>$\Delta \lambda/\lambda_{\text{eff}} = 0.35$</th>
<th>$\Delta \lambda/\lambda_{\text{eff}} = 0.47$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.0</td>
<td>3.40 $10^{-2}$</td>
<td>3.32 $10^{-2}$</td>
</tr>
<tr>
<td>0.2</td>
<td>1.8</td>
<td>3.10</td>
<td>2.99</td>
</tr>
<tr>
<td>0.3</td>
<td>1.7</td>
<td>2.78</td>
<td>2.69</td>
</tr>
<tr>
<td>0.4</td>
<td>1.6</td>
<td>2.36</td>
<td>2.24</td>
</tr>
<tr>
<td>0.5</td>
<td>1.5</td>
<td>1.86</td>
<td>1.72</td>
</tr>
<tr>
<td>0.6</td>
<td>1.4</td>
<td>1.30</td>
<td>1.15</td>
</tr>
</tbody>
</table>

Example: $m = 9$, $\lambda_{\text{eff}} = 0.43 \mu m$, $\Delta \lambda/\lambda_{\text{eff}} = 0.47$ (?) $\Rightarrow N_0 = 4215$ Hz.

$2a = 16$ cm, $\alpha = 0.5$ (nearly flat secondary) $\Rightarrow s = 5''83 = 6.5 \mu m$.

Width of transparent strips = $s \delta = 2.3 \mu m$. $1^8$ integration time gives

$\sigma_{sn} = (Q_L (4\pi a/\lambda_{\text{eff}})^2 N_0)^{-0.5} = 0''0100$ (cf. $\sigma_{sn} = 0''016$ from Table 1 in MDS).