SMALL SPOT ILLUMINATION OF OPTICAL COATINGS

by Angus MacLeod
Thin Film Center, Inc.

INTRODUCTION
In our usual calculations of the properties of our optical coatings we make use of the ideal linearly polarized plane harmonic wave. This is the simplest possible spectral component. We use this simple component for several reasons. Firstly, and of course most importantly, it is very simple. Secondly, because the effects we are dealing with are all linear, we can represent real light as a combination of such harmonic waves. We have already covered some examples, short pulses, coherence and incoherence, and oblique incidence. This article continues this discussion with a look at illumination of limited areas.

GENERAL PLANE WAVE EXPRESSION
Our first task is to derive a general expression for a plane harmonic wave traveling in any direction. How do we define a direction? When we are dealing with Cartesian coordinates a very convenient way is by means of direction cosines. Direction cosines are the cosines of the angles that the direction makes with the \( x, y \) and \( z \)-axes. By convention we write them as \( (\alpha, \beta, \gamma) \) and it is easy to see that they are also the components of a unit vector along the given direction. Thus they obey the relationship

\[
\alpha^2 + \beta^2 + \gamma^2 = 1
\]  

(1)

Now let us look at the phase factor of a plane, harmonic wave traveling in the direction given by \( (\alpha, \beta, \gamma) \). We need an expression for the phase at any point at any time. We will represent the point by its spatial coordinates \( x, y \) and \( z \) and, for convenience, take our zero phase reference when the three coordinates and the time \( t \) are all zero. The variation of phase with time is simply \( \omega t \). The distance term is a little more complicated. It will be \( 2\pi n \) times the number of wavelengths in the distance measured along the propagation direction from the reference point, the origin, to that plane of constant phase containing \( (x, y, z) \). This distance is the component along the wave direction of the length of the line to \( (x, y, z) \). Since the coordinates are actually the components of this length along the reference axes we just have to multiply them by the cosines of the angles they make with the wave direction and add. In other words, the distance is given by

\[
distance = \alpha x + \beta y + \gamma z
\]  

(2)

Those of you familiar with vector algebra will have arrived at this result much more quickly. The wavelength is given by \( \lambda / n \) where \( \lambda \) is the free space wavelength and \( n \) the refractive index. Putting it all together gives the phase factor as

\[
\exp \left[ i \left( \omega t - \frac{2\pi n}{\lambda} \left( \alpha x + \beta y + \gamma z \right) \right) \right]
\]  

(3)

SIMPLE INTERFACE
Now let us return to the simple interface that we discussed in the previous tutorial but this time we will treat it in a more general way. Figure 1 shows the direction of the incident ray and the reference axes. (See Figure 1)

The direction cosines of the incident wave will be \((\sin \theta_i, \cos \theta_i, 0)\) so that the phase factor will be

\[
\exp \left[ i \left( \omega t - \frac{2\pi n}{\lambda} \sin \theta_i \alpha x + \cos \theta_i \beta y \right) \right]
\]  

(4)

We can require that the reflected and transmitted rays are also plane harmonic waves of the form

\[
\exp \left[ i \left( \omega t - \frac{2\pi n}{\lambda} \left( \alpha x + \beta y + \gamma z \right) + \varphi \right) \right]
\]  

(5)

where the parameters are those that apply to the appropriate wave.

Our boundary conditions, the continuity of the fields, will be applied at \( z = 0 \). This continuity must exist for all time, \( t \), and for all possible values of \( x \) and \( y \). Their coefficients must, therefore, be separately identically equal. Thus \( \beta \) must be zero in both reflected and transmitted waves implying that their directions are in the plane of incidence. For the reflected wave, \( \alpha \) must be \( \sin \theta_i \) so that the angle of reflection is equal to the angle of incidence, and for the transmitted wave we must have

\[
n_i \sin \varphi = n_i \sin \varphi
\]  

(6)

which is exactly Snell's Law. In each case the coefficient of \( z \) can be found from (1).

Then, since there is no extra phase term corresponding to \( \varphi \) in the incident wave, \( \varphi \) must be zero in both reflected and transmitted waves. Note that this does not imply that there is no phase shift on reflection or transmission. Any phase shift will be determined entirely by the amplitude reflection and transmission coefficients.

At the boundary, therefore, all three phase factors are of the form

\[
\exp \left[ i \left( \omega t - \frac{2\pi n}{\lambda} \sin \theta_i \alpha x \right) \right] = \exp \left[ i \left( \omega t - \kappa_x x \right) \right]
\]  

(7)

\( \kappa_x \), that is \( 2\pi n \sin \theta_i / \lambda \), can sometimes be a more useful quantity than \( \alpha \). We can see that for an incident wave traveling in a completely general direction, the form of the expression for the phase factor at the boundary becomes

\[
\exp \left[ i \left( \omega t - \kappa_x x - \kappa_y y \right) \right]
\]  

(8)

Application to the multiple parallel interfaces of a multilayer is an obvious extension.

SMALL SPOT ILLUMINATION
Our plane harmonic wave is infinite in extent but our coatings are finite and the illuminated area may be even smaller still. A limited area of illumination is still a linear problem and we can therefore use our technique of representing the illumination in terms of a spectrum of simple components. Here, however, the illumination does not depend on time but on position. Frequency, \( \omega \), is constant and it is \( \kappa_x \) and \( \kappa_y \) that must be permitted to vary. Since this implies components with a variation in the direction of incidence, the set of components is called the angular spectrum. Analysis of a circular spot is complicated, and to keep the discussion simple we will retain the plane of incidence and expression (7), and imagine that the illumination is in the form of a strip, infinitely long in the \( y \)-direction. Any variation will be a function of \( x \) only. The analysis can readily be extended to a square spot. For the present we will also ignore polarization except that we assume it is consistently of the same orientation in all the rays.
We will consider relative phase as contained in the complex amplitude. Let $\mathcal{E}(x)$ be the complex amplitude across the strip. This amplitude will be the result of the addition, including phase, of the various components of the angular spectrum. These will all be rays with direction in the plane of incidence. The summation yields:

$$\mathcal{E}(x) = \int \mathcal{E}(\kappa_x) e^{i\kappa_x x} d\kappa_x$$

(9)

This expression can be recognized as having the form of a Fourier transform. Provided the functions satisfy certain simple conditions, Fourier's integral theorem permits us then to write:

$$\mathcal{E}(\kappa_x) = \frac{1}{2\pi} \int \mathcal{E}(x) e^{-i\kappa_x x} dx$$

(10)

where $\mathcal{E}(\kappa_x)$, the inverse Fourier transform of $\mathcal{E}(x)$, is the $x$-component of the electric field amplitude of the angular spectrum component. There will be similar expressions for the magnetic field. The infinite limits to the integration may be, at first sight, worrying. However, $\mathcal{E}(x)$ will be zero beyond the confines of the strip, and, provided $\lambda$ is rather smaller than the width of the strip, then $n \sin \theta_0$ will not be required to be excessively large. The mathematical purist will find the technique rigorously justified by Lalor[1].

Now let the strip have limits $x = \pm a/2$ and the illumination across the strip be of uniform amplitude, $\mathcal{E}_0$ and zero phase, so that $\mathcal{E}$ is real. The limits of (10) become $-a/2$ to $+a/2$ and its evaluation yields:

$$\mathcal{E}(\kappa_x) = \frac{1}{2\pi} \left[ \frac{\mathcal{E}_0 e^{i\kappa_x a}}{i\kappa_x} \right] = \mathcal{E}_0 \sin \frac{\kappa_x a}{2}$$

(11)

$\mathcal{E}_0$ being the amplitude of the normally incident component. The irradiance, or power per unit area, as a function of angle then becomes

$$\text{Irradiance} = I_0 \sin^2 \left( \frac{\kappa x}{2} \right)$$

(12)

where, strictly, this is the component of irradiance normal to the interfaces. Sinc, short for sine cardinal, is defined as

$$\text{sinc} \varphi = \begin{cases} \sin \varphi & (\varphi \neq 1) \\ \frac{\varphi}{\varphi} & (\varphi = 1) \end{cases}$$

(13)

The calculation is equivalent to Fraunhofer diffraction at a slit illuminated by a plane wave at normal incidence.

**Gaussian Beams**

The Fourier transform of a Gaussian function is another Gaussian function. The two functions:

$$\exp \left( -\frac{\mu^2 \kappa^2}{2} \right) \quad \text{and} \quad \exp \left( -\frac{x^2}{2\mu^2} \right)$$

(14)

with appropriate multiplying constants, are transform pairs. Let us replace the angular spectrum of our evenly illuminated slice by one that yields, at the same position, a Gaussian distribution of amplitude and, therefore, power. We can retain $a$ as a measure of the width of the slice but now it indicates that width at which the amplitude falls to $1/e$, and the irradiance to $1/e^2$, of its maximum value.

$$\mathcal{E}(x) \approx \exp \left( -\frac{4x^2}{a^2} \right)$$

(15)

From (14) the angular spectrum is given by

$$\mathcal{E}_0 \exp \left( -\frac{a^2 \kappa_x^2}{16} \right)$$

or, in terms of irradiance,

$$I_0 \exp \left( -\frac{a^2 \kappa_x^2}{8} \right)$$

(17)

The angle at which the amplitude falls to $1/e$, and the irradiance to $1/e^2$, of its maximum is given by

continued on page 40
Such beams are known as Gaussian. The distribution of irradiance in the angular spectra of the two beams is shown in Figure 2. Over the central part there is little difference. At larger incidence the Gaussian has a more regular shape.

\[ \kappa' = \frac{2\pi n_0 \sin \theta_0}{\lambda} = \frac{4}{a} \]  \hspace{1cm} (18)

\[ n_0 \sin \theta_0 = \frac{2\lambda}{\pi a} \]  \hspace{1cm} (19)

Figure 2. A comparison of the irradiance distributions in the angular spectra of the evenly illuminated slit (sinc\(^2\) function (12)) and the Gaussian spot (17).

All the components have zero relative phase at the origin in the center of the slit. At any point outside the slit the components can be added up to give a local resultant. It can thus be shown that the resultant planes of constant phase have a paraboloidal shape in the \( x-z \) plane. Their radius of curvature falls as the beam approaches the illuminated strip until it reaches a minimum and then rises to become infinite at the strip. At larger distances the irradiance is \( 1/e^2 \) of its central value at \( \theta_0 \).

\[ n_0 \sin \theta_0 = \frac{4\lambda}{\pi a} \]  \hspace{1cm} (20)

The corresponding value of \( \theta_0 \) is a useful measure of the beam divergence.

The minimum width is \( a \), known as the beam waist. There are many sources of more detailed information on Gaussian beams [2, 3].

Figure 3. The profile of a Gaussian beam. The bold lines are known as the asymptotes of the beam and mark the points at which the irradiance falls to \( 1/e^2 \) of the central irradiance.

**Coating Response**

The planes of constant phase in the resultant Gaussian beam are all curved except at the beam waist and this can lead to a misconception that the waist is where an interference coating must be inserted. As long as the coating is large enough to accommodate the entire beam it can be inserted anywhere.

The analysis so far shows that the response of the coating to a small spot of illumination is equivalent to its response in an illuminating cone. If a coating is tilted with respect to the direction of incidence, its characteristic moves towards shorter wavelengths. For larger tilts the characteristic may also exhibit increasing distortion and polarization splitting. A cone of illumination contains a range of incident angles and there may also be a degree of polarization in the cone. To keep the discussion simple, we will assume that the central ray of the cone is aligned to coincide with the design incidence of the coating and therefore experiences the theoretical performance. We will also assume that the divergence of the cone is small enough to avoid any
distortion of the characteristic except for a spectral shift and we will neglect polarization effects. The spot size will be the beam waist rather than the actual illuminated area. The effect of the cone will be, therefore, a smearing of the characteristic in wavelength. The degree to which this can be tolerated will depend on the sharpness of the spectral features of the coating and the tilt and divergence of the cone.

A useful technique for simplifying the analysis of tilt-induced spectral displacement uses the concept of effective index, \( n_e \). \([4, 5]\) This is the index of that single layer exhibiting the same degree of spectral shift. The effective index is within the range of the indices of the coating materials. A reasonable value for approximate results is the geometric mean of the extreme values.

We use approximations of the second order. The shift in wavelength of any feature will be such that the phase thickness of our effective layer remain constant.

\[
\frac{2\pi n_e d \cos \vartheta}{\lambda - \Delta \lambda} = \frac{2\pi n_i d}{\lambda}
\]

i.e.

\[
\frac{\lambda - \Delta \lambda}{\lambda} = \cos \vartheta = 1 - \frac{\vartheta^2}{2} = 1 - \frac{n_e^2 \vartheta^2}{2n_i^2}
\]

(21)

The quantity \( \Delta \lambda / \lambda \) can be set to be the allowable shift before the characteristic is too greatly degraded. Then, using (20), we can derive a value for the allowable spot size. At normal incidence this is:

\[
\frac{a}{\lambda} = \frac{\pi n_e}{2 \sqrt{2} \Delta \lambda}
\]

(22)

At oblique incidence it is a little more complicated. Let us suppose that the cone is tilted at an angle greater than the beam divergence then we have:

\[
\Delta \lambda = \frac{\vartheta^2_{\max} - \vartheta^2_{\min}}{2} = \frac{n_i^2}{n_e^2} (\vartheta^2_{\max} - \vartheta^2_{\min})
\]

(23)

The factor in parentheses on the right-hand side is twice the beam divergence time twice the mean tilt angle. This gives

\[
\frac{a}{\lambda} = \frac{8n_0}{\pi} \frac{\vartheta_0}{\Delta \lambda}
\]

(24)

where \( \vartheta_0 \) is expressed in radians.

A good rule of thumb when scanning a filter is that the spectral bandwidth should be not greater than one third of the width of the narrowest feature. Here we are dealing with a Gaussian function and the halfwidth is roughly one half of the width at the 1/e² points and so we can afford to increase the criterion to a width of perhaps 2/3.

Figure 4 shows a narrowband filter for 1000nm with a width of around 1nm. We can assume a value of \( n_e \) as \( 3/8 \) with air as the incident medium so that \( n_e \) is unity. Then at normal incidence the limiting \( \Delta \lambda / \lambda \) should be 2/3000 giving a minimum spot size of 0.075mm. At an angle of 10° the minimum spot size (beam waist) becomes 0.22nm. Values based on (23) are, for the figure, clearly slightly more conservative than those based on (25).

**REFERENCES**


Angus Macleod is currently Vice President of the Society of Vacuum Coaters. He was born and educated in Scotland. Then in 1979 he moved to Tucson, where he is President of Thin Film Center, Inc. and Professor Emeritus of Optical Sciences at the University of Arizona. His best-known publication is “Thin-Film Optical Filters,” now in its third edition. In 2002 he received the Society’s Nathaniel H Sugerman Memorial Award.

For further information, contact Angus Macleod, Thin Film Center, Inc. at angus@thinfilmcenter.com.

---

**BULLETIN READERS — if you have suggestions for the 2007 Technical Program in Louisville, KY, please contact one or all of the following individuals:**

- Program Chairs: Ludwik Martinu, École Polytechnique, Canada (lmartinu@polymtl.ca) and Ric Shimshock, MLD Technologies LLC (richshimshock@mld.com)
- Technical Director: Dom Mattox (dommattox@sve.org)