Optical Coating Design with the Essential Macleod

Angus Macleod and Christopher Clark
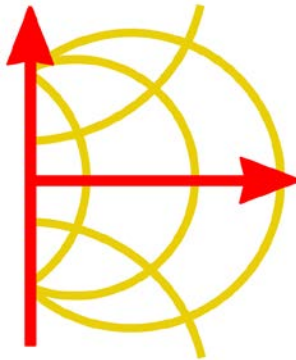
Contents

1   Introduction

Computer-aided design means exactly what it says. The computer is used to help in the design process not to take it over. The inspiration, the knowledge, the experience, the enthusiasm and the direction, are all brought to the process by the designer. However, design also includes tasks that are tedious, time-consuming, frequently repetitive and often involving such a volume of calculation as to be impossible to accomplish without assistance. The computer carries out these tasks effortlessly and tirelessly. It remembers everything that it is instructed to remember. It answers "what if" questions virtually immediately. It is almost the perfect personal assistant to the designer.

The word "almost" is used because the computer as assistant must be instructed in its tasks. Without instructions it will do nothing. The instruction may be simply the pressing of a particular key or the selection of a menu item, but it must, nevertheless, be given. Also at this current state of the art the instructions must usually be given in rather closely defined and limited ways. Otherwise they will not be understood and acted upon.

This book concentrates on the assistance that the computer can offer to the designer and on the communication between designer and computer. It is directed specifically at the Essential Macleod package and, therefore, emphasizes that particular interface. However, much of the material is more general in its application.

The manual supplied with the software is comprehensive but there is a difference between knowledge of a manual and actually being able effectively to use a software package. We can liken it to the understanding of an instruction book on driving a car and the ability to use the car to travel from A to B. The Essential Macleod software for the design and analysis of optical coatings is designed to be easy to use. There are all kinds of features that make problem solving easy and even enjoyable. No matter how comprehensive the manual, it really cannot substitute for a tutorial that covers tips, tricks, techniques, and examples. Note, however, that the manual is complete. It covers everything in the software. This book is not intended to compete with the manual but rather to supplement it. The manual tells you what the software does. This book tells you rather what you can do with the software.

The book assumes some knowledge of optical coatings. However, it starts with the fundamental definitions, sign conventions, and the like, because the structure of everything that follows depends on them. Many of the terms can be defined and expressions written in different ways that are equally valid, but also incompatible.

2.5   Coherence

The concept of coherence is important in all interference phenomena and, therefore, in thin film optics. When two ideal beams of light of identical frequency but with phase difference φ, are combined, the irradiance of the sum can be written

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\, \cos\varphi \qquad\qquad (2.36)$$

The cosine term is known as the interference term and it can be either positive or negative, depending on φ. Now let the two ideal light rays each be replaced by a bundle of rays. The interference term that results will be the sum of a large number of like terms, each with a different φ. If the interference term that results is essentially undiminished then we term this combination of rays coherent. Should the interference term disappear, then the term is incoherent. Partially coherent is the term we employ when the interference term is diminished but not eliminated. The two most important cases are the coherent case and the incoherent case and we limit this discussion to these. Although the incoherent case is simply the coherent case integrated over a sufficiently wide range of parameters to eliminate the interference effects, it is usual to calculate those components that clearly support incoherent beam combination by simply summing the irradiances.

How can we recognize when the beam combination should use the incoherent summation? Here the concept of coherence length can be very useful.

Let us express our primary light in terms of its spectral components and let the spectrum be a continuous band of wavelengths centered on λ and with a bandwidth of $\Delta\lambda$. We produce interference fringes using this light with a variable path difference. Each elemental wavelength will produce its own set of fringes and as we increase the path difference the fringes will be smeared out and the contrast will fall. The path difference at that point where they just disappear will be a measure of the coherence length. We can define the point as that where the fringes will become smeared out over the interval between fringes. In other words, a fringe of order m at the smallest wavelength will just coincide with one of order m-1 at the longest wavelength.

$$\text{Path difference} = m\left(\lambda - \frac{\Delta\lambda}{2}\right) = (m-1)\left(\lambda + \frac{\Delta\lambda}{2}\right) \qquad\qquad (2.37)$$

and some mild analysis arrives at the result

$$\text{Coherence length} = \left(m - \frac{1}{2}\right)\cdot\lambda = \frac{\lambda^2}{\Delta\lambda} \qquad\qquad (2.38)$$

This is an exceptionally important and useful result. The coherence time will be this result divided by the velocity of light. Since these are all vacuum wavelengths, the coherence length is automatically referred to vacuum and the coherence time will be given by

$$\text{Coherence time} = \frac{\lambda^2}{c\Delta\lambda} \qquad\qquad (2.39)$$

where c is 299.792458 nm/fs.

Cones of illumination also have coherence effects and we can calculate an effective coherence length that can be useful. If ϑ is the cone semiangle and the cone axis is at normal incidence then, using an approximate two-dimensional model, we can rewrite (2.37) as:

$$m\lambda = (m+1)\lambda \cos \vartheta \qquad (2.40)$$

and a similar extraction process yields:

$$\text{Coherence length} = m\lambda = \frac{\lambda \cos \vartheta}{1 - \cos \vartheta} \simeq \frac{2\lambda}{\vartheta^2} \qquad (2.41)$$

or, if ϑ is given in degree

$$\text{Coherence length} = \frac{7 \times 10^3 \lambda}{\vartheta^2} \qquad (2.42)$$

Coherence length should not be thought of as an exceedingly precise parameter. Fringes fade away rather than abruptly disappear. Nevertheless it is an exceptionally useful concept. We can estimate the coherence length in any system and if it is less than the path difference we can safely use the incoherent combination. If it is greater than the path difference we should use the coherent calculation, if necessary, integrating over the particular range of wavelengths or angles of incidence to give the partially coherent case. A coherence length that is an order of magnitude greater than twice the optical thickness of the film (remember the path difference is twice the optical thickness) will safely indicate the coherent case. We can see from (2.38) and (2.42) that for values of bandwidth of around 1nm and cone angles of 1° or so, the coherence length will normally be much greater than the path difference in our thin films (although very narrow-band filters may show a perceptible effect) while less than the path difference corresponding to a substrate of a few millimetres in thickness. Thus we shall usually treat substrates and similar components as supporting the incoherent case and our thin film coatings as supporting the coherent case.

Bandwidth and cone tools are available in the Essential Macleod as part of the calculations in a Stack document. They will be described later in Section 9.5.

$$R = \left[\frac{y_0 - \dfrac{y_H^{2x}}{y_L^{2(x-1)} y_{sub}}}{y_0 + \dfrac{y_H^{2x}}{y_L^{2(x-1)} y_{sub}}} \right]^2 \qquad (7.8)$$

Reflectance (%)

Layer Thickness (Quarterwaves)
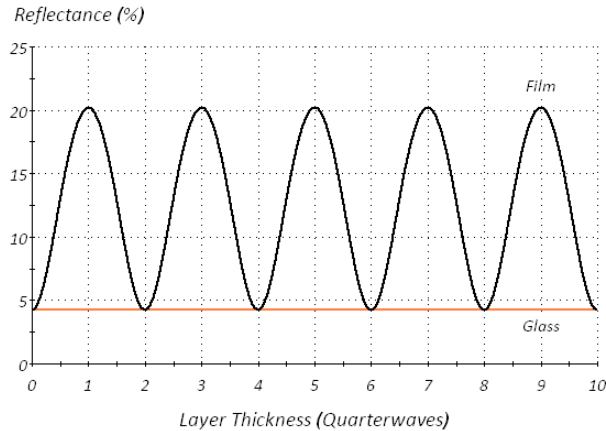
Figure 7-1. Calculated reflectance of a 2.00 index film on a substrate of index 1.52 in air of index 1.00.

7.2.2   The Admittance Diagram

We can imagine that we are gradually depositing a thin-film coating and that we measure all the time the surface admittance at the front of the growing coating. The record of the surface admittance will be a continuous one because of the continuity conditions on tangential electric and magnetic fields at any boundary. Except in the special cases of the previous section, the admittance will normally be complex, and its values will form a continuous locus in the complex plane. This locus is known as the *admittance locus* and the complete diagram as the *admittance diagram*.

Since the surface admittance of any plane parallel to the interfaces depends only on the structure to the rear of the plane, the admittance diagram can also be considered to be the record of the variation of the surface admittance of the plane as it is moved smoothly through the coating from back to front. It is essentially a graphical representation of the calculation of the properties of the coating. The admittance diagram (unless there is gain in the system) is confined to the right hand part of the complex plane.

For a single dielectric layer of characteristic admittance y_f, the locus is extremely simple. It is a circle centered on the real axis and cutting it in two points with admittances α and β, the product of which is the square of the layer characteristic admittance, which will be recognized as the quarterwave rule.

$$\alpha\beta = y_f^2 \qquad\qquad (7.9)$$

The real axis, and a circle centered on the origin and passing through the point y_f on the real axis, cut the circular locus into eighth waves.
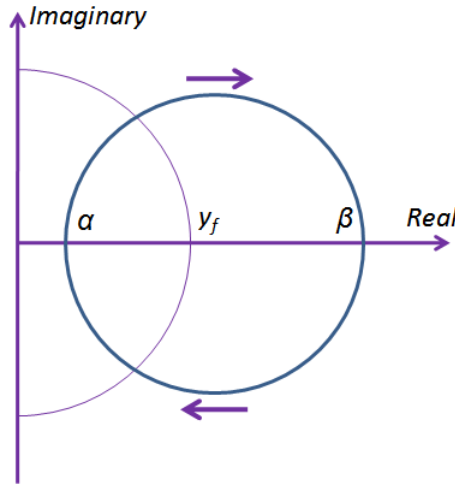
Figure 7-2. The admittance locus of a dielectric layer of admittance y_f. The locus is described clockwise. A complete circle corresponds to a halfwave while the semicircles above and below the real axis correspond to quarterwaves. The circle centered on the origin and passing through y_f cuts the locus further into eighth waves. The points of intersection with the real axis are related by the quarterwave rule, $\alpha\beta = y_f^2$.

The starting point for the circular locus must lie on the locus. If the starting point moves so that it is no longer on a particular circle, there will be a new circular locus that will pass through the new starting point and also cut the real axis according to the quarterwave rule. We can imagine a set of admittance circles that all apply to a particular material with characteristic admittance y_f, say, but each corresponding to a different starting point. Each circle will be entirely outside or entirely inside the other circles of the set. None will intersect. The point y_f on the real axis will be inside every circle, although it will not be the center of any. We call such a set of circles a nested set and describe them as nested about the point y_f. Figure 7-3 shows such a set of nested circles. In this case they correspond to material of admittance 1.38. The real axis and the blue circle centered at the origin and through 1.38 cut every circle into eighth waves. This gives a good idea of the way in which the optical thickness varies around the circles.

When absorption is introduced into a dielectric layer the loci are no longer circular but gradually spiral in towards the point $(n-ik)$ the rate of approach being greater, the greater the value of k.

Imaginary Admittance
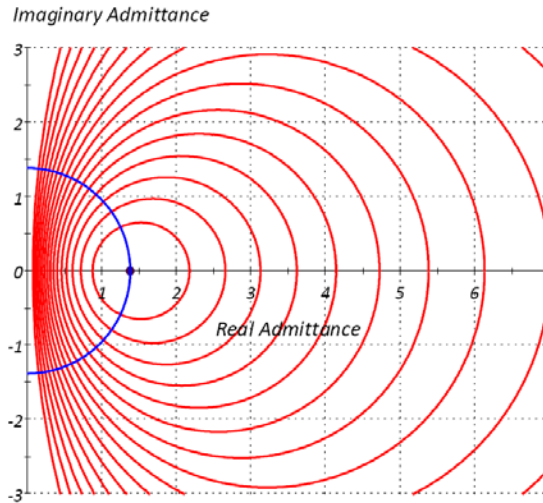
Figure 7-3. A set of possible admittance loci for dielectric material of
characteristic admittance 1.38. The admittance circles form a set that is
nested about the point 1.38 on the real axis. Each circle corresponds to
the same characteristic admittance of 1.38 but has a starting point that is
on no other circle.

For metals, the locus is a little more complicated, Figure 7-4. If the losses are zero, that is the real part of the characteristic admittance is zero, then the locus is exactly circular with center on the real axis and passing through the points k and $-k$ on the imaginary axis. If the losses are small, so that n is now finite, then the locus remains virtually circular but is slightly rotated so that it passes from the point $(-n, k)$ to the point $(n, -k)$, the part on the left-hand side of the imaginary axis being inaccessible. It is as if the pattern were rotated through a small angle about the origin. The phase thickness, $2\pi nd/\lambda$ is not a useful concept when dealing with metals. A related quantity, $2\pi kd/\lambda$, sometimes denoted by the symbol β, is a more useful parameter. Contours of constant β when n is zero are sets of circles centered on the imaginary axis and nested about the points k and $-k$ on the axis. These are shown in Figure 7-4 using thin lines. When n is rather larger, as in metals like chromium and nickel, the circular loci become distorted into spirals.

The complete admittance locus of a coating with more than one layer consists of a series of segments, each corresponding to a particular layer. These segments are attached to each other without breaks, so that the complete locus from substrate to outer surface of the coating is continuous.

A set of constant reflectance contours may be added to the admittance diagram, Figure 7-5. These contours are circular and coincide with the set of admittance circles of the incident medium material, zero reflectance at, and only at, the point y_0 on the real axis. They allow the reflectance of the coating to be readily judged. Note that the geometry of the figure implies that the extrema of reflectance coincide with the intersections of the loci with the real axis. It is also easy to see that provided the admittance of the film is greater than that of the incident medium, as in Figure 7-5, then that part of the locus

below the real axis represents falling reflectance with increasing thickness and that part above represents rising reflectance.
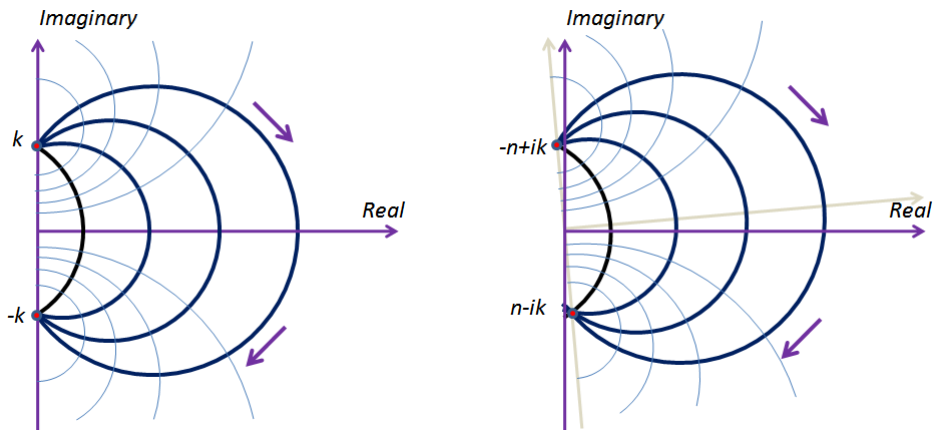
Figure 7-4. Metal admittance loci. The left-hand side shows an ideal metal with zero n. The circles intersect at k and $–k$ on the imaginary axis. The right-hand side shows a metal with small but finite n. The loci now intersect at the points $(-n+ik)$ and $(n--ik)$.
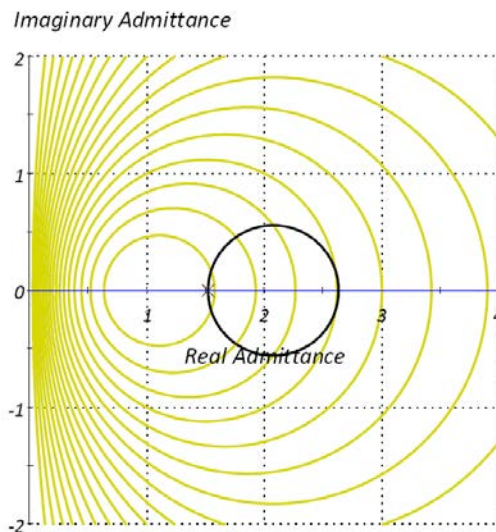
Figure 7-5. The yellow circles are the isoreflectance contours for air incident medium spaced at intervals of 5%. The black circle is the admittance locus corresponding to the design of Figure 7-1. The reflectance extrema are at the intersections with the real axis.

8.18 Rugate

Incident Medium: 1.52		Substrate: 1.52	
λ_0: 1000nm			
Symbol	Material	Optical Thickness (FWOT)	Physical Thickness (nm)
None.	1.45	-	
None	2.00	-	
Starting Formula	See Rugate Generator …		
Calculation range	250nm	2000nm	

Performance

Reflectance (%)

Wavelength (nm)
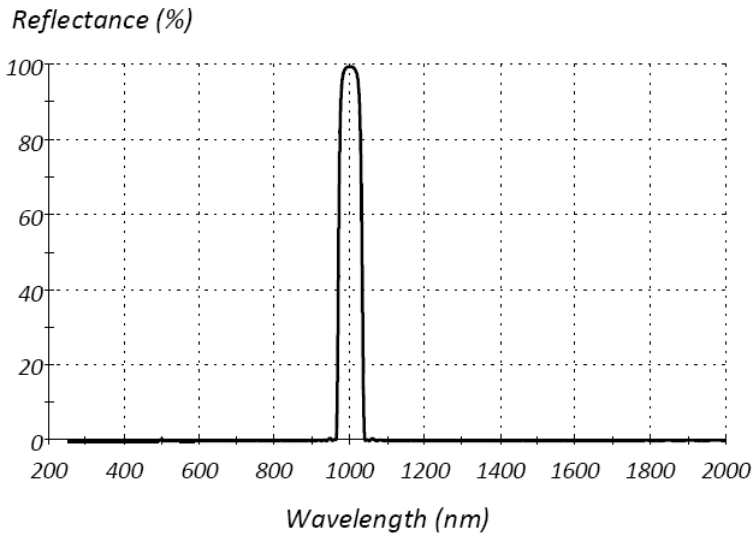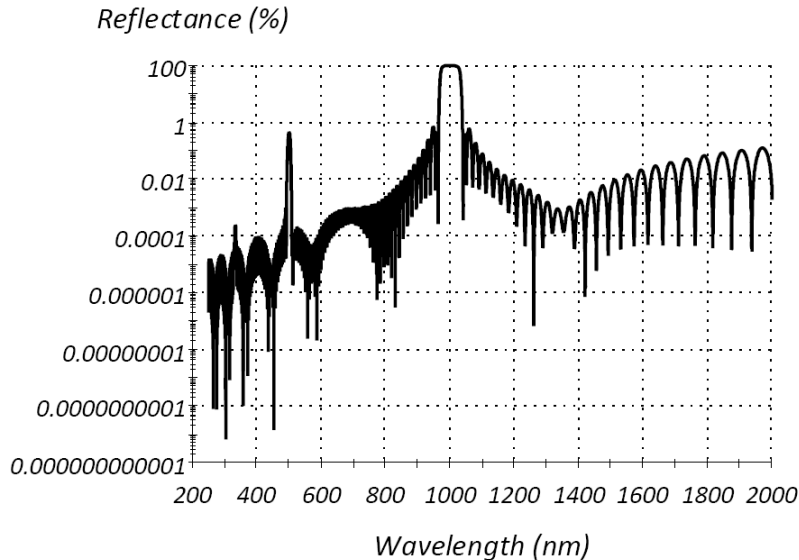
Notes

A rugate filter in the classical sense is an inhomogeneous layer with a sinusoidal variation of index through it. The characteristic of the filter is a single reflection peak, rather like a quarterwave stack with all the higher orders suppressed. The term rugate, however, is frequently applied to any inhomogeneous system and it is this broader sense that applies to the Generate Rugate… tool. There is no exact general analytical solution to the problem of the calculation of the properties of an inhomogeneous layer although there are some special cases. The calculation technique, therefore, is to use a discrete model with sufficiently thin layers. We judge the layers as sufficiently thin when further thickness reduction has no effect. Usually an optical thickness of around one tenth of the shortest wavelength to be calculated suffices. The tool permits easy variation of layer thickness. We use it here to set up a simple rugate to illustrate its use.

The rugate is divided into three parts. The first and third parts are inhomogeneous matching structures that vary smoothly in index between incident medium and rugate and substrate and rugate respectively. Here we use a sinusoidal profile and make them one halfwave thick at the longest wavelength of interest. The rugate part is a single line (reflection peak) with sinusoidal variation between a high and low index, NHigh and NLow respectively. It is apodized (that is multiplied by a function to reduce ripple) using a half sine cycle. To simplify matters we use non dispersive materials, 1.52 for the substrate, 1.52 for the incident medium and a mixture of 2.00 and 1.45 for the coating materials, corresponding roughly to a silicon oxynitride material.

The macro has been written to be easily understood so that there are more variables defined and more lines of code than might be strictly necessary. It could be made much more compact. The code is reproduced below. Default comment lines that are automatically shown in the window list predefined variables and are not shown below. Note that comments start with "!". Executable statements end with ";".

We first set up certain variables. The statements are all editable so that the variables can easily be changed if needed. RugateWavelength is the center wavelength of the rugate peak and LongestWavelength is the longest wavelength of interest. We are working in nm. Note that we specify the rugate amplitude in terms of index. The manual shows an example where the excursion of packing density is specified.

If you use the macro to generate or replace a design do make sure once the design is entered that the incident medium and substrate are set to 1.52 material. Try changing the total thickness and the number of layers and see the effect. If you want to change the incident medium to air then leave the rugate generator as it is but add a discrete layer antireflection coating to the front surface to match 1.52 to air.

Design Technique

Generate Rugate.... The code for Generate Rugate follows:

Settings:

Material	2.00
Total Thickness	30.000000
Reference Wavelength	1000.00
Void Material	1.45
Void Density	1.00000
Number of Layers	500
Angle Units	Degrees

Note that the statements should be on one line each. Some, because of the formatting for this page have spread onto two lines. Comments are prefixed by and exclamation mark (!) and are shown also in green here for greater clarity.

!---
! IMPORTANT
! NInc and NSub must not be less than NVoid.
! Void Density must be set at 1.00
! Optical Thickness Convention must be set as FWOT
!---
!
! The rugate is made up of three sections, two matching structures, one at either end
! and a central single-line rugate
! First we must define certain parameters. The rugate wavelength can be changed
! if desired.

RugateWavelength=1000;

LongestWavelength=2000;

NInc=1.52;

NSub=1.52;

NMaterial=2.00;

NVoid=1.45;

! Next the index excursion for the sine wave variation of the rugate

NHigh=1.8;

NLow=1.6;

! Now we calculate the necessary thicknesses of the matching assemblies.

MatchThickness=LongestWavelength/(2*ReferenceWavelength);

! The thickness available for the rugate is what is left.

RugateThickness=TotalThickness-2*MatchThickness;

! The three regions must be defined by logical statements. The region will be recognized

! by a value of unity.

! This means we must have a minus sign in front of the logical statements as shown.

! Remember that Thickness is a predefined variable that corresponds to the thickness at

! the center of the current layer.

Match1=-(Thickness<=MatchThickness);

Rugate=-(Thickness>MatchThickness&Thickness<TotalThickness-MatchThickness);

Match2=-(Thickness>=TotalThickness-MatchThickness);

! The required parameters are mostly given in this rugate in terms of refractive indices.

! The rugate design will be defined in terms of packing density and so we have to

! convert.

! We need the terminal packing densities at incident medium and substrate. These

! give terminal indices equal to those of incident medium and substrate respectively.

PFirst=(NInc-NVoid)/(NMaterial-NVoid);

PLast=(NSub-NVoid)/(NMaterial-NVoid);

! Next calculate the swing in packing density for the rugate.

PHigh=(NHigh-NVoid)/(NMaterial-NVoid);

PLow=(NLow-NVoid)/(NMaterial-NVoid);

PMean=(PHigh+PLow)/2;

PAmplitude=(PHigh-PLow)/2;

! The argument of the sine function that will generate the rugate is required.

! Note that a complete rugate cycle occupies one half of the rugate wavelength.

! Also the design will be defined at the reference wavelength that may be different from

! the rugate.

! We want the sine function to start at zero angle and so we have to subtract the

! thickness of the first ! matching section from the variable Thickness to obtain the

! parameter we need to insert in the argument.

RugateArgument=360*(Thickness-MatchThickness)*ReferenceWavelength*2/RugateWavelength;

! The packing density variation for the rugate is then a simple sine function.

PRugate=PAmplitude*sin(RugateArgument);

! Additional rugates at different wavelengths can be added if required.

! For example:

! PRugate=PAmplitude1*sin(RugateArgument1)+PAmplitude2*sin(RugateArgument2)

! etc.

! Now we need the apodizing function. The rugate sinusoidal variation will be

! multiplied by this.

! We use a half cycle of a sine function so that the rugate thickness corresponds to

! 180 degrees.

Apodization=sin(180*(Thickness-MatchThickness)/RugateThickness);

! Finally we set up the various parts of the entire structure. The rugate part so far is just

! a sinusoidal variation. It must be multiplied by the apodizing function and then must

! be added to the mean packing density. For the matching structures we find it easier

! to use a single quartercycle of a sine function and to square it to give the shape we need

! (equivalent to 4th and 1st quadrants).

Match1: PackingDensity=PFirst+(PMean-PFirst)*(sin(90*Thickness/MatchThickness))^2;

Match2: PackingDensity=PLast+(PMean-PLast)*(sin(90*(TotalThickness-Thickness)/MatchThickness))^2;

Rugate: PackingDensity=PMean+Apodization*PRugate;

11 Short-Pulse Phenomena in Coatings

Traditionally, coating designers have been able to rely on the steady-state nature of the effects they seek to produce. There are now laser systems, known as ultrafast, capable of generating pulses of light that are short enough for transient response to become significant. Optical communication uses trains of very short pulses that are manipulated by beam splitters in the form of multiple-cavity narrow band filters. These pulses are rather longer than those in ultrafast systems but on the other hand the effect of the multiple-cavity filters is correspondingly greater. The theory of the effect in either case is similar [1, 2].

The speed of light in vacuo is 299.792 nm/fs. A normal high reflector consisting of a quarterwave stack might be some twenty-five quarterwaves in thickness. At a wavelength of 1µm this implies a trip length for light traveling from the front to the rear of the coating and back again of 12.5µm or a trip time of around 42fs. Pulses that are around 50fs in length are now common in the ultrafast field and pulses can be an order of magnitude still shorter. It is clear that transient response of coatings must now be considered important in such applications, but the effects, in fact, can be significant even with pulses some two or three orders of magnitude longer. The idea that coating properties should have an influence on short pulses and that they might be engineered to have prescribed effects is not new. It is, however, only recently that the field has expanded and the technology advanced to the stage where the application is becoming of major importance. In the communication field the components in question are multiple-cavity narrowband filters with extremely narrow widths. The pulses are reflected backwards and forwards many times before they emerge from these components and so the transient effects can be measured in terms of picoseconds rather than femtoseconds.

11.1 Short pulses

A short pulse can be thought of as an envelope over a carrier. The carrier contains the phase information associated with the pulse and it travels at what is known as the phase velocity. The energy is obviously associated with the envelope that travels at what is known as the group velocity. The group velocity and the phase velocity are generally different, the phase velocity being the greater. Thus the carrier appears to run through the pulse envelope. A short pulse with Gaussian envelope is shown in Figure 11-1.

The pulse may also be visualized in a different way, as a collection of monochromatic component waves with a continuous distribution of frequencies over a given band. The coherent combination of these monochromatic waves yields the envelope and carrier of the alternative model. Both of the models are entirely equivalent and if we wish, we can pass from one to the other by way of a Fourier transform.

Pulse envelopes frequently have a Gaussian shape. For simplicity we can look at the temporal variation at the origin of our coordinates, $t = 0$, $z = 0$, and then

$$F(t) = Ae^{-\frac{t^2}{2\mu^2}} \tag{11.1}$$

where μ has the dimensions of time. The Fourier transform gives the frequency distribution and it is also a Gaussian function,

$$G(\omega) = \mathcal{B}e^{-\frac{\mu^2(\omega-\omega_0)^2}{2}} \tag{11.2}$$

and it gives the form of the variation of the amplitude of the component wave of frequency ω. If the time between the half-maximum points [F and G having values of $1/\sqrt{2}$ times peak] is τ and the full width of the pulse (angular) frequency distribution also at half-maximum is $2\Delta\omega$ [we will be using $\Delta\omega$ later as the frequency width from the peak to the half-peak point] then

$$\tau \cdot 2\Delta\omega = 4 \cdot \log_e 2 \tag{11.3}$$

Note that both these quantities are functions of μ. For example,

$$\tau = 2\sqrt{\log_e 2} \cdot \mu \tag{11.4}$$
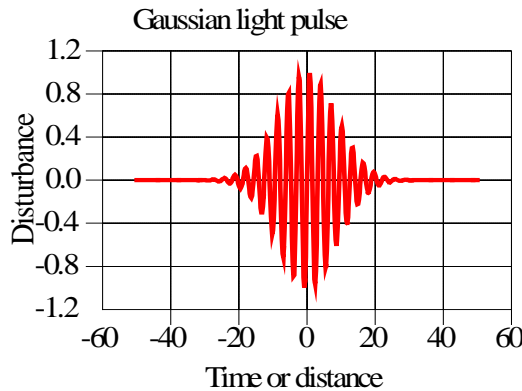
Figure 11-1. A short Gaussian-shaped pulse consisting of an envelope over a carrier of constant frequency. The carrier phase may move faster than the pulse when it will appear to run through the envelope as it travels.

The center of the pulse is the point where all of the component waves are exactly of identical phase. If all the component waves travel at the speed of light in vacuo then the phase coincidence will also travel at that speed and the center of the pulse will move with it. Similarly if all waves slow down equally then the pulse will slow down to the same extent but will otherwise be unchanged.

11.2   Group Velocity

The relative phase of the carrier within the pulse is set by the value of phase where all the component waves coincide. If the phase of the waves is zero then the carrier will have a peak exactly at the peak of the pulse. We can find the position of the pulse peak at any time by a simple procedure.

The pulse can be considered to be made up of monochromatic component waves. As these propagate the phase relationships between them will change, but if the pulse shape

is unaltered as it propagates then at any particular time there must be a distance along the path where the phase is identical for all the component waves, and this must correspond to the pulse center. We use the normal thin film convention of $(\omega\tau\text{-}\kappa z)$ in the phase factor where $\kappa = 2\pi n/\lambda$ with λ the free space wavelength. We write the component wave phase at distance z and time t as $\varphi = \varphi_0 + \Delta\varphi$. Then for coincidence of all component phases, $\Delta\varphi$ must be zero.

This condition is

$$\left(\omega_0 + \Delta\omega\right)t - \left(\kappa_0 + \Delta\kappa\right)z = \varphi_0 + \Delta\varphi$$
$$\omega_0 t - \kappa_0 z = \varphi_0$$
$$\Delta\varphi = 0 = \Delta\omega t - \Delta\kappa z \tag{11.5}$$
$$z = \frac{\Delta\omega}{\Delta\kappa}t = v_g t$$

The quantity $\Delta\omega/\Delta\kappa$ is known as the group velocity, v_g, and clearly it must remain constant if the position z is to be the same for all the component waves and the shape of the pulse unchanged.
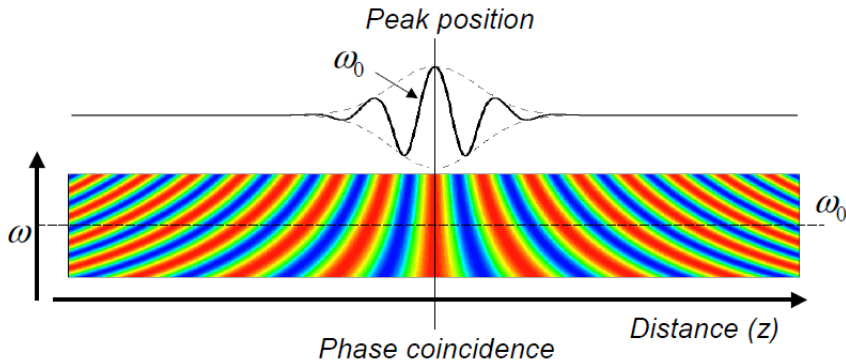
Figure 11-2. Sketch showing the component waves of the pulse with the z-axis as the direction of propagation. The sinusoidal profile of the components is modeled as a set of colored contours that can be thought of as projecting out of and into the plane of the diagram. The pulse peak coincides with the position where the phase of all the components is exactly equal and that coincidence is clearly seen in the center of the figure. In an actual case the amplitudes of the individual components will reduce following a Gaussian expression as the frequency moves further from ω_0, the carrier frequency.

An alternative visualization involves a simple diagram. We plot the z-direction horizontally and ω vertically. We reproduce the profile of the bundle of component waves making up the pulse, as a set of colored contours. As ω increases the wavelength reduces so that the spacing of the contours reduces. Contours of constant φ follow the colored bands. Provided there is one contour that runs normally across the lines, and that

can be seen in the center of the diagram, then the pulse peak will be positioned there and the pulse shape will be unchanged.

In a nondispersive medium the phase at the peak will be constant (zero in our case because of the starting condition) because all the component waves will be traveling at identical velocity even though it may be less than the velocity in free space. In a dispersive medium, the component waves travel at different velocities according to the particular value of refractive index. Provided the variation in velocities still permits a phase coincidence somewhere, then the pulse will appear there and will be unchanged in shape although the phase of the carrier wave will be altered. It is clear from equation (11.5) that the critical condition is for the group velocity to remain constant across the frequency spectrum of the pulse.

In a dispersive medium, the refractive index changes with frequency. We can calculate the group velocity in terms of this change.

$$\kappa = \frac{2\pi n(\omega)}{\lambda} = \frac{\omega n(\omega)}{c}$$

$$\frac{dk}{d\omega} = \frac{n(\omega)}{c} + \frac{\omega}{c} \cdot \frac{dn(\omega)}{d\omega}$$

$$v_g = \frac{c}{n(\omega) + \omega \cdot \dfrac{dn(\omega)}{d\omega}} \tag{11.6}$$

In a medium with normal dispersion, this is, unfortunately, not constant.

11.3   Group Velocity Dispersion

There is no guarantee that the group velocity be constant with changing frequency. If the second derivative of ω with respect to κ is non zero then there can be no phase coincidence and the pulse will be perturbed. If we limit ourselves to the second derivative then we can write the expression for the phase of an arbitrary component wave as:

$$(\omega_0 + \Delta\omega) \cdot t - \left(\kappa_0 + \Delta\omega \cdot \frac{d\kappa}{d\omega}\bigg|_0 + \frac{1}{2}(\Delta\omega)^2 \cdot \frac{d^2\kappa}{d\omega^2}\bigg|_0 \right) \cdot z = \varphi + \Delta\varphi \tag{11.7}$$

and we can immediately identify a problem. The third term in the coefficient of z is even in $\Delta\omega$ and so cannot be compensated by the other terms. A simple way of handling this problem is to pair the component corresponding to $+\Delta\omega$ with that corresponding to $-\Delta\omega$. In each case we ensure that the value of $\Delta\varphi$ is zero. This implies introducing a different time t for each member of the pair although we retain the same value of z, and, therefore, gives two equations for each $\Delta\omega$ instead of the usual one.

$$\Delta\omega \cdot t_1 - \Delta\omega \cdot \frac{d\kappa}{d\omega}\bigg|_0 \cdot z - \frac{1}{2}(\Delta\omega)^2 \cdot \frac{d^2\kappa}{d\omega^2}\bigg|_0 \cdot z = 0 \tag{11.8}$$

$$-\Delta\omega \cdot t_2 + \Delta\omega \cdot \frac{d\kappa}{d\omega}\bigg|_0 \cdot z - \frac{1}{2}(\Delta\omega)^2 \cdot \frac{d^2\kappa}{d\omega^2}\bigg|_0 \cdot z = 0 \tag{11.9}$$

Then
$$\Delta t = (t_1 - t_2) = \Delta\omega \cdot \frac{d^2\kappa}{d\omega^2} \cdot z \tag{11.10}$$

The component pairs are, therefore, displaced from each other in time and, of course, the pulse must be broadened in consequence. We could treat the $\Delta\omega$ in expression (11.10) as half the width of the frequency distribution of the basic initial pulse, so that the range of frequencies would then be from $\omega_0 - \Delta\omega$ to $\omega_0 + \Delta\omega$. The pulse would then be smeared out by the Δt in (11.10). Actually, more rigorous calculation shows that we should better use a value of $2\Delta t$ instead of just Δt

Then the new pulse width is given by

$$\tau^2 = \tau_0^2 + (2\Delta t)^2$$

If we write

$$\tau_g^2 = \frac{d^2\kappa}{d\omega^2}\bigg|_0 \cdot z$$

then,

$$\tau^2 = \tau_0^2 + (2\Delta\omega\tau_g^2)^2$$

and since $\Delta\omega = (\sqrt{\log_e 2})/\mu^2$ and $\tau_0 = 4(\sqrt{\log_e 2})\mu^2$,

$$\tau = \tau_0 \left(1 + \frac{\tau_g^4}{\tau_0^4}\right)^{1/2} \tag{11.11}$$

Note that $(\tau_g/\mu)^4$ rapidly becomes negligible when τ_g becomes less than μ but rises sharply when it is greater. This means that as long as long as τ_g is comfortably less than μ its effect can be neglected.

Figure 11-3 shows a sketch of the spectral components of the pulse in the presence of appreciable group velocity dispersion. At the center of the pulse there is no exact phase coincidence. All that can be done is to pick the best position where the phase varies least in total across the pulse. We can think of the pulse as split into two parts each of which can be thought of as representing a pulse on its own, with its own group velocity. Because these group velocities are different, there is a difference in the arrival times of

the two sub-pulses representing a broadening of the resultant. Also, as can readily be seen from the diagram, the carrier frequency is different. In the case illustrated the frequency at the front of the pulse is lower than that at the rear. The resultant pulse is said to be chirped.
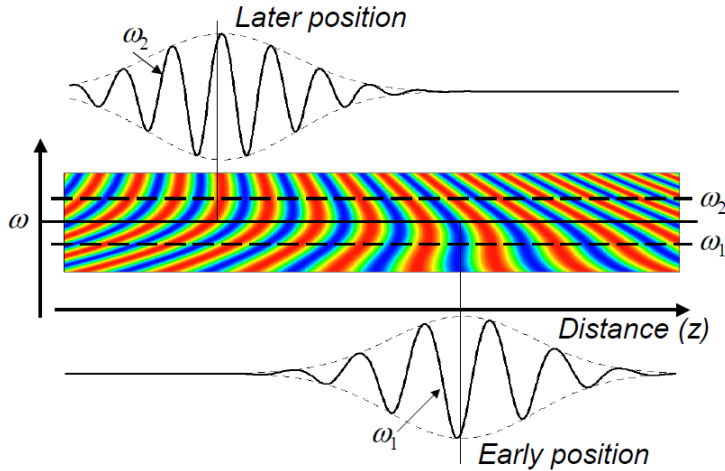

Figure 11-3. The pulse frequency distribution can now be thought of as roughly split into two parts, each of which can be considered as representing a component pulse with its own center position and carrier frequency. Since the group velocity is different for the two component pulses they separate such that one lags behind the other and the combined pulse is broadened and its carrier frequency gradually shifts from ω_1 to ω_2.
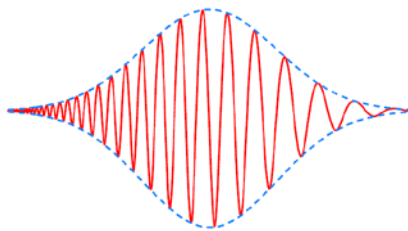

Figure 11-4. A sketch of a chirped pulse, that is, one where the carrier frequency is varying through the pulse. In this case the frequency is lower at the front than at the back and this would be the result of normal dispersion.

A more rigorous calculation of the effect of the group velocity dispersion on the pulse is a little involved. We choose the position z where we will measure the pulse. We will then treat z as a constant in the analysis. Then, at z, each individual component of the pulse can be written in the form of an amplitude that depends on the frequency together with the phase factor. This is:

Set MyTable = CreateObject("EMacleod.TableCreator")

but note that there is an alternative way of creating a table covered later.

To access an application that supports Object Linking and Embedding such as Microsoft Word we write:

Dim MyWord As Object

Set MyWord = CreateObject("Word.Basic")

Messages can then be sent to Word. See the scripting help file for more information and some examples. Obviously knowledge of the class Word is required to make use of this facility.

18.5.2   Using ThisSession and other Objects

The ThisSession object is always defined in a script. ThisSession contains methods that allow creation of most of the objects in the Essential Macleod. For example, to create a Materials object that gives access to the list of available materials and the data in any chosen material, use the GetMaterials method.

Dim myMaterials as Object

Set myMaterials = ThisSession.GetMaterials

A new table can be created using

Dim MyTable As Object

Set MyTable = ThisSession.NewTable

The table object created in this way differs from the table object created through EMacleod.TableCreator. See the section on tables below.

The ThisSession object also contains a method for opening an existing design. The path and name of the design file should be entered as a string.

Dim myDesign as Object

Set myDesign = ThisSession.OpenDesign("C:\Users\John\Documents\Thin Film Center\Designs\Four Layer AR.dds")

Note that Set is used only when an object variable is the subject of the statement. If the statement concerns a property that is a simple variable then the form is that of a straightforward assignment as in:

Dim Names() As String

Names = ThisMaterials.GetSubstrateNames

where Names is a string array and not an object.

18.5.3   Summary

Most of the various objects that can be created using either the CreateObject expression with the appropriate EMacleod class or the methods in the predefined objects are shown in the following table.