F. Gracer
R. A. Myers

Graphic Computer-assisted Design of Optical Filters

Abstract: An interactive graphic system is described in which a designer using a light pen creates an optical filter design at an IBM 2250 Graphic Console controlled by an IBM 1130 computer. The designer can observe the reflectivity (or some other property) of the filter as a function of wave number plotted on the CRT. The application was developed using an experimental graphic version of the 1130 Continuous System Modeling Program (CSMP), a general purpose, block-oriented simulation language in which the functional blocks represent the elements and organization of an analog computer. The designer has available a full set of operators for further analyzing the behavior or modifying the design of the filter. In addition to providing a highly flexible analytic tool, the system is intended to explore means for making the interactive computer terminal an important element in the inventive process.

Introduction

The process of design of a multilayer optical filter is characteristic of a wide range of problems which are particularly amenable to computer-aided solution. While *analysis* of the properties of a specified design is straightforward and can be implemented on a computer, automatic *synthesis* of a design from a desired set of requirements is very difficult. In general, the latter problem has not been solved except for limited systems, e.g., lens design, for which optimization techniques have been found.

Most filter designers formulate a trial design and then change the optical structure and parameters of the filter by successive approximation until the requirements have been met. The changes are based on the designer's insight and experience. With the conventional use of computers this process is discontinuous since the desired changes cannot be made easily at the computer. As a result fewer iterations are performed than might be desirable for the achievement of an optimal design.

A powerful design tool should respond *rapidly* to each trial design and feed back results on-line to the designer in the language of his problem. This would permit him to maintain his train of thought and respond with appropriate changes in a fluent and natural way. With this kind of interaction he could be more thorough in evaluating his ideas and exploring alternate solutions.

Most designers tend to represent a design pictorially as a block diagram with appropriate symbols having mnemonic value. The most convenient form of feedback to the designer is *also* pictorial—a graph or plot. If the computer could accept the block diagram as input and produce graphic output, the designer need not translate to and from computer format. With an *interactive* graphic system he could make graphic changes directly on the cathode ray tube, as he would with pencil and paper.

This paper is intended to show the usefulness of interactive graphic design and the ease of implementation of such a system in the design and analysis of optical filters. This application was developed using a graphic version of the 1130 Continuous System Modeling Program[1] (CSMP). This program is a general purpose analysis tool with a variety of familiar mathematical functions and it can be applied to many kinds of problems with minimal effort.

In the present application the designer can request, for example, a plot of reflectivity of the filter as a function of wavenumber. This request is made at the CRT console using a light pen and a keyboard; the resultant output is a display of the requested curve on the CRT. Changes to the structure or parameters of the filter can be made with simple light pen actions. Many different functions of wavenumber can be computed and displayed by appropriate use of program facilities.

The underlying mathematical structure of optical filter theory facilitates the use of the program for a number

F. Gracer is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598; R. A. Myers is at IBM Corporate Headquarters, Armonk, New York 10504.

of other physical problems. In particular, problems involving iterated two-port elements, such as transmission line filters,[2] can be handled with the program described in this paper.

Optical filters

When the thickness of a sheet of transparent material begins to approach the wavelength of light, interference effects become noticeable. These well-known effects are of enormous technological significance. Even without the glamorous application to lasers, thin optical films are ubiquitous in industry as, for example,[3] antireflection coatings, beam splitters, narrow-band filters, dielectric mirrors, and polarizers.

Each of these applications and each different filter within an application is a new design problem. A filter may have a single layer or more than a hundred; it may be made of any one of several dozen materials or of several different materials. The order in which these materials are used and the thickness of each layer must be specified by the designer.

The problem at hand is the following: Given a set of M layers of thickness d_i, index of refraction $N_i(\lambda)$, absorption coefficient $K_i(\lambda)$, deposited on a substrate material with index N_s and absorption coefficient K_s, what is the reflection, absorption, transmission, or phase shift, etc., as a function of wavelength for light of a given polarization incident on the stack at an angle ϕ_0 from an (infinite) medium with index of refraction N_0? The thickness of a layer may range from micrometers to meters. The procedure used to obtain the solution to this problem is described in detail in many texts[3,4] and need not be reproduced here.

We shall, however, present the important steps in the derivation, so that the generalizations of our approach will be made clear. Although our present implementation of the calculation is limited to nonabsorbing (but dispersive) films at normal incidence, from the user's viewpoint further generalization offers no problem and the software changes required are small.

A matrix approach has been taken in the algorithm, not so much because it was intrinsically suited to CSMP, but because it is the usual method for film calculations using digital computers. In essence an optical impedance is introduced for each layer, which leads to a system of equations analogous to those describing the electrical properties of transmission lines. Thus, each layer is represented by a 2×2 matrix that relates the electric and magnetic fields, **E** and **H**, at the two boundaries of the layer. By multiplying the matrices in the proper order, the fields at the two ends of the stack may be related and the various properties of the stack may be determined.

The optical impedance is defined so that it varies continuously with the thickness of the stack, even through discontinuities at a surface. The usual definition is

$$Z = \frac{E_{\text{tangential}}}{H_{\text{tangential}}}. \tag{1}$$

The characteristic impedance of the medium is continuous through the stack because of the boundary conditions on the tangential field components E_t and H_t; it may be defined as the ratio of the tangential components of electric and magnetic fields for a positive-going wave in the medium and may have two different values, one for each orientation of the (optical) plane of polarization. If the angle of incidence of the light at the interface between the kth and $(k+1)$th layers is ϕ_k, then the characteristic admittance η is

$$\eta_k = N_k/\cos\phi_k \tag{2}$$

for polarization parallel to the plane of incidence and

$$\eta_k = -N_k \cos\phi_k \tag{3}$$

for polarization perpendicular to the plane of incidence. The (possibly complex) index of refraction of the layer is given by N_k.

It is a straightforward matter to show that the fields in two adjacent layers are related by the equations

$$E_{k+1} = E_k \cos(\psi_k d_k) + i(H_k/\eta_k)\sin(\psi_k d_k) \tag{4}$$

and

$$H_{k+1} = i\eta_k E_k \sin(\psi_k d_k) + H_k \cos(\psi_k d_k), \tag{5}$$

which are seen to be equivalent to the matrix equation

$$\begin{bmatrix} E_{k+1} \\ H_{k+1} \end{bmatrix} = \begin{bmatrix} \cos(\psi_k d_k) & (i/\eta_k)\sin(\psi_k d_k) \\ i\eta_k \sin(\psi_k d_k) & \cos(\psi_k d_k) \end{bmatrix} \begin{bmatrix} E_k \\ H_k \end{bmatrix}. \tag{6}$$

Here the quantity ψ_k is given by

$$\psi_k = 2\pi N_k/\lambda \tag{7}$$

and d_k is the thickness of the layer.

The calculation starts with the layer $k = 1$ adjacent to the incident medium, although this layer is physically the last one deposited. Multiplication of the matrices (the "characteristic matrices") of the different layers in the stack relates the waves in the incident medium to those in the substrate. This method is not limited to discrete layer problems; by virtue of the continuity of the optical admittance, we can analyze a medium in which the index of refraction is continuously varying along the direction of light propagation simply by approximating it by a series of discrete layers. Also, since we use in our program a matrix representation for a two-port element, any physical system representable by this form can be analyzed.

Graphic CSMP

The Continuous System Modeling Program[1] (CSMP) is a general purpose computational tool especially useful for the solution of differential equations or design of analog equipment. The system is described elsewhere in greater detail.[5,6]

The computations to be performed are described by the user as a block diagram consisting of functional blocks connected by lines representing control flow. The functional blocks implement mathematical functions, many of which are familiar to users of analog computers (e.g., gains, integrators, summers, multipliers, etc.), enabling digital simulation of the response of an analog computer. The inclusion of other standard function blocks, such as dividers, zero-order holds, pulse generators, etc., results in a far more powerful and elegant tool. In addition there are five special purpose blocks whose functions can be tailored to the user's needs.

The computational sequence is specified by the block diagram, which describes the input/output connections of the function blocks, up to a maximum of 75 blocks. Each can have up to three inputs (i.e., "fan-in" from other blocks), three internal parameters, and one output (with unlimited "fan-out" capability).

The program has an internal fixed-interval clock and, for each clock cycle, computes the output for each block in the model after an initial sort to assure the proper sequence of calculation. The user chooses the clock increment as well as the total clock time.

Output from the program is a continuous plot of two variables, assigned to the two axes, whose scale parameters are supplied by the user. For example, clock time can be plotted on the horizontal axis and any desired function or functions on the vertical axis. In fact, if no x coordinate input is provided, clock time is used by default. Although time is commonly used in a continuous system, any meaning can be given to the independent variable and the program thus becomes a general purpose calculator. A nonlinear function of time can also be plotted as the x coordinate.

A graphic version of CSMP has been implemented using an experimental IBM 2250 Graphic Console attached to an IBM 1130 computer. In this implementation the user develops his model as a drawing on the CRT in the form of a block diagram by copying from a standard repertoire of block types, attaching blocks together with the aid of the light pen and keyboard, and revising the basic model with appropriate erasures and additions. The user can change the shape or label of a standard block as a mnemonic aid in modeling, or eliminate unneeded blocks.

In the process of drawing a block diagram, the user is provided with a set of standard functional blocks at the bottom of the screen. He positions the light pen tracking pattern, points at the standard block he wants,

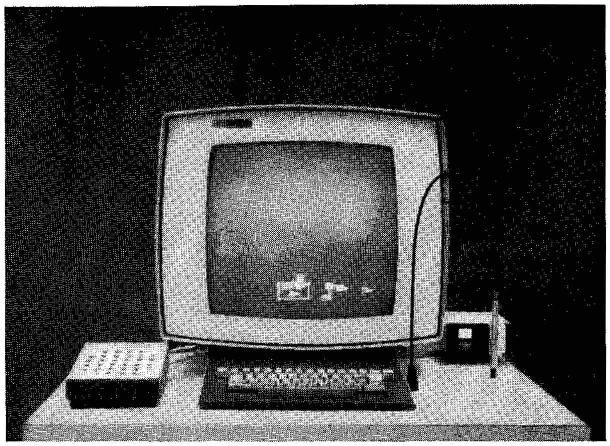


Figure 1 The 2250 Graphic Console with display of optical filter primitives.

and presses a key. A copy of the block appears at the light pen position. The tracking pattern position is updated to the output terminal of the block; if a new block is copied, the input terminal of the second block is connected to the output terminal of the first block. Connections can also be made by drawing lines between blocks with light pen and key controls.

For parameter entry the user points with the light pen at the block desired, indicates parameter entry by depressing a key, and enters the values of the parameters on the keyboard. The parameters of any block can be displayed by pointing at it.

Output from the program is provided as a plot on the 2250 screen. The output element is a functional block that simulates an oscilloscope and specifies x and y coordinate plotting values as separate input.

The system maintains disk storage for block diagrams, permitting the creation of a library. Hard copy output is available with an on-line plotter or by the use of a camera; the illustrations for this article were made with a Polaroid camera clamped to the 2250 table. An on-line listing of output values is also available, which makes it possible to obtain exact values over part or all of the range.

CSMP and optical design

An optical layer is modeled using a CSMP special function block that is defined by a user-written FORTRAN subroutine. The entire stack is represented as a series of "layer" blocks with the output of one connected to the first input of the succeeding block. The optical thickness and index of refraction of each layer are coded as parameters of its representative block. When nondispersive layers are used, the thickness of the layer entered by the designer is the "Quarter Wave Optical Thickness," or

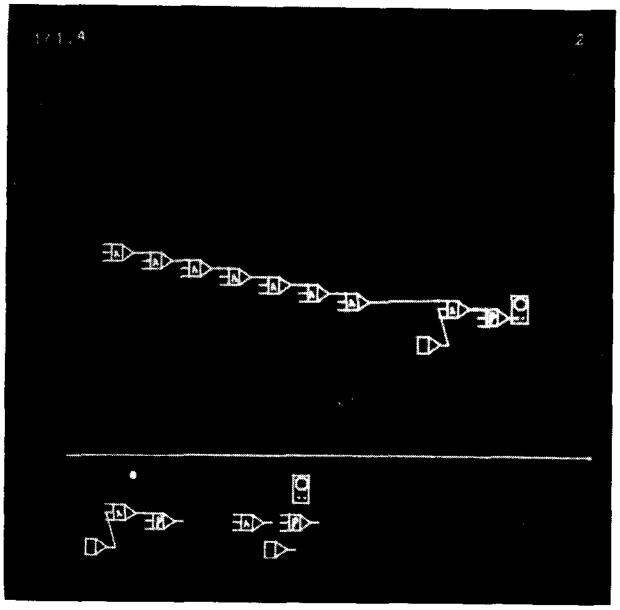


Figure 2 Block diagram of sample seven-layer stack.

QWOT, given by the product $4Nd$, where d is the physical thickness. This corresponds to a specification of a particular wavelength for which the filter would ordinarily be optimized in some way. When the layer is dispersive, the designer must enter the physical thickness d and the function $N(\lambda)$.

The basic function of a layer block is to calculate and store the cumulative characteristic matrix of the stack, as implied by Eqs. (6) and (7). The program ensures that blocks will be calculated in the proper sequence, beginning with the layer adjacent to the incident medium. When the substrate layer is reached, the characteristic matrix of the stack is available for further calculation. The desired property of the stack is then calculated and plotted on the CRT.

In the first implementation, reflectivity was the property chosen for analysis and wavenumber, the naturally occurring energy variable, was taken as the independent variable and was associated with clock time. This initial set of primitive blocks is illustrated in Fig. 1 as it appears on the 2250 console. Unused block types have been erased from the standard function set and the layer blocks have been labeled Λ for layer.

A set of three blocks, which can be copied as a single unit, has been added to the standard set to represent the substrate. Index of refraction and absorption can be specified as substrate parameters.

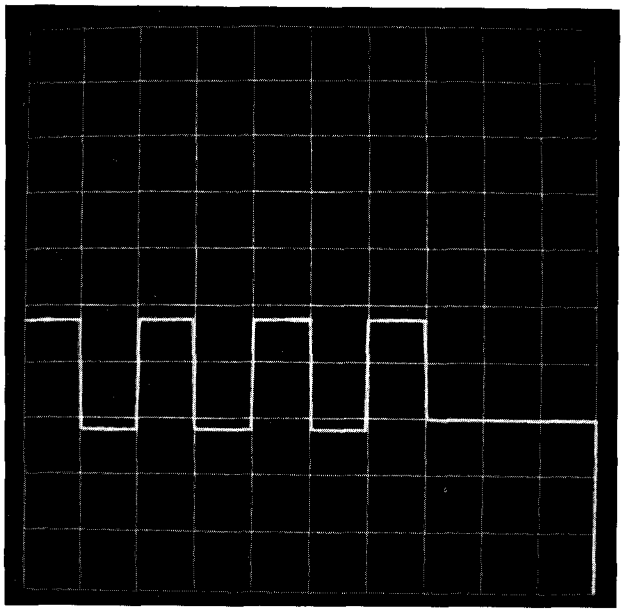
A sample stack model, as the designer would see it, is shown in Fig. 2. It consists of seven layers of equal thickness (QWOT = 1 micrometer) with alternating indices of refraction ($N_1 = N_3 = N_5 = N_7 = 2.38$; $N_2 = N_4 = N_6 = 1.4$). The incident medium is air ($N_0 = 1$, $K_0 = 0$) and the substrate is glass ($N_s = 1.5$, $K_s = 0$). The parameters are not seen unless the light pen is pointed at a layer; the numbers in the upper left hand corner of the screen are the QWOT and the N of the second layer.

To give the designer a graphic description of the stack for checking, the system provides a "profile" plot of the stack. The y axis represents the index of refraction, and the x axis, optical depth in the stack with the substrate (assumed infinitely thick) on the right hand side. The choice of profile plot (Fig. 3) or reflectivity (Fig. 4) is made by a parameter value of the function block P. The flexibility of light pen operations enhances the design process since the addition of layers anywhere in the stack is very simple and portions of the stack can be disconnected without erasure, thus permitting later reuse.

The basic program calculates reflectivity as a function of wavenumber; however, if a plot vs. wavelength is desired, the variable $1/\sigma$ can be plotted on the x axis, instead of clock time. In this case the wavelength increment $\Delta\lambda$ would not be linear since $\Delta\sigma$ is the fundamental computational increment. In addition, many *functions* of reflectivity can be calculated and plotted, simply by insertion of function elements (e.g., an integration block) between the P block and the oscilloscope.

The system also permits an optical thickness or index of refraction to vary continuously with wavenumber

Figure 3 Profile plot of the seven-layer stack. The ordinate represents the index of refraction (zero to 5 full scale); the abscissa is the optical depth in the stack (arbitrary units).



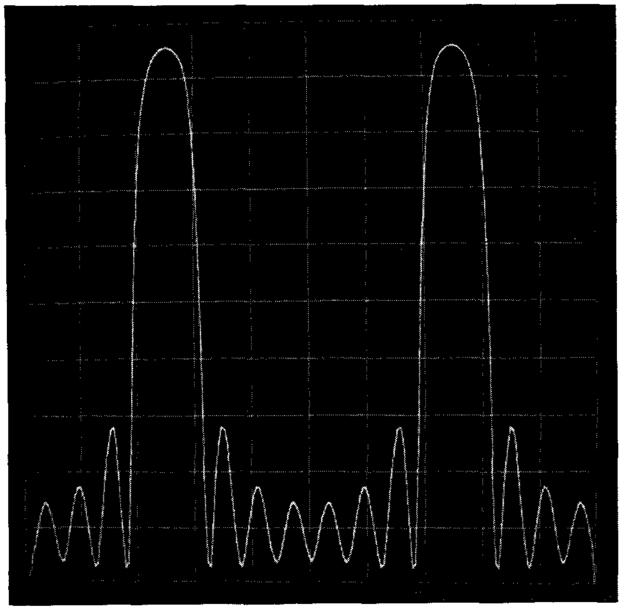


Figure 4 Reflectivity vs. wavenumber for the stack displayed in Fig. 2. The reflectivity (ordinate) scale is zero to 1; the wavenumber (abscissa) scale is zero to 40,000 cm^{-1}.

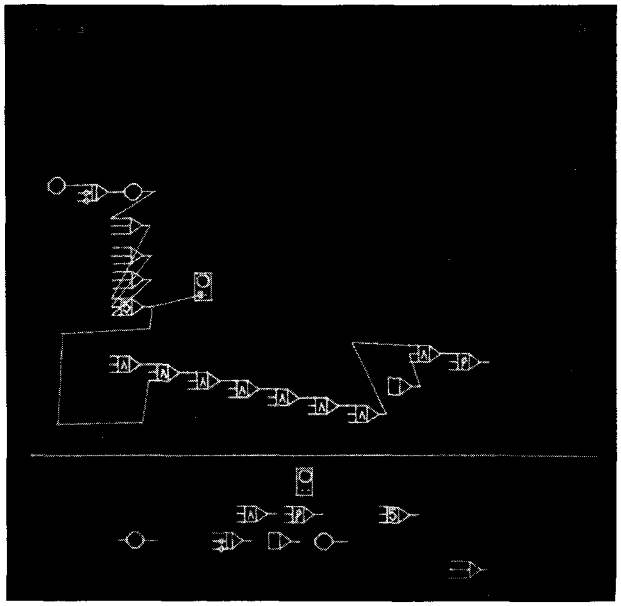


Figure 5 Block diagram for dispersion in the second layer.

Figure 6 Dispersion function. The ordinate represents the index of refraction (zero to 2 full scale); the abscissa is the wavenumber (zero to 40,000 cm^{-1} full scale).

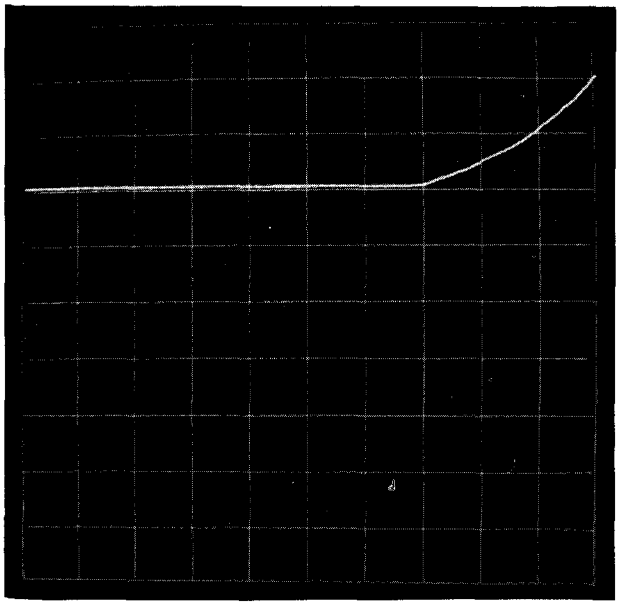
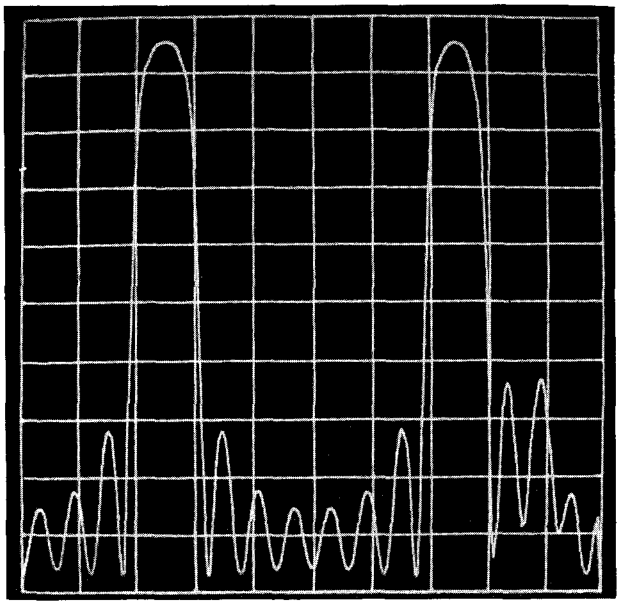


Figure 7 Reflectivity vs. wavenumber for the stack displayed in Fig. 5 (with dispersion in the second layer). The reflectivity (ordinate) scale is zero to 1; the wavenumber (abscissa) scale is zero to 40,000 cm^{-1}.



(i.e., to be dispersive) by the use of a function generator (Block 5). This function generator provides a representation of the dispersion curve with 11 function values (supplied by the designer as block parameters). Figure 5 shows the stack of Fig. 2 with the dispersion function generator applied to the dispersive layer (the second layer in the stack). The oscilloscope shown on the same figure has been positioned to generate a display of the dispersion function, which is shown in Fig. 6. The reflectivity of the stack is shown in Fig. 7.

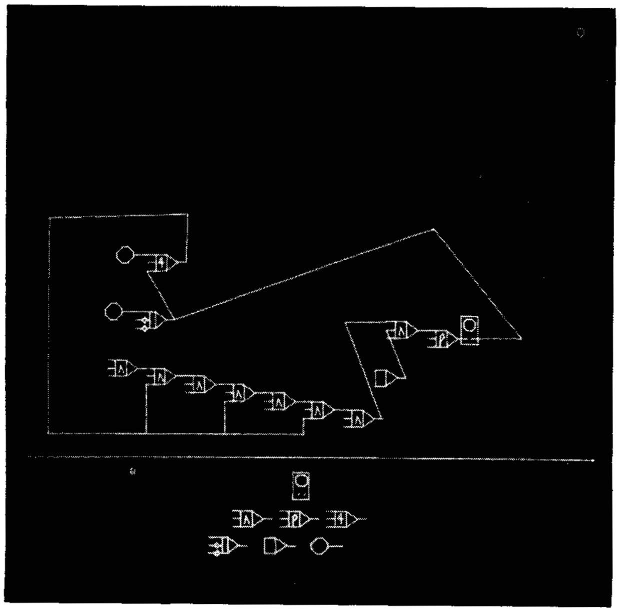


Figure 8 Block diagram of multiple-plot feature.

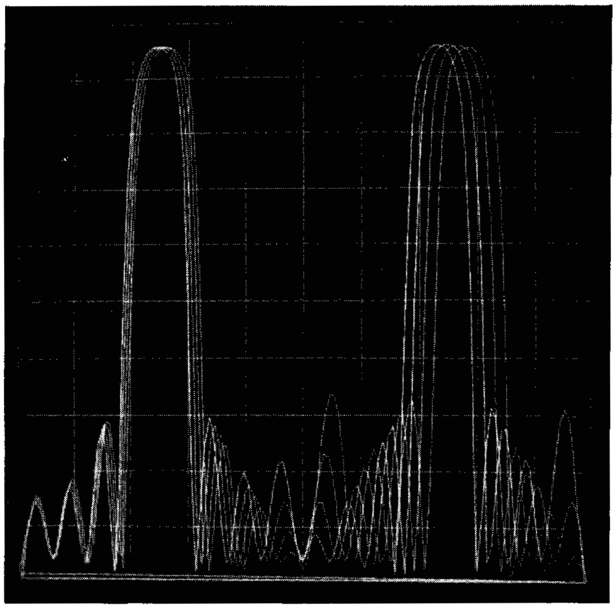


Figure 9 Multiple reflectivity curves. The reflectivity (ordinate) scale is zero to 1; the wavenumber (abscissa) scale is zero to 40,000 cm^{-1}.

Another feature permitting the plotting of multiple curves for monotonically incremented values of thickness or index of refraction is implemented as CSMP Block 4. This block provides the designer with a one-parameter family of curves characterizing a property of the stack. That is, the property (e.g., reflectivity) is calculated, one of the parameters is then incremented by a prescribed amount, and the calculation is repeated. The minimum, maximum, and incremental values are specified as block parameters. All curves are drawn as they are computed and a parameter (which may be common to several different layers, e.g., QWOT) is incremented, until all curves are displayed simultaneously. The sequence of curves drawn can be interrogated by the use of the light pen.

Figure 8 is a block diagram exhibiting the multiple-curve feature and Fig. 9 shows the reflectivity curves produced by varying the QWOT of the second layer from 0.9 to 1.1 micrometers in increments of 0.05 micrometer. The stack of Fig. 2 is used for the rest of the model so that the reflectivity curve of Fig. 4 is the same as the third curve drawn in Fig. 9.

The possibility of additional features is limited only by the needs of the designer, the ingenuity of the programmer, and (ultimately) the capacity of the computer. Considerable unused capacity is available in the implementation described (a 16 K IBM 1130) and response time is very rapid.

Further extensions

A system as flexible as the one described above lends itself to a number of ready extensions. There are two types of extensions under consideration: (a) adding computational functions by adding to the number of stack properties that can be analyzed by the program and (b) additional applications that can be handled by generalizing the approach described here.

Adding computational functions

It is clear from the preceding sections that the existing program does not exhaust even the most obvious possibilities for optical calculations. Some of the likely additions are:

Other stack properties It is often convenient to the designer, during the design process, to have available the *phase* as well as the amplitude of the reflected light or the *trace* of the characteristic matrix. These and other comparable options can be provided easily with additional coding as alternate output.

Non-normal incidence Many physical problems require that a filter be designed for non-normal incidence or even for several angles of incidence. This feature requires that the angle of incidence and the plane of polarization of the incident light be specified. The program must then compute the angle of refraction in the first and each succeeding layer (by means of Snell's law). The arguments of the trigonometric functions must be changed and two independent computations must be performed, one for each of the principal polarizations because their characteristic impedances differ.

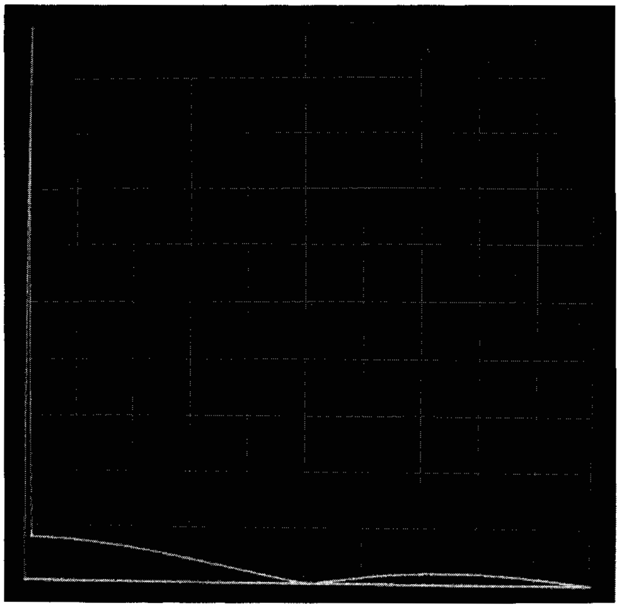


Figure 10 Fourier transform of the reflectivity curve displayed in Figure 4. The ordinate represents the amplitude of the transform (arbitrary units); the abscissa is clock time (arbitrary units).

Absorbing layers When the elements of the stack are absorbing, an absorption parameter K_i, possibly dispersive, must be specified for each layer by the filter designer. In addition, the characteristic admittance η_k, the phase factor ψ_k, and the angle of incidence ϕ_k (calculated from Snell's Law) become complex. The matrices then require complex matrix multiplication and the calculation of hyperbolic functions.

Birefringent layers For the optical filter designer, use of an anisotropic layer—one with an index of refraction that depends on both the polarization and the angle of incidence—is uncommon. Nevertheless there are cases in which anisotropy is important, e.g., the frustrated total internal reflection filter or the birefringent Fabry-Perot filter. In these problems the complication is slightly greater; two (or even three) principal indices of refraction must be specified for each birefringent layer and the eigenvectors and eigenvalues of the indicatrix must be calculated for the actual angle of incidence for each layer.

Fast Fourier transform subroutine[7] This has been adapted for the 1130 and implemented as a CSMP function block. Figure 10 represents the time domain transform of the reflectivity plotted in Fig. 4. This function could be applied, for example, to certain aspects of mode-locked lasers. (The transform of the reflectivity of the laser mirrors might be a fair approximation to the time dependence of the output.)

• *Additional applications*

The system we have described has considerably more generality than the design of new and better filters. For instance, the program already encompasses many laser cavity problems—a "layer" 5 cm long can be a model of a laser since both the dispersion and the gain (negative absorption) of the laser medium can be specified. One can then compute the resonant frequencies of a cavity with gain. By the addition of extra "layer" blocks to the model, the properties of coupled resonators can be obtained.

The program need not deal with optics at all. The mathematics is exactly that of transmission lines and CSMP (unlike optical films) could conceivably provide lumped circuit elements. Thus a variety of interesting filter design problems can be solved with our approach.

Acknowledgments

This project is due to initial stimuli from H. B. Baskin and R. V. Pole. E. Spiller and E. Lean have contributed important suggestions and we have benefited greatly from the programming assistance and encouragement of P. W. Baumeister of the University of Rochester who provided most of the special FORTRAN coding.

References

1. *1130 Continuous System Modeling Program,* H20-0209-1, IBM Data Processing Division, White Plains, N. Y.
2. See, for example, L. Brillouin, *Wave Propagation in Periodic Structures,* Dover Publications, New York, esp. pp. 193–227 (1953).
3. P. W. Baumeister, "Optical Design," *Military Standards Handbook,* MIL-HDBK-141, Chap. 20, available to contractors to the Department of Defense from the Naval Supply Depot.
4. O. Heavens, *Optical Properties of Thin Solid Films,* Butterworth Scientific Publications, London (1955).
5. R. D. Brennan and M. Y. Silberberg, "Two Continuous System Modeling Programs," *IBM Syst. J.* **6,** 242 (1967).
6. H. B. Baskin and S. P. Morse, "A Multi-Level Modeling Structure for Interactive Design," *IBM Syst. J.* **7,** 218 (1968).
7. H. Hellerman, private communication.

Received August 30, 1968