In parts 1 and 2 we surveyed the history of the understanding of optical interference and of the optical properties of materials, inseparable from the development of an understanding of the nature of light itself. The account ended with Drude's book on optics, published in German in 1900 and in English two years later [1], that set out the modern classical approach to optical phenomena. There was little on thin films in Drude's book, however. Thin film interference was not a subject of great importance at that time. Apart from some natural structures, the Lippmann emulsion was the main example of an artificial multilayer and there was no great drive for accurate calculations of its properties. Optical coatings still had to wait a few decades before accurate calculations became important. Developments in the early Twentieth Century were mainly directed towards metallic coatings. Then in the late 1930's, suddenly, optical systems needed optical coatings and industry responded.

**Early Optical Coatings**

Poisson, Fresnel and Airy had all recognized and developed the theory of multiple beam interference but there were few useful developments employing it until around the end of the Nineteenth Century when multiple beam interferometers became important in high-resolution spectroscopy. Particularly significant in the later development of optical coatings was the multiple beam interferometer conceived by Fabry and Perot [2], but its performance suffered because of its reliance on silver reflectors with their inherently high losses.

Chemically deposited silver was the common reflecting coating for mirrors, beam splitters, interferometer plates, virtually everything requiring enhanced reflectance, but quality was problematic. Freshly deposited silver exhibited a very poor outer surface that had to be polished by hand before it was suitable for optical applications. For partially transmitting coatings the polishing continued until the correct thickness was obtained, a difficult and uncertain process.

Then in 1912, Robert Wichard Pohl (1884-1976) and Peter Pringsheim (1881-1963) [3] required clean metal films for photoelectric investigations and devised a vacuum process using what they called vacuum distillation, but that we would now term thermal evaporation. They observed that with a substrate exhibiting a sufficiently high quality of surface finish, the coatings would immediately form a mirror of equal quality without any further need of polishing. At the time, little notice by the optical community seems to have been taken of this major advance and neither author appears to have realized that their short, one and a half page, paper would later be recognized as the beginning of an important industry. Both were associated with further substantial advances but in fields remote from vacuum coating.

It took a while for the new method to become established. There were scattered contributions in the literature, mostly from academics with little apparent response from industry. Then in the later 1920's and early 1930's the rate of development increased significantly. One of the pioneers at this time was John Donovan Strong (1905-1992). Strong left Kansas with an AB degree in 1926 for PhD studies at the University of Michigan. For his far infrared work he needed a mirror coating on one surface of a KBr prism, impossible with the wet chemical process. A lecture by Leonard Salomon Ornstein (1880-1941) visiting from the University of Utrecht, told him of the new process involving evaporation of silver under vacuum, and so began his pioneering work in the deposition of optical coatings. In 1930, with his PhD, he joined the California Institute of Technology on a fellowship, where he teamed with Charles Hawley Cartwright (1904-1964) who had just completed his PhD there. Together they investigated the deposition of an enormous number of metals and dielectrics [4]. By 1935 Strong had aluminized the 100 inch Hooker reflector of the Mount Wilson Observatory [5] and in 1936 he published the first paper on the use of an evaporated dielectric material as an antireflection coating [6]. His direct involvement in coatings ended with his aluminizing of the 200 inch Hale primary of the Mount Palomar Observatory in 1947. Meanwhile Cartwright had spent some time in Germany in the same Berlin laboratory as Arthur Francis Turner (1906-1996) who was working towards his doctorate with Marianus Czerny and Peter Pringsheim. Turner returned to the US and a teaching position at the Massachusetts Institute of Technology in 1935 and was joined by Cartwright in 1937, when the pair began their ground-breaking research in interference coatings including antireflection coatings and the reflector we know as the quarterwave stack [7].

Industry was at last becoming active in optical coatings and this was happening in virtually every developed country. In Germany, Alexander Smakula (1900-1983) at the Carl Zeiss company in Jena had developed, at the same time as Strong, an antireflection coating for lenses [8], but this, for several years, remained a close secret and was used primarily for military applications. Walter Heinrich Gefcken (1904-1995) at the sister company Schott in Jena, around the same time, devised a wide range of coatings including the first narrowband interference filters [9]. Like Smakula's, most of Gefcken's early advances were also kept secret [10]. Then Turner left the Massachusetts Institute of Technology and joined Bausch and Lomb in 1939 where he ran the Optical Physics Department until his retirement in 1971. By 1940, James Weir French, the chairman of Barr and Stroud in Glasgow, was able to state in a letter to Nature [11] “Today, cryolite treatment of surfaces has become a routine operation.” In fact, we know [12] that by this time Barr and Stroud were making use of the tougher magnesium fluoride rather than cryolite.

**Early Optical Thin Film Calculations**

Thus, by the end of the 1930's optical coatings represented a viable and growing industry segment. With this development came an increasing need for design and performance calculation techniques.
At the beginning of the Twentieth Century, multiple-beam interference in single layers was already well understood, but, as Drude had pointed out in his book [13], there were deficiencies in multilayer calculations. The properties of greatest interest were, and still are, those that we describe as specular. Incoherent and coherent scattering are much more complex and difficult phenomena.

Stratified media were also important in fields other than optics and there was a general lack of accurate theoretical models. In 1912, John William Strutt, Third Baron Rayleigh (1842-1919), decided to remedy the situation. Rayleigh had studied at Trinity and had followed Maxwell as the second Cavendish Professor of Physics. He was accomplished in virtually every field of physics existing at that time and had made considerable advances in optics, but to make his paper [14] generally applicable to a wider range of phenomena including optics, he biased it towards acoustics, clearly considered as the more important topic. “We may therefore for the present confine ourselves to the acoustical form, knowing that the results will admit of interpretation in numerous other cases,” and then later, “these results are at once available for the corresponding optical problems.” From the optical point of view the system was a series of lossless dielectric layers of arbitrary thickness. Rayleigh started with the partial waves in each layer and introduced general boundary conditions that with a little manipulation can be recognized as applying also to the optical case. His analysis was rather cumbersome, and, beginning at the rear surface, he quickly ran into a problem immediately familiar to any optical coating designer. “The procedure would entail no difficulty in any special case numerically given, but the algebraic expression...soon becomes complicated, unless further simplifying conditions are introduced.” Rayleigh was seeking a general analytical expression rather than an algorithm for calculation. Much of his paper is an attempt to introduce suitable simplifying assumptions. In a subsequent paper [15] he simplified matters somewhat by concentrating on optical effects and introducing the idea of a regularly stratified medium. In this we recognize the quarterwave stack [16] and some ideas of symmetrical periods, but his study was still devoid of any idea of calculation design. A primary objective of this second paper was to explain some natural phenomena such as “the remarkable coloured reflection from certain crystals of chlorate of potash described by Stokes, the colours of old decomposed glass, and probably those of some beetles and butterflies.”

The optical coating practitioners, who came a little later, mostly ignored what others had done and worked out a calculation model for themselves. Pierre Rouard (1908-1989) was typical with his method included in his doctoral thesis [17]. Rouard replaced the final film by a single equivalent surface with calculated amplitude reflection coefficient in polar form. This became the rear surface of the adjacent layer that was then treated in the same way, and similarly with subsequent layers. Cartwright and Turner at the Massachusetts Institute of Technology persuaded Arthur Hardy to ask one of his students, Monarch L. Cutler [18] to work out a method as a bachelor’s thesis, which he successfully accomplished and actually used in the design of color filters. Although he was unaware of Rouard's technique, his was not very different in that he replaced structures on either side of a film with simple interfaces exhibiting effective properties. In his preface Cutler acknowledges the help of Richard P. Feynman in devising the technique. Smakula and Gefcken, too, had calculation techniques, Gefcken's much influenced by Rayleigh's papers. Doubtless there were many other attempts at models. We have to remember that calculators were rudimentary, and much manual calculation was involved, often with the help of tables of logarithms.

Meanwhile, a completely different thread in the development of calculation techniques was forming. After obtaining his BS degree from Tufts in 1928, Philip Hagar Smith (1905-1987) joined Bell Telephone Laboratories where there was considerable interest in short wave transmission. The mathematics of transmission lines had been worked out by that time but calculation was, as usual, tedious in the extreme. Smith, a keen proponent of graphical methods, quickly, in 1931, devised a rectangular chart where standing wave ratio and wave position were linked to impedance and where lengths of transmission line were represented as circular paths. In appearance this chart was not unlike our admittance diagram for optical thin films and it shares the property that to encompass the entire possible range of parameters it must stretch to infinity. Our admittance diagram is primarily used for understanding rather than calculation and so this is not a serious defect, but Smith was looking for a calculation tool and by 1936 he had modified his chart so that all values were enclosed within a circular boundary. In his 1939 form, further improved in 1944 [19] we finally recognize the familiar form of the Smith Chart [20].

Although, like Smith, not recognized as a thin film pioneer, Sergei Alexander Schelkunoff (1897-1992) also at Bell Telephone Laboratories made what can now be recognized as a contribution of great significance. Schelkunoff was born in Russia, and was educated at the University of Moscow before being called to army service in 1917 while Russia was involved in the Great War. 1917 also marked the beginning of the revolution in Russia, a time of great disturbance with the army in complete disarray. Schelkunoff eventually succeeded in making his way east to Seattle where he gained degrees in mathematics followed by a PhD from Columbia. He then joined the research group of Western Electric that later became Bell Labs where he was internationally known for his work on virtually all aspects of electromagnetic wave propagation and where he remained until he retired. In 1938 he published a paper on the concept of impedance and its application to many different fields [21]. In part III of the paper he discussed its application to simple optical surfaces and their reflectance and transmittance both at normal and at oblique incidence.

Suddenly, in 1939, much of the world was at war. By 1941 almost all advanced nations were involved on one side or the other. Coating production, especially antireflection coatings, was enormously expanded everywhere, but publication was greatly inhibited, although not stopped completely. With the end of the war in 1945 came a flood of publications.

**Advances in Calculation Techniques**

Accurate calculation of the properties of coatings was now of higher priority. Robert Mooney wrote two papers [22, 23]. Antonin Vasicek, the leading thin film worker in Czechoslovakia, published several theoretical studies [24-26]. Doris Caballero [26] and Walter Welford [28] also contributed. Most of this work used iterative techniques that although differing in detail did not add significantly to the earlier studies, but Welford did succeed in putting his method into matrix form. Looking at most of these with hindsight, continued on page 30
Our Matrix Model

By 1950 many different accurate ways of calculating the optical properties of thin film optical multilayer coatings existed, all giving with the same input the same answer, and we still see some of them occasionally employed in the current literature. However the most common, most used, and most useful, is undoubtedly what has come to be called the matrix method.

A complete discussion of the matrix method is beyond the capacity of this article and can be found elsewhere [35]. Here we will limit ourselves to normal incidence and we will use the current symbols and terminology.

We deal with a sequence of thin films consisting of homogeneous and isotropic materials separated by a set of parallel, flat, featureless interfaces consisting of abrupt discontinuities in the optical properties. A linearly polarized plane harmonic electromagnetic wave is incident normally. The coating creates a transmitted wave along the same direction as the incident wave and a reflected wave along the opposite direction. Both reflected and transmitted waves are linearly polarized in the same plane as the incident wave. The phase factor of the wave is, by convention, of the form \( \exp[i(\omega t - kz)] \).

Materials are characterized by their complex refractive index \((n - ik)\) and characteristic admittance \(y\), the latter being the ratio of magnetic to electric field amplitudes of a propagating harmonic wave and at optical frequencies, because of the absence of direct magnetic effects, given by \((n - ik)\mathcal{G}\) where \(\mathcal{G}\) is the admittance of free space. Interfaces are characterized by their total tangential fields, \(E\) and \(H\), and the ratio \(H/E\) yields the surface admittance, \(Y\).

We set up a Cartesian coordinate system with the \(z\)-axis normal to the thin-film interfaces and positive into it. We take the positive direction of the electric field for all three waves along the \(x\)-axis and the origin of the coordinates at the front surface (the surface of incidence).

The matrix model starts with the total tangential electric and magnetic fields at the emergent interface, which are identical to the field amplitudes of the emergent wave, and transforms them into those at each successive interface, eventually reaching the incident interface. It is then a simple calculation to extract the more measurable properties of reflectance, transmittance, phase changes on reflection and transmission and similar quantities. The transformation from one interface to the next is given by the matrix multiplication:

where film \(f\) is bounded by interface \(a\) in front and interface \(b\)

\[
\begin{bmatrix}
F_a \\
H_a \\
\end{bmatrix} = 
\begin{bmatrix}
\cos \delta_f & i \sin \delta_f \\
\frac{y_f}{y_f} & \frac{y_f}{y_f} \\
\end{bmatrix}
\begin{bmatrix}
F_b \\
H_b \\
\end{bmatrix}
\]

(1)

behind and has phase thickness \(\delta\) given by \(2\pi(n - ik)d/\lambda\), \(\lambda\) being the free space wavelength and \(d\) the physical thickness. The matrix is known as the characteristic matrix of the appropriate layer. It consists only of properties \(\delta\) and \(y\) of the layer, so that each layer is completely represented by a characteristic matrix with only its own properties. The treatment can readily be extended to a multilayer of an arbitrarily large number of layers:

\[
\begin{bmatrix}
B \\
C \\
\end{bmatrix} = 
\prod_{j=1}^{n}
\begin{bmatrix}
\cos \delta_j & i \frac{\sin \delta_j}{y_j} \\
\frac{y_j}{y_j} & \frac{y_j}{y_j} \\
\end{bmatrix}
\begin{bmatrix}
1 \\
\frac{y_{\text{emergent}}}{y_{\text{emergent}}} \\
\end{bmatrix}
\]

(2)

where, as usual, since we are dealing with linear systems we have slightly simplified the expressions by normalizing the fields so that the tangential component of electric field at the rear interface is unity. The resulting electric and magnetic field amplitudes at the surface of incidence are indicated by \(B\) and \(C\) respectively. Since there is no counter-propagating wave in the emergent medium, the surface admittance of the emergent surface is \(y_{\text{emergent}}\) the characteristic admittance of the emergent medium. Reflectance and transmittance are then given by:

\[
R = \left(\frac{y_0B - C}{y_0B + C}\right)\left(\frac{y_0B - C}{y_0B + C}\right) \\
T = \frac{4y_0\Re y_{\text{emergent}}}{(y_0B + C)(y_0B + C)}
\]

(3)
where $y$, the characteristic admittance of the incident medium, must be real [36].

**Conclusion**

Why is this model preferred over all the others in existence? They all, with the same input parameters have similar volume of calculation and give the same answer, and so it is not a matter of accuracy or validity or speed. A primary reason is that the matrix method focuses on the layers, whereas most of the other techniques emphasize the interfaces. With the matrix technique we can remove, add or modify layers without impacting the rest of the coating. We can rapidly assess the effect of repeat structures. It allows us to think more clearly about the effect of changes in design. Properties like inhomogeneity can readily be modeled by replacing a homogeneous layer by a succession of very thin discrete layers - thin enough so that a decrease in their thickness makes no perceptible change in the results. These are just some of the advantages.

It is no accident that it has been our chosen model for more than 60 years.

**Acknowledgement**

I owe my grateful thanks to Don Mattox for providing me with a copy of Cutler's thesis (reference [18]).

**References**