Publications of the Astronomical Society of the Pacific, **113**:1300–1305, 2001 October

Spectral Modulation, or Ripple, in Retardation Plates for Linear and Circular Polarization

D. K. Aitken[1,2] and J. H. Hough[1]

Received 2000 December 5; accepted 2001 June 14

ABSTRACT. High-resolution spectropolarimetry sometimes suffers from a spectral modulation in polarization and position angle. It is shown that this artifact can be attributed to multiple reflections within the plates of the wave plate assembly and that the effect is inherent to the use of wave plates except at the design wavelength. The treatment given applies to strictly plane-parallel wave plate elements and complete coherence, and consequently the predicted ripple amplitudes are upper limits. It is possible that plates of poorer optical quality will not show this effect so prominently, but otherwise a simple modification to normal observing practice can be used to eliminate the ripple.

1. INTRODUCTION

A ripple in both the linear polarization fraction and position angle is sometimes observed in high-resolution spectropolarimetry (Adamson & Whittet 1995; Chrysostomou et al. 1996; Adamson et al. 2000), and it seems this effect is becoming particularly pronounced in the near-infrared as observations push to increasingly high spectral resolution. Adamson & Whittet (1995) observed such ripples in the optical (0.56–0.59 μm) and attributed the effect to the cement layer in a Savart plate analyzer. Chrysostomou et al. (1996) observed between 4.5 and 4.8 μm and Adamson et al. (1999) between 3.1 and 3.8 μm, attributing the effect to reflections between the wave plate components.

These artifacts have been largely removed from data by identifying the separate Fourier components of the ripple in the Q and U spectra and interpolating through them (e.g., Adamson & Whittet 1995). Such a procedure can be performed only on data of adequately high signal-to-noise ratio (S/N), has some subjective element, and will involve some loss of spectral information. Identification of the source of the ripple at the least gives vindication to the otherwise ad hoc empirical treatment but also, as shown below, leads to procedures that reduce or eliminate it at the point of observation.

The effect is almost certainly associated with internal multiple reflections between the parallel faces of the retardation plates, acting like a Fabry-Perot interferometer. Interference between the multiply reflected beams produces intensity modulation along the fast and slow axes of the plates, and the effect on measured polarization depends on the phase difference between these beams. For an exact first-order half-wave plate, the phase difference between the simply transmitted e and o rays is just π. Subsequent multiply reflected rays will all be delayed by 2π, and the resultant intensity modulation of all rays will be in phase with their respective simply transmitted rays. Although there will be intensity modulation, the ratio e/o will be unchanged, and the polarization fraction will be unaffected by multiple reflections in the plate.

In addition to intensity modulation of the e and o rays, there is also a modulation of retardance with the same ripple frequency. This will not produce polarization in an unpolarized source but will affect the efficiency of detection of linear polarization.

For a quarter-wave plate, however, the simply transmitted e and o rays differ by $\pi/2$ and the multiply reflected ones differ by π; here the intensity modulations of the e and o rays are now exactly out of phase, and their ratio and therefore the deduced polarization show a modulation or ripple.

Such effects have been observed on occasion with quarter-wave mica plates (e.g., Smith 1969; Weinberger & Harris 1964); a treatment of this problem is given by Bennett & Bennett (1978). Yolken, Waxler, & Kruger (1967) find that their intensity modulations are close to those given by the Bennett & Bennett treatment in the *Handbook of Optics*.

From this it might seem that multiple reflection does not pose a problem for linear polarization studies using first-order half-wave plates, apart from small effects where the retardation departs from an exact half-wave at other than the design wavelength. In the near- and mid-infrared, however, material properties often require that retarders consist of two (or more) crossed plates in which the thickness of each is greater than that required to produce a half-wave retardation but with a difference in retardation between them of one half-wave; these are often referred to as zero-order plates. Often the two plates are not in optical contact and are separated by a distance much greater than their individual thicknesses. In the absence of interference effects, there is no a priori constraint on the indi-

[1] University of Hertfordshire, Division of Physical Sciences, College Lane, Hatfield, Hertfordshire AL10 9AB, UK; dka@star.herts.asc.uk, jhh@star.herts.ac.uk.

[2] Mailing address: Garbhein House, Garbhein Road, Kinlochleven, PH50 4SE, UK.

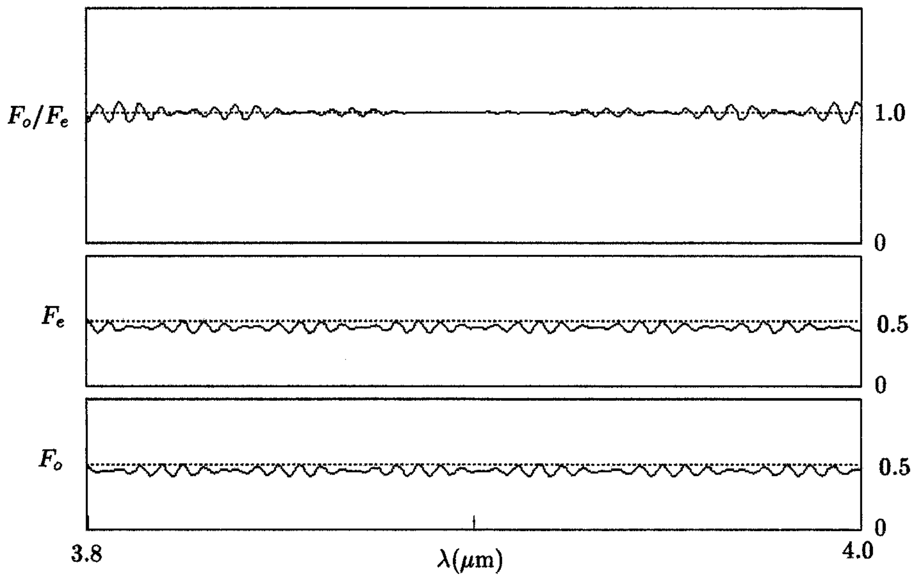

FIG. 1.—Ripple amplitudes resulting from multiple reflections in two crossed retarders in series. F_o and F_e show the output intensities from the two axes and F_o/F_e their ratio. The plates have retardance $5\lambda/2$ and $6\lambda/2$ at 3.9 μm.

vidual plate thicknesses other than that above to define the resultant retardance, and it may not always be clear what a manufacturer has supplied.

2. TWO PLATE RETARDERS

2.1. Outline

Given two plates that are physically separated and that differ in retardance by π, then the effects of interference in the ensemble need to be considered. The gap between the plates is not birefringent, so multiple reflections between the plates will affect all states of polarization equally, at least to the approximation that $n_e - n_o$ is small compared to n_e or n_o, and so will affect only the intensity measurement. However, the individual plates themselves each produce modulated *e* and *o* rays with similar modulation amplitude but different frequency, and, unless the plates are an exact multiple of half-wave, the modulations of the resultant rays will be shifted in phase, and modulation of the *e*/*o* ratio will occur and show a beat pattern from the two different modulation frequencies, as has been observed in a number of cases (e.g., Adamson et al. 1999). For a typical two-plate MgF_2 half-wave plate, the conversion of intensity to polarization can be several percent or more with ripple period ~0.005 μm.

In addition to intensity modulation of the *e* and *o* rays, there is also a modulation of retardance at the same frequency. As mentioned above, this does not produce polarization but affects only the efficiency of detection, and for MgF_2 is small, with amplitude 1°–2° and polarization efficiency of greater than 0.995.

Even if the two plates do satisfy the $m\lambda/2$ criteria, the ripple will vanish only at the design wavelength, and, while the wavelength dependence of retardance of the combination will not differ from a first-order $\lambda/2$ plate, the retardance of the individual plates will be *m* times more wavelength dependent and the ripple-free zone correspondingly reduced.

For internal multiple reflections to interfere, the plates must have faces that are plane parallel to better than a wavelength; a possible way to reduce the ripple might be to use components of less than optical quality so long as image quality and wander do not become a problem.

The effects are greatly reduced if the plates are cemented or in optical contact. Here although there is a high-frequency ripple that results from the separation of the two dielectric-air interfaces it is of lower amplitude than from separated plates, does not show a beat pattern, and vanishes at the design wavelength. In practice, compound retarders, usually achromats, show much smaller amplitude ripple than two separated plates.

Often it is not possible to cement the plates or to provide an antireflection coating, and, as explained above, making the individual plates satisfy $m\lambda/2$ eliminates the ripple only over a small wavelength range. Fortunately, the problem can be effectively eliminated in practice by noting that if the wave plate combination is rotated through 90°, the ripple pattern inverts; an extension of normal observing procedure to repeat all observations with a 90° shift and average these data will effectively eliminate the problem without increase of observing time for a given S/N.

2.2. Details

Holmes (1964) has pointed out that the usual expressions describing the performance of retardation plates ignore the ef-

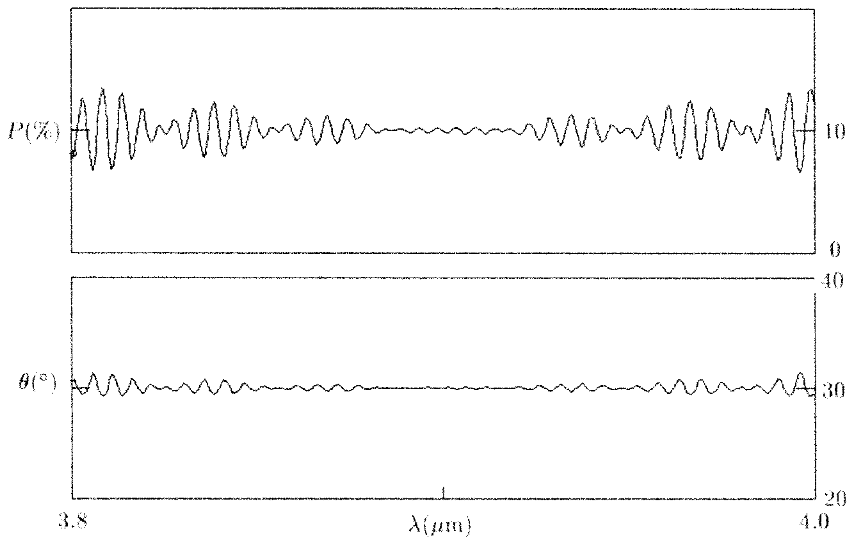

FIG. 2.—Simulated polarization measurement using the crossed plates in Fig. 1 (see text). The input polarization is 10% at a position angle of 30°.

fects of multiple reflections within the plates; he derives an exact theory of retardation plates that reveals some interesting differences from first-order theory. In a variant on standard treatments, Holmes treats the plates as a boundary value problem in electromagnetic theory.

Holmes gives the transmitted intensities of the *o* and *e* rays from a single retarder as

$$I_{e1} = \frac{I_{ei}}{1 + (K_e^2 - 1)\sin^2 \beta_e d},$$

$$I_{o1} = \frac{I_{oi}}{1 + (K_o^2 - 1)\sin^2 \beta_o d},$$

where $K_e = (n_e^2 + 1)/2n_e$ and $\beta_e = 2\pi n_e/\lambda$ and similarly for the *o* beam; λ is the vacuum wavelength, I_{ei} and I_{oi} are the components of input intensities polarized along the *e* and *o* directions, respectively, and *d* is the plate thickness.

The *e* and *o* intensities are modulated in wavelength at two close frequencies $2\beta_o d$ and $2\beta_e d$, and their ratio shows a high-frequency ripple $(\beta_o + \beta_e)d$ with a slowly varying factor of frequency $(\beta_o - \beta_e)d$; when $(\beta_o - \beta_e)d = \pi$ (the half-wave plate condition), the amplitude of the slowly varying factor is zero, and at the design wavelength there is no modulation of the *o*/*e* ratio.

The phase difference between the transmitted *e* and *o* rays is

$$\delta = \arctan \frac{K_o \tan \beta_o d - K_e \tan \beta_e d}{1 + K_o K_e \tan \beta_o d \tan \beta_e d}.$$

The value of δ oscillates about the usually quoted value of $d(\beta_o - \beta_e)$ when *K* is not equal to unity. The oscillation frequency is the same as the intensity modulation, but for a half-wave plate of MgF_2 in the near-infrared its amplitude is less than 2° and has an insignificant effect on the measurement of linear polarization.

For two plates the transmitted intensities are modulated by

$$F_e = \frac{1}{[1 + (K_e^2 - 1)\sin^2 \beta_e d_1][1 + (K_o^2 - 1)\sin^2 \beta_o d_2]}, \quad (1)$$

$$F_o = \frac{1}{[1 + (K_o^2 - 1)\sin^2 \beta_o d_1][1 + (K_e^2 - 1)\sin^2 \beta_e d_2]} \quad (2)$$

defining F_o and F_e. These intensities and their ratios are shown in Figure 1 for MgF_2 and two plates with retardation difference equal to a half-wave at 3.9 μm; here the shorter plate has a 5/2 wave retardance. While there is no modulation of the ratio at the design wavelength, the ripple amplitude increases significantly for small deviations from this. As a diagnostic, one may expect the low-frequency beats to contain *n* high-frequency components if the retardance is $n\lambda/2$.

In typical observing procedures, either the ratios of the transmitted intensities at wave plate positions of 0° and $(I_o/I_e)_0/(I_o/I_e)_{45}$ or the difference between observations at 0° and 45° is used to define one of the Stokes parameters. If the analyzer axes are parallel to the wave plate axes, then the intensities presented to the analyzers will be as in Figure 1,

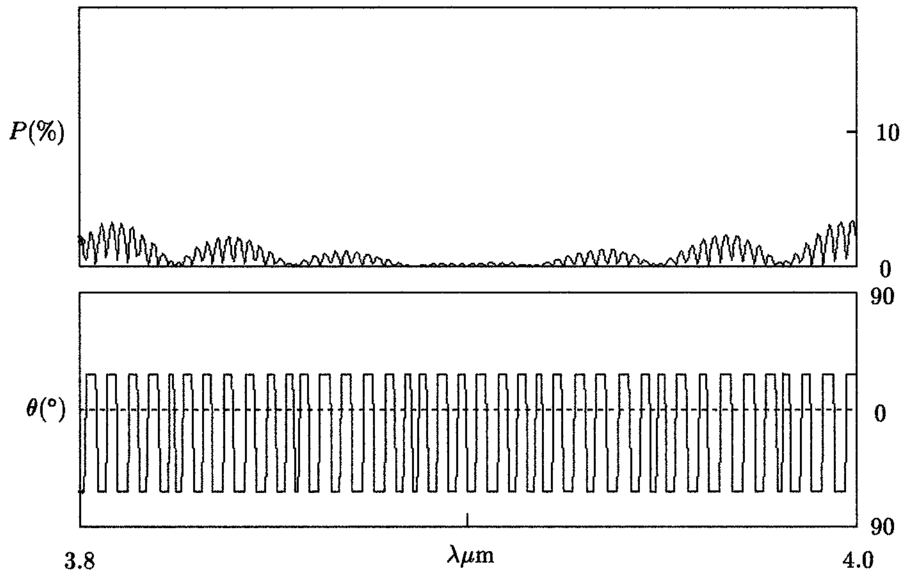

FIG. 3.—Same as Fig. 2 but for an input polarization of zero.

but at other angles the analyzer transmissions will be linear combinations of equations (1) and (2). If the angle between the wave plate and analyzer axes is ϕ, the intensities transmitted by the analyzer(s), denoted by I_a (and I_b), will be

$$2I_a = F_e(1 + \cos 2\phi) + F_o(1 - \cos 2\phi), \quad (3)$$

$$2I_b = F_o(1 + \cos 2\phi) + F_e(1 - \cos 2\phi). \quad (4)$$

At $\phi = 45°$ there will be equal contributions to both a and b analyzers, and their ratio (or difference) will show no ripple. Nevertheless, in any (minimal, e.g., ϕ confined to a $\pi/2$ range) sequence needed to define the Stokes parameters, some effect of the ripple will appear in the polarization at wavelengths that differ from the design wavelength. A standard procedure is to determine the ratios (e.g., Adamson et al. 1999; Tinbergen 1996)

$$R_q^2 = \frac{(I_a/I_b)_0}{(I_a/I_b)_{45}},$$

$$R_u^2 = \frac{(I_a/I_b)_{22.5}}{(I_a/I_b)_{67.5}}$$

and from them find

$$q = (R_q - 1)/(R_q + 1),$$

$$u = (R_u - 1)/(R_u + 1)$$

and hence p and θ in the usual way. Using equations (3) and (4) for $\phi = 0°$ and 45°, we find that the ratio R_q^2 is modulated by the factor F_e/F_o.

If the input polarization is zero or small, R_q^2 will oscillate about unity and q about zero, so the polarization ripple frequency will be effectively doubled.

Multiplying the analyzer inputs (eqs. [3] and [4]) by the input intensities

$$I_a = \frac{1 + p\cos(2\theta - 4\phi)}{2},$$

$$I_b = \frac{1 - p\cos(2\theta - 4\phi)}{2},$$

where the interference effects arising from the difference in intensity in the o and e rays are neglected (good to better than $p/10$), we then find, from $\phi = 0$ and 45°,

$$R_Q^2 = \left(\frac{1 + p\cos 2\theta}{1 - p\cos 2\theta}\right)^2 \frac{F_e}{F_o},$$

which is the expected ratio modulated by the factor F_e/F_o.

Figure 2 shows the effect of continuing this process at $\phi =$ 22°.5 and 67°.5 for two crossed plates of MgF_2 of 5 and 6 half-wave retardance at 3.9 μm. The input polarization is 10% at position angle 30°. The plates are assumed to be sufficiently separated that interference effects from reflections between the separate plates can be neglected; those between adjacent faces cannot in any case affect the resultant polarization since the gap between the plates is not birefringent. It is perhaps surprising

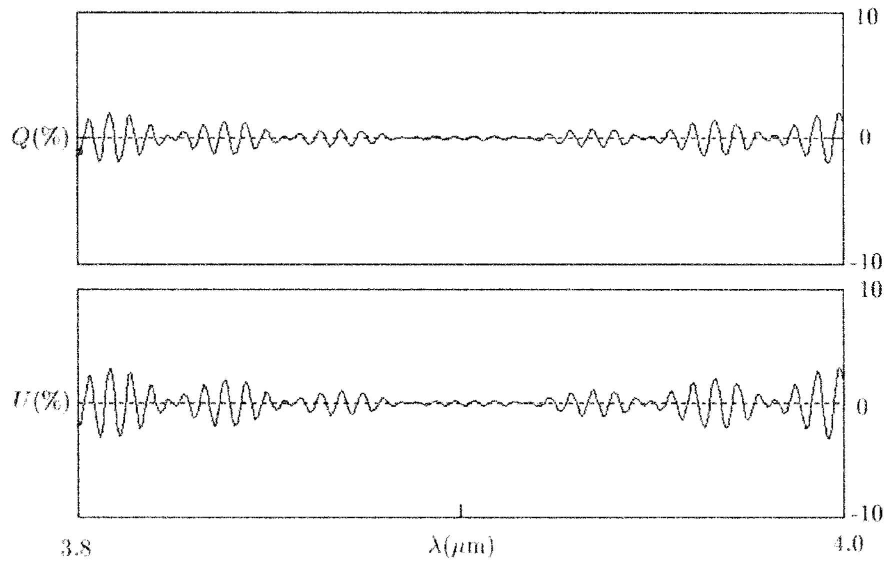

FIG. 4.—Spectra of Q and U corresponding to Fig. 3.

that MgF_2, with a relatively small refractive index, can give such large modulation, which approaches $\pm 10\%$ if the separate plates have retardance of an odd number of quarter waves.

If the input polarization is smaller than the wave plate modulation, then the Stokes parameters show a ripple excursion through zero, causing position angle flips through 90° and introducing an effective doubling of the polarization ripple frequency. Figure 3 shows this effect for an input polarization of zero and the individual contributions from Q and U in Figure 4. In most observational situations, this would be interpreted as poor S/N data on account of the chaotic behavior of the position angle.

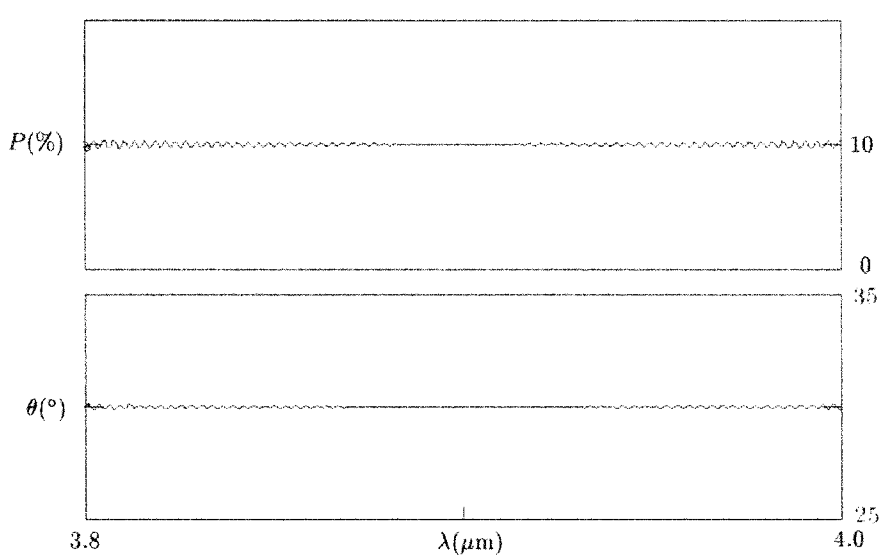

FIG. 5.—Same as Fig. 2 except that the two plates are in optical contact.

Figure 5 shows the results obtained for the same configuration if the two plates are cemented; the input polarization is 10% as in Figure 2. The amplitude is now much smaller, and, while it vanishes at the design wavelength, the beats are absent. The faster ripple frequency is the result of the combined length of the two plates in this case.

If the four sets of observations used to define Q and U are repeated from 90° to 157°.5 and the results from observations at ϕ and $\phi + 90°$ averaged, then, from equations (3) and (4), we see that

$$2I_{a(0+90)} = (1 + p\cos 2\theta)(F_o + F_e),$$

$$2I_{b(0+90)} = (1 - p\cos 2\theta)(F_o + F_e), \tag{5}$$

$$2I_{a(45+135)} = (1 - p\cos 2\theta)(F_o + F_e),$$

$$2I_{b(45+135)} = (1 + p\cos 2\theta)(F_o + F_e). \tag{6}$$

The factor $(F_o + F_e)$ is common, so

$$R_q^2 = \left(\frac{1 + p\cos 2\theta}{1 - p\cos 2\theta}\right)^2,$$

and the ripple contribution from the wave plate has cancelled.

For a single analyzer, it is usual to take the difference in observations at 0° and 45° to define Q (intensity) and their sum to define total intensity. Again, if we average readings separated by 90°, we find from equations (5) and (6) that

$$Q = I_{a(0+90)} - I_{a(45+135)} = 2p\cos 2\theta(F_e + F_o),$$

$$I_{\text{tot}} = I_{a(0+90)} + I_{a(45+135)} = 2(F_e + F_o),$$

so the fractional Stokes q is again free from ripple.

Some spectropolarimetric procedures sample at other intervals of ϕ than $\pi/8$ and deduce the Stokes parameters by fitting sinusoids of period $\pi/2$ to the data. The ripple problem can be similarly avoided by averaging data points separated by $\pi/2$ in ϕ.

Although the ripple amplitudes are more severe for a $\lambda/4$ retarder, a typical observing procedure, if the retarder axes are aligned with the analyzer axes, is to sample at 0° and 90°. Taking the ratio would yield Stokes V with the ripple cancelled. More generally, additional observations at 45° and 135°, to compensate for unknown axis alignment, will similarly cancel wave plate ripple.

The treatment given here assumes that the individual plates have faces plane and parallel to $\lambda/4$. If this criterion is exceeded, the ripple pattern becomes more complex depending on the detailed form of the faces and above a few wavelengths deviation effectively vanishes. The restrictions on the mutual parallelism of the elements of the plate assembly are much less critical since only reflections in the birefringent gaps affect the ratio e/o. Furthermore, note that in all of the above we have assumed, as in Holmes (1964), that coherence is maintained over the path differences. Because of these effects, the ripple amplitudes are expected to be upper limits.

3. CONCLUSIONS

It is shown that the high-frequency ripple sometimes seen in high-resolution spectropolarimetry is the result of multiple refections within the birefringent plates of the wave plate. While it seems that it might be difficult to produce plates that do not show this effect at high resolution, a simple extension to observing procedure is shown to eliminate it.

We wish to thank Bob Goodrich for useful comments and suggestions.

REFERENCES

Adamson, A., & Whittet, D. C. B. 1995, ApJ, 448, L49
Adamson, A. J., Whittet, D. C. B., Chrysostomou, A., Hough, J. H., Aitken, D. K., Wright, G. S., & Roche, P. F. 1999, ApJ, 512, 224
Bennett, J. M., & Bennett, H. E. 1978, in Handbook of Optics (New York: McGraw-Hill), 102
Chrysostomou, A., Hough, J. H., Whittet, D. C. B., Aitken, D. K., Roche, P. F., & Lazarian, A. 1996, ApJ, 465, L61
Holmes, D. A. 1964, J. Opt. Soc. Am., 54, 115
Smith, P. H. 1969, Proc. Symp. Recent Dev. Ellipsometry, Surf. Sci., 16, 31
Tinbergen, J. 1996, in Astronomical Polarimetry, ed. J. Tinbergen (Cambridge: Cambridge Univ. Press), 100
Weinberger, H., & Harris, J. 1964, J. Opt. Sci. Am., 54, 552
Yolken, H. T., Waxler, R. M., & Kruger, J. 1967, J. Opt. Soc. Am., 57, 283