

# Error Analysis for Dual-Beam Optical Linear Polarimetry<sup>1</sup>

FERDINANDO PATAT AND MARTINO ROMANIELLO

European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748, Garching bei München, Germany; fpatat@eso.org, mromanie@eso.org

Received 2004 September 24; accepted 2005 August 29; published 2006 January 3

**ABSTRACT.** In this paper we present an error analysis for polarimetric data obtained with dual-beam instruments. After recalling the basic concepts, we introduce the analytical expressions for the uncertainties of polarization degree and angle. These are then compared with the results of Monte Carlo simulations, which are also used to briefly discuss the statistical bias. We then approach the problem of background subtraction and the errors introduced by an imperfect Wollaston prism, flat-fielding, and retarder plate defects. Finally, we investigate the effects of instrumental polarization and propose a simple test to detect and characterize it. The application of this method to real VLT-FORS1 data has shown the presence of a spurious polarization that is of the order of  $\sim 1.5\%$  at the edges of the field of view. The cause of this problem has been identified as the presence of rather curved lenses in the collimator, combined with the incomplete removal of reflections by the coatings. This problem is probably common to all focal-reducer instruments equipped with a polarimetric mode. An additional spurious and asymmetric polarization field, whose cause is still unclear, is visible in the  $B$  band.

*Online material:* color figures

## 1. INTRODUCTION

Performing polarimetry basically means measuring flux differences along different electric field oscillation planes. In ground-based astronomy, this becomes a particularly difficult task, due to the variable atmospheric conditions, which make it difficult to detect the relatively low polarization degrees that characterize most astronomical sources (a few percent; see, e.g., Leroy 2000). These fluctuations in fact introduce flux variations among different polarization directions, which can be eventually mistaken for genuine polarization effects.

This problem has been solved in a number of different ways, reviewed by Tinbergen (1996), to which we refer the reader for a detailed description. In this paper, we focus on the so-called dual-beam configuration, which is the most popular one for instruments currently mounted at large telescopes. Despite new technologies, the basic concept of astronomical dual-beam polarimeters (see, e.g., Appenzeller 1967; Scarrot et al. 1983) has remained unchanged. A mask is placed on the focal plane, preventing image (or spectra) overlap, followed by a Wollaston prism, which splits the incoming beam into two rays that are characterized by orthogonal polarization states and are separated by a suitable angular throw. The rotation of the polarization plane is usually achieved with the introduction of a turnable retarder plate (half- or quarter-wave for linear and circular polarization, respectively) just before the Wollaston prism (see, e.g., Schmidt et al. 1992). Recently, new solutions

have been proposed to fully solve the problem in a single exposure (see Oliva 1997; Pernechele et al. 2003, for an example application), but so far they have been implemented in a few cases only.

Alternatives to Wollaston-based systems have been devised. These are mainly based on the charge transfer in CCDs, which allows an *on-chip* storage of two different polarization states that are obtained by rotating a polarization modulator. After the pioneering work of McLean et al. (1983), this technique, originally proposed by P. Stockman, has been successfully applied in a number of instruments (McLean 1997).

In this work, we address the most relevant problems connected to two-beam polarimetric observations and data reduction. The paper is organized as follows. In § 2 we introduce the basic concepts of the problem, and in § 3 we recall the analytical expressions for the uncertainties of polarization degree and angle, which are then compared to Monte Carlo simulations in § 4. In the same section we also recap the basics of polarization bias. Section 5 deals with the effects of background on polarization measurements, and § 6 covers flat-fielding issues. Section 7 is devoted to Wollaston prism deviations from the ideal. The consequences of retarder plate defects are addressed in § 8, while the effects of postanalyzer optics are discussed in § 9. Section 10 is dedicated to instrumental polarization, and § 11 deals with the case of VLT-FORS1. Finally, in § 12 we discuss and summarize our results.

## 2. BASIC CONCEPTS

The polarization state of incoming light can be described through a Stokes vector  $S(I, Q, U, V)$ ; see, e.g., Chandrasekhar

---

<sup>1</sup> This paper is partially based on observations made with ESO telescope at Paranal Observatory, under programs 066.A-0397, 69.C-0579, 069.D-0461, and 072.A-0025.

(1950). Its components, also known as Stokes parameters, have the following meanings:  $I$  is the intensity,  $Q$  and  $U$  describe the linear polarization, and  $V$  is the circular polarization. Linear polarization degree  $P$  and polarization angle  $\chi$  are related to the Stokes parameters as:

$$P = \frac{\sqrt{Q^2 + U^2}}{I} \equiv \sqrt{\bar{Q}^2 + \bar{U}^2}, \quad (1)$$

$$\chi = \frac{1}{2} \arctan \frac{U}{Q}, \quad (2)$$

where we have introduced the normalized Stokes parameters  $\bar{Q} = Q/I$  and  $\bar{U} = U/I$ . The above relations can be easily inverted to yield

$$\bar{Q} = P \cos 2\chi, \quad \bar{U} = P \sin 2\chi. \quad (3)$$

Finally, the degree of circular polarization, not discussed in this paper, is simply  $P_c = \bar{V} \equiv V/I$ . For the sake of clarity, we set  $V = 0$  and neglect all circular polarization effects throughout the paper.

The ideal measurement system for linear polarization is composed of a half-wave retarder plate (HWP) followed by the analyzer, which is a Wollaston prism (WP) producing two beams with orthogonal directions of polarization. In general, each of these elements can be treated as a mathematical operator that acts on the input Stokes vector  $\mathcal{S}$  (see, e.g., Shurcliff 1962; Goldstein 2003). What one measures on the detector are the intensities in the ordinary and extraordinary beams at a given HWP angle  $\theta_i$ , which are related to the Stokes parameters by

$$\begin{aligned} f_{O,i} &= \frac{1}{2} [I + Q \cos 4\theta_i + U \sin 4\theta_i], \\ f_{E,i} &= \frac{1}{2} [I - Q \cos 4\theta_i - U \sin 4\theta_i]. \end{aligned} \quad (4)$$

If the observations are carried out using  $N$  positions for the HWP, the whole problem of computing  $I$ ,  $Q$ , and  $U$  reduces to the solution of the  $2N$  linear equations system given by equation (4). It is clear that given three unknowns ( $I$ ,  $Q$ , and  $U$ ), at least  $N = 2$  HWP position angles have to be used.

Introducing the normalized flux differences  $F_i$ ,

$$F_i \equiv \frac{f_{O,i} - f_{E,i}}{f_{O,i} + f_{E,i}}, \quad (5)$$

and noting that  $f_{O,i} + f_{E,i} = I$ , equation (4) reduces to the  $N$  equation

$$F_i = \bar{Q} \cos 4\theta_i + \bar{U} \sin 4\theta_i = P \cos(4\theta_i - 2\chi). \quad (6)$$

We note that each  $F$ -parameter is totally determined by a single observation and is therefore independent of changes in

sky conditions. It is also worth mentioning that there are alternative approaches to the normalized flux ratios. One example can be found in Miller et al. (1987).

In principle, one can use any set of HWP angles to solve the problem, but it is easy to show that adopting a constant step  $\Delta\theta = \pi/8$  is the optimal choice. In fact, besides minimizing the errors of the Stokes parameters, this choice makes the solution of equation (6) trivial:

$$\bar{Q} = \frac{2}{N} \sum_{i=0}^{N-1} F_i \cos\left(\frac{\pi}{2} i\right), \quad (7)$$

$$\bar{U} = \frac{2}{N} \sum_{i=0}^{N-1} F_i \sin\left(\frac{\pi}{2} i\right). \quad (8)$$

Finally, it prevents ‘‘power leakage’’ (see, e.g., Press et al. 1999) when performing a Fourier analysis (see below).

In the ideal case, the normalized flux differences  $F_i$  obey equation (6), which is a pure cosinusoid. Since all possible effects introduced by the HWP must reproduce after a full revolution, it is natural to consider them as harmonics of a fundamental function whose period is  $2\pi$ .

Therefore, if  $\theta_i = \pi i/8$ , equation (6) can be rewritten as the following Fourier series:

$$F_i = a_0 + \sum_{i=1}^{N/2} a_k \sin\left(k \frac{2\pi i}{N}\right) + b_k \cos\left(k \frac{2\pi i}{N}\right),$$

where the Fourier coefficients are given by

$$\begin{aligned} a_0 &= \frac{1}{N} \sum_{i=0}^{N-1} F_i, \\ a_k &= \frac{2}{N} \sum_{i=0}^{N-1} F_i \cos\left(k \frac{2\pi i}{N}\right), \\ b_k &= \frac{2}{N} \sum_{i=0}^{N-1} F_i \sin\left(k \frac{2\pi i}{N}\right), \end{aligned} \quad (9)$$

which are valid for  $N = 4, 8, 12$ , and  $16$ . Comparing equation (9) with equation (6), it is clear that the polarization signal is carried by the  $k = N/4$  harmonic. In a quasi-ideal case, all Fourier coefficients are expected to be small compared with  $a_{N/4}$  and  $b_{N/4}$ , and deviations from this behavior could arise from a number of effects. For such an approach to the error analysis, and for the meaning of the various harmonics, the reader is referred to Fendt et al. (1996). Here we just note that the  $a_0$  term, which should be rigorously null in the ideal case, is related to the deviations of the WP from the ideal result (see § 7).

In general, a Fourier analysis is meaningful when  $N = 16$ , and it can reveal possible problems directly related to HWP quality (cf. § 8). In most cases, however, for practical reasons, one typically uses  $N = 4$ , and in that case, a different error treatment is required.

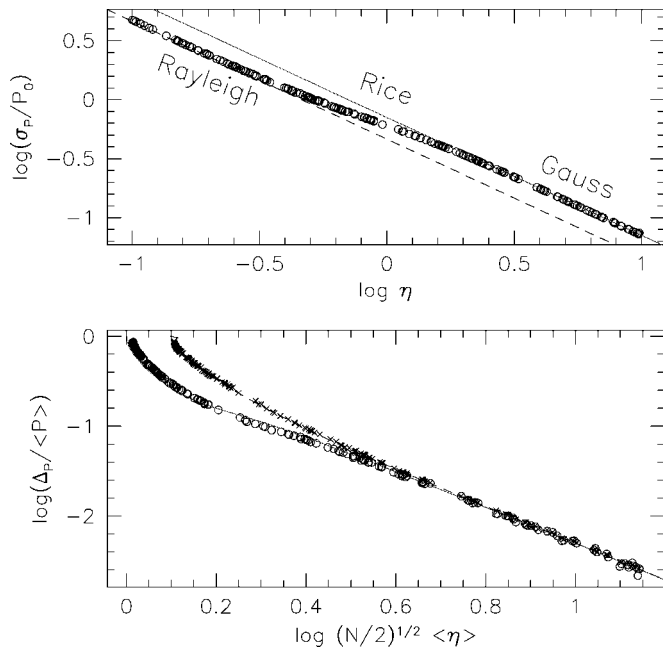


FIG. 1.—*Top*: Comparison between the rms error on the polarization degree from MC simulations (*circles*) and eq. (10). The dashed line traces the expected rms error for the Rayleigh distribution (see text). *Bottom*: Bias estimated using the mode (*circles*) and the average (*crosses*). For comparison, the solid curve traces the Wardle & Kronberg (1974) solution for the mode, while the dashed line shows the Sparks & Axon (1999) solution for the average. [See the electronic edition of the PASP for a color version of this figure.]

### 3. ANALYTICAL ERROR ANALYSIS

Under the assumption that all relevant quantities are distributed according to Gaussian laws, one can analytically derive simple expressions for the corresponding errors of the final results. As we see in § 4, this assumption is not always correct, and when this happens, a numerical treatment is required in order to test the analytical results and their range of validity. Assuming that the background level is the same in the ordinary and extraordinary beams, and that the readout noise can be neglected, the analytical expression for the absolute error of  $P$  can be readily derived (see, e.g., Miller et al. 1987) by propagating the various errors through the relevant equations<sup>2</sup>:

$$\sigma_p = \frac{1}{\sqrt{N/2}(S/N)}, \quad (10)$$

where  $S/N$  is the signal-to-noise ratio of the intensity image ( $f_o + f_E$ ). The  $S/N$  one expects to achieve in the polarization

<sup>2</sup> Here we consider photon shot noise as the only source of random error. Another potential source is represented by the mispositioning of the HWP with respect to the optimal angles. However, as analytical solutions and numerical simulations show, with the typical positioning accuracy currently attainable ( $<1^\circ$ ), the associated error of the polarization degree and angle is negligible.

degree,  $(S/N)_p = P/\sigma_p$ , is simply given by

$$(S/N)_p = \sqrt{N/2}P(S/N).$$

As for the error of  $\chi$ , this is given by

$$\sigma_\chi = \frac{1}{2\sqrt{N/2}P(S/N)} \equiv \frac{\sigma_p}{2P}, \quad (11)$$

from which it is clear that at variance with the degree of polarization, the accuracy of the polarization angle does depend on the intrinsic polarization degree.

## 4. MONTE CARLO SIMULATIONS

The analytical treatment presented in § 3 relies on the assumption that all relevant variables obey Gaussian statistics. Numerical simulations are required in order to derive more realistic distributions and to verify the validity of the analytical results. One can easily implement the concepts we have developed until now in a Monte Carlo (MC) code, which also allows higher sophistication, such as the inclusion of Poissonian noise. With this tool, one can readily investigate the effects of non-Gaussian distributions of the derived quantities, the most important of which is the systematic error of the polarization, as first pointed out by Serkowski (1958).

### 4.1. Linear Polarization Bias

Due to the various noise sources, the vector components  $\bar{Q}$  and  $\bar{U}$  are normally distributed, but since  $P$  is defined as the quadrature sum of  $\bar{Q}$  and  $\bar{U}$ , the statistical errors always add in the positive direction, leading to a systematic increase of the estimated polarization degree, thus introducing a bias. The problem was addressed by several authors, using both analytical and numerical methods (Serkowski 1958; Wardle & Kronberg 1974; Simmons & Stewart 1985; Clarke & Stewart 1986; Sparks & Axon 1999). We refer the reader to those papers for a detailed description of the problem; here we recall just the basic concepts and apply them to our case.

The polarization bias is usually quantified using a robust estimator, supposedly giving a statistically significant representation of the observed value, which is then compared with the input polarization in order to derive the systematic correction. Different choices have been adopted (see Sparks & Axon 1999). Following the considerations by Wardle & Kronberg (1974) we have adopted here the mode  $\langle P \rangle$  of the distribution in order to estimate the bias, which we therefore define as  $\Delta P = \langle P \rangle - P_0$ , where  $P_0$  is the input polarization. Once applied to the observed data, the bias correction  $\Delta P$  tends to restore the symmetry of the deviation distribution (see, e.g., Sparks & Axon 1999, their Fig. 4).

In Figure 1 we show the results of our MC simulations for the estimated rms error of the polarization  $\sigma_p$  (*top panel*) and polarization bias  $\Delta P$  (*bottom*) for  $N = 4$ . Following

Sparks & Axon (1999), we have used  $\eta \equiv P_0(S/N)$  and  $\langle \eta \rangle \equiv \langle P \rangle(S/N)$  as independent variables in our plots. As we had anticipated,  $\sigma_p$  follows the analytical prediction of equation (10) when  $\eta > 2$ . For lower values of  $\eta$ ,  $\sigma_p$  tends to be systematically smaller than the analytical prediction, and it converges to the value expected for the Rayleigh distribution (*dashed line*), which becomes a very good approximation for  $\eta < 0.5$ , provided that  $S/N > 3$ . For intermediate values of  $\eta$ , the distribution is described by a Rice function (Rice 1944). In conclusion, one can safely use the analytical solution given by equation (10) for  $\eta \geq 2$  only.

As for the polarization bias, we have plotted it as a function of measurable quantities, namely the S/N and the *observed* polarization  $\langle P \rangle$ .

The results of our MC simulations, as plotted in the bottom panel of Figure 1, are in good agreement with the analytical solution found by Wardle & Kronberg (1974) for the same statistical estimator. For comparison, we have also plotted the results obtained when the average is adopted (*crosses*). For  $(N/2)^{1/2} \langle \eta \rangle > 4$ , the relation between  $\log(\Delta P/\langle P \rangle)$  and  $\log([N/2]^{1/2} \langle \eta \rangle)$  is well approximated by a linear law. A least-squares fit gives the result

$$P_0 = \langle P \rangle \left[ 1 - \left( \frac{0.62}{\sqrt{N/2} \langle \eta \rangle} \right)^{1.92} \right],$$

which can be used to correct the observed polarization values according to the input S/N, measured polarization, and number of HWP positions. In general, the bias effect is present even at reasonably high values of S/N when the polarization is small and  $\sigma_p$  and  $\Delta P$  tend to become similar, so that the systematic bias correction is comparable to the random uncertainty of the polarization. This is better seen in Figure 2, where we have plotted the ratio  $\Delta P$  over  $\sigma_p$  as a function of  $P_0/\sigma_p \equiv (S/N)_p$  deduced from our simulations. As anticipated, for low values of  $(S/N)_p$ , the ratio between  $\Delta P$  and  $\sigma_p$  tends to unity, with some variations among different estimators. For  $(S/N)_p \geq 3$ , the bias correction is less than 10% of the expected accuracy, and it is therefore negligible. Moreover, above that threshold, all estimators give practically identical results.

In Figure 2 we have plotted for comparison the function computed by Simmons & Stewart (1985), who have used a maximum likelihood estimator in order to evaluate the bias. As these authors have shown, this is the best estimator for  $(S/N)_p \leq 0.7$ , while for  $(S/N)_p > 0.7$  the mode first used by Wardle & Kronberg (1974) should be used.

## 5. THE EFFECTS OF THE BACKGROUND

Until now we have assumed that one is able to perfectly subtract the background contribution. This is most likely the case when performing polarimetric measurements on pointlike sources, since in that situation local background subtraction is in most cases straightforward.

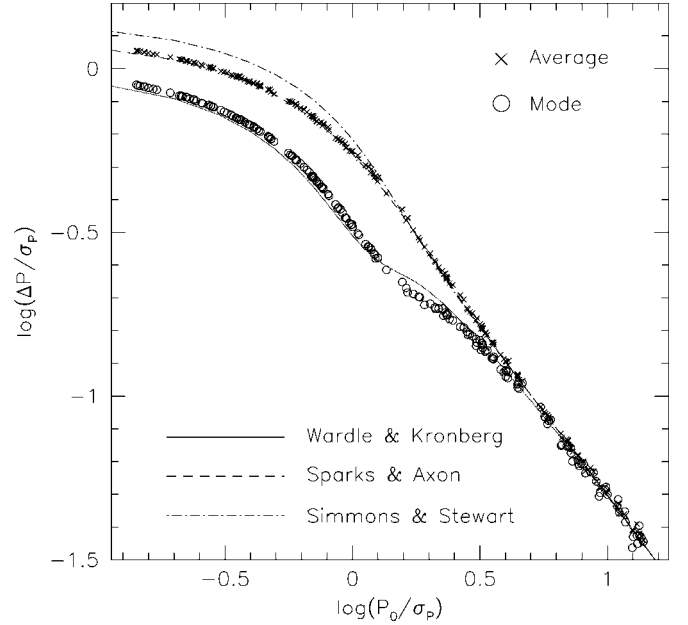


FIG. 2.—Comparison between systematic bias  $\Delta P$  and random error  $\sigma_p$  for different estimators as a function of input polarization S/N. [See the electronic edition of the PASP for a color version of this figure.]

We note that the background, whatever its nature is, must be subtracted before calculating normalized Stokes parameters (see also Tinbergen 1996), so that possible background polarization can be *vectorially* removed.

If we assume that the object is characterized by  $P_o$  and  $\chi_o$ , and the background by  $P_b$  and  $\chi_b$ , the two polarization fields can be expressed using Stokes vectors defined as  $S_o(I_o, I_o P_o \cos 2\chi_o, I_o P_o \sin 2\chi_o)$  and  $S_b(I_b, I_b P_b \cos 2\chi_b, I_b P_b \sin 2\chi_b)$ , where we have neglected any circular polarization. Since Stokes vectors are additive (see, e.g., Chandrasekhar 1950), the resulting polarization field is described by  $S = S_o + S_b$ , and therefore the total polarization is given by the formula

$$P = \frac{I_o P_o}{I_o + I_b} \sqrt{1 + r^2 + 2r \cos [2(\chi_o - \chi_b)]},$$

where  $r = (I_b P_b)/(I_o P_o)$ ; i.e., the ratio between the polarized fluxes of background and object. The corresponding polarization angle is

$$\chi = \frac{1}{2} \arctan \frac{\sin 2\chi_o + r \sin 2\chi_b}{\cos 2\chi_o + r \cos 2\chi_b}.$$

Clearly, the background is going to significantly influence the object when  $r \approx 1$ . For  $r \sim 1$ , one can write

$$P \approx (P_o/\sqrt{2}) \sqrt{1 + \cos [2(\chi_o - \chi_b)]},$$

which implies that for comparable polarized fluxes, the re-

sulting polarization is nulled when the polarization fields are *perpendicular* ( $|\chi_o - \chi_b| = \pi/2$ ).

## 6. FLAT-FIELDING

One of the basic problems in reducing the data produced by dual-beam instruments is the flat-fielding. Because image splitting occurs after the focal mask, collimator, and the HWP, one would in principle need to obtain flat exposures with all optical components in the light path. Unfortunately, in all practical conditions, this introduces strong artificial effects, due to the strong polarization typical of flat-field sources (either twilight sky or internal screens). In principle, one can reduce this effect using the continuous rotation of the HWP as a depolarizer. This is implemented, for example, in EFOSC2, currently mounted at the ESO 3.6 m telescope (Patat 1999), and it is effective only if the HWP rotation time is much shorter than the required exposure time. The depolarizing effect can also be achieved by averaging flats taken with the same set of HWP angles used for the scientific exposures. In fact, with the use of the optimal angle set (see § 2), one has

$$\sum_{i=0}^{N-1} f_{o,i} = \frac{N}{2} I$$

and a similar expression for  $f_{e,i}$ , which do not contain any polarization information. The problem is that this is true only if the source is stable in intensity, which is surely not the case for the twilight sky, and probably not really true for most lamps.

An alternative solution (at least for imaging) is the use of a set of twilight flats obtained without HWP and WP. While on the one hand this eliminates source polarization, on the other hand it does not allow for a proper flat-field correction. In fact, while the pixel-to-pixel variations are properly taken into account, the large-scale patterns are not, due to the splitting of the beam, which maps a given focal plane area into two different regions of the post-WP optics and the detector. Moreover, these calibrations do not carry any information about possible spatial effects introduced by the HWP and the WP. However, as the simulations show, this problem becomes milder if some redundancy is introduced. For example, if one uses  $N = 4$  HWP positions, the ordinary and extraordinary rays will just swap when the angle differs by  $\pi/4$  within each of the two redundant pairs ( $f_o^i = -f_e^{i+2}$ ; see eq. [4]). This tends to cancel out the flat-field effect and becomes more efficient if the maximum redundancy ( $N = 16$ ) is used. However, it must be noted that time-dependent effects, such as fringing, may affect the redundant pairs in a different way, therefore decreasing the cancellation efficiency.

## 7. EFFECTS OF A NONIDEAL WOLLASTON PRISM

So far we have assumed that our system is described by equation (4); i.e., that the Wollaston prism splits incoming un-

polarized light into identical fractions. A deviation from this ideal behavior can be described by the introduction of a new parameter  $t$  in equation (4), which can be reformulated as

$$\begin{aligned} f_{o,i} &= t[I + Q \cos 4\theta_i + U \sin 4\theta_i], \\ f_{e,i} &= (1-t)[I - Q \cos 4\theta_i - U \sin 4\theta_i]. \end{aligned} \quad (12)$$

An ideal system is obtained for  $t = \frac{1}{2}$ . Now, for unpolarized light ( $Q = U = 0$ ), these new equations give  $f_{o,i} = tI$  and  $f_{e,i} = (1-t)I$  for all HWP angles, so that all normalized flux differences turn out to be identical (i.e.,  $F_i = 2t - 1$ ). Therefore, the value of  $t$  can be directly estimated observing an unpolarized source.

In the simplest situation, where  $N = 2$ , neglecting the presence of the  $t$  term would lead to a spurious polarization degree  $P = \sqrt{2}(2t - 1)$ , with a polarization angle  $\chi = \pi/8$ . It is interesting to note that this is not the case, for example, when  $N = 4$ . In fact in that situation, because all redundant  $F$ 's are identical, equations (7) and (8) would correctly yield null Stokes parameters.

The problem is more complicated when the incoming light is polarized, since the normalized flux differences are no longer a linear combination of  $\bar{Q}$  and  $\bar{U}$ , as one can verify from equation (12):

$$F_i = \frac{\kappa + \bar{Q} \cos 4\theta_i + \bar{U} \sin 4\theta_i}{1 + \kappa \bar{Q} \cos 4\theta_i + \kappa \bar{U} \sin 4\theta_i}, \quad (13)$$

where we have set  $\kappa = 2t - 1$  ( $|\kappa| \leq 1$ ,  $\kappa = 0$  in the ideal case). If  $\kappa$  is known, one can correct the observed  $f_o$  and  $f_e$  by dividing them by  $2t$  and  $2(1-t)$ , respectively, before following the procedure adopted for the ideal case. Of course, if the source has a known polarization (e.g., a polarized standard), one can use this information together with the observed  $F$  ratios to derive  $\kappa$  for each HWP angle, according to the relation

$$\kappa_i = \frac{\bar{Q} \cos 4\theta_i + \bar{U} \sin 4\theta_i - F_i}{F_i(\bar{Q} \cos 4\theta_i + \bar{U} \sin 4\theta_i)}.$$

If the input polarization is unknown, then one can in principle derive  $\kappa$  from the observations themselves, provided that  $N \geq 4$ . In fact, after introducing the parameter

$$g_j = \frac{1 + F_{j-1}F_{j+1}}{F_{j-1} + F_{j+1}} \quad (14)$$

and using equation (13), it is easy to demonstrate that

$$\kappa_j = g_j \pm \sqrt{g_j^2 - 1}, \quad (15)$$

where  $j = 1, \dots, N/2 - 2$  and the positive sign refers to the case  $g_j \leq 0$ . For instance, for  $N = 4$ , one can determine two

independent estimates of  $\kappa$ , which can be averaged to improve on the accuracy.

It is interesting to note that when  $P \ll 1$ , equation (13) can be approximated as  $F_i \approx \kappa + \bar{Q} \cos 4\theta_i + \bar{U} \sin 4\theta_i$ . If  $N$  is a multiple of 4 (i.e., if the  $F$  function is sampled for an integer number of periods  $\pi/2$ ), one has

$$\kappa \approx \frac{1}{N} \sum_{i=0}^{N-1} F_i \equiv a_0, \quad (16)$$

which clarifies the meaning of the  $a_0$  term in the Fourier series (see eq. [9]).

It is worth mentioning that the redundancy in the  $F$  parameters does eliminate the effects of a nonideal WP, to a large extent. For example, a blind application of equations (7) and (8) to the case of  $N = 4$  gives the result

$$\frac{2}{N} \sum_{i=1}^{N-1} F_i \cos(4\theta_i) = \bar{Q} \frac{1 - \kappa^2}{1 - \kappa^2 \bar{Q}^2}$$

and a similar expression for  $\bar{U}$ . If the polarization is small ( $P_0 \leq 0.1$ ), we have that  $\kappa^2 \bar{Q}^2 \ll 1$ , and therefore the application of the procedure for an ideal case in a nonideal situation would lead to a value of  $\bar{Q}$ , which is  $(1 - \kappa^2)$  times smaller than the real one. Since the same is true for  $\bar{U}$ , the resulting polarization  $P$  will also be  $(1 - \kappa^2)$  times smaller than the input value, while the polarization angle remains unchanged. For example, if  $|\kappa| \leq 0.2$ ,  $\sigma_p/P$  is less than 4%.

Equation (12) describes a particular case only, in which the incoming unpolarized flux is distributed into two fractions,  $t$  and  $(1 - t)$ . More generally, one should replace the term  $(1 - t)$  with an independent parameter  $s$  so that the fraction of light split by the WP in the ordinary and extraordinary rays become uncorrelated. Using the same procedure, it is easy to demonstrate that one can estimate the ratio  $\kappa_j = (t - s)/(t + s)$  still using equation (15).

Another effect we have investigated is the possibility that the difference in polarization direction between the ordinary and the extraordinary rays of the WP is not  $\pi/2$ . If we call  $\Delta\alpha$  the deviation from this ideal angle, then using the general expression of the Mueller matrix for a linear polarizer (see, e.g., Goldstein 2003) and deriving the expressions for the normalized flux ratios, one gets

$$F_i \approx \frac{\bar{Q}}{2} [\cos(4\theta_i - 2\Delta\alpha) + \cos 4\theta_i] + \frac{\bar{U}}{2} [\sin(4\theta_i - 2\Delta\alpha) + \sin 4\theta_i],$$

where the approximation is valid under the assumption that  $P \ll 1$ . From of this expression, it is easy to conclude that for

reasonably small values of  $\Delta\alpha$  ( $\leq 10^\circ$ ), the implied errors are of the order of 0.05% on the polarization degree and  $5^\circ$  on the polarization angle, irrespective of the number of HWP positions used.

## 8. HWP DEFECTS

In the ideal case, the normalized flux differences are modulated by the HWP rotation according to equation (6), which is a pure cosinusoid. If defects such as dirt or inhomogeneously distributed dust are present on the HWP, one can expect spurious flux modulations that are unrelated to the polarization of the incoming light and that can reduce the performance of the instrument. As a consequence, error estimates based on pure photon statistics are systematically smaller than the actual errors.

These kinds of problems can be investigated with the aid of Fourier analysis, following the procedure we have outlined at the end of § 2. This method becomes particularly effective when the observations are taken sampling the full HWP angle range (i.e.,  $2\pi$ ), which, given the choice of the optimal angle set  $\theta_i = \pi i/8$ , implies  $N = 16$  retarder plate positions. Under these circumstances, one can determine the Fourier coefficients  $a_k$  and  $b_k$  for the first eight harmonics, the fundamental harmonic ( $k = 1$ ) being related to local transparency fluctuations, which repeat themselves after a full revolution, like dirt or dust. By definition (e.g., eq. [6]), the fourth harmonic is directly related to the linear polarization (i.e.,  $a_4 \equiv \bar{Q}$  and  $b_4 \equiv \bar{U}$ ). All other harmonics, with the only remarkable exception being the second one ( $k = 2$ ), are simply overtones of harmonics with lower frequencies, and include part of the noise generated by the photon statistics, which is present at all frequencies and is therefore indicated as white noise. For this reason, the global random error is often estimated as the signal carried by the harmonics with  $k = 3, 5-8$  (see, e.g., Fendt et al. 1996, their Appendix A.3).

The second harmonic deserves a separate discussion. Ideally, the HWP operates as a pure rotator of the input Stokes vector, with the advantage that one does not need to rotate the whole instrument in order to analyze different polarization planes. In the real case, in which the HWP is usually constructed using birefringent materials, it is affected by so-called pleochroism. This is a wavelength-dependent variation of the transmission that takes place when the direction of the incoming light is changed with respect to the crystal lattice. Because of the way the HWP is manufactured, the crystals have an axial symmetry, which gives a period of  $\pi$ . Therefore, this effect is seen as the  $k = 2$  component.

In Figure 3 we show a real case in which we have applied this analysis to archival data obtained with FORS1, which is currently mounted at the Cassegrain focus of the ESO VLT 8.2 m telescope (Szeifert 2002).

A bright (and supposedly unpolarized) star was observed using  $N = 16$  HWP positions. First, a Fourier analysis indi-

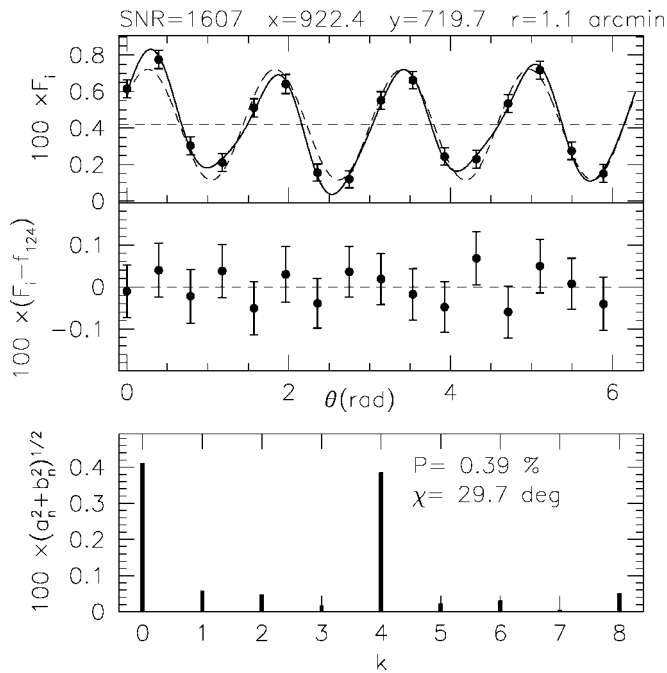


FIG. 3.—Example of Fourier analysis applied to archival VLT-FORS1 observations of a bright star in the V passband (see text). *Top*: Normalized flux differences. Partial reconstructions using eight harmonics (*solid curve*) and the fourth harmonic only (*dashed curve*). The dashed horizontal line is placed at the average of  $F$  values ( $a_0$ ). *Middle*: Residuals after subtracting the  $k = 1, 2$ , and  $4$  components. *Bottom*: Harmonics power spectrum.

icates the presence of a small deviation of the WP from the ideal ( $\kappa \approx a_0 \approx 4.1 \times 10^{-3}$ ; see § 7). Second, a clear polarization is detected at the level of about 0.4%, while all other components are smaller than 0.05% (this polarization is actually an instrumental effect present in FORS1; see § 10). The effective significance of harmonics other than  $k = 4$  can be judged on the basis of the expected errors of the Fourier coefficients. For example, using the expression  $a_k$ , one finds that

$$\sigma_{a_k} = \frac{2}{N(S/N)} \sqrt{\sum_{i=1}^N \cos^2 k \frac{2\pi i}{N}}.$$

With this kind of analysis, one can see that in the example of Figure 3,  $a_k$  and  $b_k$  are consistent with a null value for  $k = 3, 5-7$  (see center panel). As for the  $k = 1$  and  $2$  harmonics, the Fourier coefficients are non-null at a  $2\sigma$  level. Since for the test star it was  $S/N \sim 1600$ , it is clear that to detect  $k = 1$  and  $2$  harmonics of this amplitude (0.05%), a  $S/N \geq 3000$  is required.

It is important to note that the presence of these non-null harmonics is implicitly corrected for when one has a sufficient number of HWP positions covering the maximum period  $2\pi$ . In the most common case, in which  $N = 4$  angles spaced by  $\pi/8$  are used, one can derive the fundamental (i.e., linear po-

larization, period  $\pi/2$ ) and the first overtone (period  $\pi/4$ ) only. The latter corresponds to the  $k = 8$  component of the  $N = 16$  cases, which therefore carries the high-frequency information only. As a consequence, if other harmonics are present, they are not properly removed and contribute to the final error, effectively setting the maximum accuracy one can achieve, irrespective of the S/N. Numerical simulations performed assuming a virtually infinite S/N show that in the presence of  $k = 1$  and  $2$  components, the use of  $N = 4$  HWP angles leads to systematic errors that are of the same order of amplitude as the two harmonics. From this and the example reported in Figure 3, we estimate that the absolute maximum accuracy attainable with FORS1 using  $N = 4$  is of the order of 0.05%.

Another typical problem that affects the retarder plates is the chromatic dependence of the angle zero point. This is usually measured by means of a Glan-Thompson prism, and it can change by more than  $5^\circ$  across the optical wavelength range. The computed polarization angle can be corrected by simply adding the HWP angle offset for the relevant wavelength (or effective wavelength, in the case of broadband imaging); see, for example, Szeifert (2002).

Finally, we have investigated the effects produced by a deviation  $\Delta\beta$  from the nominal phase retardance of an HWP ( $\pi$ ). Using the general expression of the Mueller matrix for a retarder (see, e.g., Goldstein 2003), the normalized flux ratios turn out to be

$$F_i = \bar{Q}[\cos 4\theta_i - \sin^2 2\theta_i(1 - \cos \Delta\beta)]$$

$$+ \frac{1}{2}\bar{U} \sin 4\theta_i(1 + \cos \Delta\beta) + \bar{V} \sin 2\theta_i \sin \Delta\beta,$$

from which it is clear that the measured linear polarization depends on the circular polarization of the input signal. For  $V = 0$  and  $\Delta\beta \leq 10^\circ$ , the corresponding absolute error of the polarization degree is less than 0.05%, while the outcome on the polarization angle is negligible. For  $V \neq 0$ , the exact effect depends on the ratio between the degrees of circular and linear polarization. For example, for  $\Delta = 10^\circ$  and  $\bar{Q} = \bar{U} = \bar{V} = 0.01$ , the absolute error of the computed polarization degree is about 0.1% for  $N = 4$ , which decreases to 0.01% for  $N = 16$ . It is worthwhile noting that this defect would be detected by a Fourier analysis as a component with a period  $\pi$  and an intensity of  $|\bar{V} \sin \Delta|$ .

## 9. EFFECTS OF POSTANALYZER OPTICS

Typically, the analyzer is followed by additional optics, such as filters, grisms, and camera lenses, which, due to their possible tilt with respect to the optical axis, may behave as poor linear polarizers. In the most probable case, in which the polarization is produced by transmission (see also § 10), the properties of postanalyzer (PA) components can be described by the ap-

proximate Mueller matrix

$$M_{\text{PA}} \approx A \begin{pmatrix} 1 & Bc & Bs & 0 \\ Bc & 1 & 0 & 0 \\ Bs & 0 & C & 0 \\ 0 & 0 & 0 & C \end{pmatrix},$$

where  $A \approx 1$ ,  $C \approx 1$ ,  $c = \cos 2\varphi$ ,  $s = \sin 2\varphi$ , and  $\varphi$  is the polarization angle (which can change across the field of view), while  $B$  is related to the polarization degree introduced by the PA optics. This expression can be deduced from the general formulation (see Keller 2002, eq. [4.63]) after applying the usual matrix rotation (see, e.g., Keller 2002, eq. [2.5]). If  $S_0 = (I_0, Q_0, U_0, V_0)$  is the input Stokes vector, the effect of PA optics can be evaluated by computing the Stokes vectors that correspond to the ordinary and extraordinary beams produced by the WP, transforming them using the operator  $M_{\text{PA}}$ , and using the resulting intensity components to compute the normalized flux differences  $F_i$ . After simple calculations, one arrives at the expression

$$F_i \approx B \cos 2\varphi + \bar{Q}_0 \cos 4\theta_i + \bar{U}_0 \sin \theta_i,$$

where we have assumed that  $|B| \ll 1$ , i.e., that the linear polarization induced by the PA optics is small. Given this expression, it is clear that the redundancy in the HWP positions ( $N = 4, 8$ , and  $16$ ) eliminates this problem, since the additive term  $B \cos 2\varphi$  is not modulated by the HWP rotation, while for  $N = 2$  the derived polarization degree and angle would be affected, possibly severely. If the optimal HWP angle set has been used, it is easy to verify that  $B \cos 2\varphi = \sum_{i=0}^{N-1} F_i/N$ , which is identical to equation (16). This means that in a first approximation, it is not possible to distinguish between an imperfect WP and the presence of polarization in the PA optics. Therefore, the fact that  $a_0 \approx 0.4\%$  in Figure 3 can actually be attributed to both kinds of problems. The PA optics effect definitely becomes stronger when these include highly tilted components, such as gratings. This is very well illustrated by the two examples in Figure 4, where we show the results obtained using VLT-FORS1 archival data of a highly polarized star (Vela 1 95,  $\alpha = 09^{\text{h}}06^{\text{m}}00^{\text{s}}$ ,  $\delta = -47^{\circ}19'00''$ ) and an unpolarized star (WD 1615–154,  $\alpha = 16^{\text{h}}17^{\text{m}}55^{\text{s}}$ ,  $\delta = -15^{\circ}35'51''$ ),<sup>3</sup> which were observed on the optical axis, where the instrumental polarization is known to be null (Szeifert 2002).

In both cases, the polarization degree deduced using  $N = 2$  (*center panels*) is markedly different from that derived with  $N = 4$  (*top*), and the deviation is particularly severe for the unpolarized object. As is finally apparent, the resulting values of  $B \cos 2\varphi$  show a strong wavelength dependency and are higher than 5% at about 800 nm. It is interesting to note that  $B \cos 2\varphi \sim 0$  at about 450 nm; i.e., at the wavelength where the antireflection coatings are optimized (see § 10). This fact,

together with the marked wavelength dependency and the much lower level seen in broadband imaging (see Fig. 3), strongly suggests that the effect seen in Figure 4 is indeed produced by the tilted surfaces of the grism.

## 10. EFFECTS OF INSTRUMENTAL POLARIZATION

So far, we have assumed that all optics preceding the analyzer do not introduce any polarization. Of course, this is not generally true (see, e.g., Tinbergen 1996; Leroy 2000 for a general introduction to the subject).

To show the effect of instrumental polarization, we assume that the preanalyzer optics, which include telescope mirrors, collimator, HWP, and so on, introduce an artificial polarization, which depends on the position in the field. For the sake of simplicity, we assume that these optics act as a nonperfect linear polarizer characterized by a position-dependent polarization degree  $p(x, y)$  and polarization angle  $\varphi(x, y)$ . This can be described by the Mueller matrix

$$M_I(x, y) = \frac{1}{1+p} \begin{pmatrix} 1 & pc & ps \\ pc & 1 - ps^2 & psc \\ ps & psc & 1 - pc^2 \end{pmatrix},$$

where we have neglected circular polarization and have set  $s = \sin 2\varphi$  and  $c = \cos 2\varphi$ . For  $p = 0$ , one obtains a totally transparent component, while  $p = 1$  gives an ideal linear polarizer.

If  $S_0(I_0, Q_0, U_0)$  is the Stokes vector describing the input polarization state, it will be transformed by preanalyzer optics into the vector  $S_1 = M_I \cdot S_0$  before entering the analyzer:

$$\begin{aligned} I_1(1+p) &= I_0 + pcQ_0 + psU_0, \\ Q_1(1+p) &= Q_0 + pcI_0 - ps^2Q_0 + pcsU_0, \\ U_1(1+p) &= U_0 + psI_0 + pcsQ_0 - pc^2U_0, \end{aligned} \quad (17)$$

and therefore the measurements would lead to  $S_1$ , which would then need to be corrected for the instrumental effect inverting equation (17), provided that  $p(x, y)$  and  $\varphi(x, y)$  are known.

Of course, if the observed source is known to be unpolarized,  $p$  and  $\varphi$  can be derived immediately, for example by placing a single target on different positions of the field of view, or observing an unpolarized stellar field.

If the source is polarized, the problem becomes much more complicated, since equation (17) is nonlinear in  $c$  and  $s$ . The solution can be simplified by assuming that  $p \ll 1$  and  $P_0 \ll 1$ , which is a reasonable hypothesis in most real cases, since instrumental polarization is typically less than a few percent. In

<sup>3</sup> See <http://www.eso.org/instruments/fors/inst/pola.html>.



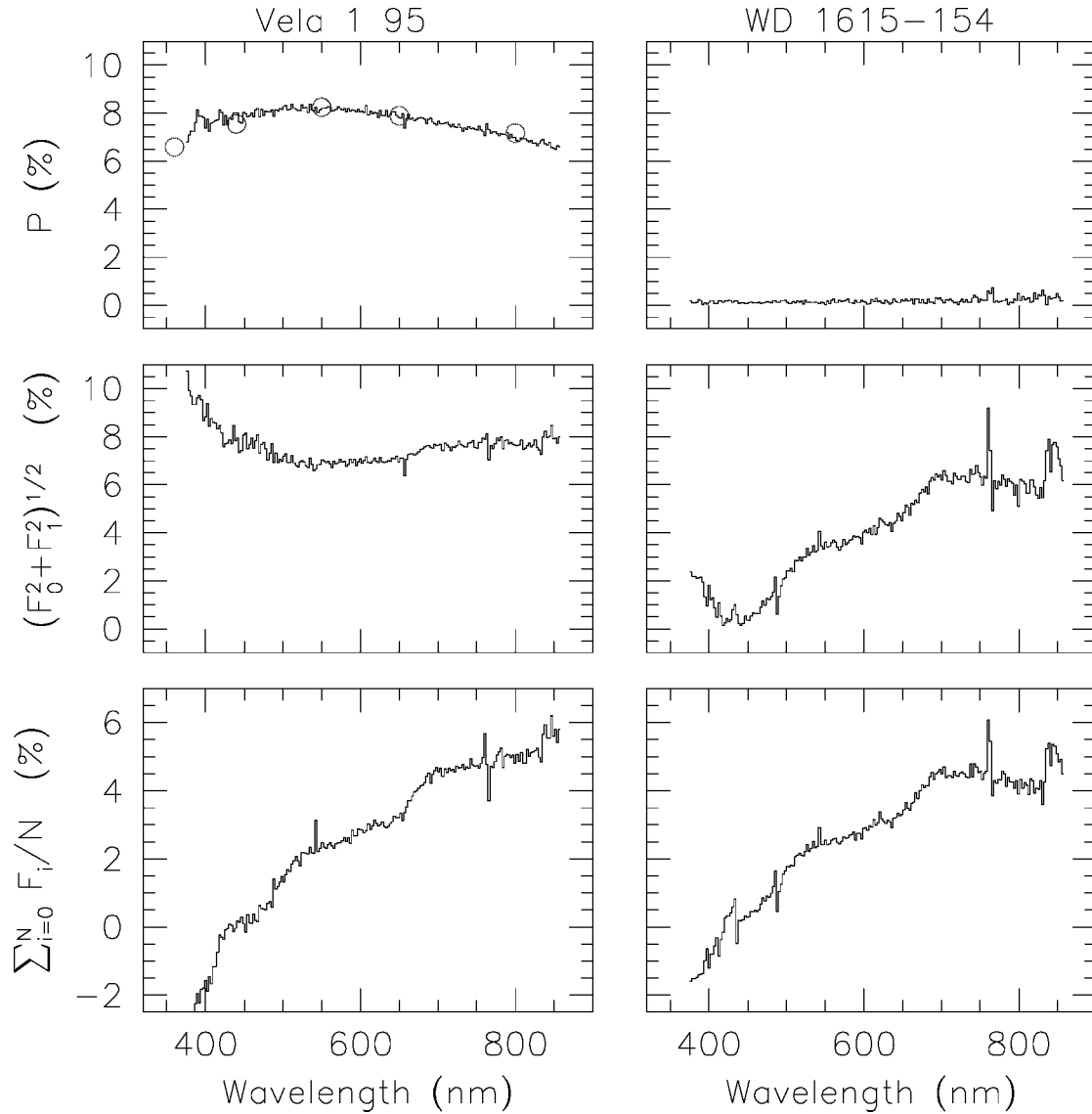


FIG. 4.—VLT-FORS1 observations of Vela 1 95 (left) and WD 1615–154 (right). The plots show the linear polarization derived with  $N = 4$  HWP positions (top),  $N = 2$  (middle) and  $\sum_{i=0}^N F_i/N$  for  $N = 4$  (bottom; see text for more details). The original spectra were obtained with the 300V grism and a slit of 1"; for presentation, they have been binned to 25 Å. The open circles in the top left panel mark the broadband polarimetric measurements for Vela 1 95 (UBVRI, from left to right). [See the electronic edition of the PASP for a color version of this figure.]

this case, equation (17) can be rewritten as

$$\begin{aligned}
 I_1 &\approx I_0, \\
 Q_1 &\approx Q_0 + pI_0 \cos 2\varphi, \\
 U_1 &\approx U_0 + pI_0 \sin 2\varphi.
 \end{aligned}
 \tag{18}$$

It is important to note that the instrumental polarization is not removed by the local background subtraction. Moreover, it is independent of the object's intensity; in fact, using the

previous expressions, one can verify that

$$P = \sqrt{P_0^2 + p^2 + 2P_0p \cos [2(\chi_0 - \varphi)]},$$

where  $P_0$  and  $\chi_0$  are the input polarization degree and angle, respectively. From this expression, it is clear that when  $P_0 \gg p$ , it is also  $P \approx P_0$ , while in the case that object and instrumental polarization are comparable ( $p \approx P_0$ ), the observed polarization is approximately given by

$$P \approx \sqrt{2P_0} \sqrt{1 + \cos [2(\chi_0 - \varphi)]},$$

which, according to the value of  $(\chi_0 - \varphi)$ , gives values that range from 0 to  $2P_0$ . It is important to note that the main difference between instrumental polarization and a polarized background is that the latter is effective only when the background is  $\geq I$  (see § 5), while the former acts regardless of the object intensity; what counts is its polarization.

With the aid of these approximate expressions (eq. [18]), one can easily evaluate the instrumental polarization, provided that the input polarization field is known and the observed source covers a large fraction of the instrument field of view. In fact, solving equation (18) for  $\varphi$  and  $p$  yields

$$\tan 2\varphi \approx \frac{U_1 - U_0}{Q_1 - Q_0}$$

and

$$p_1 \approx \frac{Q_1 - Q_0}{I_0 \cos 2\varphi}, \quad p_2 \approx \frac{U_1 - U_0}{I_0 \sin 2\varphi},$$

where  $p_1$  and  $p_2$  are two independent estimates of  $p$  that can be averaged to increase the accuracy.

As it is well known, the night sky shows a polarization that varies according to the ecliptic and Galactic coordinates (see Leinert et al. 1998 for an extensive review). It is mostly dominated by zodiacal light polarization, which reaches its minimum, below a few percent, at the antisolar position (Roach & Gordon 1973). Since this is not expected to vary on scales of a few arcminutes, in principle, relatively empty fields represent suitable targets for panoramic polarization tests, provided that the S/N per spatial resolution element is of the order of several thousand.

## 11. THE CASE OF FORS1 AT THE ESO VLT

As an example application, we have performed a test using real data obtained with FORS1 at the ESO VLT. In this instrument, the polarimetric mode is achieved by inserting into the beam a superachromatic HWP and a WP that has a throw of about  $22''$  (Szeifert 2002).

We have identified in the ESO archive three sets of data obtained in rather empty fields in  $B$ ,  $V$ , and  $I$  passbands. Table 1 lists the equatorial coordinates ( $\alpha$ ,  $\delta$ ), ecliptic longitude and latitude ( $\lambda$ ,  $\beta$ ), heliocentric longitude ( $\lambda - \lambda_\odot$ ), and degree of sky polarization and angle ( $P_{\text{sky}}$ ,  $\chi_{\text{sky}}$ ) for the different fields. In all three cases, the S/N achieved on the sky background  $I_{\text{sky}}$  in the combined images is larger than  $S/N \approx 200 \text{ pixel}^{-1}$ . With such a signal and for a typical 1% polarization, the bias effect is expected to be small (see Fig. 1), and the rms error of the polarization degree, according to equation (10), is of the order of 0.3%, while the uncertainty of  $\chi$  is about  $9^\circ$  (see eq. [11]). In order to further increase the accuracy and to allow for outlier rejection, we have computed a clipped average in

TABLE 1  
DATA FOR THE EMPTY FIELDS

Filter	$\alpha$ (J2000.0)	$\delta$ (J2000.0)	$\lambda$ (deg)	$\beta$ (deg)	$\lambda - \lambda_\odot$ (deg)	$P_{\text{sky}}$ (%)	$\chi_{\text{sky}}$ (deg)
$B$ .....	03 32 17	-27 44 24	-41.1	-45.1	128.1	6.45	+68.4
$V$ .....	13 58 03	-31 22 21	218.6	-18.1	-159.7	1.74	-53.8
$I$ .....	20 36 08	-13 06 39	308.0	+5.3	-110.0	1.34	+68.7

NOTE.—Units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds.

$30 \times 30$  pixel bins, which, given the FORS1 detector scale ( $0.2 \text{ pixel}^{-1}$ ), translates into an angular resolution of  $6''$ .

Since the instrumental polarization on the optical axis, measured with unpolarized standard stars, is smaller than 0.03% (Szeifert 2002), we are confident that the sky background polarization field ( $P_{\text{sky}}$ ,  $\chi_{\text{sky}}$ ) measured close to that area is not affected by spurious effects (the values are reported in the last two columns of Table 1). Therefore, we can easily compute  $p(x,y)$  and  $\varphi(x,y)$ , using the method previously outlined, where  $I_0 = I_{\text{sky}}$ ,  $Q_0 = I_{\text{sky}} P_{\text{sky}} \cos 2\chi_{\text{sky}}$ , and  $U_0 = I_{\text{sky}} P_{\text{sky}} \sin 2\chi_{\text{sky}}$ . The results of these calculations are presented in Figures 5, 7, and 9. With the remarkable exception of the  $B$  band, the instrumental polarization of FORS1 shows a quasi-symmetric radial pattern. For example, for the  $V$  filter, the instrumental polarization remains below 0.1% within  $1'$  of the geometrical center of the detector, while it grows to  $\sim 0.6\%$  at  $3'$ , to reach the maximum (i.e.,  $\sim 1.4\%$ ) at the corners of the field of view.

This is illustrated in the top panels of Figures 6, 8, and 10, where we have plotted the estimated instrumental polarization for each  $30 \times 30$  pixel bin as a function of its average distance  $r$  from the center. The deviation from a perfect central symmetry is distinctly shown by the dispersion of the points, which is larger than the measurement error. Particularly marked is the case of the  $B$  band, which shows a strong azimuthal dependence and thus deserves a separate discussion (see § 11.1). For the  $V$  band, there is a systematic deviation from central symmetry for a polar angle  $\alpha = \arctan(y/x)$  between  $10^\circ$  and  $80^\circ$ . This region is probably disturbed by the presence of a saturated star and a reflection caused by the HWP, visible in the input images. Excluding these points, a linear least-squares fit to the  $V$  data gives

$$p(r) = 0.012r + 0.046r^2 + 0.002r^3,$$

where  $r$  is expressed in arcminutes (Fig. 8, *solid curve*). The  $I$  band shows the smoothest behavior, and the observations are described very well by the polynomial

$$p(r) = -0.017r + 0.105r^2 - 0.006r^3.$$

The absolute rms deviations shown by the data from the best fits are of the order of 0.05%, so that in these two passbands the spurious polarization can be corrected with an accuracy that is comparable to that dictated by the photon statistics. In both

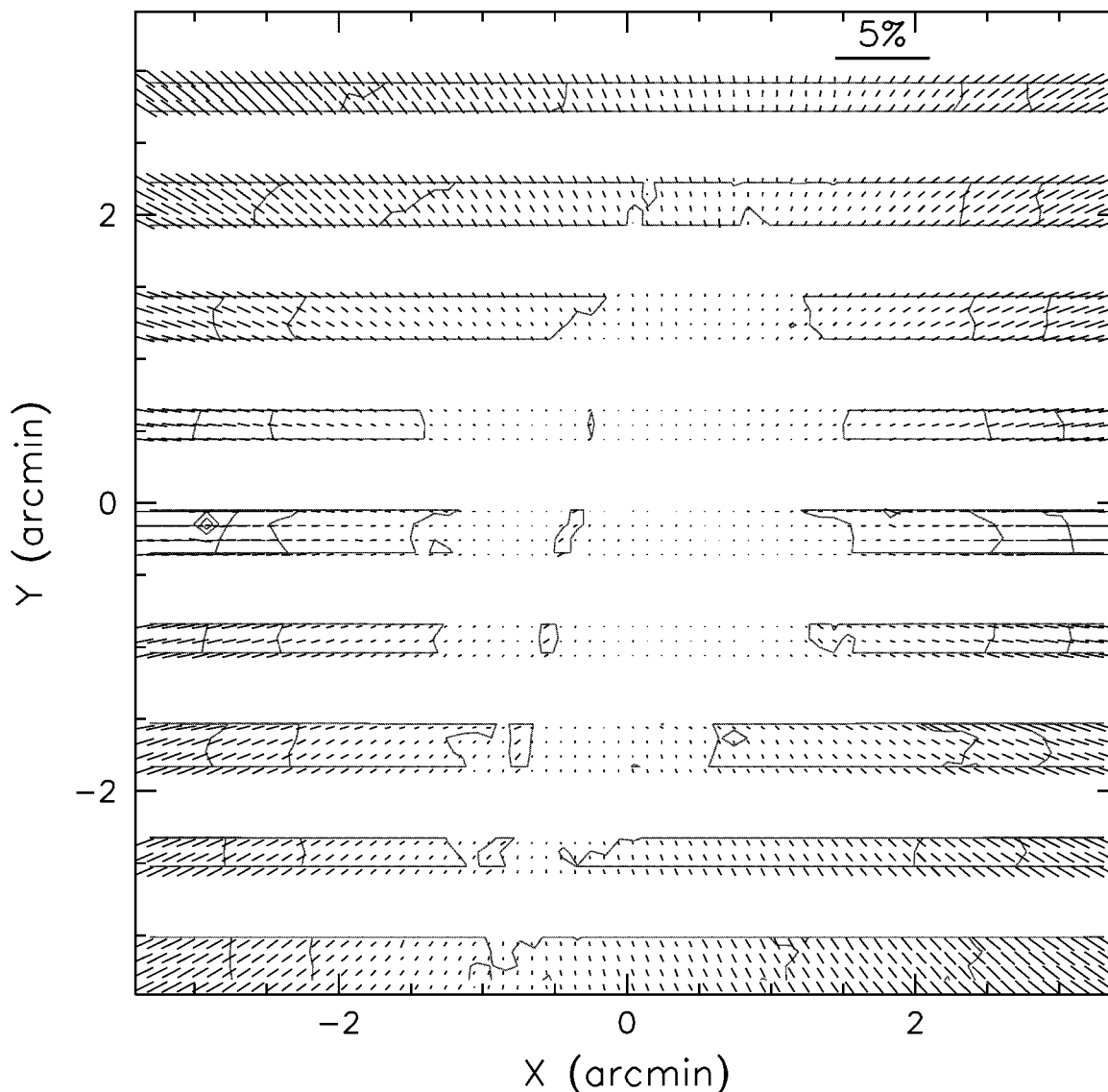


FIG. 5.—FORS1 instrumental polarization map in the  $B$  band. The contours trace 0.3%, 0.6%, and 0.9% polarization levels. Coordinates, expressed in arcminutes, refer to the geometrical center of the detector. [See the electronic edition of the *PASP* for a color version of this figure.]

cases, but especially in the  $I$  band, the pattern is remarkably radial, as shown in the bottom panels of Figures 8 and 10, where we have plotted  $\varphi$  as a function of polar angle  $\alpha$ .

In order to verify these results for the  $V$  filter, we have carried out a test observing an unpolarized standard star placed in the lower right corner of the detector. Measured polarization was  $P = 0.92\% \pm 0.04\%$  and  $\chi = -48^\circ \pm 1.4^\circ$ , while according to the previous analysis, the expected instrumental polarization in that position is  $p = 0.96\%$  and  $\varphi = -51.9^\circ$ , which are in very good agreement with each other (see also Fig. 7, *lower right corner*).

Once the instrumental polarization is mapped, one can correct for it using the approximate equation (18), which holds

when  $p$  and  $P_0$  are small, and only if the instrumental polarization is produced by a linear polarizer preceding the analyzer.

### 11.1. The Cause of Instrumental Polarization in FORS1

As we have seen, the spurious polarization detected in FORS1 in  $V$  and  $I$  passbands has a clear central symmetry, is null on the optical axis (Szeifert 2002), shows a radial pattern, and grows with the distance from the optical axis. All these facts suggest that this must be generated by the optics that precede the analyzer (i.e., within the collimator). In fact, when a light beam enters an optical interface along a nonnormal direction, the component of the transmitted beam perpendicular

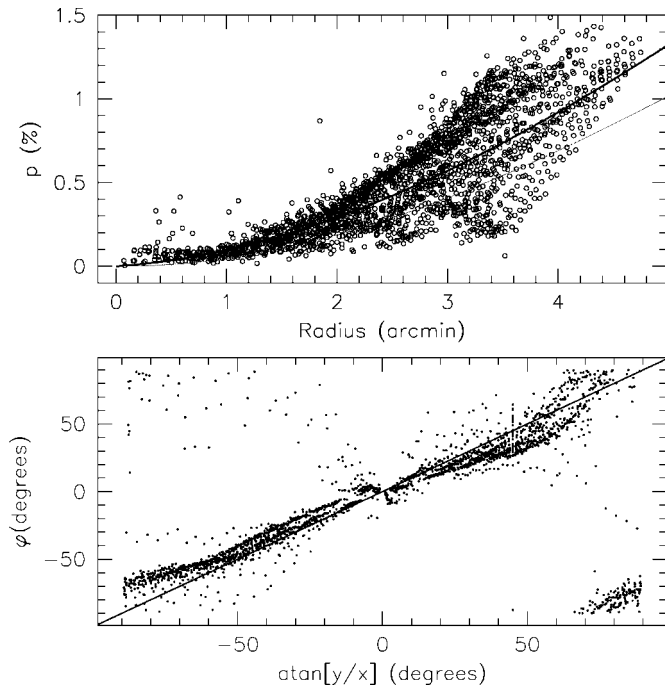


FIG. 6.—*Top*: FORS1 instrumental polarization radial profile for the *B* band. Each point is the result of a  $30 \times 30$  pixel binning in the original images. Radius, expressed in arcminutes, is computed from the geometrical center of the detector. The thick line traces a linear least-squares fit, while the thin line is the polarimetric ray-tracing prediction. *Bottom*: Instrumental polarization angle as a function of pixel polar angle. The solid line is not a fit to the data, but rather has unit slope and zero intercept.

to the plane of incidence is attenuated, according to Fresnel equations (see, e.g., Born & Wolff 1980). As a consequence, the emerging beam is linearly polarized in a direction that is parallel to the plane of incidence. If the surface is curved, as it is in the case of lenses, the incidence angle quickly increases, moving away from the optical axis, and this in turn produces an increase in the induced polarization. The effect becomes more pronounced if the lens is strongly curved; i.e., if the curvature radius is comparable to its diameter. Of course, on the optical axis, the incidence angle is null, so no polarization is produced. Therefore, at least from a qualitative point of view, the polarization induced by transmission has all the features necessary to explain the observed pattern.

The polarization induced by transmission can be easily evaluated using the appropriate expression for the corresponding Mueller matrix (see Keller 2002, eq. [4.63]). For a typical refraction index  $n = 1.5$  and an incidence angle of  $30^\circ$ , refraction through an uncoated glass would produce a polarization of about  $B = 1.7\%$  per optical surface. This polarization is usually reduced in a drastic way (i.e., down to  $0.1\%–0.2\%$ ) by antireflection (AR) coatings. Nevertheless, since the effect of multiple surfaces is roughly additive, in the presence of numerous and rather curved lenses, residual polarization can

be nonnegligible. Another important aspect is that this mechanism has no effect on circular polarization, supported by the fact that no instrumental circular polarization has been measured in FORS1 (Bagnulo et al. 2002).

We have run polarization ray-tracing simulations, including telescope mirrors, collimator lenses, and AR coatings. This kind of calculation allows one to describe in detail the optical system, taking into account partial polarization cancellation produced by symmetries within the optical beams, and the depolarizing effect of AR coatings.

The standard resolution collimator of FORS1 contains three lenses and a doublet, all treated with a single-layer  $\text{MgF}_2$  quarter-wave AR coating at 450 nm (Seifert 1994). The ray-tracing calculations (Avila 2005) show that the polarization induced by transmission is indeed not totally removed by the AR coatings. For *V* and *I* filters, a best fit to the simulated data gives a radial dependence that is very similar to the results we have derived from the experimental data. The deviation reaches maximum at the edges of the field of view, where the ray-tracing model gives a polarization that is  $\sim 0.08\%$  and  $\sim 0.05\%$  smaller than what is actually observed in *V* and *I*, respectively (see also Figs. 8 and 10, *thin curves, top panel*). Possible explanations for this small discrepancy can be attributed to imperfections in the AR coatings, and to the effects of nonorthogonal incidence on the HWP.

In principle, since single-layer AR coatings are optimized for one specific wavelength (450 nm, in the case of FORS1), the residual polarization is expected to be higher at other wavelengths. Simulations have been run in order to sample the wavelength range 400–900 nm, and these show that the expected wavelength dependency can be very well approximated by the linear relation

$$\frac{p(\lambda)}{p(550)} \approx 0.02 + 1.73 \times 10^{-3} \lambda,$$

where  $\lambda$  is expressed in nm. This relation predicts reasonably accurately the  $\sim 40\%$  relative increase we indeed see passing from *V* to the *I* passband, and it can thus be safely used to predict the effect in the *R* band.

According to the simulations, one would expect that the spurious polarization in *B* is about 25% smaller than in *V*. But as we have already mentioned, this passband shows a rather weird behavior and does not conform to the model predictions. In fact, the polarization pattern strongly deviates from central symmetry, displaying a marked azimuthal dependence (Fig. 5, *left panel*). This becomes more evident in the radial profile presented in Figure 6 (*top*): the purely radial dependence is clearly disturbed by an asymmetric field. In some directions, the polarization field grows much faster than in others, producing a great spread in the observed data, and in most of the cases, the observed polarization is larger than that predicted by the polarization ray-tracing (*thin solid curve*). The deviations from a centrally symmetric pattern reach up to 0.5% (see Fig. 11), making

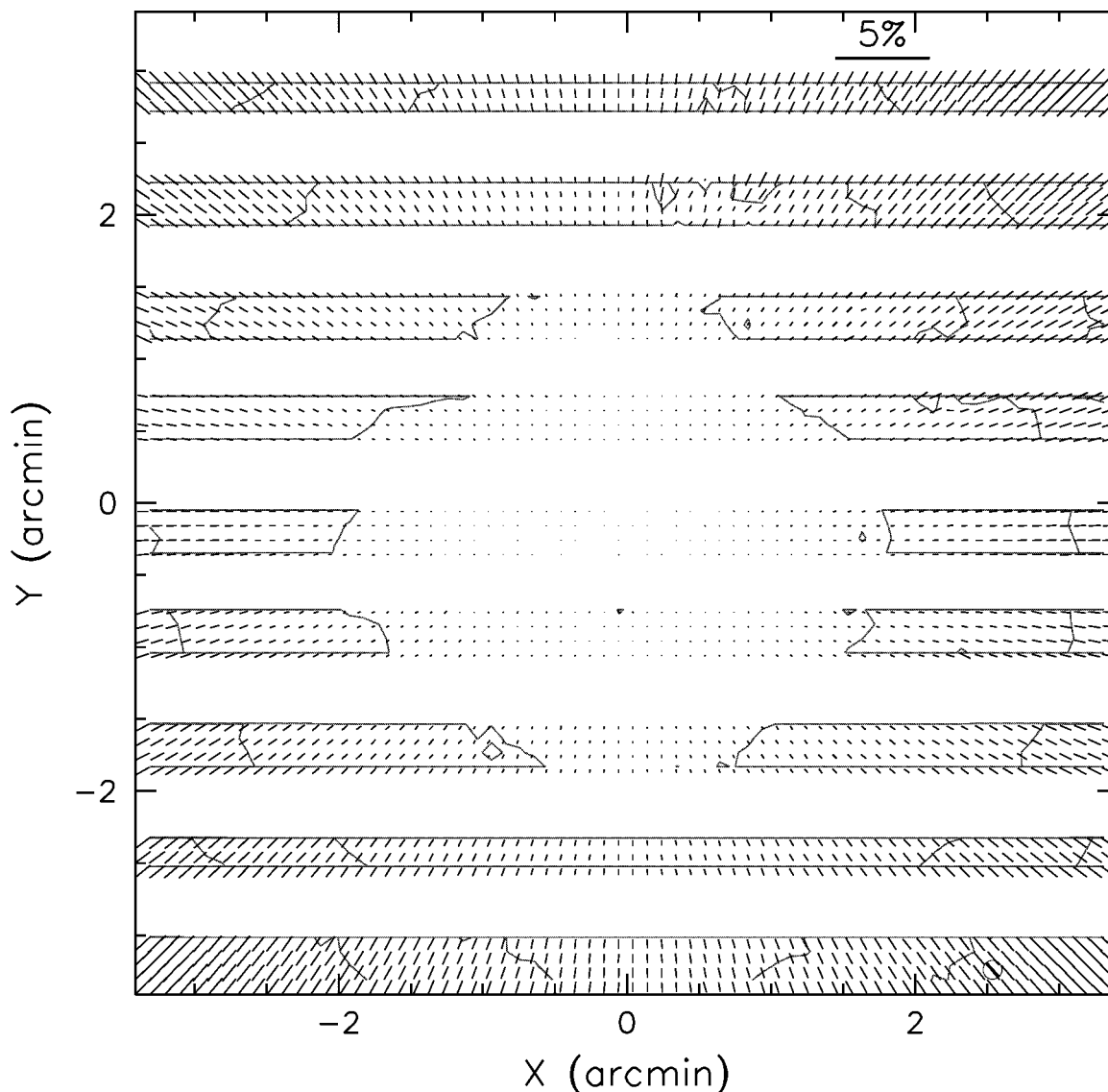


FIG. 7.—Same as Fig. 5, but for the  $V$  band. The dark segment marked by a circle in the lower right corner indicates the values obtained from an unpolarized standard star. [See the electronic edition of the *PASP* for a color version of this figure.]

the correction in the  $B$  band quite difficult, and certainly not feasible using simple smooth functions, as in the cases of  $V$  and  $I$ . Rather, a much more accurate correction can be obtained by interpolating the map of Figure 5 at the required field position. We must note, however, that a rigorous correction for this secondary effect will be possible only once its physical cause is identified and its mathematical description is formulated.

We have tried to reproduce the observed behavior by introducing defects in the system, such as a weak linear polarization from the HWP, and the presence of linear polarization in the post-analyzer optics. In both cases, the effect is completely different from what we see in the  $B$  band. Therefore, the physical reason for this phenomenon is still unclear (see also the discus-

sion in § 12). What we can say here is that the deviation from central symmetry is also present, although to a much smaller extent, in the  $V$  band. This is shown by the contours at constant polarization, which are clearly box-shaped (see Fig. 7), while in the  $I$  band they are practically circular (see Fig. 9). The conclusion is that this additional effect, whatever its origin is, becomes more severe at shorter wavelengths.

We must note that our method relies heavily on the assumptions that the instrumental polarization is null on the optical axis and that the sky background polarization is constant across the field of view. In fact, the night-sky polarization is not very well studied. The only extensive analysis we could find in the literature is that of Wolstencroft & Bandermann

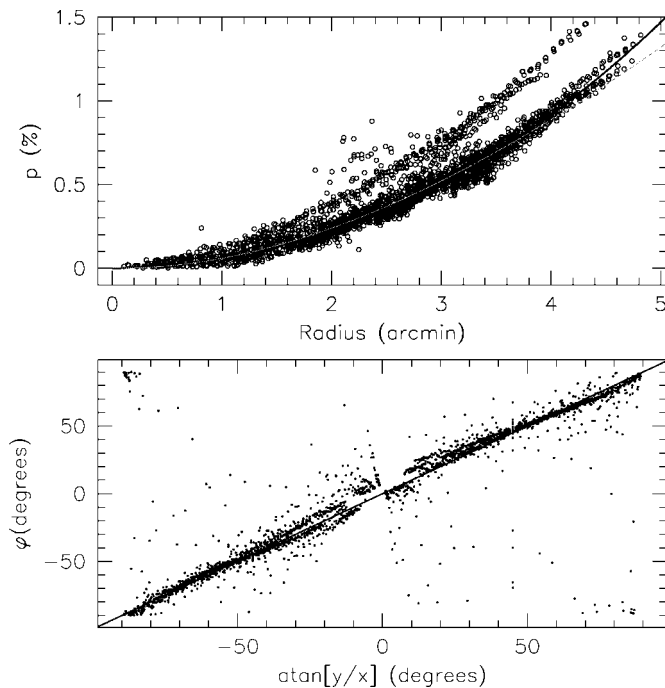


FIG. 8.—Same as Fig. 6, but for the *V* band.

(1973), who concluded that the polarization structure varies in scale from a few degrees to about  $30^\circ$ ; i.e., on scales that are much larger than the field of view of FORS1 ( $6.8 \times 6.8$ ). In order to explain the deviations we observe from a centrally symmetric pattern, one would need a variation in the sky polarization of the order of 0.5%, on a scale of a few arcminutes. Even though this seems to be quite a large gradient, in principle we cannot exclude it. Only further tests will clarify the nature of the effect we see in the *B* band.

## 12. DISCUSSION AND CONCLUSIONS

Dual-beam polarizers coupled to two-dimensional arrays provide a tool with which to perform panoramic imaging polarimetry and multiobject spectropolarimetry. For these instruments, the problem of atmospheric fluctuations is solved by obtaining simultaneous measurements of two orthogonal polarization states. Of course the use of a WP has also some drawbacks, such as the flat-fielding issue discussed in § 6. With the exception of this one feature, data reduction and analysis are totally similar to other polarimetric systems, as we have shown with both analytical and numerical approaches (§§ 3 and 4).

When the targets of study are extended and cover a large fraction of the field of view, accurate background subtraction becomes an issue whose effects we have investigated in § 5. This is particularly important when the background is not the simple sky background, but rather has a complicated structure. This is the case, for instance, for a faint supernova projected onto a galactic spiral arm.

Another problem that may reduce the performance of a dual-beam polarimeter is the imperfect behavior of the WP. In § 7 we discuss this issue and present a test to determine possible deviations from the ideal case. As an example, we have applied it to the FORS1 archive data described in § 10. Using an object-free region roughly in the center of the field of view, we have used equations (14) and (15) to compute  $t$  (see eq. [12]), which turns out to be  $t = 0.502 \pm 0.001$ , i.e., perfectly compatible with the value derived from the Fourier analysis (§ 8). As we have shown, the redundancy introduced by having  $N \geq 4$  strongly reduces this problem, even in the cases where  $t$  differs by about 10% from the ideal case ( $t = 0.5$ ). This is also the case for the presence of linear polarization in the postanalyzer optics, whose effects are practically eliminated by the redundancy (§ 9).

Finally, we have addressed the instrumental polarization issue, described its consequences, and proposed an easy test to detect any spurious effect, with rather high accuracy (§ 10). As an example, we have applied it to archival FORS1 data and have detected an instrumental polarization pattern that is roughly centrally symmetric (for *V* and *I*) and has a radial dependency. The presence of this spurious polarization affects all objects placed at distances larger than 1.5 from the optical axis, with intrinsic polarizations of a few percent or less. The problem becomes particularly severe when  $p \approx P$ , in which case the measured Stokes parameters can be very wrong. For objects filling most of the field of view, there will always be regions that are affected by this problem. Moreover, the correct sky background estimate, which is absolutely necessary to recover the intrinsic object field in the outer parts of the Galaxy, becomes impossible if the instrumental polarization is not taken care of properly. The spurious field must be removed before one is able to estimate the background contribution. Both our data and ray-tracing simulations show that the effect is wavelength dependent. In the case of FORS1, a strong deviation from central symmetry is seen in the *B* band, and we have interpreted this as a signature of an additional effect that has yet to be explained and is not included in the ray-tracing simulations that, in contrast, accurately reproduce the observed data in *V* and *I*.

One possible source of asymmetric instrumental polarization is the unrelieved stress birefringence in the optical glasses, due to thermal strain and mechanical loading (see, e.g., Theocaris & Gdoutos 1979). This phenomenon is known to introduce a retardance that can in turn change the polarization status of incoming polarized light (the effect is null if the light is unpolarized). Since the incoming radiation is certainly polarized by FORS1 in a differential way across the field of view, this would also imply that the secondary effect should be weaker where the centrally symmetric component is smaller. The fact that this is indeed the case (see Fig. 11), and also that the retardance is expected to grow faster than  $\lambda^{-1}$ , seems to suggest that this is a plausible explanation for the asymmetric component. If this is indeed the case, then it is not possible to correct the measured linear polarization just by vector-sub-

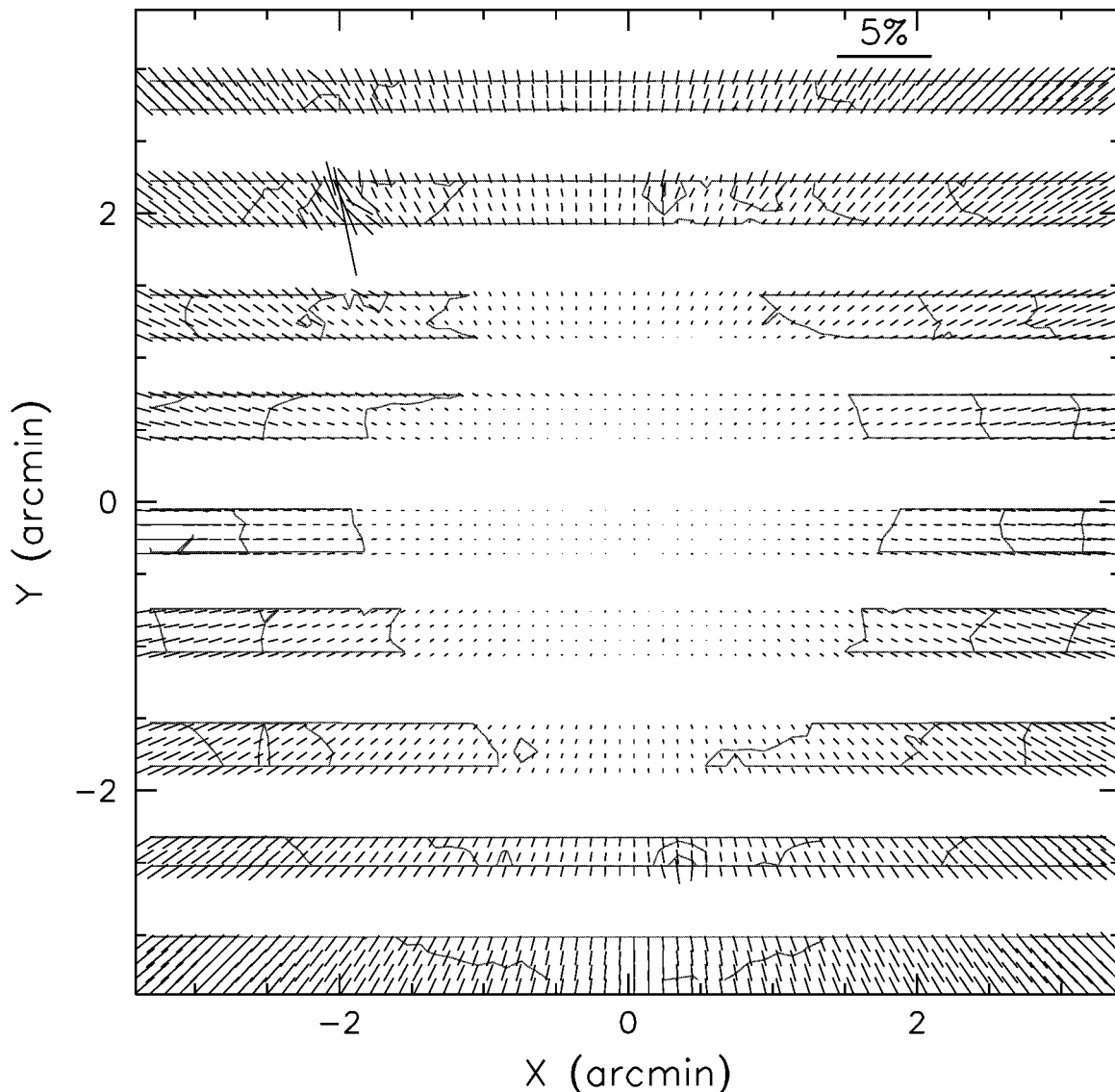


FIG. 9.—Same as Fig. 5, but for the *I* band. [See the electronic edition of the *PASP* for a color version of this figure.]

tracting the residual field (like the one shown in Fig. 11), simply because the effect of retardance depends on the polarization state of the incoming light. This requires a more sophisticated treatment that is necessarily based on the exact knowledge of the physical mechanism and its mathematical description through Mueller matrix formalism.

In general, instrumental polarization induced by transmission is most likely common to all focal reducers equipped with a polarimetric mode. While the overall pattern should be a general feature of these instruments, the exact radial dependence may change according to the optical design and the curvature of the lenses. The method we have described in this paper provides an accurate way to characterize the instrument, as well as a tool to correct for this effect.

This paper is partially based on observations made with ESO telescopes at Paranal Observatory under programs 066.A-0397, 69.C-0579, 069.D-0461, and 072.A-0025. The authors would like to thank T. Szeifert for his kind support and collaboration, J. Walsh and S. Bagnulo for interesting discussions, G. Ruprecht and W. Seifert for providing us with the optical design specifications of FORS1, R. Tommasini for introducing us to polarimetric ray-tracing, S. D'Odorico and H. Dekker for their kind support, and G. Avila for his polarimetric ray-tracing calculations. Special thanks go to C. Keller for his invaluable advice, clarifications, and help during the analysis of the instrumental polarization of FORS1. Finally, we would like to thank an anonymous referee for her/his comments and suggestions, which helped to improve the quality of the paper.

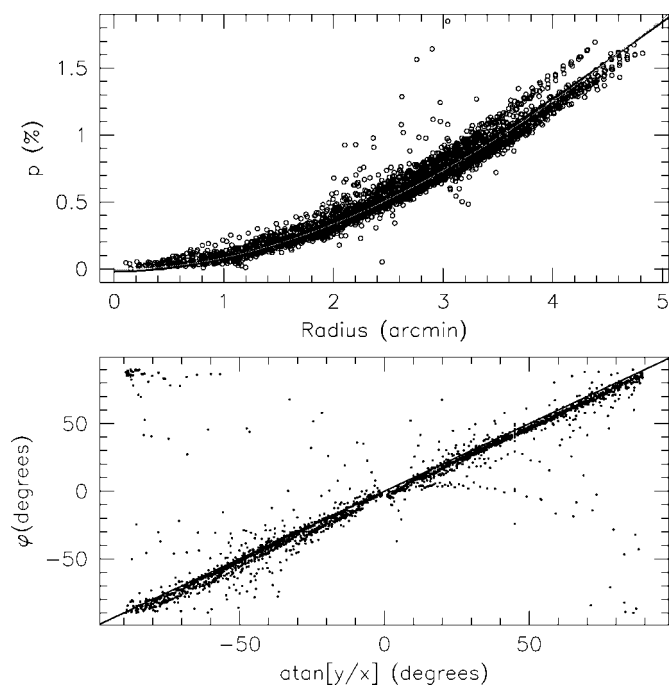


FIG. 10.—Same as Fig. 6, but for the *I* band. [See the electronic edition of the *PASP* for a color version of this figure.]

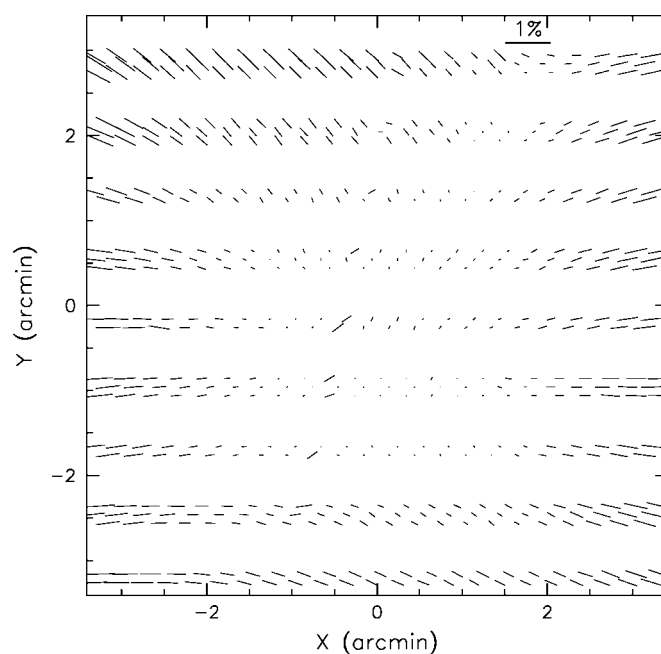


FIG. 11.—Residual field obtained subtracting the ray-tracing model from the observed polarization in the *B* band.

## REFERENCES

- Appenzeller, I. 1967, *PASP*, 79, 136
- Avila, G. 2005, ESO Tech. Note INS-05/01 (Garching: ESO)
- Bagnulo, S., Szeifert, T., Wade, G. A., Landstreet, J. D., & Mathys, G. 2002, *A&A*, 389, 191
- Born, M., & Wolf, E. 1980, *Principles of Optics* (6th ed.; New York: Pergamon Press)
- Chandrasekhar, S. 1950, *Radiative Transfer* (Oxford: Oxford University Press)
- Clarke, D., & Stewart, B. G. 1986, *Vistas Astron.*, 29, 27
- Fendt, C., Beck, R., Lesch, H., & Neninger, N. 1996, *A&A*, 308, 713
- Goldstein, D. 2003, *Polarized Light* (2nd ed.; New York: Marcel Dekker)
- Keller, C. U. 2002, in *Astrophysical Spectropolarimetry* (Cambridge: Cambridge Univ. Press), 303
- Leinert, Ch., et al. 1998, *A&AS*, 127, 1
- Leroy, J. L. 2000, *Polarization of Light and Astronomical Observation* (Singapore: Gordon and Breach)
- McLean, I. S. 1997, *Electronic Imaging in Astronomy* (Chichester: Wiley)
- McLean, I. S., Aspin, C., & Reitsema, H. 1983, *Nature*, 304, 243M
- Miller, J. S., Robinson, L. B., & Goodrich, R. W. 1987, in *Instrumentation for Ground-Based Optical Astronomy*, ed. L. B. Robinson (New York: Springer), 157
- Oliva, E. 1997, *A&AS*, 123, 589
- Patat, F. 1999, *EFOC2 User's Manual* (LSO-MAN-ESO-36100-0004; Garching: ESO)
- Pernechele, C., Giro, E., & Fantinel, D. 2003, *Proc. SPIE*, 4843, 156
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1999, *Numerical Recipes in C: the Art of Scientific Computing* (2nd ed.; Cambridge: Cambridge Univ. Press)
- Rice, S. O. 1944, *Bell System Tech. J.*, 24, 96
- Roach, F. E., & Gordon, J. L. 1973, *The Light of the Night Sky* (Dordrecht: Reidel)
- Scarrot, S. M., Warren-Smith, R. F., Pallister, W. S., Axon, D. J., & Bingham, R. G. 1983, *MNRAS*, 204, 1163
- Schmidt, G. D., Stockman, H. S., & Smith, P. S. 1992, *ApJ*, 398, L57
- Seifert, W. 1994, *FORS Final Design Report: Optics* (VLT-TRE-VIC-13110-0011; Garching: ESO)
- Serkowski, K. 1958, *Acta Astron.*, 8, 135
- Shurcliff, W. A. 1962, *Polarized Light: Production and Use* (Cambridge: Harvard Univ. Press)
- Simmons, J. F. L., & Stewart, B. G. 1985, *A&A*, 142, 100
- Sparks, W. B., & Axon, D. J. 1999, *PASP*, 111, 1298
- Szeifert, T. 2002, *FORS1+2 User's Manual* (VLT-MAN-ESO-13100-1543; Garching: ESO), Issue 2.3
- Theocaris, P. S., & Gdoutos, E. E. 1979, *Matrix Theory of Photoelasticity* (Berlin: Springer), Chap. 2
- Tinbergen, J. 1996, *Astronomical Polarimetry* (Cambridge: Cambridge Univ. Press)
- Wardle, J. F. C., & Kronberg, P. P. 1974, *ApJ*, 194, 249
- Wolstencroft, R. D., & Bandermann, L. W. 1973, *MNRAS*, 163, 229