Phase measuring algorithm for extraction of isochromatics of photoelastic fringe patterns

J. A. Quiroga and A. González-Cano

In recent years phase-measuring techniques have been applied to the problem of extracting information of photoelastic data. We present a new phase-measuring algorithm for extraction of the isochromatics of photoelastic fringe patterns. The algorithm permits the extraction of the isochromatic phase with almost no influence from the isoclinics, thus avoiding the usual problems of low-modulation areas associated with isoclinics. The isochromatic phase map obtained with this algorithm is well suited for a full separation of the stress components in a sample. The algorithm can be used with any commercial diffuse-light circular polariscope. © 1997 Optical Society of America

Key words: Photoelasticity, phase-measuring algorithms, isochromatics.

1. Introduction

Photoelasticity is a well-established technique for stress analysis, and it has a wide range of industrial and research applications.[1] Recently several methods of analyzing photoelastic fringe patterns by means of phase-measuring techniques have been presented.[2–10]

The main problem of these techniques when they are applied to photoelasticity is that the isoclinic and isochromatic fringe patterns are completely mixed. The modulation of the isoclinic phase map depends on the isochromatics, and vice versa. For that reason the unwrapping of the corresponding phase maps is difficult because of the appearance of logical inconsistencies associated with low-modulation areas.

The methods for the isoclinic calculation proposed in the literature solve these problems adequately. In this case low-modulation areas associated with the isochromatic fringes appear when monochromatic light is used. The use of a white-light source allows us to overcome these problems.[2,3]

These kinds of low-modulation problems are more difficult to solve in the isochromatic calculation.

The methods based on phase-measuring techniques that exist in the literature present several problems. The method of Brown and Sullivan[4] can be applied only when there is half an isochromatic fringe in the field of interest; therefore the stresses must be small, the photoelastic constant of the studied material must be low, or the thickness of the sample must not be great. In general, it is difficult to obtain samples with these characteristics.

The isochromatic calculation proposed by Buckberry and Towers[5] and Wernicke *et al.*[6] presents problems of sign ambiguity in the isochromatic parameter. Buckberry and Towers use three different wavelengths, taking six images for each wavelength. Wernicke *et al.* mark interactively the regions where the sign changes and they perform the corresponding correction. Asundi[7] presents the application of phase-measuring techniques to the Tardi compensation method, but the isochromatics are calculated only for a fixed value of the isoclinic parameter. The method of Carazo-Álvarez *et al.*[2] and Patterson and Wang[8] is based on an equation that presents some problems when applied to practical situations, as we show later. Also, the technique of Carazo-Álvarez *et al.* and the method of Sarma *et al.*[9] for different reasons require nonstressed reference samples so they are difficult to use with frozen stressed samples.

We present here a new technique for the isochromatic calculation that uses only one wavelength where no sign ambiguity appears. It requires a set of eight (or six) images. The use of this technique reduces the transmission to the isochromatic calculation of the errors that are produced in the isoclinic computation. The technique consists of two four-

The authors are with the Departamento de Óptica, Facultad de Ciencias Físicas, Universidad Complutense, Ciudad Universitaria s/n, 28040 Madrid, Spain. When this research was performed, J. A. Quiroga was with Steinbichler Optotechnik GmbH, Am Bauhof 4, DE-83115 Neubeuern, Germany.
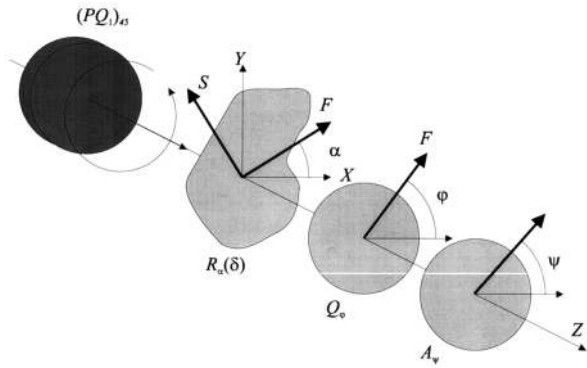
Fig. 1. Arrangement of optical elements for the general configuration of a circular polariscope. F and S stand for fast and slow axes, respectively. P, A, and Q stand for polarizer, analyzer, and quarter-wave plate, respectively, and $R_\alpha(\delta)$ for the stressed sample taken as a retardation plate of retardation δ and whose fast axis is at an angle α with the x axis. The polarizer and the first quarter-wave plate form an angle of 45° with respect to each other; therefore the light incident upon the sample is circularly polarized.

step phase-shift algorithms together with a weighted-averaging process. No reference samples are required. The algorithm can be used with any commercial diffuse-light circular polariscope.

2. Experimental Setup

For the isochromatic calculation a diffuse-light circular polariscope is used (Fig. 1). The sample—denoted by $R_\alpha(\delta)$—is placed between the two quarter-wave plates and it is taken as a retardation plate whose fast axis makes an angle α with the x axis and whose retardation is δ (δ and α have a value that is different for each point of the sample). The physical origin of the retardation δ is the stress-induced optical anisotropy of the sample. For a given wavelength, this retardation is proportional to the difference of principal stresses for each point.

The intensity output of the circular polariscope can be obtained with Jones calculus and, when circularly polarized light is incident upon the sample, is given by

$$I = 1 - \sin 2(\psi - \varphi)\cos \delta - \sin 2(\varphi - \alpha)$$
$$\times \cos 2(\psi - \varphi)\sin \delta, \qquad (1)$$

where φ and ψ are the angles that the second quarter-wave plate and the analyzer, respectively, form with the x axis.[2]

For phase-measuring techniques, the angles α and δ play the role of the phase of interferograms, which we extract with those techniques. We call isoclinics to the loci of points where α is constant and, in the same way, we call isochromatics to the loci of points with δ constant. In this way we speak of the isoclinic and the isochromatic phase maps.

The isoclinic parameter α can be determined by any of the algorithms present in the literature. We use the method of Refs. 2 and 3 to compute this parameter, and we take it as known for the isochromatic calculation.

We used a commercial diffuse-light Tiedemann polariscope with no modifications. It was computer controlled and each element could be rotated independently. The image-processing system consisted of a Sony Model AVC-D5CE CCD camera and a Data Translation Model DT-2861 image-processing board. All the components of the system were controlled by the FRAMES program on a 486-33-MHz PC. The functions for the isoclinic and isochromatic computing were written in C programming language as modules of this program.[11]

3. Isochromatic Computation

We present here a new algorithm for the extraction of the isochromatic parameter δ. This algorithm is a modification of the one in Refs. 2 and 3. Monochromatic light illumination (sodium discharge lamp) was used.

Unlike isoclinics, isochromatics depend on the wavelength of the light source and on the photoelastic constant and the thickness of the sample. In this way isochromatics do not change with a rotation of the whole polariscope.

For simplicity we establish the following notation: We denote by P, Q, and A the polarizer, the quarter-wave plate, and the analyzer, respectively; and we specify the orientation of the elements by writing as a subscript the angle with respect to the x axis formed by the transmission axis of the polarizers or by the fast axis of the retardation plates. For example, by P_{90} we indicate a polarizer whose transmission axis is perpendicular to the chosen x axis. We express the elements in the order in which the light finds them. The sample is taken, as we said before, as a retardation plate of retardation δ and whose fast axis forms an angle of α with the x axis. For brevity, the sample will not be included in the expression of the configurations of the polariscope. Therefore by $P_{90}Q_{45}Q_{(-45)}A_0$ we mean: (1) a polarizer at 90°, (2) a quarter-wave plate with fast axis at 45°, (3) the sample (not explicitly specified), (4) a quarter-wave plate with fast axis at −45°, and (5) an analyzer at 0°.

First we use four different configurations for the polariscope to obtain four intensity distributions. The configurations and their output intensities are summarized in Table 1. We see that there are two configurations for which the cos δ term can be isolated, namely the numbers 3 and 4 of Table 1, called circular dark field and circular bright field, respectively. The problem is that no configuration generates the term sin δ alone. This fact makes it impossible to perform a simple phase-stepping algorithm for the isochromatic calculation.

We denote by W the so-called wrapping operator.[12] In this way, if ϕ is a continuous phase map with a free range of values, $W(\phi) = \phi + 2k\pi$, with k integer, in such a way that $-\pi \leq W(\phi) \leq \pi$. As we have said, we consider as known the isoclinic parameter α. If a four-step phase-measuring algorithm is used to determine it, we will have $W(4\alpha)$. We can unwrap the isoclinic phase and afterwards compute cos 2α.

Table 1. Polariscope Configurations and Output Intensities of First Four Images for Isochromatic Calculation

No.	Configuration	Output Intensity
1	$P_{90}Q_{45}Q_{45}A_{(-45)}$	$I_1 = \frac{1}{2}(1 + \cos 2\alpha \sin \delta)$
2	$P_{90}Q_{45}Q_{(-45)}A_{45}$	$I_2 = \frac{1}{2}(1 - \cos 2\alpha \sin \delta)$
3	$P_{90}Q_{45}Q_{(-45)}A_0$	$I_3 = \frac{1}{2}(1 - \cos \delta)$
4	$P_{90}Q_{45}Q_{45}A_0$	$I_4 = \frac{1}{2}(1 + \cos \delta)$

Table 2. Polariscope Configurations and Output Intensities of Last Four Images for Isochromatic Calculation

No.	Configuration	Output Intensity
5	$P_{(-45)}Q_{90}Q_{90}A_0$	$I_5 = \frac{1}{2}(1 + \sin 2\alpha \sin \delta)$
6	$P_{(-45)}Q_{90}Q_0A_{90}$	$I_6 = \frac{1}{2}(1 - \sin 2\alpha \sin \delta)$
7	$P_{(-45)}Q_{90}Q_0A_{45}$	$I_7 = \frac{1}{2}(1 - \cos \delta)$
8	$P_{(-45)}Q_{90}Q_{90}A_{45}$	$I_8 = \frac{1}{2}(1 + \cos \delta)$

With this information we can calculate $\sin \delta$ and $\cos \delta$ by

$$\sin \delta_{\mathrm{I}} = \frac{1}{\cos 2\alpha} (I_1 - I_2), \qquad (2)$$

$$\cos \delta_{\mathrm{I}} = I_4 - I_3. \qquad (3)$$

From these,

$$W(\delta_{\mathrm{I}}) = \arctan\left(\frac{I_1 - I_2}{I_4 - I_3} \cdot \frac{1}{\cos 2\alpha}\right). \qquad (4)$$

With this equation we compute one wrapped estimation of δ, $W(\delta_{\mathrm{I}})$. We denote by δ_{I} one estimation of the "true" isochromatic phase δ used in all the equations of Table 1.

If the intensities and, therefore, the differences $I_1 - I_2$ and $I_4 - I_3$ were continuous, the problem of the calculation of the isochromatic phase map would be solved with Eq. (4). However, the images represented by I_i are usually digitized so their values are discrete (typically 8 bits). If $\cos 2\alpha$ is small, the difference $I_1 - I_2$ is also small; therefore, for the regions where the values of $\cos 2\alpha$ are low, we calculate δ_{I} by the quotient of two small discrete numbers. Then the error of the calculation is high, thus causing regions of low modulation in the phase map. These regions of low modulation are associated with failures and problems in the phase map, usually in the form of fringe breaks and fluctuations in the phase values, that make almost impossible the phase unwrapping of $W(\delta_{\mathrm{I}})$ (necessary for any further quantitative analysis). These regions of low modulation associated with small values of $\cos 2\alpha$ cannot be avoided because they are the consequence of the calculation method.

We can calculate mathematically the modulation function resulting from the isoclinics. The process of obtaining $W(\delta_{\mathrm{I}})$ is not exactly a four-step algorithm because of the division by $\cos 2\alpha$, but we can use the general results for phase-measuring algorithms to obtain the modulation. This is directly proportional to the square root of the sum of the squares of the denominator and the numerator of Eq. (4) (Ref. 13), that is

$$m_{\mathrm{I}} \propto [(I_2 - I_1)^2 + (I_4 - I_3)^2 \cos^2 2\alpha]^{1/2} = |\cos 2\alpha|, \qquad (5)$$

where m_{I} represents the modulation of the phase map $W(\delta_I)$. Equation (5) states that the modulation is directly proportional to the absolute value of $\cos 2\alpha$, as we discussed before.

To overcome the problem of low modulation, we obtain four new images that produce a new estimation of δ that we denote by $W(\delta_{\mathrm{II}})$, whose modulation is dependent not on $\cos 2\alpha$ but on $\sin 2\alpha$. In this way the low-modulation areas for both estimations are in opposition. We can achieve this by rotating the whole polariscope by an angle of $\beta = 45°$. After this global rotation, we repeat the four previous configurations with respect to the new global orientation. For these configurations the terms in $\cos 2\alpha$ are replaced by terms in $\sin 2\alpha$. This can be checked with Jones calculus or simply with substituting α by $\alpha - 45$ in the equations of Table 1. The output intensities of the new configurations are summarized in Table 2.

As in the previous calculation for δ_{I}, we can calculate again $\sin \delta$ and $\cos \delta$ if we know α:

$$\sin \delta_{\mathrm{II}} = \frac{1}{\sin 2\alpha} (I_5 - I_6), \qquad (6)$$

$$\cos \delta_{\mathrm{II}} = I_8 - I_7. \qquad (7)$$

And then

$$W(\delta_{\mathrm{II}}) = \arctan\left(\frac{I_5 - I_6}{I_8 - I_7} \cdot \frac{1}{\sin 2\alpha}\right). \qquad (8)$$

Therefore from Eq. (8) we can make a second estimation $W(\delta_{\mathrm{II}})$ of the phase map associated with the "true" isochromatic phase δ.

As before, we find that $W(\delta_{\mathrm{II}})$ has low-modulation regions associated with regions with low values of $\sin 2\alpha$. If we calculate the modulation of the phase map $W(\delta_{\mathrm{II}})$ from Eq. (8), we obtain

$$m_{\mathrm{II}} \propto [(I_6 - I_5)^2 + (I_8 - I_7)^2 \sin^2 2\alpha]^{1/2} = |\sin 2\alpha|. \qquad (9)$$

Again, m_{II} represents the modulation of $W(\delta_{\mathrm{II}})$ resulting from the isoclinics. Now the modulation term is proportional to $\sin 2\alpha$, as expected.

Equations (4) and (8) were previously obtained by Patterson and Wang[8] and were used by Carazo-Álvarez et al.[2] As we discussed, the two estimations for $W(\delta)$ are mathematically the same if continuous variables are used in the arctangent calculation. But because this is not the case in practical applications, we must face the problem of combining both estimations to produce a logically consistent phase map for δ.

Because the sine and the cosine are in quadrature, the regions of low values for $\cos 2\alpha$ coincide with re-

gions with high values for the sin 2α and vice versa. This implies that the regions of low modulation (high error) of $W(\delta_I)$ and $W(\delta_{II})$ are in opposition. Therefore for every point in the region of interest we have two estimations for $W(\delta)$, with different degrees of error depending on the region. We must decide how to combine these estimations to get the best possible values for $W(\delta)$. In the region with high values for any $\cos 2\alpha$ or $\sin 2\alpha$ it is easy to make a decision: we choose the estimation with the highest value of modulation. The problem arises in the zones of transition between regions of high modulation for $W(\delta_I)$ and $W(\delta_{II})$.

There are two possible ways to perform the fusion of both phase maps: the first is to work with the phase maps $W(\delta_I)$ and $W(\delta_{II})$ and the second is to work directly with the intensities I_i, $i = 1, \ldots, 8$. For a perfect acquisition process and for perfect components of the polariscope (quarter-wave plates and polarizers), both approximations must give the same result. However, some problems appear when a real measuring process is considered.

The theoretical intensities I_7 and I_8 are equal to I_3 and I_4, respectively, so the cosine terms of Eqs. (3) and (7) must be equal. But owing to nonuniformities in the plates, the isochromatics change slightly when the whole polariscope is rotated, producing in this way small differences between both cosine terms. Also electronic noise is present in the process of image acquisition. In addition, we have the above-mentioned problems of low modulation resulting from the isoclinics.

For all these reasons the estimations $W(\delta_I)$ and $W(\delta_{II})$ are good results for $W(\delta)$ in their respective regions of high modulation, but they do not give the same results in the overlapping region. This is shown schematically in Fig. 2. We can see that the fusion of estimations is a low-pass filtering process in the transition zone; the fusion algorithm must take this into account.

To perform the fusion, Franz *et al.*[3] produced two binary masks based in the thresholding of the modulation functions $|\cos 2\alpha|$ and $|\sin 2\alpha|$. Both binary masks were complementary and divided the area of interest into several regions. For each region the value for the isochromatic phase was determined by these binary masks. This procedure has serious drawbacks. First, the election of the threshold is a difficult decision. Second, the two sheets of the phase map do not fit along the borders of the regions defined by the masks, with scratches appearing in the resulting phase map, especially in the areas containing a phase jump. This poor fit produces logical inconsistencies in the form of broken fringes that make impossible the further phase-unwrapping process.

In a first stage, we tried to modify this method by using a weighted average of $W(\delta_I)$ and $W(\delta_{II})$. The obvious candidates for weighting functions are the modulation functions of both phase maps, that is, $|\cos$
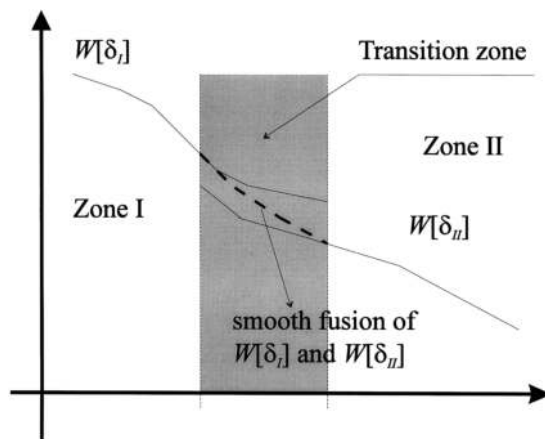

Fig. 2. Schematic representation of the behavior of the estimations I and II of the wrapped isochromatic phase, $W(\delta)$. A low-pass filtering is needed to combine smoothly $W(\delta_I)$ and $W(\delta_{II})$ in the transition zone between the high-modulation areas.

$2\alpha|$ and $|\sin 2\alpha|$, which can be expressed mathematically:

$$W(\delta) = \frac{W(\delta_I)|\cos 2\alpha|^n + W(\delta_{II})|\sin 2\alpha|^n}{|\cos 2\alpha|^n + |\sin 2\alpha|^n}, \quad (10)$$

where n is the order of the averaging.

For $n = 0$, we have just the mean value of the estimations. If $W(\delta_I)$ and $W(\delta_{II})$ were continuous functions, Eq. (10) would be good enough to assure a correct fusion of the estimations. But, by definition, $W(\delta_I)$ and $W(\delta_{II})$ are discontinuous functions with phase jumps. In the regions with no phase jumps, Eq. (10) will work well, but in the regions with phase jumps, Eq. (10) will smooth the phase jumps and even destroy them, making the correct unwrapping of $W(\delta)$ impossible. This process is equivalent to trying to filter a phase map with a moving average filter: the noise will be suppressed but so also will be the phase jumps. We could try to unwrap $W(\delta_I)$ and $W(\delta_{II})$ to produce the continuous phase maps associated with each, but, owing to the low-modulation areas that they contain, this is not possible. Equation (10) can be used only if less than half an isochromatic fringe is present in the field of view because, in this case, $W(\delta_I)$ and $W(\delta_{II})$ are continuous.

An example of this fusion procedure is shown in Fig. 3. The object is an epoxy resin Araldit disk under diametrical compression. In this case, Eq. (10) was applied with $n = 2$. The white arrows point to the broken fringes produced by the low-pass effect depicted above. In the regions where $W(\delta_I)$ and $W(\delta_{II})$ are smooth, that is, with no fringes, the fusion works well. It can be seen that these broken fringes make impossible the correct unwrapping of the isochromatic phase map.

From this discussion we see that we cannot obtain a good estimation of the isochromatic phase working with the phase maps because of the discontinuities they contain. Therefore we must look for some continuous functions related to the isochromatic phase.
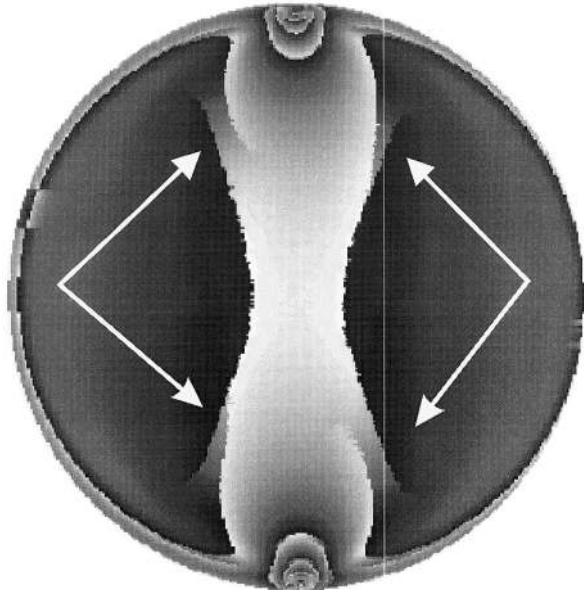
Fig. 3. Isochromatic phase map for a diametrically loaded disk, calculated with Eq. (10) with $n = 2$. The arrows point to the broken fringes produced by the low-pass effect of this equation.
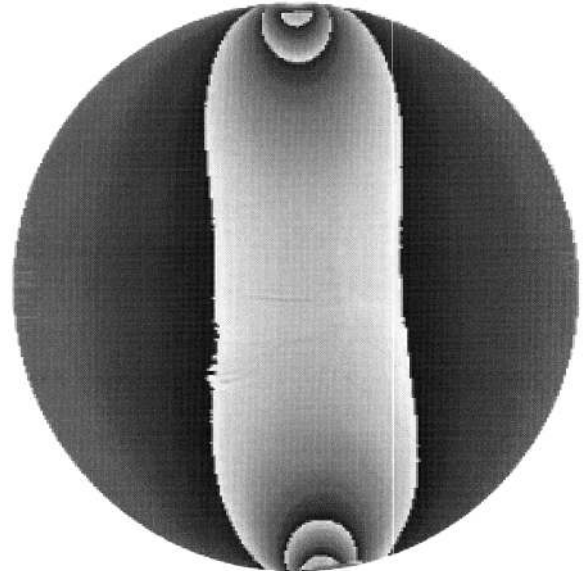

Fig. 4. Isochromatic phase map for the object of Fig. 3 calculated with Eq. (13). The resulting phase map is valid for unwrapping and further processing.

These are the intensities I_i of Tables 1 and 2. They are smooth and continuous, with no breaks, and they are also direct measurements.

In consequence, we developed a new algorithm that uses these intensities directly. We use as weighting functions $\sin 2\alpha$ and $\cos 2\alpha$ and form the following products:

$$(I_1 - I_2)\cos 2\alpha + (I_5 - I_6)\sin 2\alpha = \sin \delta, \quad (11)$$

$$\frac{1}{2}[(I_4 - I_3) + (I_8 - I_7)] = \cos \delta, \quad (12)$$

and, then,

$$W(\delta) = \arctan\left\{\frac{(I_1 - I_2)\cos 2\alpha + (I_5 - I_6)\sin 2\alpha}{\frac{1}{2}[(I_4 - I_3) + (I_8 - I_7)]}\right\}. \quad (13)$$

We see that, because $(I_4 - I_3)$ and $(I_8 - I_7)$ are theoretically equal, we could obtain $W(\delta)$ only with six images (not taking into account the last two intensities). But, owing to the above-mentioned problems of the image acquisition process, we have confirmed that, although this suppression is possible, it is not convenient because the acquisition of two more images is a fast procedure and maintaining the eight images makes the algorithm more robust and stable in the presence of any noise.

Now, because the intensities are smooth functions, the estimation works well over the whole area of interest. No low-modulation areas appear because we always have a high-modulation term in the numerator of Eq. (13). In this way the weighted average of Eq. (13) acts over the intensities, producing an isochromatic phase map well suited for unwrapping and further processing. As to its accuracy, this al-
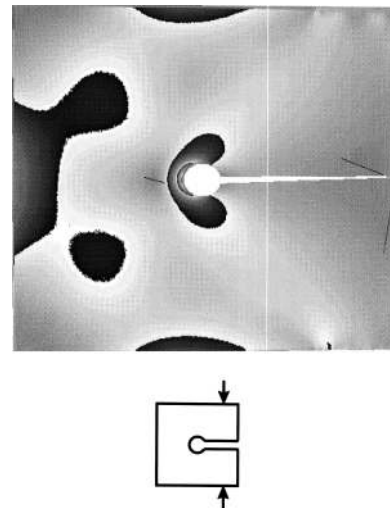

Fig. 5. Results of application of the algorithm to a rectangular plate with a hole and a cut. The diagram shows the compression force applied.

gorithm shares the main characteristics of other phase-measuring algorithms,[13] with the additional advantage that possible errors in the determination of the isoclinic parameter do not propagate into the determination of isochromatics.

An example of the application of Eq. (13) is shown in Fig. 4. It can be seen that the resulting phase map is appropriate for phase unwrapping and further processing.

The example of a disk is illustrative but perhaps too academic. We have used a more complicated sample to check the performance of the algorithm; the results are shown in Fig. 5. The sample is a rectangular plate with a hole and a cut that joins the hole with an edge of the plate. A compression force

was applied in the extremes of the plate in the points marked with arrows in the area of the diagram of Fig. 5. We can see that neither fringe breaks nor low-modulation areas resulting from isoclinics appear in the isochromatic phase map. This shows that the algorithm is well suited for realistic samples.

With Eq. (13) we can achieve a correct estimation of $W(\delta)$ without sign ambiguity and with any number of isochromatic fringes in the field of view (obviously below the Nyquist limit of the detector).

4. Conclusions

We have presented a new algorithm for isochromatic computation that works directly with eight intensity distributions, not with phase maps. With this algorithm, no sign ambiguity in the determination of the isochromatic parameter δ appears and any number of isochromatic fringes can be present in the field of view. Also, only one wavelength is needed for the isochromatic calculation. No reference sample is needed. The algorithm is implementable in any commercial polariscope (polariscopes are usually equipped with a white and/or a monochromatic light source), without the necessity for any additional elements.

References

1. M. M. Frocht, *Photoelasticity* (Wiley, New York, 1941 and 1948).
2. J. Carazo-Álvarez, S. J. Haake, and E. A. Patterson, "Completely automated photoelastic fringe analysis," Opt. Laser Eng. **21,** 133–149 (1994).
3. T. Franz, A. Maidhof, and J. Sun, "Verfahren und Vorrichtung zur Bestimmung der Isochromatenwerte in der Spannungsoptik," German patent DE-195 03 851 A1 (10 August 1995).
4. G. M. Brown and J. L. Sullivan, "The computer aided holophotoelastic method: theory and experiment," in *Proceedings of SEM conference on Hologram Interferometry and Speckle Metrology*, Baltimore, 5–8 November 1990 (Society for Experimental Mechanics, Bethel, Conn., 1990), pp. 102–109.
5. C. Buckberry and D. Towers, "Automatic analysis of isochromatic and isoclinic fringes in photoelasticity using phase-measuring techniques," Meas. Sci. Technol. **6,** 1227–1235 (1995).
6. R. Wernicke, K. P. Gründer, J. Munschau, and W. Winde, "A computer-aided photoelastic measuring system," in *Fringe '93, Proceedings of the 2nd International Workshop on Automatic of Fringe Patterns*, W. Jüptner and W. Osten, eds. (Akademie Verlag, Berlin, 1993), pp. 276–281.
7. A. Asundi, "Phase shifting in photoelasticity," Exp. Tech. **17,** 19–23 (1993).
8. E. A. Patterson and Z. F. Wang, "Towards full field automated photoelastic analysis of complex components," Strain **27,** 49–53 (1991).
9. A. V. S. S. R. Sarma, S. A. Pillai, G. Subramanian, and T. K. Varadan, "Computerized image processing for whole-field determination of isoclinics and isochromatics," Exp. Mech. **32,** 24–29 (1992).
10. Y. Otani, T. Shimada, T. Yoshizawa, and N. Umeda, "Two-dimensional birefringence measurement using the phase shifting technique," Opt. Eng. **33,** 1604–1609 (1994).
11. FRAMES-DT 2.1 *Reference Guide* (Steinbichler Optotechnik GmbH, Neubeuern, Germany, 1995).
12. D. Ghiglia, G. A. Mastin, and L. A. Romero, "Cellular-automata method for phase unwrapping," J. Opt. Soc. Am. A **4,** 267–279 (1987).
13. J. E. Greivenkamp and J. H. Bruning, "Phase-shifting interferometers," in *Optical Shop Testing*, 2nd ed., D. Malacara, ed. (Wiley, New York, 1992), pp. 522ff.