Phase retrieval errors in standard Fourier fringe analysis of digitally sampled model interferograms

Zlatko Vučić and Jadranko Gladić

We investigate the reliability of phase retrieval by use of the fringe Fourier analysis method for measuring the displacements of facets during the growth of equilibriumlike-shaped crystals. The mean phase change between two successive interferometric images contains an inherent error that emerges from the noninteger number of fringes in the image field. The magnitude of the retrieved phase error of the ideal fringe pattern is investigated as a function of spatial carrier frequency, of the initial phase setting, and of the deviation of the number of fringes from the nearest integer value. The suggested modified algorithm suppresses the error more than threefold.

OCIS codes: 100.2000, 100.2650, 100.5070.

1. Introduction

Considerably improved resolution (down to 2 nm) in measurements of facet displacement has been presented in experiments on the equilibrium crystal shape of ^{4}He single crystals at temperatures of 2 to 250 mK in which the frontal growth rate of an atomically smooth *c* facet is measured by use of a two-beam laser interferometer.[1] As a result, a new facet growth mode has been detected, primarily owing to a successful application of digital laser interferometry and of the appropriate computer-assisted fringe Fourier analysis (FFA) method.[2] The method originally proposed by Takeda *et al.*[3] is based on a fast Fourier transformation (FFT) method including Fourier spectrum shifting and filtering. The method is hereafter called carrier frequency removal and filtering (CFRF) because of its characteristic features.

Silver and copper chalcogenides turned out to be good high-temperature counterpart example systems[4] for equilibrium crystal shape as well as in investigations of the crystal growth mechanism.[5] We have grown almost centimeter-sized spherical equilibrium crystal shape single crystals of cuprous selenide[6,7] that have symmetrically arranged, well-developed (111) facets on their surfaces near 800 K. By using facets as highly reflecting objects, we aim to employ digital laser interferometry, combined with FFA,[2] including recent improvements,[8] to measure the facet's frontal nanometer displacement as a function of time, searching for the possible presence of this new growth mode.

To establish an efficient tool for fringe pattern (FP) phase retrieval,[1–3,8] by implementation of the corresponding FFA algorithm, limited strictly to the CFRF method, we first tested the method itself by applying it to an artificial, ideal fringe pattern. Using the method as described in Kostianovski *et al.*[2] revealed the errors in the retrieved phase compared with the initially set phase.

The sources of these errors had been recognized in spectrum analysis of temporal multifrequency signals sampled discretely in time.[9] As the FFT method is employed, the so-called leakage effect necessarily occurs for those frequency components that are not integer multiples of $(N\Delta t)^{-1}$, i.e., of the $(N/2)$th fraction of the Nyquist frequency. The leakage effect means that the intensity, rather than being localized in a relatively narrow maximum, leaks into its neighborhood. The FFA applied to the spatial FP seems to reveal a similar effect, consequently yielding the error in the retrieved phase value.[2] It was stated in Ref. 2 that the error is a direct consequence of the discreteness of the image-sensing device (CCD camera) combined with a noninteger number of fringes within the image field. However, the magnitude of the error and its dependence on initial parameters have not been thoroughly analyzed.[2,8] The leakage effect was discussed recently when a Fourier transform pro-

The authors are with the Institute of Physics, Bijenička cesta 46, P. O. Box 304, 10001 Zagreb, Croatia. Z. Vučić's e-mail address is vucic@ifs.hr; J. Gladić's is gladic@ifs.hr.

Received 20 April 2005; revised manuscript received 10 July 2005; accepted 22 July 2005.

0003-6935/05/326940-08$15.00/0

filometry method with improved space resolution was proposed as a shape quality-control method.[10,11]

The error that occurs in the retrieved phase owing to the leakage effect can be successfully removed[9–11] by numerical elimination of the leakage effect, i.e., by iterative solution of a set of nonlinear equations in Fourier space, assuming that the contributions to the Fourier spectrum from FP nonidealities have been efficiently suppressed. In this way the unknown parameters, the exact noninteger spatial frequency, and the corresponding real and imaginary parts of the complex phase for each interferometric image can be found.

In the process of improving the precision of the (x, y) facet displacement measurements in the z direction it becomes increasingly important to establish the highest resolution possible when the CFRF method is employed.[1,2,8] It is even more important to estimate the reliability of the phase retrieval. Our primary concern is the evaluation of the reliability, i.e., the dependence of the difference between the retrieved and the initially set phases on the relevant parameters (density of fringes, image field size N, deviation of the number of fringes from the nearest integer value, initial phase setting). We confine ourselves strictly to the CFRF method applied to digitally recorded, discretely sampled ideal, i.e., noiseless, FP.

First we define the model two-dimensional fringe pattern, representing it by its one-dimensional projection for further analysis. The CFRF method[2,8] is described afterward in steps that reveal the reason for the appearance of the leakage effect. The formalism for obtaining the Fourier spectrum of the discretely sampled model FP is outlined, describing the spectrum's prominent features. We then discuss the removal of the carrier frequency by a shift of the Fourier spectrum and subsequent filtering with an appropriate Gaussian window, stressing the role of the Gaussian window form. An appropriate fully adjustable Gaussian window form is proposed. The inverse Fourier transformation is then applied to the shifted and filtered spectrum, showing analytically that the conventional expression for the phase field extraction from the function obtained is not accurate enough. The sources, magnitude, and functional dependence of errors in the phase thus retrieved are presented. At the end of Section 4 we offer a corrected algorithm that decreases the influence of the leakage effect by more than threefold.

Further enhancement of the reliability of the phase retrieval depends, up to certain limits, on the extent to which one is ready to sacrifice the (x, y) spatial resolution.[12]

2. Fringe-Pattern Fourier Analysis

The crystal growth experiments and specific experimental conditions that have motivated the present investigation have already been described in detail.[8] The most prominent feature is a high frequency of recording interferograms imposed to keep the phase changes between successive frames from being larger than π. At low and medium growth rates the changes greater than π might be caused by mechanical vibrations of the crystal holder[6] as well as by changes in the refractive index caused by thermal fluctuations of the surrounding medium. Vibrations and thermal fluctuations also influence the spatial carrier frequency (wave vector $\mathbf{Q}$) of the FP in real interferometric experiments.

One of the refinements of the CFRF method[2] achieved by Lovrić *et al.*[8] yields the arbitrarily accurate determination of the noninteger spatial carrier frequency by use of the sampling theorem within the method of FP analysis. The CFRF method necessarily involves as accurate a knowledge as possible of the spatial carrier frequency, and so for the procedure described below it has been assumed that the refinements have been carried out, i.e., that the spatial carrier frequency has been determined by the use of either a sampling theorem or a numeric iterative method in Fourier space.[9,10]

The phase extraction or the phase retrieval procedure applied to the interferometric image is separated into five steps in what follows: They are a FFT, a shift of the Fourier spectrum, filtering of the shifted Fourier spectrum, an inverse fast Fourier transform (IFFT), and phase extraction.

A. Interferometric Image Specifications

We are usually dealing with two-dimensional FPs that have been recorded with a CCD camera, structured as a matrix of a large number of pixels each of which has a 2^l intensity resolution (l is the number of bits). The FP is a set of straight parallel fringes such that the two-dimensional field intensity $i(\mathbf{r})$ is modulated along wave vector $\mathbf{Q}$, the magnitude and direction of which are selected by fine tilting of the reference mirror of the interferometer:

$$i(\mathbf{r}) = a(\mathbf{r}) + 2b(\mathbf{r})\cos[\mathbf{Q}\mathbf{r} + \Phi(\mathbf{r})]. \quad (1)$$

Assuming the absence of any disturbances or imperfections, one may represent Eq. (1) by its one-dimensional projection by setting the x axis to be perpendicular to the interferometric fringes:

$$i(x) = a + 2b\cos(Qx + \Phi). \quad (2)$$

Moreover, in doing so we have chosen that the magnitudes of background intensity $a(\mathbf{r})$, the amplitude of modulation $b(\mathbf{r})$, wave vector $\mathbf{Q}$, and phase $\Phi(\mathbf{r})$ of the modulated signal are constant and are coordinate ($\mathbf{r}$ or x) independent. Phase Φ is a measure of a position of the fringe pattern relative to the arbitrarily chosen origin, usually the position of the image-sensing device (CCD camera). If two successive images that have the same origin yield different phase

values, the difference of their averaged phases $\Delta\bar{\Phi}$ is related to the averaged frontal facet displacement $\Delta\bar{z}$ according to the well-known relation

$$\Delta\bar{z} = \Delta\bar{\Phi}\,\frac{\lambda}{4\pi\cos(\theta/2)}, \tag{3}$$

where λ is the laser beam's wavelength and θ is the angle between the incoming and the reflecting laser beams from an atomically smooth facet.

We chose a one-dimensional formalism because of the simplicity and clarity of the final results, knowing that the extension to two dimensions, both analytically and computationally, is straightforward. One should keep in mind that the field intensity in Eq. (2) is in fact a series of N intensities separately recorded by pixels with 6–12 bit intensity resolution.[13]

It is convenient to write Eq. (2) in the complex form

$$i(n) = a + b\,\exp\left\{j2\pi\left[\frac{(k_Q + D_Q)n}{N} + \phi\right]\right\} + b\,\exp\left\{-j2\pi\left[\frac{(k_Q + D_Q)n}{N} + \phi\right]\right\}, \tag{4}$$

By inserting Eq. (4) into Eq. (5) one obtains

$$C(k) = \frac{b}{N}\exp(-j2\pi\phi) \times \sum_{n=-(N/2)}^{(N/2)-1}\exp\left[-j\,\frac{2\pi}{N}\,(k + k_Q + D_Q)n\right] + \frac{a}{N}\sum_{n=-(N/2)}^{(N/2)-1}\exp\left(-j\,\frac{2\pi}{N}\,kn\right) + \frac{b}{N}\exp(j2\pi\phi) \times \sum_{n=-(N/2)}^{(N/2)-1}\exp\left[-j\,\frac{2\pi}{N}\,(k - k_Q - D_Q)n\right]. \tag{6}$$

All three terms contain finite geometrical series of the type

$$\sum_{n=-(N/2)}^{(N/2)-1}\exp\left(-j\,\frac{2\pi}{N}\,pn\right) = \exp\left(j\,\frac{\pi}{N}\,p\right)\mathrm{sinc}(p), \qquad \mathrm{sinc}(p) = \frac{\sin(\pi p)}{\sin[(\pi/N)p]}. \tag{7}$$

Thus one obtains the final Fourier spectrum

$$C(k) = \begin{cases} \dfrac{b}{N}\exp(-j2\pi\phi)\exp\left[j\,\dfrac{\pi}{N}\,(k + k_Q + D_Q)\right]\mathrm{sinc}(k + k_Q + D_Q) & \text{first negative maximum} \\ \dfrac{a}{N}\exp\left(j\,\dfrac{\pi}{N}\,k\right)\mathrm{sinc}(k) & \text{central maximum,} \\ \dfrac{b}{N}\exp(j2\pi\phi)\exp\left[j\,\dfrac{\pi}{N}\,(k - k_Q - D_Q)\right]\mathrm{sinc}(k - k_Q - D_Q) & \text{first positive maximum} \end{cases} \tag{8}$$

$$Q = \frac{2\pi}{Nd_p}(k_Q + D_Q), \quad x = nd_p, \quad \Phi = 2\pi\phi, \tag{4a}$$

$$-\frac{N}{2} \le n \le \frac{N}{2} - 1, \quad 0 \le k_Q \le \frac{N}{2} - 1, \quad -0.5 < D_Q \le 0.5. \tag{4b}$$

Note that spatial carrier frequency Q is expressed [Eqs. (4a) and (4b)] as the sum of an integer (k_Q) and a noninteger (D_Q) component in units of discrete Fourier space.

B. Step 1: Fast Fourier Transforms

The spatial frequency spectrum of a one-dimensional interferogram is obtained by Fourier transformation of Eq. (4), as was first suggested by Takeda *et al.*[3]:

$$C(k) = \mathrm{FFT}[i(n)] = \frac{1}{N}\sum_{n=-(N/2)}^{(N/2)-1} i(n)\exp\left(-j2\pi\,\frac{k}{N}\,n\right), \quad -\frac{N}{2} \le k \le \frac{N}{2} - 1. \tag{5}$$

in which each of the maxima is separately marked, as by Kostianovski *et al.*[2] Obviously the intensity distribution in Fourier space, i.e., the Fourier spectrum $C(k)$, is entirely related to the behavior of the sinc function. When the number of fringes contained within the field of N pixels is an integer ($D_Q = 0$), no intensity leakage to nearby spatial frequencies occurs. The spectrum is everywhere equal to zero, except at three maxima, $k = \pm k_Q$ and $k = 0$. For a noninteger number of fringes ($D_Q \neq 0$), each of the terms that represents the first side maximum (both positive and negative) is spread out over all the reciprocal space (see Fig. 1). The effect is so pronounced that the two complex conjugated maxima overlap at each point of the inverse space, irrespective of their position $\pm(k_Q + D_Q)$. The strength of the overlap increases as D_Q increases. Note particularly (Fig. 1) that there is also a noticeable intensity cutoff at the Nyquist frequency as long as $D_Q \neq 0$.

C. Step 2: Fourier Spectrum Shift and Removal of Carrier Frequency

To extract the phase of the interference signal, the following step shifts the FFT spectrum to place one of

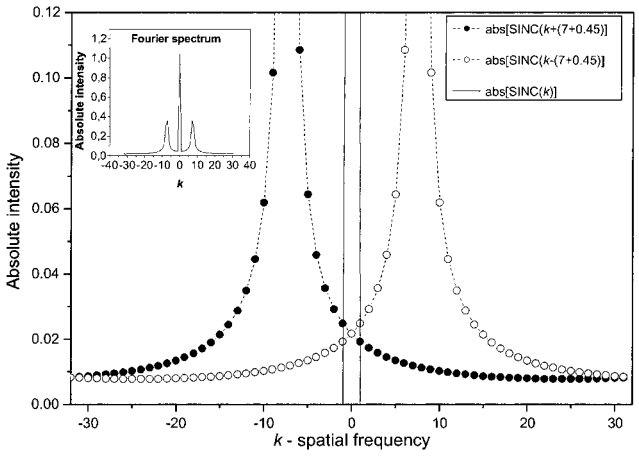

Fig. 1. Demonstration of the leakage effect, enlarged. Fourier spectra with separately depicted first-order maxima (positive ○ and negative ●) originating from the sinc $[k \pm (7 + 0.45)]$ functions, and narrow central maximum related to the sinc(k) function. Inset, sum of all three functions with a full amplitude scale. Parameters are $N = 64$, $b/a = 0.5$, $\alpha = 3$, $k_Q = 7$, and $D_Q = 0.45$.

the side maxima to $k = 0$, i.e., to wipe out the carrier frequency.

The position of the first positive maximum, which is to be shifted, is generally a noninteger number $(k_Q + D_Q)$ in Fourier space [see Eqs. (4a)]. Because the signal is recorded discretely, the shift can be done only by an integer number k_Q. Therefore the shifted positive maximum will be left in an off-center position by D_Q. By applying the shift to Eq. (8), i.e., by replacing $k \rightarrow k + k_Q$, one obtains

$$\bar{C}(k) = \frac{b}{N}\exp\left[-j2\pi\left(\phi - \frac{D_Q}{2N}\right)\right]\exp\left[j\frac{\pi}{N}(k + 2k_Q)\right] \times \operatorname{sinc}(k + 2k_Q + D_Q) + \frac{a}{N}\exp\left[j\frac{\pi}{N} \times (k + k_Q)\right]\operatorname{sinc}(k + k_Q) + \frac{b}{N}\exp\left[j2\pi \times \left(\phi - \frac{D_Q}{2N}\right)\right]\exp\left(j\frac{\pi}{N}k\right)\operatorname{sinc}(k - D_Q), \quad (9)$$

where $\bar{C}(k)$ denotes the shifted FFT of the FP.

D. Step 3: Filtering

One has to isolate the shifted first positive maximum, now located at a position near $k = 0$, from the rest of the spectrum to extract the phase. One accomplishes the isolation by applying the multiplication of $\bar{C}(k)$ by a relatively narrow rectangular function (window) centered at the origin $(k = 0)$ and apodized by some appropriate function. We use a window apodized by a Gaussian function [Gaussian window (GW)] as a filter, as suggested by Ruutu *et al.*[1] and Kostianovski *et al.*[2]:

$$\mathrm{GW} = \exp(-Bk^2). \quad (10)$$

The constant B is inversely proportional to the square of the GW's width and will be specified below. The role of the GW is to wipe out the entire spectrum except the first positive maximum (now being off the origin by D_Q) but without introducing new spectral features. In particular, it has to reduce the intensity of the shifted, former central maximum, now located at $-k_Q$, to the level of relative irrelevance. The former first negative maximum should be automatically wiped out too, because, as shifted, it is twice as distant $(-2k_Q)$ as the former central maximum. Generally, the GW is convenient as a filter because under a FFT–IFFT it is transformed into a Gaussian function alternating only its width (narrow to wide and vice versa) without introducing any new spectral features.

E. Determination of the Gaussian Window Width

The requirement for the GW is to suppress the former central maximum, at least to a level comparable with noise, i.e., $10^{-\alpha}$ $(\alpha \geq 3)$. Consequently this means that the width of the GW must be adjusted to act equally on the former central maximum at $-k_Q$, irrespective of the particular magnitude of k_Q.

It is relatively easy to determine k_Q and D_Q for each interferogram by using one of the known methods.[8,9] In real experiments the k_Q value will remain constant during a single experimental measurement (10^7 images), whereas D_Q may oscillate from image to image for as much as half of the integer.[8] In the worst case, according to definition (4b), two neighboring integer components of the spatial carrier frequencies (k_Q and $k_Q + 1$) may appear.

By suppressing the former central maximum we have removed only a part of the problem: Namely, as long as k_Q is 1 to $N/4 - 1$, the closest feature to the first maximum in Fourier space that has to be suppressed is the former central maximum. More-distant localized features are automatically wiped out. As soon as k_Q takes a value from $N/4 - 1$ to $N/2 - 1$, the closest feature that influences the further procedure results becomes a sudden jump in intensity, formerly located at the Nyquist spatial frequency (Fig. 1). That is, as long as $D_Q \neq 0$, the sinc functions extend throughout space, with the significant cutoff at $N/2 - 1$. We therefore introduce a distinction from the result of Kostianovski *et al.*[2] by requiring that the GW suppress not only the former central maximum, now at $-k_Q$, but also the intensity cutoff at the position of the former edge of reciprocal space, now at $(N/2 - 1 - k_Q)$. Thus, to prevent the influence of the intensity jump as well, we specify the width of the GW that incorporates both distances k_Q and $(N/2 - 1 - k_Q)$ through their product. The width becomes symmetric about $k_Q = N/4 - 1/2$; it is narrowest for $k_Q = 1$ and $N/2 - 2$ and widest for $k_Q = N/4 - 1$ and $N/4$ and is thus fully adjustable to the spectrum's specific spatial carrier frequency.

The final expression for the GW, including the normalization factor and the described dependence of the

width on k_Q, is given as

$$\mathrm{GW} = \exp\left(-\left\{\alpha \ln(10)\,\frac{[(N/2)-2]^2}{{k_Q}^2[(N/2)-1-k_Q]^2}\right\}k^2\right). \tag{11}$$

The criterion for defining the k_Q dependence of the GW width in Eq. (11) is related to the two requirements. The first is met through the normalization factor $(N/2-2)^2$ to ensure the required GW strength in the most unfavorable cases, $k_Q = 1$ and $k_Q = N/2-2$, where the maximal overlap with the central maximum or the intensity jump, respectively, takes place. The second is answered by the functional dependence of the GW width on the product containing k_Q. It is taken to be linear to ensure a minimal reduction of the spatial resolution.

Note that in a real experiment k_Q either is constant or oscillates between two neighboring integers and is thus known during the entire experiment, so that the GW is determined (by a single number or two numbers) as soon as we choose the tolerance level, i.e., the magnitude of $\alpha \geq 3$.

F. Steps 4 and 5: Inverse Fast Fourier Transformation and Phase Extraction

Because the GW is defined to suppress the intensity of the former central maximum as much as we choose, it will also erase the rest of the spectrum that is more distant from the GW's location, i.e., from the origin, $k = 0$. Then Eq. (9) is transformed to

$$\begin{aligned}\bar{C}_W(k) = \frac{b}{N}\Bigg\{&\exp\left[-j2\pi\left(\phi - \frac{D_Q}{2N}\right)\right]\exp\left[j\,\frac{\pi}{N}\,(k + 2k_Q)\right]\mathrm{sinc}(k + 2k_Q + D_Q)\exp(-Bk^2)\\ &+ \exp\left[j2\pi\left(\phi - \frac{D_Q}{2N}\right)\right]\exp\left(j\,\frac{\pi}{N}\,k\right)\mathrm{sinc}(k - D_Q)\exp(-Bk^2)\Bigg\} + \mathrm{occ(FCM)},\end{aligned} \tag{12}$$

where B stands for the expression in braces in Eq. (11) and the occ(FCM) contains the remainder of the former central maximum (FCM) intensity.

Equation (12) shows that, in spite of the expectation of extracting information contained in the first positive maximum only [second term in Eq. (12)], there is an inevitable, although much smaller, contribution from the former first negative maximum [first term in Eq. (12)]. At this step we differ significantly from the analysis presented by Kostianovski *et al.*[2] in which the term that belongs to the first negative maximum [the first term in Eq. (12)] is completely neglected in the phase extraction procedure. Equation (12) yields clear evidence of the leakage effect, which is in turn a direct consequence of $D_Q \neq 0$. No matter how narrow a GW filter one uses, there is always (unless $D_Q = 0$) an intensity contribution from the first negative maximum at $k = 0, \pm 1, \ldots$, which causes the intrinsic error of the retrieved phase.

The next step is the application of the IFFT to the shifted and filtered spectrum $\bar{C}_W(k)$ to obtain a complex function $I(n)$ in real space from which the phase can be extracted. The usual expression for the phase field extraction from $I(n)$presented in the following equation, which concludes the conventional CFRF method of phase retrieval, is no longer accurate enough:

$$\phi(n) = \frac{D_Q}{2N} + \frac{1}{2\pi}\arctan\left\{\frac{\mathrm{Im}[I(n)]}{\mathrm{Re}[I(n)]}\right\}, \tag{13}$$

$$\begin{aligned} I(n) &= \mathrm{IFFT}\left[\bar{C}_W(k)\right]\\ &= \frac{1}{N}\sum_{k=-(N/2)}^{(N/2)-1}\bar{C}_W(k)\exp\left(j\,\frac{2\pi}{N}\,kn\right).\end{aligned} \tag{14}$$

As can be seen from Eq. (12), in the complex plane $I(n)$ can be regarded as composed of two vectors [two terms in Eq. (12); occ(FCM) is ignored here and further on] that have considerably mutually different moduli and meet at an angle of $\pi - 2\phi$. The phase of the sum (the retrieved phase) is different from the initial phase, except when $\phi = 0, \pi/2$.

This remark should not be misidentified with the similar one given by Kostianovski *et al.*[2] in which the influence of the first negative shifted maximum is completely ignored, while the residuum of the shifted central maximum, after a GW application, is taken into account.

As can be seen from Eq. (13), there is a linear, spatial-carrier-frequency-independent phase correction $[D_Q/(2N)]$ discussed by Lovrić *et al.*[8] For given D_Q this correction is smaller the higher the image field dimension N is.

3. Results and Discussion

The main disadvantage of the FFA-with-CFRF method is the x–y resolution problem: Namely, in each retrieved x–y phase field, besides the real initial modulation there is an artificial one, too. The artificial one emerges from the incomplete shift of the first positive maximum toward the origin of Fourier space being D_Q off center. The amplitude of the artificial modulation is dependent on the chosen GW strength–width just as it is for the real component. Thus one cannot seriously decrease its relative influence on the retrieved phase field by playing with the filter.

If, for example, D_Q were to stay the same in two successive images, the phase field difference would yield only a real change of the facet's position including details that are limited by normal x–y resolution, which in turn is filter-width dependent.[12] Unfortunately, during the long sequence of measurements D_Q oscillates from image to image with an average amplitude that can be any number from several percent to 50% of an integer, depending on the experimental conditions, i.e., the quality of the interferometric images or signal-to-noise ratio in Fourier space. There-

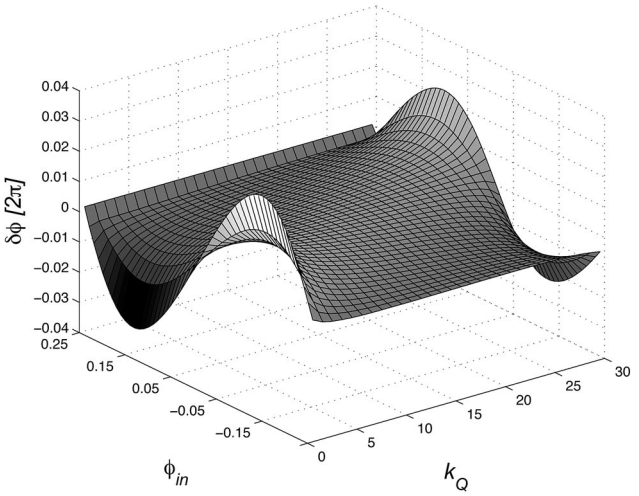

Fig. 2. Difference between the phase retrieved from Eq. (13) and the initially set phase (the retrieved phase error), in units of 2π, as a function of spatial carrier frequency (k_Q) and initially set phase (ϕ_{in}). Parameters are $N = 64$, $b/a = 0.5$, $\alpha = 3$, $k_Q = 7$, and $D_Q = 0.5$.

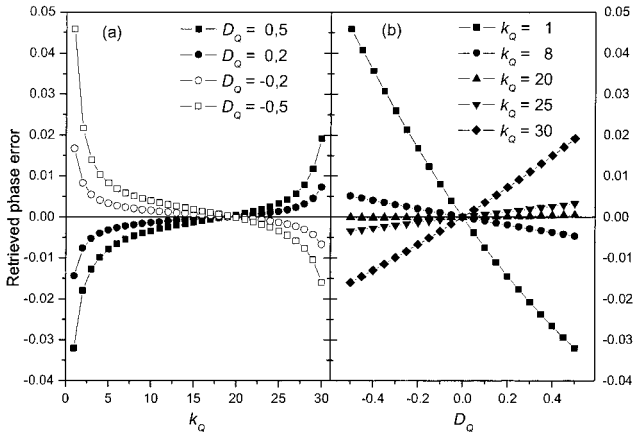

Fig. 3. Difference between the phase retrieved from Eq. (13) and the initially set phase (the retrieved phase error), in units of 2π, as a function of (a) spatial carrier frequency (k_Q) for several values D_Q within the $(-0.5, 0.5)$ interval and (b) the noninteger component of spatial carrier frequency (D_Q) for several k_Q values. Parameters are $N = 64$, $b/a = 0.5$, $\alpha = 3$, and $\phi_{in} = 0.15$.

fore, for this method, the x–y resolution expectation should always be related to the D_Q oscillation amplitude during the parts of the experiment in which the experimental conditions stay unchanged. The magnitude of the D_Q oscillation amplitude has a direct effect on the x–y resolution of the phase field difference. The higher D_Q is, the coarser the x–y resolution is.

The frontal facet z displacement, however, extracted from the difference between two successive averaged phase fields is much less influenced.

In this paper we deal with an ideal FP in which we have specified the initial phase field ϕ_{in} as independent of n (along the image field), to study consequences of changes of D_Q in the most unfavorable circumstances. It appears that the single retrieved phase field $\phi(n)$ is artificially modulated, so, to find the retrieved phase error, phase field averaging is obligatory. Therefore, sacrificing the x–y resolution, we calculated the mean value $\bar{\phi}$ from Eq. (13) by averaging $\phi(n)$ over the entire image field, N.

Before the averaging, one of the appropriate simple unwrapping methods[14] is utilized because, depending on the initial phase value and GW strength, phase wrapping may appear in the $(-0.25,\ 0.25)$ or, in terms of Φ, in the $(-\pi/2,\ \pi/2)$ interval.

When the realistic phase field or shape reconstruction is important, either the method of local phase retrieval[10] should be used or high-quality interferometric images should be provided.

The error in the retrieved phase, averaged over the phase field, relative to the initially set phase $[\delta\bar{\phi}(k_Q, \phi_{in})]$ obtained by evaluation of Eqs. (12) and (13), assuming the maximum realistic departure of Q from the integer value (i.e., that $D_Q = +0.5$), $N = 64$, $b/a = 1/2$, and $\alpha = 3$, as a function of the initially set phase value ϕ_{in} and of an integer component of the spatial carrier frequency, k_Q, is shown in Fig. 2. The maximum error is found when $k_Q = 1$, where $\phi_{in} = \pm 0.15$ has the value $0.032 \times 2\pi$ (0.064π), which is equivalent to 10 nm in the frontal facet displacement [Eq. (3) when $\lambda = 632.8$ nm is used for the He–Ne laser].

The dependence of the retrieved phase error on the noninteger D_Q value is shown in Fig. 3 for the worst case of an initially set phase value for which the error reaches the maximum value ($\phi_{in} = 0.15$). It can be seen that at any k_Q the error is decreasing as D_Q tends to zero.

Note that the phase error surface $\delta\bar{\phi}(k_Q, \phi_{in})$ (Fig. 2) is symmetric with respect to the axis at $\phi_{in} = 0$. The area characterized by the smallest errors, which has a value of not more than $0.0045 \times 2\pi$ ($\sim$1.5 nm), is located in the relatively wide middle part of the spatial frequency field ($10 < k_Q < 22$). Also, note that what really matters in the displacement measurements is the change in D_Q between two successive images, which, in low-quality interferometric images, may take a value of as much as half of an integer.

4. Corrected Phase Extraction

The right-hand side of Eq. (13) as the conventional expression for phase retrieval hides the dominant part of the phase retrieval error. This becomes evident as soon as the structure of $I(n)$ [Eq. (14)] is closely examined. By inserting Eq. (12) into (14) and omitting the third term, occ(FCM), we obtain

$$
\begin{aligned}
I(n) = \frac{b}{N}\Bigg\{ & \exp\left[-j2\pi\left(\phi - \frac{D_Q}{2N}\right)\right] \\
& \times \sum_{k=-(N/2)}^{(N/2)-1} \exp\left(j\,\frac{2\pi}{N}\,nk\right)\exp\left[j\,\frac{\pi}{N}\,(k+2k_Q)\right] \\
& \times \mathrm{sinc}(k + 2k_Q + D_Q)\mathrm{GW}(k) \\
& + \exp\left[j2\pi\left(\phi - \frac{D_Q}{2N}\right)\right] \sum_{k=-(N/2)}^{(N/2)-D} \exp\left(j\,\frac{2\pi}{N}\,\mathrm{nk}\right) \\
& \times \exp\left(j\,\frac{\pi}{N}\,k\right)\mathrm{sinc}(k - D_Q)\mathrm{GW}(k)\Bigg\}. \qquad (15)
\end{aligned}
$$

Sums in both terms of Eq. (15) are just the inverse Fourier transformations of two mutually similar complex functions $O(k)$ and $K(k)$ sited at different positions in Fourier space: one at D_Q near the origin and the other at $-(2k_Q + D_Q)$. We shall denote these sums by the functions IFO(n) and IFK(n), respectively:

$$I(n) = \frac{b}{N}\left(\exp\left\{-j2\pi\left[\phi(n) - \frac{D_Q}{2N}\right]\right\}\mathrm{IFO}(n) + \exp\left\{j2\pi\left[\phi(n) - \frac{D_Q}{2N}\right]\right\}\mathrm{IFK}(n)\right), \quad (16)$$

$$O(k) = \exp\left(j\frac{\pi}{N}k\right)\mathrm{sinc}(k - D_Q)\exp(-Bk^2),$$

$$K(k) = \exp\left[j\frac{\pi}{N}(k + 2k_Q)\right]\mathrm{sinc}(k + 2k_Q + D_Q) \times \exp(-Bk^2). \quad (17)$$

Functions $O(k)$ and $K(k)$ are products of three functions–matrices that are dependent on image dimension N, spatial carrier frequency Q ($Q = k_Q + D_Q$), and the width of the GW [Eq. (11)], which in turn depends on the chosen tolerance α. The accuracy of the retrieved phase depends on the magnitude of D_Q. Irrespective of the D_Q value, for the same GW strength the phase can be determined more accurately if Eq. (16) is disentangled or solved as a linear system of equations for the sine and the cosine of the corrected phase $[\phi(n) - D_Q/(2N)]$ by use of real and imaginary parts of particular functions in Eq. (16). The final phase retrieval now involves a more complicated, but more accurate, expression:

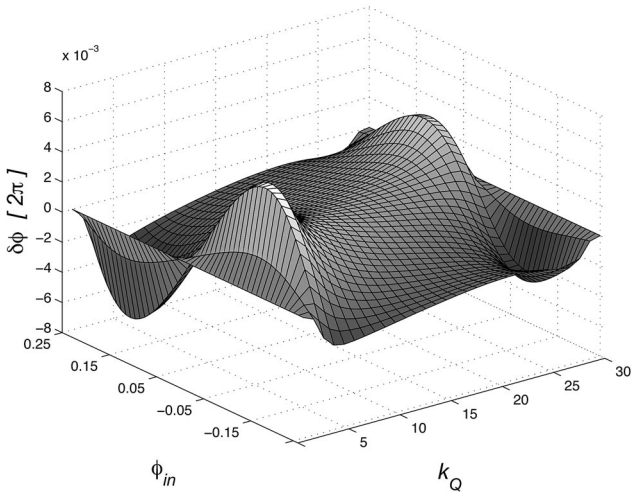

Fig. 4. Difference between the phase retrieved from Eq. (18) and the initially set phase (suppressed error), in units of 2π, as a function of spatial carrier frequency (k_Q) and the initially set phase (ϕ_{in}) value. Parameters are $N = 64$, $b/a = 0.5$, $\alpha = 3$, $k_Q = 7$, and $D_Q = 0.5$.

the spatial carrier frequency and on the initial phase value [Eq. (18)] for maximum $D_Q = +0.5$ is shown in Fig. 4. Note that N is taken to be 64 (the number usually used in our experiments) and that the ratio b/a [Eq. (4)] is taken to be 1/2. Also note that the computation is done with a GW width [Eq. (11)] with $\alpha = 3$. One can further suppress the error by increasing α.

The extreme possibility is to take a window so narrow that the spectral intensity in [Eq. (12)] as well as its IFFT in [Eq. (15)] retains only the $k = 0$ term. Because the k dependence is lost, the n dependence of the phase field $I(n)$ makes no sense. For this case Eq. (18) is transformed into

$$\tan 2\pi\left[\phi(n) - \frac{D_Q}{2N}\right] = \frac{\{\mathrm{Re}[\mathrm{IFO}(n)] + \mathrm{Re}[\mathrm{IFK}(n)]\}\mathrm{Im}[I(n)] - \{\mathrm{Im}[\mathrm{IFO}(n)] + \mathrm{Im}[\mathrm{IFK}(n)]\}\mathrm{Re}[I(n)]}{\{\mathrm{Re}[\mathrm{IFO}(n)] - \mathrm{Re}[\mathrm{IFK}(n)]\}\mathrm{Re}[I(n)] + \{\mathrm{Im}[\mathrm{IFO}(n)] - \mathrm{Im}[\mathrm{IFK}(n)]\}\mathrm{Im}[I(n)]}. \quad (18)$$

It is obvious that, instead of a single inverse Fourier transform [Eq. (13)], one should perform the additional two transforms (IFO and IFK), which are defined as soon as spatial carrier frequency Q is determined. Note that in this one-dimensional case $O(k)$ and $K(k)$ are one-dimensional matrices.

Using this new, corrected expression for the retrieved phase, we obtain results in which the error is more than three times smaller than when Eq. (13) is used. The improved dependence of the error on

$$\tan 2\pi\left[\phi - \frac{D_Q}{2N}\right] = \frac{\frac{\mathrm{Im}[I_w(0)]}{\mathrm{Re}[I_w(0)]}\left[1 + 2\frac{\tan(\pi D_Q/N)}{\tan(2\pi k_Q/N)}\right] - \tan\frac{\pi D_Q}{N}}{1 - \frac{\mathrm{Im}[I_w(0)]}{\mathrm{Re}[I_w(0)]}\tan\frac{\pi D_Q}{N}}. \quad (19)$$

Note that, within the standard calculation algo-

rithm already used in Eqs. (12) and (18), $I_w(0)$ stands for the matrix obtained in Eq. (15) but with a modified GW filter that is independent of the spatial carrier frequency, GW = $\exp[-(\alpha \geq 3) \times \ln(10) \times k^2]$. Keeping the values of the other parameters the same as in Figs. 3 and 4, we obtain that the magnitude of the retrieved phase error has dropped below $10^{-5}\pi$. Equation (19) is not of practical use, but rather it serves to reveal the terms that dominate the dependence on the parameters: k_Q, D_Q, and N. It is obvious that, when D_Q tends to zero, Eq. (19) becomes equal to the conventional Eq. (13).

5. Conclusions

In this paper we have investigated the effects of use of a specific[2,3,8] interferometric fringe Fourier analysis on the magnitude of the retrieved phase relative to the phase initially set for a perfect sinusoidal interferometric fringe pattern. A FFA, relying on a fast Fourier transformation technique, involves the removal of the integer part of the spatial carrier frequency by shifting the positive first-order maximum of the Fourier spectrum close to the origin and subsequent filtering to isolate and extract the phase field of the FP. The retrieved phase as a field average yields the information about displacement (the state) of the reflecting surface under observation by use of digital laser interferometry.

The filter that we used, the rectangular window apodized by a Gaussian function, has been successfully designed to be of universal use, fulfilling its task independently of specific parameters of interferometric images (other than density of fringes), unchanged during the entire experiment and restricted only by an imposed accuracy tolerance, i.e., by the level of suppression of the Fourier spectrum outside the first-order maximum.

We have shown that the retrieved phase error inevitably emerges as a result of the leakage effect, which in turn occurs as a rule, to a varying extent, in the Fourier spectrum of any interferometric images. The leakage effect originates from discrete sampling of the interference fringes (image-sensing elements, pixels in a CCD camera) and a noninteger number of fringes in the field of view containing N pixels. The algorithm usually used for the phase retrieval yields an error that ranges from 1.5 to 10 nm, depending on the density of interferometric fringes, from medium ($N/4$) to low (and high), respectively.

Based on a detailed inspection of the Fourier spectrum equation, we are able to offer an improved last part of the usual algorithm based on the corresponding disentanglement of real and imaginary parts of the inverse Fourier transform equation to extract the phase that is closer to the value set initially. We have shown that by this adapted algorithm the retrieved phase error may be reduced more than threefold, thus ranging from 0.4 to 2 nm, for medium to low density fringes, respectively. Note that the adapted algorithm is the last step in the procedure in which the FP is first refined, in the sense of the recommendations by Lovrić *et al.*[8] or Berryman *et al.*,[15] and then Fourier transformed and consequently shifted and filtered to extract the part of the Fourier spectra that carries information about the phase.

We acknowledge gratefully the financial support of the Ministry of Science, Education and Sport of the Republic of Croatia. The authors thank D. Lovrić and N. Demoli for their useful comments and suggestions, and student V. Cipan for his technical help.

References

1. J. P. Ruutu, P. J. Hakonen, A. V. Babkin, A. Zu. Parshin, and G. Tvalashvili, "Growth of ^{4}He crystals at mK temperatures," J. Low Temp. Phys. **112,** 117–164 (1998).
2. S. Kostianovski, S. G. Lipson, and E. N. Ribak, "Interference microscopy and Fourier fringe analysis applied to measuring spatial refractive–index distribution," Appl. Opt. **32,** 4744–4755 (1993).
3. M. Takeda, H. Ina, and S. Kobayashi, "Fourier–transform method of fringe—pattern analysis for computer–based topography and interferometry," J. Opt. Soc. Am. **72,** 156–160 (1982).
4. T. Ohachi and I. Taniguchi, "Roughening transition for the ionic–electronic mixed superionic conductor α-Ag_2S," J. Cryst. Growth **65,** 84–88 (1983).
5. Z. Vučić and J. Gladić, "Shape relaxation during equilibrium-like growth of spherical cuprous selenide single crystals," Fiz. A (Zagreb) **9**(1), 9–26 (2000).
6. Z. Vučić and J. Gladić, "Growth rate of equilibriumlike-shaped single crystals of superionic conductor cuprous selenide," J. Cryst. Growth **205,** 136–152 (1999).
7. J. Gladić, Z. Vučić, and D. Lovrić, "Critical behaviour of the curved region near the 111-facet edge of equilibrium shape cuprous selenide large single crystals," J. Cryst. Growth **242,** 517–532 (2002).
8. D. Lovrić, Z. Vučić, J. Gladić, N. Demoli, S. Mitrović, and M. Milas, "Refined Fourier-transform method of analysis of full two-dimensional digitized interferograms," Appl. Opt. **42,** 1477–1484 (2003).
9. H. Renders, J. Schoukens, and G. Vilain, "High-accuracy spectrum analysis of sampled discrete frequency signals by analytical leakage compensation," IEEE Trans. Instrum. Meas. **33,** 287–292 (1984).
10. S. Vanlanduit, J. Vanherzeele, P. Guillaume, B. Cauberghe, and P. Verboven, "Fourier fringe processing by use of an interpolated Fourier-transform technique," Appl. Opt. **43,** 5206–5213 (2004).
11. J. Vanherzeele, P. Guillaume, and S. Vanlanduit, "Fourier fringe processing using a regressive Fourier-transform technique," Opt. Lasers Eng. **43,** 645–658 (2005).
12. R. Vander, S. G. Lipson, and I. Leizerson, "Fourier fringe analysis with improved spatial resolution," Appl. Opt. **42,** 6830–6837 (2003).
13. O. A. Skydan, F. Lilley, M. J. Lalor, and D. R. Barton, "Quantization error of CCD cameras and their influence on phase calculation in fringe pattern analysis," Appl. Opt. **42,** 5302–5307 (2003).
14. D. W. Robinson, "Phase unwrapping methods," in *Interferogram Analysis: Digital Fringe Pattern Measurement Techniques*, D. W. Robinson and G. T. Reid, eds. (Institute of Physics Publishing, Bristol, 1993), Chap. 6, pp. 194–229.
15. F. Berryman, P. Pynsent, and J. Cubillo, "The effect of windowing in Fourier transform profilometry applied to noisy images," Opt. Lasers Eng. **41,** 815–825 (2004).