Chapter 1
A General Mathematical Model to Retrieve Displacement Information from Fringe Patterns

C.A. Sciammarella and L. Lamberti

Abstract The extraction of the displacement field and its derivatives from fringe patterns entails the following steps: (1) information inscription; (2) data recovery; (3) data processing; (4) data analysis. Phase information is a powerful representation of the information contained in a signal. In a previous work, the above mentioned steps were formulated and discussed for a 1D signal, indicating that the extension to 2-D was a non trivial process. Proceeding along the same line of thought when one moves from the one dimension to two dimensions it is necessary to consider a 3D abstract space to generate the additional dimension that can handle the analysis of 2D signals and simultaneously extend the Hilbert transform to 2D. In this study the basic theory developed in the preceding reference is further elaborated to produce a version of the monogenic function yielding the necessary answers to the previously described processes. The monogenic signal, a 3D vector in a Cartesian complex space, is graphically represented by a Poincare sphere which provides a generalization of the Hilbert transform to a 2D version of what is called the generalized Hilbert transform or Riesz transform. These theoretical derivations are supported by the actual application of the theory and corresponding algorithms to 2D fringe patterns and by comparing the obtained results with known results.

Keywords 2D signals • Displacement and strain determination • Generalized Hilbert (Riesz) transform • Poincare sphere

1.1 Introduction

In [1], the present authors developed a one dimensional mathematical model of fringe patterns analysis based on the general Theory of Signal Analysis. This paper now deals with a generalization of the one dimension model derivations to 2-D. The extension to a higher dimension requires the review of some basic concepts of image signal analysis. To simplify the derivations we will consider the signal analysis on plane surfaces. The extension to general surfaces in the space requires further developments that cannot be covered on a single paper.

The information to be decoded is recorded as level of gray in a 2D sensor through a device composed of optical and electronic circuits commanded by software, a measure of the light intensity of the imaged field. At this point the details of the process of data generation will set aside and the paper will concentrate in the process of information extraction. The recorded levels of gray must be converted into data that provide displacement fields and the displacement derivatives in the case of deformed bodies or geometrical parameters and their derivatives. In [1], it is shown that data conversion in one dimension requires the description of gray levels in terms of 2D complex functions (analytical functions) that lead to the introduction of the concept of phasor:

$$\overrightarrow{\mathbf{I_{sp}}}(x) = I_{sp}(x)e^{2\pi j \phi(x)} \tag{1.1}$$

The symbol $\Rightarrow$ indicates a vector in the complex plane. A phasor in the complex plane is characterized by two separate pieces of information: amplitude related to the light intensity at the considered point and a phase representing the optical path followed by the recorded wave front from a selected reference point where the phase is assumed to be zero. The classical definition of phase in optics is,

---

C.A. Sciammarella
Department of Mechanical, Materials and Aerospace Engineering, Illinois Institute of Technology, 10 SW 32nd St., Chicago, IL 60616, USA

L. Lamberti (✉)
Dipartimento Meccanica, Matematica e Management, Politecnico di Bari, Viale Japigia 182, Bari 70126, Italy
e-mail: luciano.lamberti@poliba.it

$$\boldsymbol{\phi}(x) = \frac{2\pi\boldsymbol{\delta}(x)}{p} \tag{1.2}$$

where $\boldsymbol{\delta}(x)$ is the optical path and p is pitch of the sinusoidal function, unit of measure utilized to evaluate a path length and convert distances into an angle.

The optical path length of the light arriving at an image is given by,

$$\boldsymbol{\delta}(x) = \int_0^{S_0} n(x)\, dx \tag{1.3}$$

where n(x) is the index of refraction of the medium along the path followed by the light from a certain reference point to another point following a trajectory.

A question than can arise is: Why to begin with the review of the phase concept? The answer to this question is found in [2]: the phase concept is a fundamental tool to develop a consistent theory of image analysis. There is another important aspect to the concept of phase, the definition of local phase implicit in Eq. (1.1) and the more general concept of global phase expressed by Eq. (1.3). The phase concept is associated with the notion of vector. When one introduces the definition of local phase or phase at a point for a 1D signal, one introduces an additional dimension to the mathematical model required to associate one dimensional functions with the phase concept. This additional dimension corresponds not to the actual space but to the complex plane. It is a fundamental concept in the Gabor's analytic signal theory [3], basic starting point of many developments in Signal Analysis and in Optics. A complementary development to the analytic signal theory in one dimension is the Hilbert transform [4] that converts cosines into sines and is a unitary transform that changes the phase of the signal of π/2, leaving the signal amplitude unchanged. The Hilbert transform takes the original signal, a level of gray or intensity in some scale, and associates the gray level with an analytical function:

$$\mathbf{I_{sp}}(x) = I_p(x) + I_q(x)\mathbf{j} \tag{1.4}$$

where the symbol $\mathbf{j}$ is the imaginary versor, $I_p(x)$ is the recorded signal (in-phase signal) and $I_q(x)$ is the in-quadrature signal that provides the phase,

$$\boldsymbol{\phi}(x) = arctg\frac{I_q(x)}{I_p(x)} \tag{1.5}$$

and

$$\left\| \overset{\Rightarrow}{\mathbf{I_{sp}}}(x) \right\| = \sqrt{I_p^2 + I_q^2} \tag{1.6}$$

where the double bar symbol indicates the modulus of the vector in the complex plane.

A fundamental property of the Hilbert transform is to provide a definition of local phase concept that is ancillary to the definition provided by Eq. (1.3) but applies to a single point of the gray level continuum of a one-dimensional signal provided that the gray levels are smooth functions with smooth derivatives in $\mathbf{R}^2$ that symbolizes the 2D continuum. The preceding conditions are ideal conditions that are not satisfied by actual signals. Recorded signals are inherently stochastic, hence in actual applications it is necessary to apply to the recorded gray levels smoothing procedures to approximate with certain error the theoretical ideal continuum signal. One should keep in mind these two separated aspects of the local phase definition, the theory behind this definition that is a consequence of the continuum theory and the procedures needed to implement applications of the mathematical model to actual experimental signals. In the literature of analysis of actual optical signals there is a very extensive treatment of the subject of separating stochastic and deterministic information. In this section and following sections the emphasis is on the continuum model, the stochastic aspect will be introduced later on in the paper. This is a very important simplification for the subject matter of the paper, fringe pattern information retrieval. Later on we will indicate the impact of the assumption of continuity in the handling of actual stochastic signals.

The aim of the current paper is to extend the derivations presented in the framework of a one dimensional model continuum model, [1], to a two dimensional continuum case.

1.2   Two Dimensional Sinusoidal Functions

The next step in this process is to define the properties of two dimensional sinusoidal functions, a generalization of one dimensional sinusoidal functions utilized in the one dimensional continuum [1]. Figure 1.1a shows a two dimensional sinusoidal signal, it has an amplitude and a period p as is the case in one dimension but has an additional degree of freedom, the local orientation. Figure 1.1b illustrates the 2D sinusoid as a signal in 2-D. The yellow line shows a line of equal intensity (phase); the normal **n** provides the orientation of the signal, angle θ, and the vector **r** identifies a point of phase φ in the uniform field of the 2D sinusoidal signal. As shown by Eq. (1.2), the phase is computed with respect to a selected point O (center of coordinates) and is evaluated—Eq. (1.3)—as an angle that provides the number of cycles of the unit of measure p, a rational number n. The red line corresponds to points of equal number of cycles, since as the orientation of the vector **r** changes, it also changes the projected pitch p that is the unit measure to convert distances into angles. Comparing a 2D signal with a 1D sinusoidal signal, as mentioned before, there is an additional degree of freedom, the angle θ (see Fig. 1.1b).

The considerations that follow are very important because the information that we want to retrieve is connected with a model, the continuum mechanics of solids that has its basis on the differential geometry approach to the continuum deformation with specific requirements for the signal and its successive derivatives. What this last sentence means is: specific requirements are imposed on the signal and its derivatives.

The information that we want to obtain is a tensorial field that requires in the case of orthogonal Cartesian Coordinates a family of two orthogonal carrier fringes illustrated in Fig. 1.2. The vertical fringes and the horizontal fringes are represented in the frequency plane of the Fourier Transform (FT) by power spectrum dots whose coordinates are for example of the form $(f_x, f_y) = (10, 0)$ for the x-axis, and the coordinates of the point in the negative frequencies are $(f_x, f_y) = (-10, 0)$, that is a reflection with respect to the vertical axis. In analogous fashion, for the horizontal fringes we have $(f_x, f_y) = (0, 10)$ and for the negative frequency $(f_x, f_y) = (0, -10)$. In Fig. 1.2, the system of coordinates is selected as a left-handed system according with the usual practice in image analysis literature as opposed to Fig. 1.1 where a right-handed reference system is used.
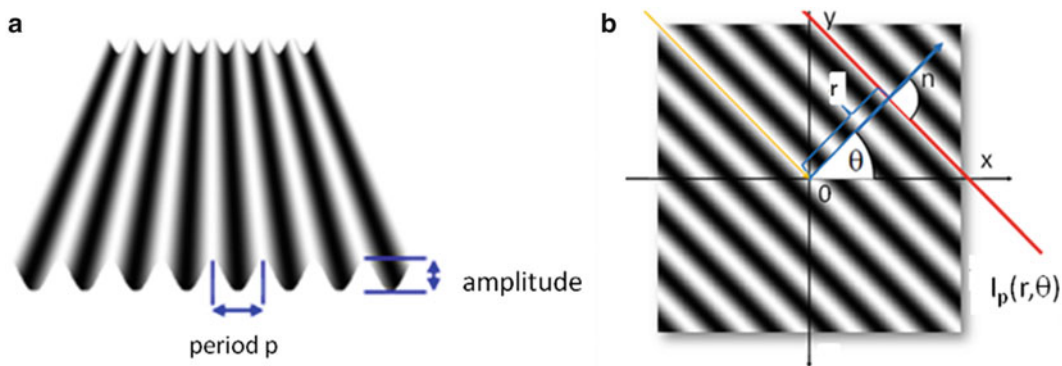

Fig. 1.1 (**a**) 2D sinusoidal signal; (**b**) additional parameter θ to define a two dimensional signal
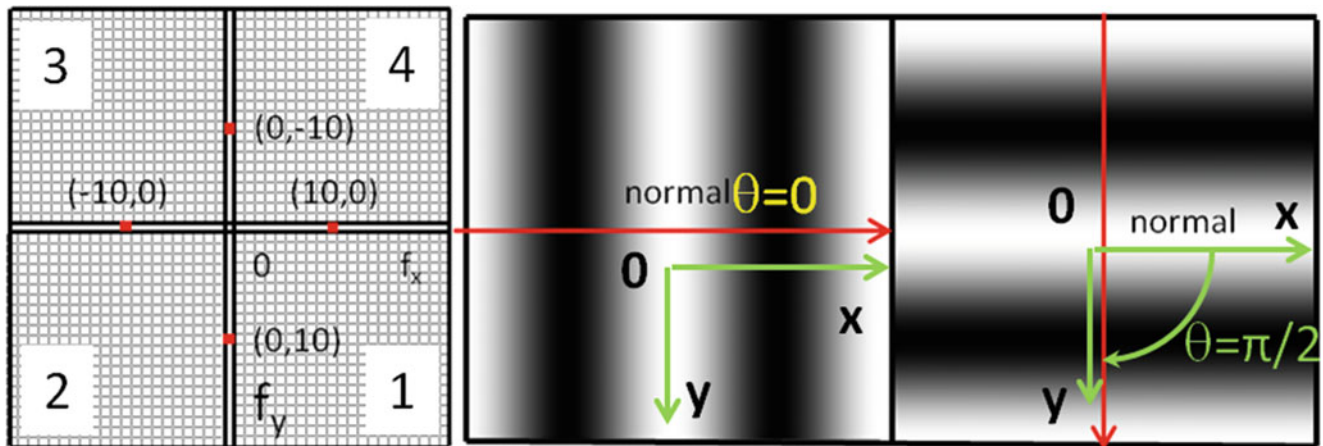

Fig. 1.2 Representation of a 2D cosinusoidal even signal when $\theta = 0$ and $\theta = \pi/2$ and the corresponding representation in the frequency plane represented by a sensor with square pixels

If we return to the concept of phase defined in Eq. (1.3) and compute the phase of a point defined by the vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ in the direction of the normal $\mathbf{n}$ (Fig. 1.1b) it follows

$$\boldsymbol{\phi}(\mathbf{r}, \boldsymbol{\theta}) = \frac{2\boldsymbol{\pi}\|\mathbf{r}\|}{p} \tag{1.7}$$

where the double bar indicate the modulus of the vector in agreement with Eq. (1.3).

To understand the developments that follow, it is necessary to come back to the concept of local phase that can be introduced [1] via the Hilbert transform. The concept of local phase is a fundamental step in the whole process described in this paper and will be dealt with later on in the paper since it involves the transition between the ideal continuum and the actual recorded stochastic signals.

The extension of local phase to 2D sinusoidal signals includes an additional degree of freedom, the angle θ indicated in Fig. 1.1b. The concept of phase requires a 2D vectorial field since it is associated with a vectorial function. The information captured by a sensor is given by levels of gray, a scalar quantity. This scalar function in the case of a 1D signal is connected to a 2D scalar potential in the complex plane that will be called $\dot{V}$; the symbol • indicates that the scalar potential is associated with a given point of coordinate x in the one dimensional continuum $\mathcal{L}^1$, that has a certain reference zero point from where the coordinate x is computed. In Eq. (1.4), the complex notation of [4] is utilized to represent an analytical function, for a more general approach in view to the extension to 3-D. A complex plane defined by the versors $\overset{\Rightarrow}{\mathbf{i}}$ and $\overset{\Rightarrow}{\mathbf{j}}$ is introduced, thus avoiding the utilization of quaternions that are the extension of the complex notation beyond 2-D.

Returning to the complex plane required to introduce the concept of local phase, the gradient of the scalar potential is given by

$$grad \ \dot{V} = \mathbf{G}_2 = \frac{\partial \dot{V}}{\partial x_c} \overset{\Rightarrow}{\mathbf{i}} + \frac{\partial \dot{V}}{\partial y_c} \overset{\Rightarrow}{\mathbf{j}} \tag{1.8}$$

where: $\overset{\Rightarrow}{\mathbf{i}}$ and $\overset{\Rightarrow}{\mathbf{j}}$ are the versors in the complex plane (introducing different symbols from $\mathbf{i}$ and $\mathbf{j}$ that represent the versors in the physical space); x_c and y_c are the coordinates in the complex plane; the subscript "2" indicates 2D gradient vector in the complex space. The sinusoidal signal is represented by gray levels defined by a function of the form,

$$\dot{V}^e = I_p \cos\left(\frac{2\pi}{p}x + \phi_0\right) \tag{1.9}$$

The upper script "e" expresses the fact that the selected function is a cosine, an even function. It is possible to see that the local phase depends on the selection of the phase at the reference point.

Computing the dot product of the ∇ operator with the vector $\mathbf{G_2}$, the divergence of the field is obtained as:

$$\nabla \cdot \mathbf{G}_2 = \frac{\partial^2 \dot{V}}{\partial x_c{}^2} + \frac{\partial^2 \dot{V}}{\partial y_c{}^2} \tag{1.10}$$

Calling $\dot{V}_{xc} = \frac{\partial \dot{V}}{\partial x_c}$ and $\dot{V}_{yc} = \frac{\partial \dot{V}}{\partial y_c}$, and computing the vector product "×" of the ∇ operator with the $\mathbf{G_2}$ vector, it follows:

$$\nabla \times \mathbf{G}_2 = \left(\frac{\partial \dot{V}_{yc}}{\partial x_c} - \frac{\partial \dot{V}_{xc}}{\partial y_c}\right) \overset{\Rightarrow}{\mathbf{k}} \tag{1.11}$$

Since the field is a scalar field, the divergence of the field is zero and the rotor is also zero. Two equations can be derived:

$$\frac{\partial^2 \dot{V}}{\partial x_c{}^2} + \frac{\partial^2 \dot{V}}{\partial y_c{}^2} = 0 \tag{1.12}$$

$$\frac{\partial \dot{V}_{yc}}{\partial x_c} - \frac{\partial \dot{V}_{xc}}{\partial y_c} = 0 \tag{1.13}$$

These equations mean that the potential function in the complex plane must satisfy the Cauchy-Riemann equations,

$$\frac{\partial \dot{V}_{xc}}{\partial x_c} = -\frac{\partial \dot{V}_{yc}}{\partial y_c} \tag{1.14}$$

$$\frac{\partial \dot{V}_{xc}}{\partial y_c} = \frac{\partial \dot{V}y_c}{\partial x_c} \tag{1.15}$$

Equation (1.12) implies that the gray level potential $\dot{V}$ to define a local phase must be a solution of Laplace's equation in the complex plane. The solutions of the Laplace's equation are part of the theory of potentials; these solutions are known to be harmonic functions. The field is conservative and the vectorial field is the gradient of a potential scalar field.

The meaning of Eqs. (1.14) and (1.15) is that, in order to define a local phase, successive derivatives of gray levels must satisfy the above conditions. Furthermore, considering the full complex field, these equations are the conditions for $\dot{V}(\boldsymbol{\rho}_c) \cdot d\boldsymbol{\rho}_c$, where $\boldsymbol{\rho}_c = x_c \overrightarrow{\mathbf{i}} + y_c \overrightarrow{\mathbf{j}}$ is a given direction in the complex plane, to be an exact differential or, in other words, that is a potential such that the integral of the field is independent of the pathway followed. This conclusion leads to the complex function,

$$z(x) = \dot{V}^e(x) + \overrightarrow{j}\ \dot{V}^o(x) \tag{1.16}$$

where $\dot{V}^o(x)$ represents the odd component of the signal.

Through Eq. (1.16) one gets the connection between the Hilbert transform, holomorphic functions and the levels of gray as a potential function leading to the definition of a local phase. For example, if $\dot{V}^e$ is of the form given by Eq. (1.8), through the Hilbert transform we will obtain,

$$\dot{V}^o(x) = I_q \sin\left(\frac{2\pi}{p}x + \phi_0\right) \tag{1.17}$$

Each one of these derivatives can be computed from the information recorded in the image sensor. For each point of the $\mathcal{L}^1$ domain, one can plot the gray levels as $\dot{V}$ (x). From Eq. (1.8) $\frac{\partial \dot{V}}{\partial x_c} \equiv \frac{\partial V(x)}{\partial x}$ and $\frac{\partial \dot{V}(x_c)}{\partial y_c} \equiv \frac{\partial \dot{V}(x)}{\partial y}$ can be obtained and, finally, complementary derivatives are obtained from Eqs. (1.13) and (1.14). In summary, to represent the deformation of a continuous field the derivatives must satisfy the above relationships for a one dimensional signal. However, recorded signals will be contaminated by different signals that we designate as noise. Whatever processes that are applied to the signal to remove noise they must get successive derivatives satisfying the above conditions.

All previous developments correspond to gray levels in one dimension. To introduce the definition of local phase for the 2D sinusoidal signal shown in Fig. 1.1 it is necessary to resort to a 3D complex space. Figure 1.1 showed a 2D cosinusoidal function which has the same parameters as a 1D signal but also additional parameter, the orientation θ. The normal **n** to the fringe trajectory shown in Fig. 1.1 provides the orientation of the signal at a given point of the physical space and the angle θ defines the orientation of the segment of curve with respect to a selected reference system. Some notations that will be useful in the developments that follow are now introduced. The unit normal to fringes in a point of a sinusoidal signal (Fig. 1.1b) is,

$$\mathbf{n} = x \cos\theta\mathbf{i} + y\sin\theta\mathbf{j} \tag{1.18}$$

The above relationship is converted into cycles per unit length by multiplying Eq. (1.18) by 2π/p. Introducing the concept of wave vector for the sinusoidal signal, it can be written:

$$\mathbf{k} = \frac{2\pi}{p}\cos\theta\mathbf{i} + \frac{2\pi}{p}\sin\theta\mathbf{j} = k_x\mathbf{i} + k_y\mathbf{j} \tag{1.19}$$

The wave vector is an alternative way to define the orientation of a segment of a sinusoidal signal in 2-D and relates it to the projections of the trajectory into the reference axis x–y. From Eq. (1.19), the local value of θ is given by

$$\theta(k) = arctg\frac{k_y}{\sqrt{k_x^2 + k_y^2}} \tag{1.20}$$

At every given point of a cosinusoidal fringe field defined by the position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ there is a phasor represented by Eq. (1.1) and the local orientation of the signal defined by the angle θ.

The addition of one more parameter, the angle θ requires to extend the definition of the gray levels as a potential scalar function $\dot{V}^r$ in a 2D space, where $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ is an upper script that indicates that the potential corresponds to a point in $\mathcal{R}^2$, the 2D continuum.

Following the same steps applied in two dimensions and recalling that levels of gray are scalar quantities,

$$grad\,\dot{V}^r = \mathbf{G}_3(\mathbf{r}) = \frac{\partial \dot{V}^r}{\partial x_c}\vec{\mathbf{i}} + \frac{\partial \dot{V}^r}{\partial y_c}\vec{\mathbf{j}} + \frac{\partial \dot{V}^r}{\partial z_c}\vec{\mathbf{k}} \tag{1.21}$$

where the versors $\vec{\mathbf{i}}$, $\vec{\mathbf{j}}$ and $\vec{\mathbf{k}}$ indicate a Cartesian coordinates system in a 3D complex space, the subscript "3" indicates that one is dealing with a 3D vector in the complex space.

The divergence of the field is determined as:

$$\nabla \bullet G_3(\mathbf{r}) = \frac{\partial^2 \dot{V}^r}{\partial x_c{}^2} + \frac{\partial^2 \dot{V}^r}{\partial y_c{}^2} + \frac{\partial^2 \dot{V}^r}{\partial z_c{}^2} \tag{1.22}$$

Calling $\dot{V}_{x_c}^r = \frac{\partial \dot{V}^r}{\partial x_c}$, $\dot{V}_{y_c}^r = \frac{\partial \dot{V}^r}{\partial y_c}$, $\dot{V}_{zc}^r = \frac{\partial \dot{V}^r}{\partial z_c}$, and computing the vectorial product of the ∇ operator with the vector $G_3(r)$, it follows:

$$\nabla \times \mathbf{G}_3(\mathbf{r}) = \left(\frac{\partial \dot{V}_{zc}^r}{\partial y_c} - \frac{\partial \dot{V}_{yc}^r}{\partial z_c}\right)\vec{\mathbf{i}} + \left(\frac{\partial \dot{V}_{xc}^r}{\partial z_c} - \frac{\partial \dot{V}_{zc}^r}{\partial x_c}\right)\vec{\mathbf{j}} + \left(\frac{\partial \dot{V}_{yc}^r}{\partial x_c} - \frac{\partial \dot{V}_{xc}^r}{y_c}\right)\vec{\mathbf{k}} \tag{1.23}$$

Since we are dealing with a scalar potential, the divergence is zero. Hence, it can be written:

$$\frac{\partial^2 \dot{V}^r}{\partial x_c{}^2} + \frac{\partial^2 \dot{V}^r}{\partial y_c{}^2} + \frac{\partial^2 \dot{V}^r}{\partial z_c{}^2} = 0 \tag{1.24}$$

Equation (1.24) indicates that the potential $\dot{V}^r$ satisfies the Laplace's equation in the complex 3D space. The meaning of this equation is the same as for two dimensions. The fact that the rotor is zero implies,

$$\frac{\partial \dot{V}_{xc}^r}{\partial y_c} = \frac{\partial \dot{V}_{yc}^r}{\partial z_c} = \frac{\partial^2 \dot{V}^r}{\partial y_c \partial z_c} \tag{1.25}$$

$$\frac{\partial \dot{V}_{xc}^r}{\partial z_c} = \frac{\partial \dot{V}_{zc}^r}{\partial x_c} = \frac{\partial^2 \dot{V}}{\partial x_c \partial z_c} \tag{1.26}$$

$$\frac{\partial \dot{V}_{yc}^r}{\partial x_c} = \frac{\partial \dot{V}_{xc}^r}{\partial y_c} = \frac{\partial^2 \dot{V}^r}{\partial x_c \partial y_c} \tag{1.27}$$

Equations (1.26)–(1.28) are the conditions for the existence of a scalar potential in the 3D complex space and are equivalent to the Cauchy-Riemann conditions in the two dimensional case.

The above derivations indicate that the information contained in a 2D cosinusoidal fringe pattern is described mathematically by a conservative 3D vectorial field in the complex space. Similarly with the one dimensional case the derivatives that appear in the preceding developments can be computed in the 2D real space as recorded in the sensor. The difference with the one dimensional case is now that the information is in the form of a Monge's type surface where the gray level is of the form,

$$\dot{\vec{V}^r}(z_c) = F(x_c, y_c) \tag{1.28}$$

where $F(x_c, y_c)$ indicates a 2D function.

From Eq. (1.28), it is possible to get all the derivatives that appear in the preceding developments of the 3D complex field proceeding in a similar way to that utilized in the one dimensional case.

Since we are dealing with Continuum Mechanics problems operating on tensorial entities, a family of orthogonal cosinusoidal signals must be defined in Cartesian coordinates as shown in Fig. 1.2. The orthogonal modulated fringe patterns project the displacement vectors in two orthogonal directions. These projections are not independent from each other since they are tied together by the compatibility conditions of the continuum and involve also the Eqs. (1.25)–(1.27) that both systems of fringes must satisfy at the same points of the image.

Conclusions similar to the one dimensional case can be obtained: the successive derivatives of the gray levels must satisfy the above conditions to define a scalar potential. Hence, the passage from the actual signals to the continuum signals requires operations that must enforce the above conditions as close as it may be feasible. This conclusion is very important because the change of the orientation of the fringes is related to the curvature of the fringes, the larger is the local change of orientation the more important is the effect of the orientation on the derivatives of the gray levels function.

1.3 The Monogenic 2D Signal

The extension of the one dimension approach of signal analysis to multiple dimensions has been the object of a large number of papers (see, for example, [5–8] and the references cited therein). This study will apply the complex Riesz transform presented in the preceding publications. In these four publications are introduced the required arguments to create a transform equivalent of the Hilbert transform in a multidimensional space. To achieve this purpose, the concept of monogenic function is introduced. The original derivation of the monogenic signal concept has its foundations on the algebra of quaternions that is connected to Lie algebra isomorphisms. In this paper, a variation of the original arguments is introduced. The derivations fit the mappings originally developed by Poincare and that for the particular field considered in this study, a 2D flat field, are graphically represented by a Poincare sphere [8] that it is utilized in the field of birefringent optics and in photoelasticity to define the different forms of polarization.

The relationship of Poincare sphere and the concept of phase of the components of polarized light has been the object of several publications (see, for example, [9–11]). The connection between the preceding applications of the Poincare sphere and the phase concept and the current version introduced in this paper is a subject of a great deal of interest but is beyond the purpose of this paper. The motivation in the current version follows from the isomorphism pointed out in [12]. It has been shown that in order to introduce the concept of phase in one dimensional signals, it is necessary to resort to a 2D vectorial field, similarly for 2D signals it is necessary to introduce a vector field in the 3D complex space. The 3D phasor representing the gray levels in 2D has an amplitude that corresponds to the intensity of the signal, a phase that corresponds to the optical path information, and introduces a new variable that corresponds to the orientation of the 2D sinusoid in the physical plane defined by the normal $\mathbf{n}$ Eq. (1.18), a function of the angle θ defined in Fig. 1.1.

Figure 1.3 illustrates the Poincare sphere notation. The vector amplitude is defined by the following components: (a) the components of the gray levels I_x, I_y associated with the versors $\vec{\vec{i}}$ and $\vec{\vec{j}}$, respectively; (b) to these two components it is added a third component I_q, corresponding to the versor $\vec{\vec{k}}$.

The complex amplitude vector in the complex space is given by,

$$I_{sp} = I_x \vec{\vec{i}} + I_y \vec{\vec{j}} + I_q \vec{\vec{k}} \tag{1.29}$$

and corresponds to the radius of a Poincare sphere shown in Fig. 1.3a. This sphere represents the local phase and amplitude at a point in the 2D continuum of gray levels. In the coordinate plane $\vec{\vec{i}}$, $\vec{\vec{j}}$, sphere equator, Eq. (1.29) becomes:

$$I_p = I_x \vec{\vec{i}} + I_y \vec{\vec{j}} \tag{1.30}$$

From Fig. 1.3, it follows:

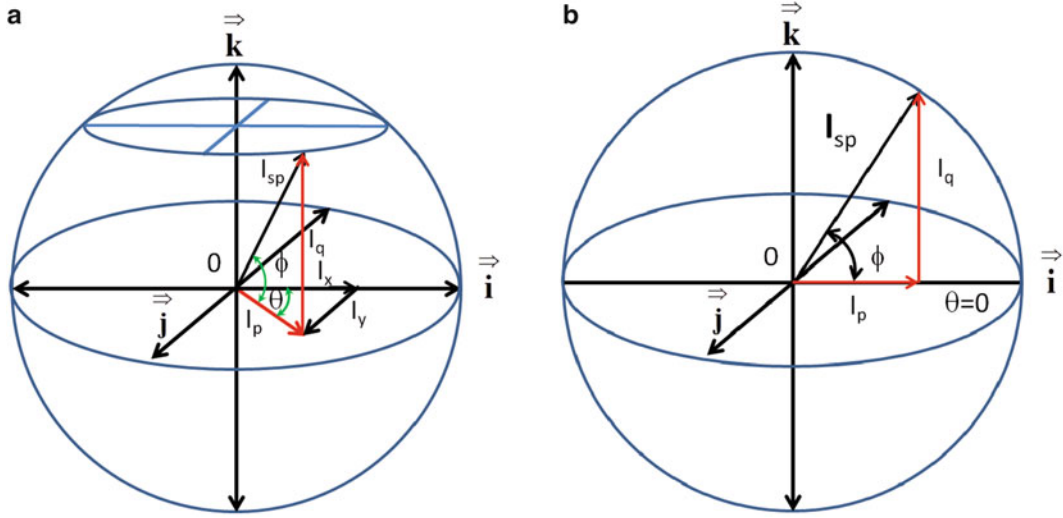$$I_x = \left\| \mathbf{I}_p \right\| \cos \theta \tag{1.31}$$

Fig. 1.3 (**a**) Poincare sphere of the levels of gray representing light intensities; (**b**) Poincare sphere of the levels of gray for $\theta = 0$

$$I_y = \left\|\mathbf{I}_p\right\| \sin \theta \tag{1.32}$$

Figure 1.3b shows that the following relationship applies:

$$\tan \phi = \frac{I_q}{I_p} \tag{1.33}$$

There are two in-quadrature quantities I_p and I_q and a phase ϕ that defines a local phase for a signal of orientation θ in the 2D space. The above derived equations lead to the following relationships between the intensities, defining I_{sp} as the modulus of the vector $\mathbf{I}_{sp}$.

$$I_x = I_{sp} \cos \phi \cos \theta \tag{1.34}$$

$$I_y = I_{sp} \cos \phi \sin \theta \tag{1.35}$$

$$I_q = I_{sp} \sin \phi \tag{1.36}$$

Finally, the monogenic signal can be represented by,

$$\mathbf{M}_s^r = I_{sp} \left[\cos \phi \cos \theta \, \vec{\mathbf{i}} + \cos \phi \sin \theta \, \vec{\mathbf{j}} + \sin \phi \, \vec{\mathbf{k}} \right] \tag{1.37}$$

The upper script indicates that it corresponds to a point **r** of the 2D continuum.

The angle θ defines the longitude of the point under consideration referred to the $\vec{\mathbf{i}} - \vec{\mathbf{k}}$ plane in the complex space. The angle ϕ is the latitude of the point with respect to the equatorial plane and provides the local phase associated with the actual signal.

The above derived relationships from local gray levels at a given point of a 2D image provide local orientation and the local phase. If a sinusoidal signal is such that the normal $\mathbf{n} \equiv \mathbf{i}$, then $\theta = 0$ and the corresponding representation in the Poincare sphere is shown in Fig. 1.3b. It can be seen from Fig. 1.2, for the vertical fringes that measure horizontal displacements the angle of the normal is $\theta = 0$. This case is depicted by the Poincare sphere of Fig. 1.3b.

An alternative space can be considered replacing the light intensities by the frequencies. Defining the energy of the vector $\mathbf{I}_{sp}$ in the frequency space as

$$I_{sp}^2 = I_x^2 + I_y^2 + I_q^2 \tag{1.38}$$

and taking into consideration the FT energy theorem, the following equation holds true in the frequency space:

$$f_{sp} = \sqrt{f_x^2 + f_y^2 + f_q^2} \qquad (1.39)$$

The Poincare sphere in the intensity space depicted in Fig. 1.3a can be transformed into the Poincare sphere of the frequency space by replacing intensities by the corresponding frequencies. All the derivations made for the 3D complex space that are related to the intensities are valid also for the frequencies.

1.4   The Riesz Transform

All the quantities that define the Poincare sphere can be obtained directly by applying the generalized Hilbert transform or Riesz transform defined in [5–7]. The Riesz transform can be computed in the physical space or in the frequency space defined by the analytic function theory [4].

Equation (1.37) provides the monogenic vector corresponding to a given point of the $\mathscr{R}^2$ gray levels continuum represented graphically by a Poincare sphere in a 3D complex space. The monogenic function vector I_{sp} has three components I_x, I_y and I_q, and its position in space is defined by two angles, θ and ϕ, a total of five unknown quantities. These quantities are related by Eqs. (1.34)–(1.36). Since of these three equations only two are independent, only three quantities (i.e. I_x, I_y and I_q) must be determined while I_p is the level of gray captured by the sensor.

The Riesz transform of gray levels of an image in the spatial domain associates with each point of the continuum two orthogonal convolution kernels (Chap. 4 of [7]):

$$h_x(r) = \frac{x}{2\pi(x^2 + y^2)^{\frac{3}{2}}} \qquad (1.40)$$

$$h_y(r) = \frac{y}{2\pi(x^2 + y^2)^{\frac{3}{2}}} \qquad (1.41)$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$. These kernels yield:

$$I_x(r) = \frac{x}{2\pi(x^2 + y^2)^{\frac{3}{2}}} ** I_p(r) \qquad (1.42)$$

$$I_y(r) = \frac{y}{2\pi(x^2 + y^2)^{\frac{3}{2}}} ** I_p(r) \qquad (1.43)$$

In the above equations, $I_p(r)$ is the gray level at $\mathbf{r}$ while the ** symbol [13] denotes a 2D convolution in physical space. These kernels satisfy the following relationships with the corresponding quantities in the frequency space:

$$h_x(\mathbf{r}) \rightarrow H_{f_x}(\mathbf{f}_r) = \frac{f_x}{\sqrt{f_x^2 + f_y^2}} \qquad (1.44)$$

$$h_y(\mathbf{r}) \rightarrow H_{f_y}(\mathbf{f}_r) = \frac{f_y}{\sqrt{f_x^2 + f_y^2}} \qquad (1.45)$$

where the symbols H_{f_x} and H_{f_y} indicate Hilbert transform in 2-D as it is called in [7]. The generic denomination multidimensional Hilbert transform is used in place of the Riesz transform and the frequency space corresponds to the analytic frequency space [4].

It should be noted that the operators H_{f_x} and H_{f_y} of Eqs. (1.44) and (1.45) define the $\cos\theta$ and $\sin\theta$ terms in the frequency plane, consistently with Eqs. (1.34) and (1.35). Then

$$\mathbf{\theta}(r) = arctg \, \frac{I_y(r)}{I_x(r)} \tag{1.46}$$

The orientation of the signal is known through the angle θ.

Rotating the coordinate axis by the amount θ and calling the rotated coordinate x',

$$I_q\left(x'\right) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{I_p(\mathbf{\eta})}{x' - \mathbf{\eta}} d\mathbf{\eta} = \frac{1}{\pi}\left(I_p\left(x'\right) * \frac{1}{x'}\right) \tag{1.47}$$

where PV indicates the principal value and η is the dummy variable of integration.

Equations (1.44)–(1.47) allow obtaining the monogenic function vector Eq. (1.37). All the above introduced computations are performed in the physical space. These computations can also be performed in the analytical functions frequency space [4].

1.5 Retrieval of the Monogenic Vector

First we will recall some properties of the FT that are of interest in the current analysis and shed light on the relationship between the FT and the Riesz transform. Both of these two transforms provide signals in-quadrature by utilizing different algorithms that however are closely related to each other since theoretically the outputs must be the same. In practice, due to the numerical processes involved obtained results may differ.

The windowed FT provides the signals in-quadrature in 2-D. Let us recall some basic steps of the process of computing in-quadrature signals. The FT has symmetry properties,

$$\widetilde{F}\left(f_x, f_y\right) = \widetilde{F}^*\left(-f_x, -f_y\right) \tag{1.48}$$

$$\widetilde{F}\left(-f_x, f_y\right) = \widetilde{F}^*\left(f_x, -f_y\right) \tag{1.49}$$

Furthermore it also satisfies the condition of separable functions,

$$\widetilde{F}\left(f_x, f_y\right) = \widetilde{F}_1\left(f_x\right) \widetilde{F}_2\left(f_y\right) \tag{1.50}$$

In Fig. 1.4, a fringe pattern is plotted and its corresponding FT in the frequency space. It is possible see that the fringe pattern in the third quadrant of angle $\theta' = 180 + \theta$ corresponds the same components $(-5, -6)$ as to a fringe pattern of angle $\theta(5, 6)$ but with signs changed. It satisfies Eq. (1.48), the components in the third quadrant are complex conjugates of the

Fig. 1.4 Symmetry conditions of the FT in the frequency space (**a**) for a 2D pattern (**b**)
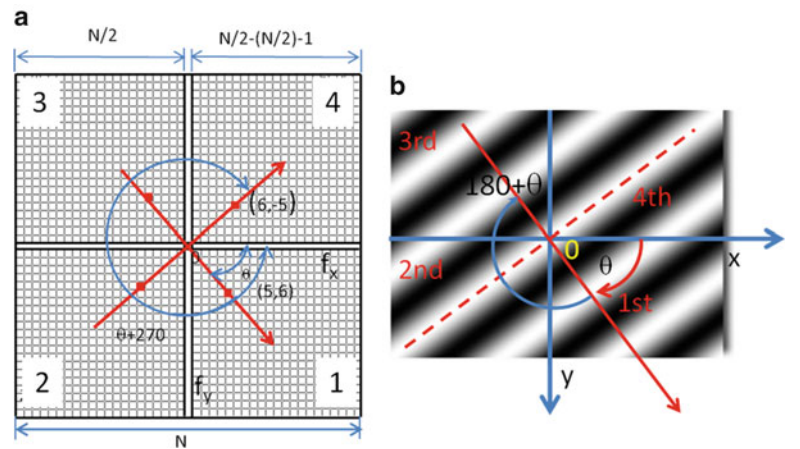
components in the first quadrant. Similar conclusion can be reached if we have a system of fringes in the second quadrant, that is the dashed red line represents the normal to the pattern with components $(-6, 5)$ and the corresponding system of fringes in the fourth quadrant $(6, -5)$ satisfying Eq. (1.49).

The FT components are complex numbers $(a_{ij} + jb_{ij})$ that can be arranged in a square matrix where "i" indicates the rows and "j" the columns. The carrier orientation is given by the angle θ of the normal to the carrier direction. In the case of a vertical carrier, it holds $\theta = 0$; in the case of a horizontal carrier, it holds $\theta = \pi/2$. In general, θ goes from 0 to $\boldsymbol{\theta \Rightarrow \pm\pi}$. Because the signal satisfies Eq. (1.50) the computation can be carried separately along rows and columns.

Because of the symmetry conditions there are only $N^2/2$ independent components to obtain the FT. There are additional components that are real, the term a_{oo} that corresponds to $f_o = 0$ (the background term is a constant), and the terms corresponding to the Nyquist frequency $f = N/2$. Then, the number of independent coefficients to compute is $N^2/2$ and two real coefficients a_{oo} and the Nyquist frequency $f_{N/2}$.
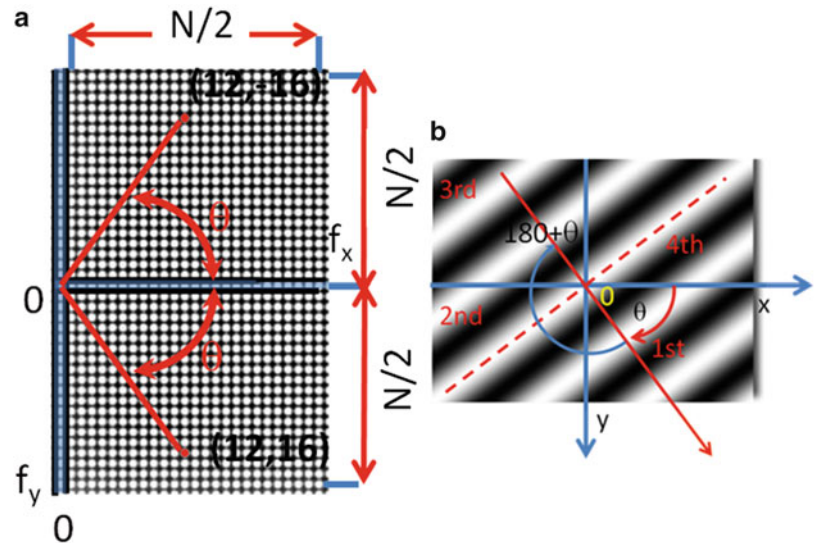
The Riesz transform is an alternative to the FT computational method to obtain the in-quadrature signals and, as pointed out before theoretically, the final output must be the same independently of the computation path followed. First it must be realized that the background term is eliminated from the computation in the Hilbert transform since it is a singularity at the origin of coordinates of the frequency space.

Figure 1.5 represents the actual physical space and the corresponding analytical function frequency space (notice the difference between the frequency space of the FT and the frequency space of the analytic functions). Physical space quadrants 1 and 3 are represented in the frequency analytical space by the lower half space $f_x, f_y > 0$, while quadrants 2 and 4 are represented by the $f_x > 0, f_y < 0$ upper half space. The corresponding coefficients of the Riesz transform are complex quantities $Z(f_x, f_y) = (a_{ij} + jb_{ij})$; their magnitude is twice in magnitude of the coefficients in the FT transform because of the single side transform as illustrated in Fig. 1.5. The total number of independent coefficients is then $N^2/2$, because of the Nyquist transform symmetry that is the coefficients of the upper half space are the complex conjugated of the lower half space. Similarly to the FT property of separability stated by Eq. (1.50), coefficients can be computed by rows and columns. The angle θ goes from 0 to $\boldsymbol{\theta \Rightarrow \pm\pi}$ and can be computed by means of Eqs. (1.44) and (1.45). The monogenic signal as defined in this paper allows us to compute all the parameters that define a 2D sinusoidal signal.

Figure 1.6 shows all the coefficients that it is necessary to compute in the frequency space to obtain the phase ϕ of a sinusoidal signal in the frequency space utilizing the monogenic signal concept. The above presented derivations emphasize the fact that the same objective can be reached by utilizing the FT.

There are many theoretical reasons for introducing the concept of analytic signals, or in other words, for eliminating the negative frequencies of the real and imaginary components of the spectrum of real signals. From the practical point of view, it has been pointed out that the negative spectrum has in essence the same information than the positive parts due to the conjugate symmetry previously mentioned. The elimination of the negative frequencies has also an impact in the efficiency signal processing by reducing the required bandwidth.

Fig. 1.5 (**a**) Frequency space for the analytical function; (**b**) physical space

$f_x > 0$

Real							Imaginary						
a_{00}	a_{10}	a_{20}	a_{30}	$\dots$	$a_{510\,0}$	$a_{511\,0}$	b_{00}	b_{10}	b_{20}	b_{30}	$\dots$	$b_{510\,0}$	$b_{511\,0}$
a_{01}	a_{11}	a_{21}	a_{31}	$\dots$	$a_{510\,1}$	$a_{511\,1}$	b_{01}	b_{11}	b_{21}	b_{31}	$\dots$	$b_{510\,1}$	$b_{511\,1}$
a_{02}	a_{12}	a_{22}	a_{31}	$\dots$	$a_{510\,2}$	$a_{511\,2}$	b_{02}	b_{12}	b_{22}	b_{31}	$\dots$	$b_{510\,2}$	$b_{511\,2}$
a_{03}	a_{13}	a_{23}	a_{31}	$\dots$	$a_{510\,3}$	$a_{511\,3}$	b_{03}	b_{13}	b_{23}	b_{31}	$\dots$	$b_{510\,3}$	$b_{511\,3}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a_{0\,510}$	$a_{1\,510}$	$a_{2\,510}$	a_{31}	$\dots$	$a_{510\,510}$	$a_{511\,510}$	$b_{0\,510}$	$b_{1\,510}$	$b_{2\,510}$	b_{31}	$\dots$	$b_{510\,510}$	$b_{511\,510}$
$a_{0\,511}$	$a_{1\,511}$	$a_{2\,511}$	a_{31}	$\dots$	$a_{510\,511}$	$a_{511\,511}$	$b_{0\,511}$	$b_{1\,511}$	$b_{2\,511}$	b_{31}	$\dots$	$b_{510\,511}$	$b_{511\,511}$

$f_y > 0$ $f_y < 0$

a_{00}	a_{10}	a_{20}	a_{30}	$\dots$	$a_{510\,0}$	$a_{511\,0}$	$-b_{00}$	$-b_{10}$	$-b_{20}$	b_{30}	$\dots$	$-b_{510\,0}$	$-b_{511\,0}$
a_{01}	a_{11}	a_{21}	a_{31}	$\dots$	$a_{510\,1}$	$a_{511\,1}$	$-b_{01}$	$-b_{11}$	$-b_{21}$	b_{31}	$\dots$	$-b_{510\,1}$	$-b_{511\,1}$
a_{02}	a_{12}	a_{22}	a_{31}	$\dots$	$a_{510\,2}$	$a_{511\,2}$	$-b_{02}$	$-b_{12}$	$-b_{22}$	b_{31}	$\dots$	$-b_{510\,2}$	$-b_{511\,2}$
a_{03}	a_{13}	a_{23}	a_{31}	$\dots$	$a_{510\,3}$	$a_{511\,3}$	$-b_{03}$	$-b_{13}$	$-b_{23}$	b_{31}	$\dots$	$-b_{510\,3}$	$-b_{511\,3}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a_{0\,510}$	$a_{1\,510}$	$a_{2\,510}$	a_{31}	$\dots$	$a_{510\,510}$	$a_{511\,510}$	$-b_{0\,510}$	$-b_{1\,510}$	$-b_{2\,510}$	b_{31}	$\dots$	$-b_{510\,510}$	$-b_{511\,510}$
$a_{0\,511}$	$a_{1\,511}$	$a_{2\,511}$	a_{31}	$\dots$	$a_{510\,511}$	$a_{511\,511}$	$-b_{0\,511}$	$-b_{1\,511}$	$-b_{2\,511}$	b_{31}	$\dots$	$-b_{510\,511}$	$-b_{511\,511}$

$f_x > 0$

Fig. 1.6 Table of coefficients in the half space assuming that the image has 1024×1024 pixels. The columns outlined in *red* will correspond to the coefficients of the zero order that are not present in the Riesz transform

1.6   Generalized Hilbert Transform for Signal Phase Retrieval

The phase shifter or Hilbert filter is an ideal symbolic operator capable of leaving signal amplitudes unchanged and introducing phase shifts of π/2. The ideal operator can be approximated in different ways. In the preceding section, it is outlined an approach in the frequency space of analytic signals with one sided spectrum. The operations can be carried out in the physical space or in the frequency space. In both cases, the Hilbert transformers are special class of filters. The difference between alternative filters can be quantified by the operational efficiency minimizing the number of operations required and minimizing phase and amplitude errors resulting from numerical operations. There are two basic types of filters: (a) infinite impulse response filters (IIR filters) whose response does not become zero, pass a certain point but continues indefinitely; (b) finite impulse filters whose response becomes zero at certain point. FIR filters are preferred filters for Hilbert transform operations, for a comprehensive discussion in this topic see [14, 15].

For reasons that will be explained further on in the manuscript we are interested in a shifted frequency modulated function that is in a function whose spectrum is shifted of a certain amount in the frequency space. We have a real function of the form (Fig. 1.1),

$$I(x) = I_p \cos \phi(x) = \frac{I_p}{2}\left[e^{j\phi(x)} + e^{-j\phi(x)} \right] \tag{1.51}$$

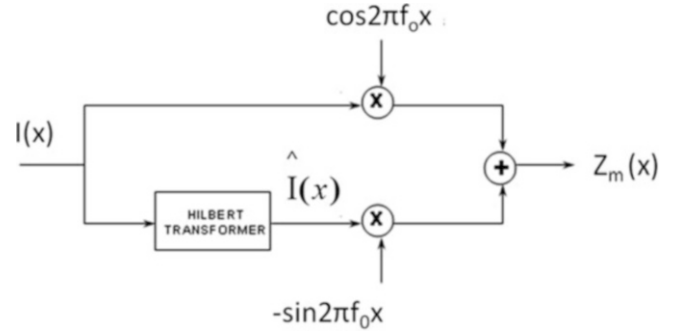The corresponding in-quadrature signal is

$$\hat{I}(x) = I_q \sin \phi(x) = \frac{I_q}{2j}\left[e^{j\phi(x)} - e^{-j\phi(x)} \right] \tag{1.52}$$

The shifted modulated analytic function is defined as [4]:

$$z_m(x) = \frac{I_p}{2}\left[e^{j\phi(x)} + e^{-j\phi(x)} \right] e^{-j2\pi f_0 x} + j\frac{I_q}{2}\left[e^{j\phi(x)} - e^{-j\phi(x)} \right] e^{-j2\pi f_0 x} \tag{1.53}$$

Fig. 1.7 Process to generate the single side frequency modulated bandpass signal that provides shifted signals in-quadrature

This expression yields,

$$z_m(x) = \frac{I_p}{2} \cos[\phi(x) - 2\pi f_0 x] + j\frac{I_q}{2} \sin[\phi(x) + 2\pi f_0 x] \tag{1.54}$$

Calling

$$P(x) = \frac{I_p}{2} \cos[\phi(x) - 2\pi f_0 x] \tag{1.55}$$

$$Q(x) = \frac{I_q}{2} \sin[\phi(x) + 2\pi f_0 x] \tag{1.56}$$

The function P(x) is the in-phase component while the Q(x) function is the in-quadrature component of the frequency modulated function $z_m(x)$.

$$z_m(x) = P(x) + jQ(x) \tag{1.57}$$

The net effect is to displace the frequency spectrum of the frequency f_o in the positive sense. Each component of the input experiences a shift that is proportional to its frequency and the overall final effect is to produce a translation of the FT of the original signal frequency spectrum by the frequency f_o. In order to implement the single band frequency modulation previously described and illustrated in the flow chart of Fig. 1.7, it is necessary to compute the Hilbert transform of I(x) (Eq. (1.51)). Afterwards, both I(x) and Î(x) must be modulated to obtain P(x) and Q(X), that is the signals in-quadrature but shifted to the frequency f_o. Algorithms that approximate the Hilbert Transformer, such as the Parks-McClellan FIR filter design technique, have been developed and can be found in MATLAB Signal Processing Toolbox™.

1.7 Transition from the Continuum to Actual Signals

To this point the continuum aspect of fringe pattern analysis has been dealt with in some detail. It has been pointed out the importance of phase in image analysis and certainly the fundamental role that it plays on fringe pattern studies. The two different concepts of phase, global and local have been introduced. The next step is to deal with the local phase and add to the continuum approach the statistical tools that are required to bridge the gap between the continuum model and the actual signals that are captured. The impact of local phase analysis in fringe pattern is in itself a huge topic. It is a well known fact that the FT is a powerful tool in global phase analysis but does not provide details of the local phase structure. A similar conclusion can be arrived for the analytical frequency space.

Reasoning in the FT frequency space, consider a 2D signal with energy [4],

$$E = \iint_\infty \|f(x, y)\|^2 dx dy \tag{1.58}$$

Assume that the values of its Fourier Transform (FT) are limited to a region A on the FT plane and in the physical plane the corresponding values of the function f(x, y) are limited to a region B. We further assume that the product A × B is also a small quantity. It is possible to define the energy ratio,

$$\boldsymbol{\alpha} = \iint_B \frac{\|f(x,y)\|^2 dxdy}{E} \tag{1.59}$$

Under the above assumption, f(x, y) is a slowly varying function in B. Then, it can be written in the local coordinate system,

$$\boldsymbol{\alpha} = \iint_B \frac{\|f(x,y)\|^2 dxdy}{E} \approx \frac{\|f(0,0)\|^2 B_0}{E} \tag{1.60}$$

where B_o is the considered local area. The optimum condition will be to maximize $\|f(0, 0)/E\|$. This problem is related to the classical problem in signal analysis of the localization of a signal in space and in frequency. There is a limit to the resolution that can be achieved. The Heisenberg Uncertainty Principle of the FT puts a limit to the resolution in a rectangular local area,

$$\Delta_x \Delta_f \geq \frac{1}{4\boldsymbol{\pi}} \tag{1.61}$$

where Δ_x is the uncertainty in the spatial coordinates and Δ_f is the uncertainty in the frequency value of the signal. An answer to the optimization of the approximated energy ratio defined in Eq. (1.60) [4] is to multiply f(x, y) by a window function. It is well known that the lower limit of Eq. (1.61) is obtained when the window function is a Gaussian function.

In the present case, the sensor gives $I_p(x, y)$ and it is necessary to locally smooth this function by introducing a band pass filter that will render the approximated energy ratio defined in Eq. (1.60) optimum. Hence, in the previously derived equations the recorded $I_p(x, y)$ can be replaced by:

$$I_p^b(x,y) = I_p ** g(x,y) \tag{1.62}$$

where the upper script b indicates a band pass filtered version of the recorded signal and g(x, y) is a selected band pass filter. In [7], a Poisson's and conjugate Poisson's kernels filters are utilized. However, there are many other alternative filters that can be utilized depending on the noise present in the analyzed signals. The derivation of the optimization of Eq. (1.60) suggests a Gaussian smoothing filter to be used for Eq. (1.51): in [7], Gaussian filters are introduced.

At this point one must go back to the derivations made in Sect. 1.2. The concept of the gray levels as a scalar potential to be valid requires that the Cauchy-Riemann conditions in 1-D and in 2-D must be satisfied at each point of the image. This implies that the gray levels and their successive derivatives are required to satisfy the continuity conditions expressed by Eqs. (1.14), (1.15), and (1.25)–(1.27), respectively. In [7], there are extensive statistical considerations to derive optimum values of orientation, phase and amplitude of signals using the Poisson's distribution that involve gray levels and successive derivatives. In [8], there are similar derivations utilizing Gaussians filters. In the case of fringe patterns, the solution is simpler at least for the small deformation theory of the continuum that requires continuous derivatives up to the third order.

Figure 1.8 represents a Gaussian filter in the physical space and in the frequency space.

Equation (1.38) relates both these filters.

$$I(ax, ay) \leftrightarrow \frac{1}{\|a\|} \tilde{I}\left(\frac{f_x}{a}, \frac{f_y}{a}\right) \tag{1.63}$$

Equation (1.63) gives the relationship between the components of the Gaussian filter in the physical space and the frequency space; the scale factor "a" is multiplicative in the physical space and it divides in the frequency domain. This is an important result because it indicates that in the process of filtering there is a scale effect that is a key in obtaining a satisfactory smoothing of the recorded signals. The scale depends on the gradients of the displacement function; large gradients require large scale factors in the physical space.
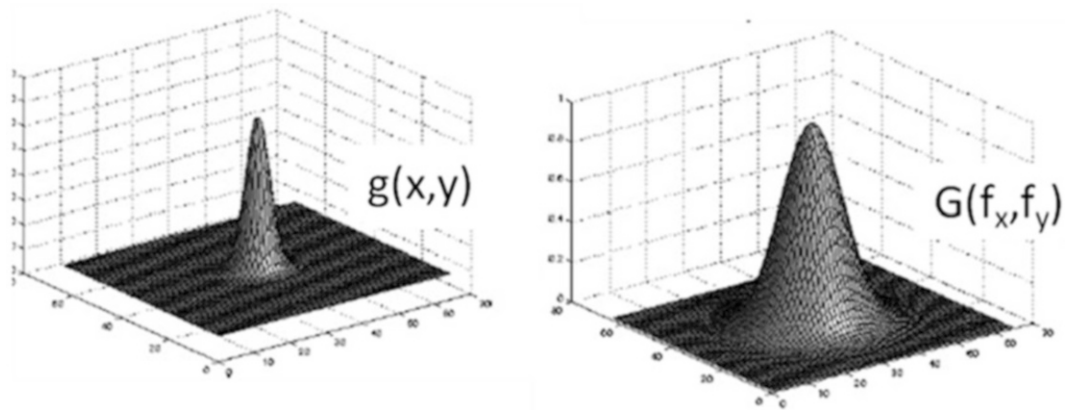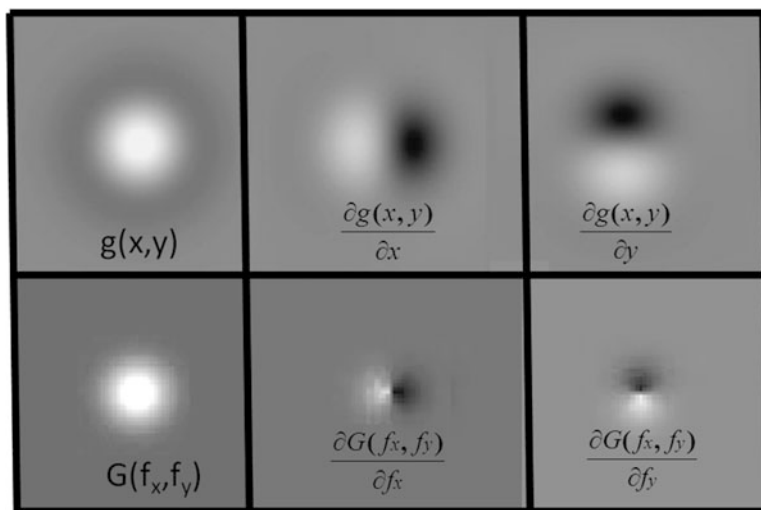
Fig. 1.8 Gaussian filters in the physical space and in the frequency space

Fig. 1.9 Gaussian filters and corresponding derivatives in the space domain and in the frequency domain

The 2D Gaussian filter represented in Fig. 1.8 is given by,

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{x^2+y^2}{2\sigma^2}} \tag{1.64}$$

In this representation, it is assumed that the mean is $(0, 0)$ while $\sigma = 1$ is the standard deviation of the function. The Gaussian filter acts a point spread distribution and the effect in the smoothed function is determined by the selected value for σ. If σ is small, in the spatial domain a large number of frequencies will contribute to the filtered value of the signal and vice versa.

Figure 1.9 shows both the Gaussian function in the coordinate's domain and in the frequency domain as well as the corresponding derivatives. Gaussian derivatives in 2D are separable functions equivalent to a convolution in the x-direction and in the y-direction.

$$\frac{\partial g(x, y)}{\partial x} * *f(x, y) = \left(\frac{-x}{\sqrt{2\pi}\sigma^3} e^{-\frac{x^2}{2\sigma^2}}\right) * \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}}\right) *f(x, y) \tag{1.65}$$

Since we have discrete values of the function, it is necessary to produce discrete approximations to the Gaussian function. The function approaches zero at about 3σ, hence the kernel size is limited to this size.

1.8   Process of Phase Recovery

The preceding section covered the analysis of fringe patterns as 2D signals. One of the important concepts in this analysis is the definition of local phase. While the local phase concept is clear in the continuum theory, in the passage from the continuum to the actual signals one should note that the concept of local phase depends of the selected widows to smooth real signals.

In [1], it was concluded that in order to satisfy the quasi-harmonic condition that allows in 1D signals the utilization of in-quadrature signals for the computation of local phase, it is required to encode the displacement signal into a high frequency carrier. A similar argument must be introduced for 2D signals. In order to achieve this objective, we must consider again the concept of local phase, that in the general theory of 1D signal analysis is known as the definition of the instantaneous frequency, subject that has been the object of extensive studies [16, 17]. The phase computation is a point-wise operation that provides the local phase through the computation of the arc tan function and can be derived by the introduction of the concept of analytic functions [4]. Within the analytic function theory one has signals that are amplitude modulated or frequency modulated. Fringe patterns are both amplitude and frequency modulated signals. This creates a very difficult problem because amplitude and simultaneously phase modulated signals do not have a uniquely defined analytic signal representation [16]. The amplitude and the phase have their own frequency spectra and these spectra can overlap. In [1], it is observed that harmonics providing different optical information besides amplitude and phase can also overlap in the frequency space with amplitude and phase information. As a consequence of these facts, phase and amplitude recovery information is not possible unless steps are taken to minimize the effect of the mixing of different harmonics.

Two important tools are utilized to get solution to these two problems of real signals. The Bedrosian-Nuttal's theorems [18–20] provide solutions to the overlapping problem. If the Bedrosian-Nuttal's theorems are satisfied, the amplitude of the general phasor A(**x**) and the phase ϕ(**x**) are separated in the frequency space. The analytical function theory can provide the in-quadrature components that lead to a real signal with a well defined local frequency,

$$f(x) = \frac{1}{2\pi} \frac{d}{dx} \arg z(x) \tag{1.66}$$

The possibility of satisfying the Bedrosian-Nuttal's theorems is related to the bandwidth of the involved signals. To get an intuitive picture of the problem of defining local phase at a given point of an image, let us return to the one dimensional signal. The simple harmonic function has been traditionally utilized to define instantaneous phase. A periodic motion is represented by a body that moves with constant speed along a circular path (see Fig. 1.10),

$$S(x) = A_0 \cos \phi(x) \tag{1.67}$$

where A_o is the radius of the circle, ϕ(**x**) = ωx, ω is the constant angular frequency connected with the spatial frequency $f_x = \omega/2\pi$.

In [1], it is concluded that for patterns where the amplitude modulation frequencies are much smaller than the frequencies of the phase modulation $f_a \ll f_\phi$, the in-quadrature signals provide accurate values of the phase. This conclusion is proven by the analysis of 1D signals extracted from moiré patterns both computer generated with known frequencies and actual optically produced moiré patterns. It is necessary to extend this conclusion to 2D signals,

Fig. 1.10 Simple harmonic motion as a model to define instantaneous frequency
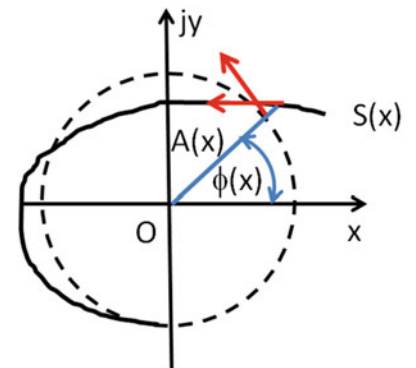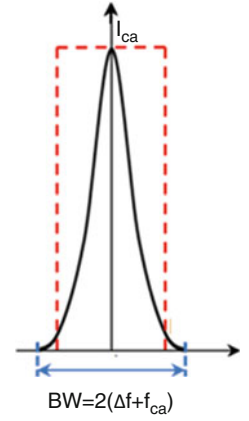
Fig. 1.11 Local power
spectrum of the signal

$$S_{mc}(x, y_c) = I_1 \cos \left[\frac{2\pi}{p_c}x + \phi(x, y_{ca})\right] \tag{1.68}$$

In Eq. (1.68), it is assumed that we are operating along rows of the signal matrix and a carrier is introduced in the x-direction and has amplitude I_1 and the pitch is p_c. The y_c indicates the corresponding constant values of the y-coordinate. Now, if we consider a local region, the FT of the carrier is a pair of pulses at frequency $f_c = 1/p_c$. The frequency f_{ca} is selected such that,

$$f_{ca} \gg f_s \tag{1.69}$$

where f_s is the frequency of the signal $\phi(x, y_c)$ in the local region. Equation (1.69) indicates that locally the signal frequency is much smaller than the carrier frequency.

Figure 1.11 represents the spectrum of the signal plus carrier at a given location, Δf is the local change of frequency of the signal postulated to be such that $\Delta f \ll f_{ca}$ in order to have a narrow band signal. Through the Bedrosian-Nuttal's theorems, amplitude and phase of signals can be assumed to be independent if their respective spectra are separated in the frequency space. Hence, the closer the signal is to a narrow band condition, the more likely is that the Hilbert-transform analytical signal provides support to the validity of Eq. (1.67). This provides an accurate value for the local frequency. Looking from the point of view of Fig. 1.10, A(X) becomes a constant and,

$$\phi(\mathbf{x}) = \omega x \tag{1.70}$$

where ω is a constant angular frequency.

1.9 Application Examples

In the preceding sections two alternative ways of fringe pattern analysis have been discussed and the theory behind these two methods dealt with in some detail. There are many other approaches to fringe pattern processing, such as for example, the Gabor transform and a generalization of the Gabor transform through the concept of wavelet transforms that introduce more refined filters systems and that can be particularly useful for noisy signals. Actually recorded fringe patterns are processed and the results obtained with the different methods compared.

A disk under diametrical compression, u–v displacement patterns and their derivatives are chosen as an example of application of the FFT, the 2D Hilbert transform, the Gabor transform and the Morlet transform. Table 1.1 provides the parameters defining the utilized specimen mechanical properties and the applied load.

The recorded patterns of Fig. 1.12 were obtained with double illumination speckle interferometry. A preliminary analysis of the fringe patterns indicates that there are four orders in the u-pattern (Fig. 1.12a) and 17 orders in the v-pattern (Fig. 1.12c). Therefore, the horizontal diameter of the disk will expand by about 5 µm under the action of the applied load while the vertical diameter of the specimen will contract by about 21 µm. The computer generated patterns have been obtained utilizing a solution taken from the theory of elasticity [21]. The difference between theory and experiment is due to the modality of load application. The theoretical model assumes a point loading that generates a singular point in the theory

Table 1.1 Details of tested specimen

Material	Aluminum
Diameter	60 mm
Thickness	6.35 mm
Young's modulus	70 GPa
Poisson's ratio	$\nu = 0.336$
Pitch	1.222 μm
Applied load	3300 N
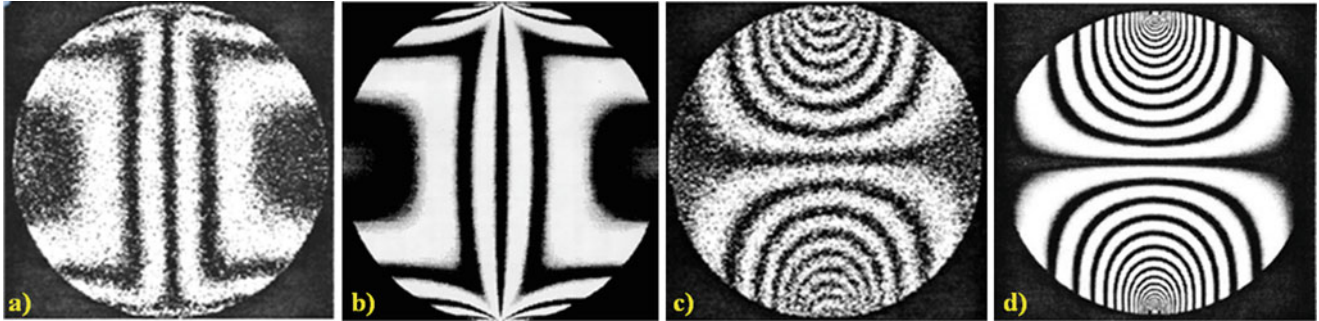
Fig. 1.12 (**a**) Recorded u-pattern; (**b**) computer generated u-pattern; (**c**) recorded v-pattern; (**d**) computer generated v-pattern
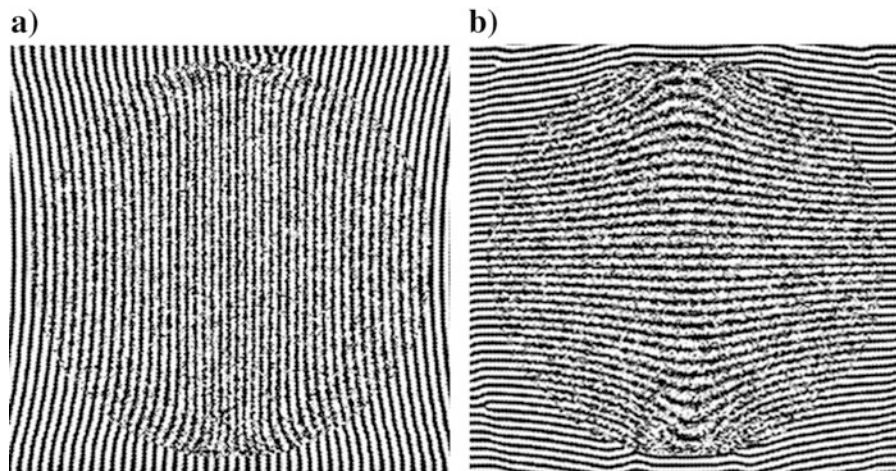
Fig. 1.13 u-extended (**a**) and v-extended (**b**) carrier fringes for the patterns of Fig. 1.12a, c

of elasticity solution while the actual loading is the result of the contact stresses between the loading application bar and the specimen. The fringe pattern analysis was not done directly on the recorded u–v patterns but in patterns where carrier fringes were introduced according to the theory presented in Sect. 1.6 (see Fig. 1.13).

There is an additional reason that is related to the adopted processing algorithms. In order to simplify the algorithm design, processing algorithms are derived under the assumption that computations are performed in a domain that asymptotically mimics an infinite domain. If an algorithm is developed for a finite region, this is a possible alternative, one has to design algorithms capable of changing as a function of the distance of the considered point to the boundaries (this distance depends on the domain geometry), which is not a simple task. An alternative choice is to keep the format of corresponding to an infinite domain and utilizing extended fringes to cover the full image area so that when one reaches the boundaries of the specimen either internal of external, the effect of the finite domain is removed [22]. The fringe extension is numerically feasible and converges faster if a high frequency carrier is present. The boundary problem is also present in both in the Gabor transform and in the wavelet transform. In the wavelet transform the edge problems present additional complications due to the different scales involved in the computation.

The above outlined path helps recovering the frequency information since it reduces the frequency spectrum to a narrower local band, resulting in a less populated energy matrix. This step is particularly useful for the Morlet wavelet transform since it reduces the number of scales and voices, thus reducing the amount of information that it is necessary to process. In this way the total amount of work required to compute frequencies is reduced and the accuracy of the results is enhanced.

In order to fully understand the comparison of different methods presented in this section, it is necessary to recall some fundamental relationships that arise from considering fringe patterns that encode displacement information as frequency modulated signals. The utilization of the energy representation of a signal in the coordinates-frequency space [23] provides powerful procedures to retrieve derivatives of displacements directly from fringe patterns without resorting to the differentiation of displacements. This alternative also simplifies the determination of displacements replacing complex unwrapping procedures by integration of the derivatives [24], a robust operation in the presence of noise. In view of the arguments presented in the preceding paragraphs, the comparison of results will be done between methods that provide displacements and methods that provide directly derivatives.

Figure 1.14 compares the theoretical distributions of u–v displacements (Fig. 1.14a) with the experimental values obtained from the FFT (Fig. 1.14b) and the experimental values obtained from the 2D Hilbert transform (Fig. 1.14c). The plots of Fig. 1.14b have been obtained starting from the extended fringe patterns of Fig. 1.13, utilizing the windowed FFT; the applied window is a Gaussian window like that displayed in Figs. 1.8 and 1.9. The plots of Fig. 1.14c have been obtained utilizing the 2D Hilbert transform described in this study. In the case of 2D Hilbert transform, displacement maps are presented in fashion of color contours; axes of symmetry of the disk are represented in order to highlight any asymmetry of the u–v fields.

Because of the pointed difference between the theory of elasticity solution and the experimental realization of this model, a FE solution was computed with the general purpose finite element software ANSYS. An elliptical distribution contact stress was assumed, the load contact width was determined from the recorded fringe patterns. Because of symmetry, only one half of the disk was modeled. Figure 1.15 shows the FE model and the u–v displacement maps computed by ANSYS; color bars for each displacement component are included in the figure. In order to obtain a mesh independent solution, convergence analysis was carried out and finally it was selected a mesh with 9242 elements (more specifically, in Fig. 1.15a, there are 9170 four-node plane elements, 60 target elements on the disk surface that may come in contact with 12 contact elements on the loading surface) connected by 9272 nodes. The FE model includes also auxiliary regions to apply the load to the disk and to support the specimen in the loading frame.

It can be verified that all these results are in good agreement. For example, the maximum u-displacement along the horizontal axis of the disk was determined as follows from experimental measurements and computations: 2.520 μm for the windowed FFT, 2.543 μm for the Hilbert transform, and 2.573 μm for ANSYS (similar to theoretical solution), respectively. A more detailed comparison is presented in Fig. 1.16 which shows the distribution of u-displacements measured or computed along the horizontal diameter: experimental data matched well the numerical simulations over the whole region of interest.

As far as it concerns the total v-displacement undergone by the vertical diameter of the disk, it can be seen that ANSYS predicted 20.7 μm (similar to theoretical solution) while the relative displacement extracted from the fringe pattern with windowed FFT and 2D Hilbert transform are about 18 μm. Both algorithms captured the asymmetry of the v-displacement pattern with respect to the horizontal diameter of the disk but could not reconstruct the whole pattern in the regions where the specimen is loaded (top) and fixed (bottom) as concentration of fringes caused a loss of resolution.

Figure 1.17a–d display the maps of strains ε_x and ε_y obtained directly from the fringe patterns of Fig. 1.12 utilizing the Gabor transform and the Morlet wavelet transform. It can be seen that the two patterns are very similar in their outlines.

Figure 1.18 shows the maps of ε_x and ε_y strains computed utilizing the patterns of Fig. 1.12 and the differentiation in the frequency space [25]. That is, derivatives are obtained directly from fringe patterns without going through the displacements similarly to what is done in the case of the Gabor and the Morlet wavelets. It is possible to see that distributions agree well in their outlines.

For a more detailed analysis, the distributions of strain ε_x along the horizontal axis obtained with the Gabor transform, the Morlet wavelet, the windowed FFT, the 2-D Hilbert transform and the finite element results are plotted in Fig. 1.19. Since values of contour lines do not match, data at the same locations along the control path are extracted. The agreement between the different methods is excellent (maximum strain is 190 με for Gabor/Morlet, 185 με for FFT, 193.5 με for Hilbert 2D and 201 με for finite elements) in spite of the fact that different numerical techniques and algorithms have been applied to obtain the corresponding values.

There is an exception that can be observed in the values of ε_y of Fig. 1.18 near the point of application of the load. In this region there is a large gradient of the displacements and this results in a concentration of fringes. If one processes the data and utilizes a single scale upon FT, filtering the field of the displacements is distorted because of the lack of spatial resolution in the region. Enlarging the scale of the contact region is possible to eliminate this error. The error is not present in the Gabor transform and in the Morlet wavelet because of a more comprehensive filtering resulting from the utilization of multiple scales.
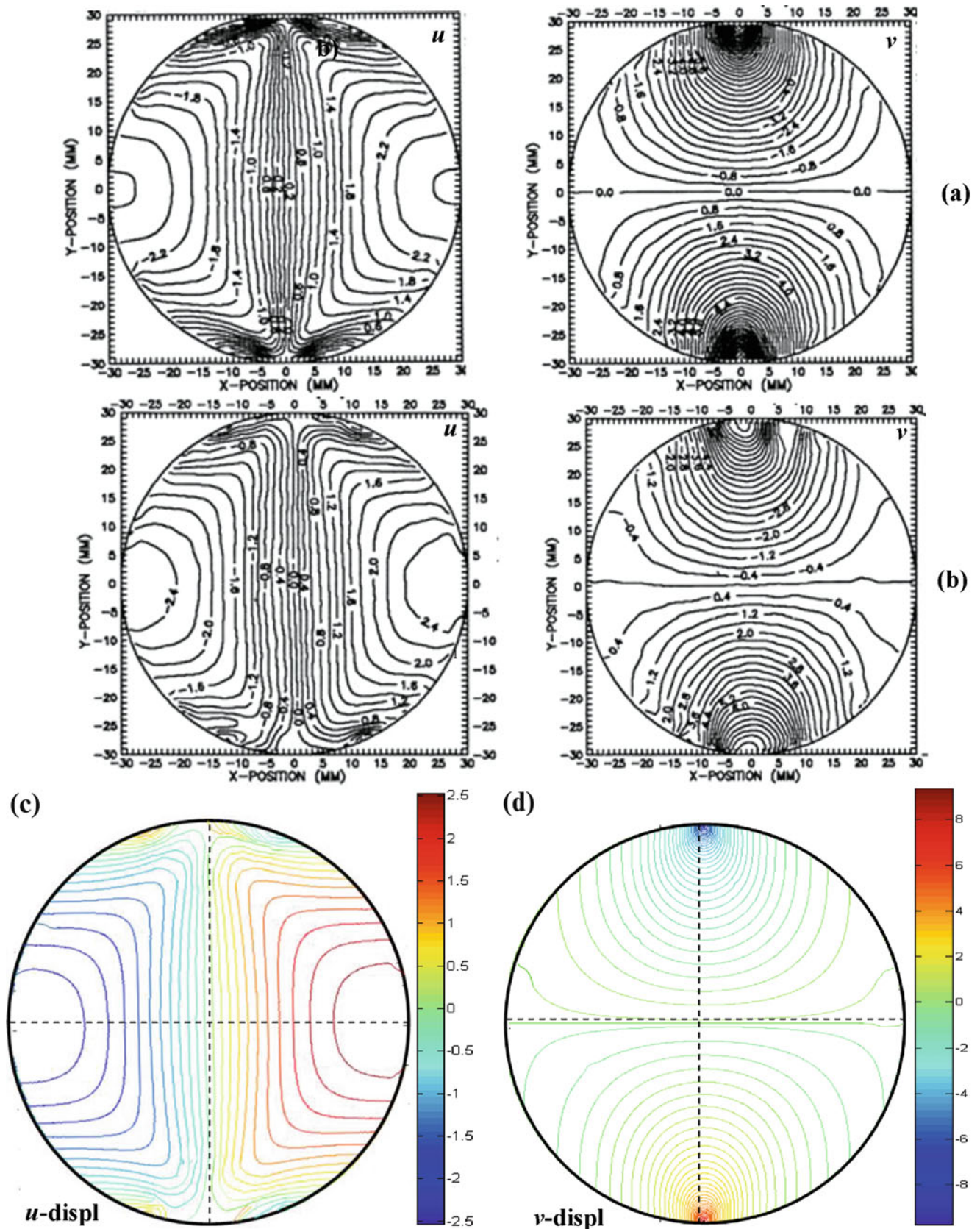
Fig. 1.14 (**a**) Theoretical *u–v* displacements (elasticity solutions from [21]); (**b**) experimental *u–v* displacements, windowed FFT [22]; (**c, d**) experimental u–v displacements, 2D Hilbert transform (present study). All scales are expressed in microns
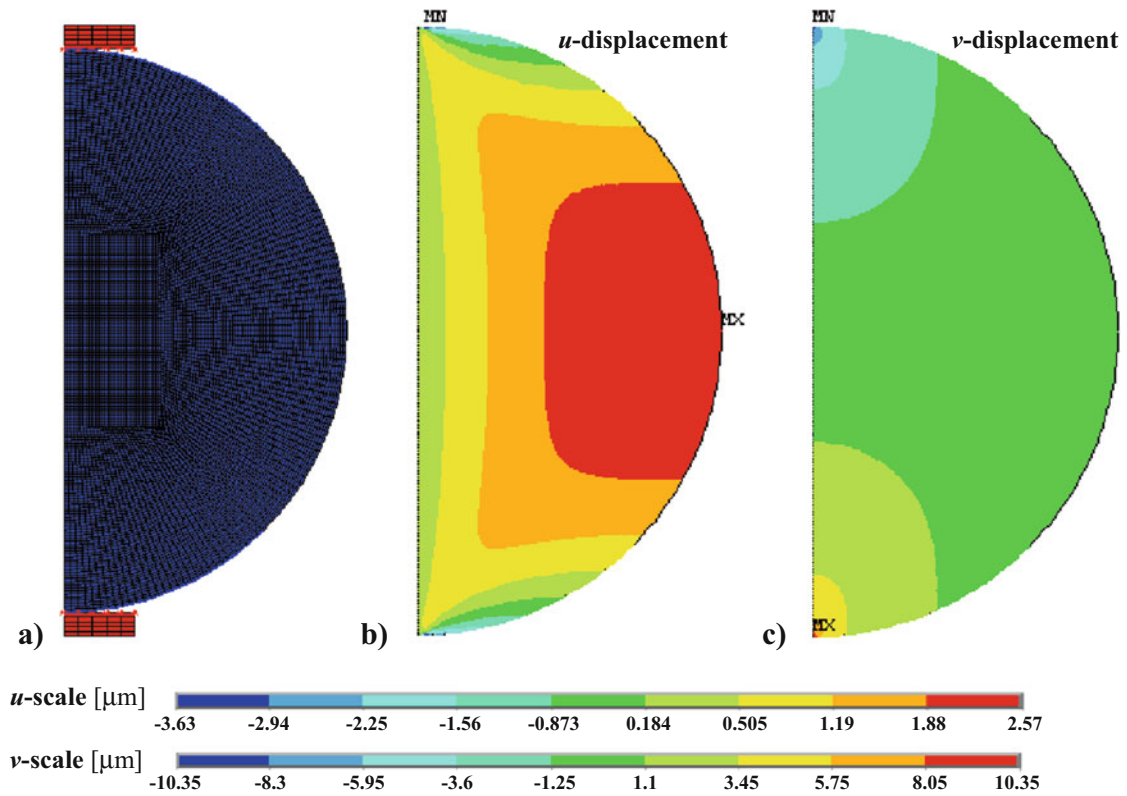
u-scale [μm]

| -3.63 | -2.94 | -2.25 | -1.56 | -0.873 | 0.184 | 0.505 | 1.19 | 1.88 | 2.57 |

v-scale [μm]

| -10.35 | -8.3 | -5.95 | -3.6 | -1.25 | 1.1 | 3.45 | 5.75 | 8.05 | 10.35 |

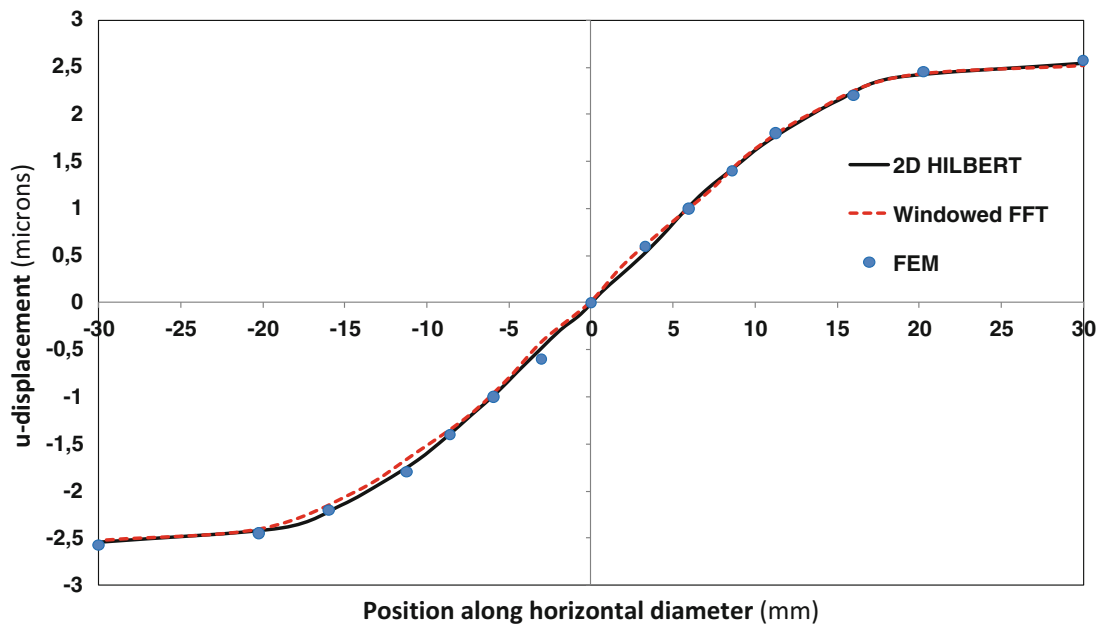**Fig. 1.15** Finite element model (**a**) and u-v displacement fields (**b-c**) computed by ANSYS for the disk under diametrical compression

Fig. 1.16 Comparison of displacement fields obtained from fringe processing techniques and FE simulation for the horizontal diameter of the disk
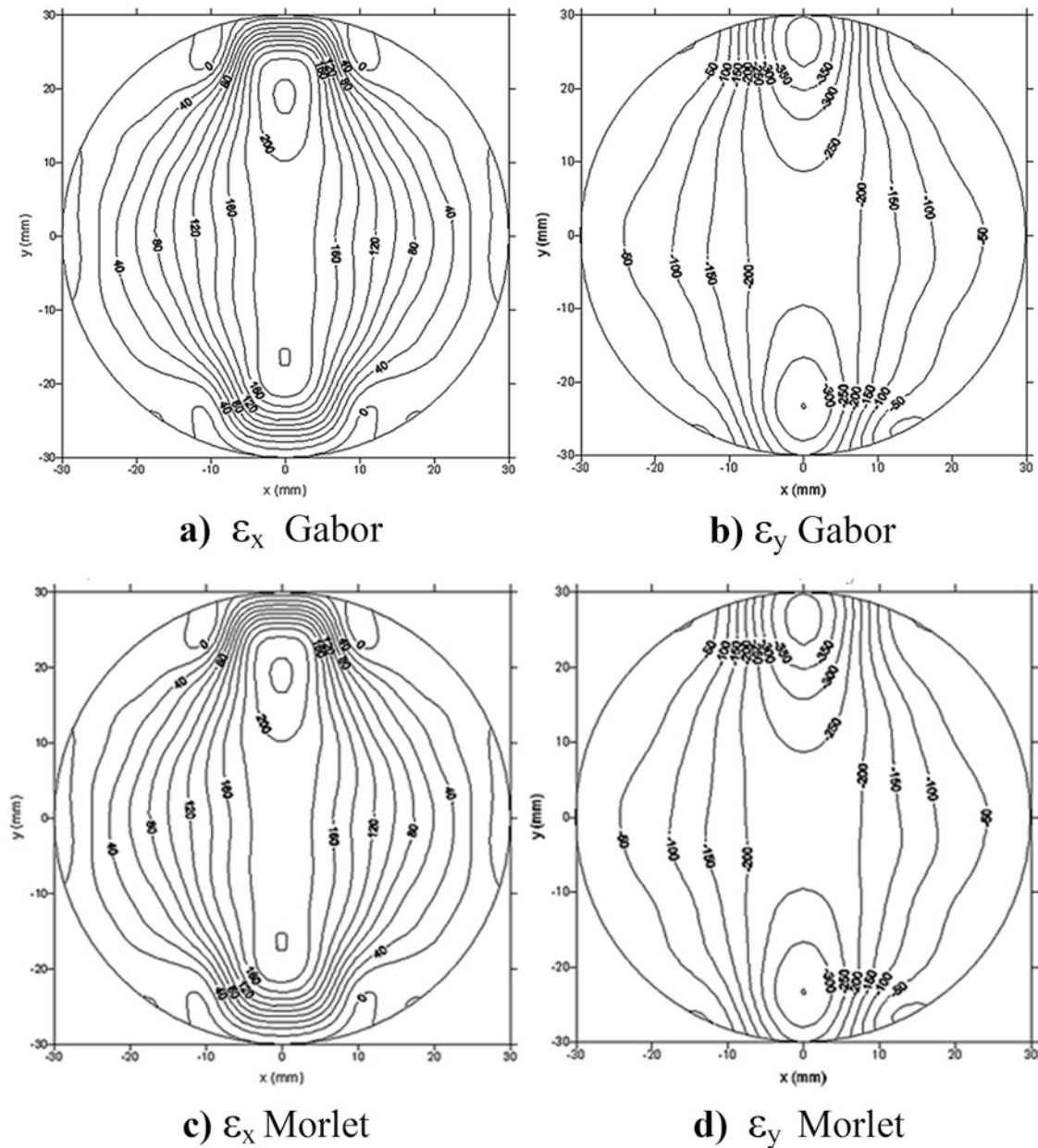
Fig. 1.17 Strains of the disk under diametrical compression utilizing the Gabor transform and the Morlet wavelet transform. (**a**) ε_x Gabor. (**b**) ε_y Gabor. (**c**) ε_x Morlet. (**d**) ε_y Morlet

1.10 Summary and Conclusions

Fringe pattern analysis is a sub discipline that is contained in the more general discipline of signal analysis. Early developments of signal analysis were done in applications to electrical engineering and communications engineering, thus basically concentrated in one dimensional signals. Image analysis has brought the generalization of this body of knowledge to two dimensions and more recently to 3-D. The fundamental idea of monogenic functions is to enclose all the mathematical derivations that apply to signals in general in a comprehensive multiple dimensional approach. An example of this comprehensive approach is found in [5, 8]. These references deal with a more general problem than the one dealt in the present publication. They deal with processes of statistical optimization in the computation of components of signals such as orientation θ, phase ϕ and signal amplitude A of the general monogenic signal defined by Eq. (1.37) and in many cases solutions of problems encountered in one dimensional signals [16, 17].
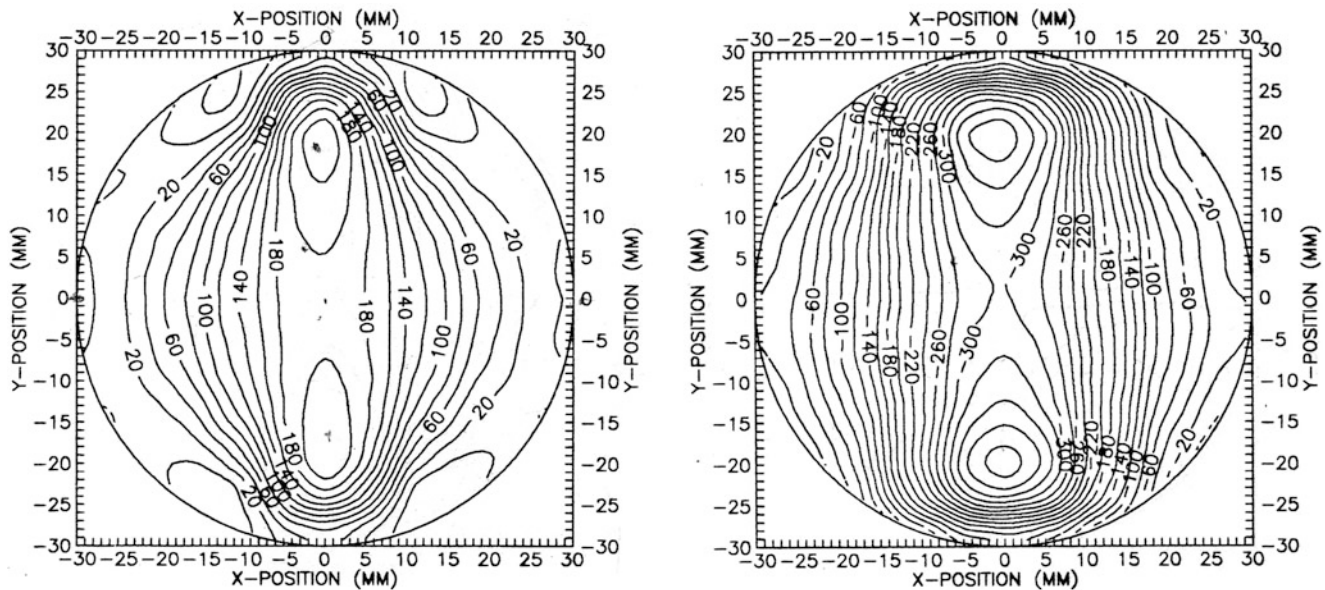
Fig. 1.18 Strains ε_x and ε_y of the disk under diametrical compression obtained from the FFT by differentiation in the frequency space [25]
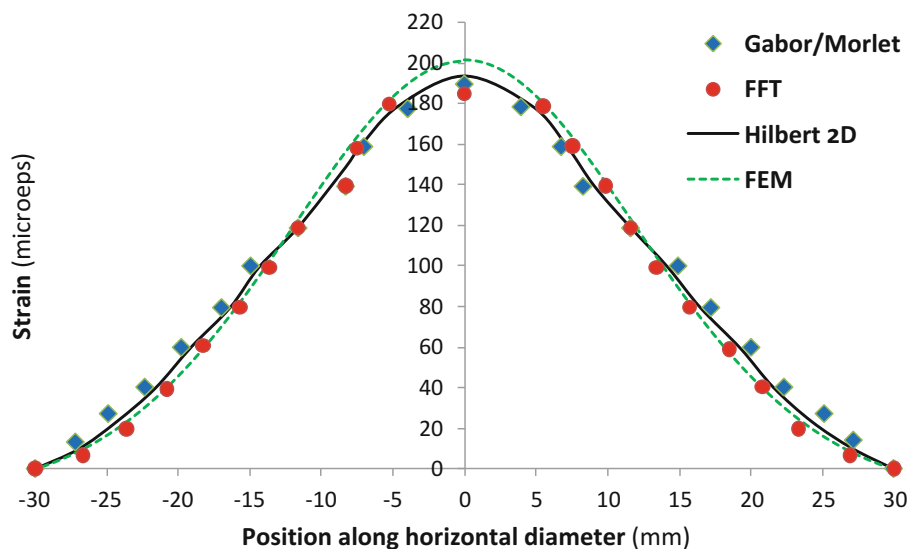

Fig. 1.19 Comparison of strains ε_x obtained by Gabor transform—Morlet wavelet and windowed FFT with finite element computation (control path corresponds to the horizontal diameter of the disk)

The current paper was devoted to put into the general theory of monogenic function, 2D fringe pattern analysis. This effort has an interesting sequel, the Poincare sphere well known to photoelasticians appears again on a different topic, the retrieval of displacement fields and their derivatives. This finding opens a new and interesting field of research in view of the fact that the phase concept has been attached in the past to the Poincare sphere [9–11]. As part of these developments the relationship between the FFT transform and the generalized Hilbert transform is analyzed. An important conclusion arrived to in [1], the equivalence of the multi phase method and the in-quadrature method, is extended to two dimensional signals. A more detailed analysis of these questions is presented in the discussion of the local phase concept closely related to the narrow band condition that makes more likely that the generalized Hilbert-transform to be a valid representation of the captured signals. Finally, the role of other transforms (Gabor transform and wavelet transform) in fringe pattern analysis is considered. The analysis of an actual fringe pattern, the fringes corresponding to a disk under diametrical compression by all the above described methods provides results that are numerically very close and also close to the finite element solution in spite of the quite different algorithms utilized in each case. This is also a very valuable outcome that provides a comparison of different approaches from the point of view of reliability and accuracy of the final results.

References

1. Sciammarella, C.A., Lamberti, L.: Mathematical models utilized in the retrieval of displacement information encoded in fringe patterns. Opt. Lasers Eng. **77**, 100–111 (2016)
2. Oppenheim, A.V., Lim, J.S.: The importance of phase in signals. Proc. IEEE **69**(5), 529–541 (1981)
3. Gabor, D.: Theory of communications. Part 1: the analysis of information. J. Inst. Electr. Eng. **93**(26), 429–441 (1946)
4. Papoulis, A.: Systems and Transforms with Applications to Optics. McGraw-Hill, New York (1968)
5. Felsberg, M.: Low-level image processing with the structure multivector. PhD Dissertation D-24098, Christian-Albrechts University of Kiel, Germany (2002)
6. Felsberg, M., Sommer, F.: The monogenic signal. IEEE Trans. Signal Process. **49**(12), 3136–3144 (2001)
7. Wietzke, L., Sommer, G.: The 2D analytic signal. Technical Report No. 0802, Christian-Albrechts University of Kiel, Germany (2008)
8. Langley, K., Anderson, S.J.: The Riesz transform and simultaneous representations of phase, energy and orientation in spatial vision. Vision Res. **50**(17), 1748–1765 (2010)
9. Pancharatnam, S.: Generalized theory of interference, and its applications—part I: coherent pencils. Proc. Indian Acad. Sci. A **44**, 247–262 (1956)
10. Berry, M.V.: The adiabatic phase and Pancharatnam's phase for polarized light. J. Mod. Opt. **34**(11), 1401–1407 (1987)
11. Gutierrez-Vega, J.C.: Pancharatnam–Berry phase of optical systems. Opt. Lett. **36**(7), 1143–1145 (2011)
12. Holm, D.D.: Geometric mechanics II: the poincare sphere $S^2 \subset s^3$. In: Lecture Notes. Imperial College, London, UK (2013)
13. Gaskill, J.D.: Linear Systems, Fourier Transforms, and Optics. Wiley, New York (1978)
14. Rabiner, L.R., Gold, B.: Theory and Application of Digital Signal Processing. Prentice-Hall International, London (1975)
15. Troncoso Romero, D.E., Dolocek, G.J.: Digital FIR Hilbert Transformers: Fundamentals and Efficient Design Methods. Intech Open, Rijeka (2012)
16. Boashash, B.: Estimating and interpreting the instantaneous frequency signal. Part 1: fundamentals. Proc. Inst. IEEE **80**(4), 520–538 (1992)
17. Boashash, B.: Estimating and interpreting the instantaneous frequency signal. Part 2: algorithms and applications. Proc. IEEE **80**(4), 540–568 (1992)
18. Bedrosian, E.: A product theorem for Hilbert transforms. Memorandum RM-3439-PM. The Rand Corporation, Santa Monica, CA (1962)
19. Bedrosian, E.: Hilbert transform of bandpass functions. Proc. IEEE **51**, 868–869 (1963)
20. Nuttall, A.H., Bedrosian, E.: On the quadrature approximation to the Hilbert transform of modulated signals. Proc. IEEE **54**(10), 1458–1459 (1966)
21. Timoshenko, S.: Theory of Elasticity. McGraw-Hill, New York (1970)
22. Sciammarella, C.A., Bhat, G.: Two-dimensional Fourier transform methods for fringe pattern analysis. In: Proceedings of the VII International Congress on Experimental Mechanics, vol. II, Las Vegas, NV (1992)
23. Sciammarella, C.A., Kim, T.: Frequency modulation interpretation of fringes and computation of strains. Exp. Mech. **45**(5), 393–403 (2005)
24. Sciammarella, C.A., Lamberti, L.: Basic models supporting experimental mechanics of deformations, geometrical representations, connections among different techniques. Meccanica **50**(2), 367–387 (2015)
25. Sciammarella, C.A., Narayanan, R.: The determination of the components of the strain tensor in holographic interferometry. Exp. Mech. **24**(4), 257–264 (1984)