DEFINITION OF THE OPTICAL WAVE PHASE AND A MULTIDIMENSIONAL ANALYTIC SIGNAL

V.A. Tartakovski

Institute of Atmospheric Optics,
Siberian Branch of the Russian Academy of Sciences, Tomsk
Received December 5, 1996

The multidimensional phase modulation of a light wave is performed in adaptive optics, probably for the first time. In this connection exist the necessity of formal definition of a phase in the spatiotemporal area. With that end in view the concept of the analytic signal is developed in this paper.

1. INTRODUCTION

The representation of an oscillation or wave process in the form of two functions namely the amplitude and the phase requires a non–contradictory definition of these concepts. The wave equation or equation, describing the oscillation, does not contain such a definition. Therefore some additional reasoning are necessary. There are only few papers[2,3,5–8,9,12,14,15–17,20,26–31] in which the definition problems of the amplitude and the phase are considered for the one–dimensional oscillation.

There are various ways of introducing the amplitude and the phase. Analysis of these ways was conducted in Refs. 6, 8, 31. The conclusions can be formulated as follows:

– All definitions give identical result for harmonic oscillations, the sine and cosine, and this is a necessary condition of the definition correctness.

– The results of various ways of definition do not coincide for narrow-band signals. This discrepancy reduces with the reduction of relative width of the signal frequency band.

– The most general definition of the amplitude and the phase can be introduced using the Gabor's analytic signal (AS).[2]

For a given real function $U(x)$ the analytic signal $W(x)$ is a complex function,

$$W(x) = U(x) + iV(x) ,$$

$$V(x) = \frac{1}{p}\ \text{v.p.} \int\limits_{-\infty}^{\infty} \frac{U(s)}{x - s} ds = \underset{x}{\mathbf{H}} U(x). \qquad (1)$$

Here the improper integral is determined as the Cauchy principal value (v.p.) when s tends to infinite and x equals □s. $V(x)$ is the imaginary part of the AS. It is Hilbert transform of its real part $U(x)$, and the operator of the Hilbert transformation over argument x is $\underset{x}{\mathbf{H}}$. Then the amplitude and the phase are calculated in a usual way

$$a(x) = \sqrt{U^2(x) + V^2(x)} , \quad \varphi(x) = \arctan \frac{V(x)}{U(x)} .$$

In Ref. 31 it is shown, that the operator $\mathbf{H}$ is the unique linear operator, for which the following equality holds:

$$\underset{x}{\mathbf{H}} \cos(\alpha_c x + \varphi_o) = \sin(\alpha_c x + \varphi_o) ,$$

where $\alpha_c > 0$ and φ_o are unknown constants, having the meaning of the carrier frequency and the initial phase. Therefore, the definition of the amplitude and the phase is made in the same way for signals with different frequency spectra.

The Hilbert transform is equivalent to multiplication in the frequency domain by the sign function:

$$\underset{x}{\mathbf{H}} U(x) = \int\limits_{-\infty}^{\infty} \text{sgn}\, \alpha e^{-i\alpha x} d\alpha \int\limits_{-\infty}^{\infty} U(y) e^{i\alpha x} dy . \qquad (2)$$

therefore, the Fourier transform of an analytic signal $W(x)$ occupies only one half of the frequency axis thus being the one–sided or causal transform.

The important property of AS application for the theory of modulation[26] is the possibility of separating out the amplitude–modulated signal at low-modulation frequency using the formula

$$\underset{x}{\mathbf{H}} \Omega(x) U(x) = \Omega(x) \underset{x}{\mathbf{H}} U(x) , \qquad (3)$$

here $\Omega(x)$ and $U(x)$ are real functions, their Fourier transforms do not overlap in the frequency domain, and $U(x)$ exists at higher–frequencies than $\Omega(x)$. Other properties of the Hilbert transform are discussed in Refs. 4, 6, 8, 15.

The optical wave is the four–dimensional function and it creates a problem for application of the AS.[25,31] There arise questions, for which of the coordinates the Hilbert transform should be done, when it is possible and in what features various analytic signals, thus arising, will correspond to one another. The preservation of uniqueness of the phase definition causes the necessity of generalization of the analytic signal concept to the multidimensional case.

0235-6880/97/03 189-09 $02.00

In this paper we numerically study the definition of the amplitude and phase by the analytic signal and compare the amplitude and the phase of the normal random process, determined by the AS when $\alpha_c > \Delta\alpha$, with the amplitude and the phase of the same process at moving its spectral bands to zero frequency and the subsequent exchange.

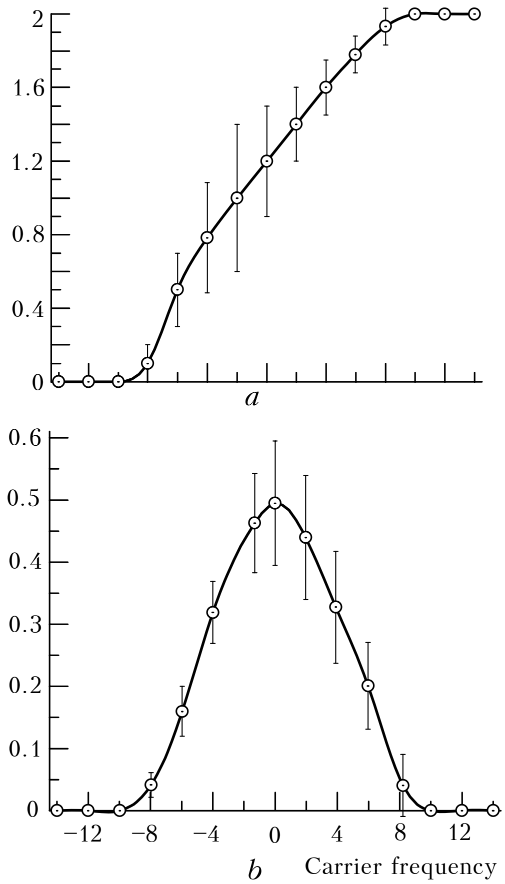

FIG. 1. Estimates of the discrepancies ε_φ (Fig. 1a) and ε_a (Fig. 1b) between the phase and the amplitude of a complex function determined in the experiment and calculated using the analytic signal as functions of the carrier frequency. The rms deviations are shown as the confidence intervals.

Spectral bands occupied, in the frequency domain, the intervals of 20 read–outs, with the total number of read–outs of this process N being equal to 256. The carrier frequency $\alpha_c \in [-14, +14]$, and the Nyquist frequency is equal to 128. The estimates are calculated as the average quotient of the norms in the space $L_2(N)$ by the formulas

$$\varepsilon_a = \sqrt{\sum_{i=1}^{N} (a_i - \tilde{a}_i)^2 / \sum_{i=1}^{N} a_i^2}\ ,$$

$$\varepsilon_\varphi = \sqrt{\sum_{i=1}^{N} (\varphi_i - \tilde{\varphi}_i)^2 / \sum_{i=1}^{N} \varphi_i^2}\ ,$$

where a_i and φ_i are the read–outs of the amplitude and the phase of the initial process, and $\tilde{a}_i$, $\tilde{j}_i$ are thecorresponding read–outs of the process with a varying position of the spectral bands.

The results of the experiment are shown in Fig. 1. It is seen, that the estimate ε_a is equal to zero, when the spectral bands do not involve the point $\alpha = 0$. Hence the amplitude is invariant relative to sign of the carrier frequency, whereas the phase changes its sign and the estimate ε_φ changes its value from zero to two. If the point $\alpha = 0$ falls inside a spectral band, the estimate ε_a differs from zero, and ε_φ takes an intermediate value between zero and two.

Thus, the possibility of using the analytic signal becomes problematic for the amplitude and phase definition, when $\alpha_c \sim \Delta\alpha$. Such situations usually take place in the interferometry, when the interference fringes essentially change their widths and curvature, when they have a ring shape.

2. THE FUNCTION CLASS FOR A MODEL REPRESENTATION OF THE WAVE

The basic property of the mathematical model, i.e. its applicability to research, is in many respects defined by properties of the functions from which model is composed. When choosing a class of functions for representation of the physical process we shall pay attention to two circumstances. The first is a univalent representation of the wave function and the interferogram by a discrete series of read–outs, necessary for making numerical analysis. Secondly, the existence of the Hilbert transform, by which the AS is introduced.

The representation of functions in the form of a discrete series of read–outs is performed on the basis of Kotel'nikov theorem if the function has the finite Fourier transform.

Under the Paley–Wiener theorem for quadratic integrable functions or the Paley–Wiener–Schwarz theorem,[15] if the spectrum has a singularity in the form of $\delta-$ functions and their derivatives, the functions having a finite spectrum are entire analytical functions of the exponential type (EFET). Moreover, if one considers only functions limited on the real axis, such EFET will be the function of a class "A", Ref. 4 or a class "B", Ref. 22.

Because of the EFET limitedness its Fourier transform or the spectrum are absolutely integrable including the case, when this EFET is not quadratically integrable. The reverse statement is also true and as the Hilbert transform of EFET does not infringe the absolute integrability of its spectrum, the Hilbert transform also will be a limited function. The Hilbert transform exists also for the functions with continuous derivative, and for a more wide class of functions, which satisfy the Hölder condition,[31] but because of the necessity of discrete representation these classes of functions are inapplicable here.

Let us now find the consequences of using the entire function for the representation of an interference pattern. The theorem from Ref. 1 allows the difference module between a real continuous function $U(x)$ and its approximation EFET from the class "A" to be an arbitrary small value for all x, if

$$\limsup_{p \to \infty} \frac{\log M(p)}{\log p} = 0,$$

where

$$M(p) = \max |U'(x)| \quad \text{at } |x| < p \,. \tag{4}$$

From this expression restrictions on the growth of the amplitude and the frequency of the interference fringes follow. This growth should not be more or less, than any power of x at $x \to \pm \infty$.

It is known *a priori* that the width of the spatial frequency spectrum and the temporal spectrum of a parabolic and quasi–monochromatic wave function $U(x, y, z, t)$ is narrow. That is $\Delta\omega / \omega_c \ll 1$, $\Delta\alpha / \alpha_c \ll 1$, where ω_c is the carrier frequency, and $\Delta\omega$ is the half–width of a temporal spectrum; α_c is the carrier frequency, and $\Delta\alpha$ is the half–width of a spatial frequency spectrum.

These properties of a light wave can be most naturally expressed, if $U(x, y, z, t)$ is the entire exponential function of each variable. Thus the physical properties namely the monochromaticity and the parabolicity are transferred to the wave approximation $U(x, y, z, t)$. The approximation is not already the exact solution of a wave equation. However there is no necessity in such a solution since the concepts of the amplitude and the phase exist only in connection with their measurement or the definition and do not follow from a wave equation.

3. THE WAYS TO INTRODUCE THE ANALYTIC SIGNAL

Let us now consider the particular solution of a scalar wave equation, describing the propagation of a quasimonochromatic wave in a homogeneous medium along the positive direction of the z-axis, which can be written as

$$W(x, y, z, t) =$$

$$= \int_0^\infty d\omega \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\alpha, \beta, \omega) \exp i(\alpha x + \beta y + \gamma z - \omega t) \, d\alpha d\beta \,, \tag{5}$$

where $\alpha, \beta, \gamma = (k^2 - \alpha^2 - \beta^2)^{1/2}$ are the spatial frequencies, ω is the temporal frequency, and $k = 2\pi / \lambda = \omega / c$ is the modulus of the wave vector, c is the speed of light, $S(\alpha, \beta, \omega)$ is a spatiotemporal spectrum in the plane $z = 0$. The sign of the spatial frequency γ is chosen, as it is known,[10,18] from the condition of damping of the evanescent waves at $z \to +\infty$, and the sign of the frequency ω from the condition that the wave front moves in the same direction.

Such a representation is sufficient for describing light propagation process in an optical system, for example, in an interferometer. The frequencies γ and ω do not change the sign, therefore, the function $W(x, y, z, t)$ is the AS of the variables z and t. The invariance condition of the sign can be violated because of the occurrence of some waves, reflected from surfaces of the optical system and propagated toward to the main wave. However, shall consider the amplitude of the reflected wave to be negligible, since the optical system allows the separation of the reflected wave from the direct wave in the spatial frequency region α, β.

The Analytic Signal on a Plane

It is interesting to analyze fields in a recording plane, normal to the z–axis. When the variables z and t are fixed, the spectrum $S(\alpha, \beta, \omega)$ in the general case and for all ω will be localized around of the coordinate origin in the plane $\alpha\beta$, and consequently the function $W(x, y, z_o, t_o)$ is not the AS in some cross–sections of the plane xy.

Let us turn the recording plane around the x–axis and y–axis at the point $(0, 0, z_o)$ at some angle θ normally to the vector $\{\eta, \zeta, \sqrt{1 - \eta^2 - \zeta^2}\}$.

In this new plane $p(x, y, z)$ coordinate z already is not fixed, it changes according to the equation

$$z = z_o - \frac{\eta x}{(1 - \eta^2 - \zeta^2)} - \frac{\zeta y}{(1 - \eta^2 - \zeta^2)}$$

by substituting it in Eq. (5), we obtain

$$W(p, t) =$$

$$= \int_0^\infty d\omega \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\alpha, \beta, \omega) \exp i[(\alpha)x + (\beta)y + \gamma z_o - \omega t) \, d\alpha d\beta, \tag{6}$$

where

$$(\alpha) = \alpha - \gamma\eta/\sqrt{1 - \eta^2 - \zeta^2} \,,$$

$$(\beta) = \beta - \gamma\zeta/\sqrt{1 - \eta^2 - \zeta^2} \,.$$

The angle θ and the projections η and ζ of the normal vector, connected with it, can be chosen of such a size, that the factors (α), (β) at the variables x and y do not change the sign, for example, at

$$\eta = \zeta < -b/k \,, \tag{7}$$

where $b = \max(|\alpha|, |\beta|)$, for $S(\alpha,\beta,\omega) \neq 0$. Therefore, the functions $\cos(\alpha)x$ and $\sin(\alpha)x$, $\cos(\beta)y$ and $\sin(\beta)y$ are connected with each other by the Hilbert transform over x and y accordingly, and the function $W(x, y, z, t)$ in the plane $p(x, y, z)$ is the analytic signal of x and y, Eq. (6).

Let us estimate numerically the angle θ, necessary to obtain a given resolution in the recording plane, for $\lambda = 0.63 \cdot 10^{-3}$ mm, $k \approx 10^4$ mm^{-1}, at $\zeta = \eta$, having in mind the ratio $\eta = 2\pi \cos\theta/\lambda = 2\pi/\tau$, where τ is a spatial period of the wave in this plane. The results are given in Table I.

TABLE I.

η/π, mm^{-1}	≈ 2000	≈ 400	200	100
k^2/η^2	≈ 2.6	≈ 62	≈ 256	1000
τ, mm	0.001	0.005	0.01	≈ 0.02
θ, °	≈ 51	≈ 83	≈ 86	≈ 88

From this table it is seen, that the angle θ is close to 90° within the framework of the geometrical optics. Such an angle can be realized in the optical systems for a large number of applications, for example, in the interference testing of an optical surface. Experimentally the angle θ is chosen to be of such a value, that the light beam, incident on a recording plane, is on the one side from some normal to this plane. It provides the causality condition for the angular spectrum of the function $W[p(x, y, z), t]$.

The Analytic Signal on a Line

Let the parametric equations of a line $l(t)$ in the three–dimensional space be

$$\begin{cases} x = x_o + v_x(t)t \ , \\ y = y_o + v_y(t)t \ , \\ z = z_o + v_z(t)t \ . \end{cases}$$

Let us elucidate under which conditions the wave is the analytic signal on this line. Having substituted the parametric equations in Eq. (5), we obtain that

$$W[l(t)] = \int\limits_0^{\infty} d\omega \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} S(\alpha, \beta, \omega) \exp i[(\circ)t - \omega t + \varphi_{\alpha\beta}] d\alpha\, d\beta, \quad (8)$$

where

$$(\circ) = \alpha v_x(t) + \beta v_y(t) + \gamma v_z(t), \quad \varphi_{\alpha\beta} = \alpha x_o + \beta y_o + \gamma z_o \ .$$

Let us assume that $\eta = \max(|\alpha|, |\beta|, |\gamma|)$ at $S(\alpha, \beta, \omega) \neq 0$, $v = \max(|v_x(t)|, |v_y(t)|, |v_z(t)|)$ and then we find that $\max|(\circ)| < 3\eta v$. Let us show using a numerical example, that $q = 3v\eta/\omega_c = 3v\cos\theta/c < 1$, at $\Delta\omega/\omega_c \ll 1$, where ω_c is the carrier frequency of the temporal spectrum, $\Delta\omega$ is a half–width of the temporal spectrum. The initial data and the results for a wave scanning in a frame are shown in Table II.

TABLE II.

Size of the frame, cm	1×1
Number of lines in the frame	3×10^4
Scanning time of the frame, sec	1×10^{-3}
θ, °	51
Speed of light in the vacuum, cm/sec	3×10^{10}
The common length of the scan, cm	3×10^4
Scanning speed v, cm/sec	3×10^7
$\cos\theta$	0.63
$q \approx$	2×10^{-3}

From the estimate for q obtained follows the inequality $(\circ) - \omega_c < 0$. It is obvious, that some acceptable scanning parameters of the frame can also be selected for the interval $(\omega_c \pm \Delta\omega)$.

Thus, the factor at the variable t in Eq. (8) does not change the sign, and the function $W[l(t)]$ is the analytic signal on a line with the parameter t.

The Analytic Signal in the Interferogram

Not being an analytic signal, the field can receive this property due to interference. Let us consider a two–dimensional interference pattern $G(x, y)$ of the analyzed field with the unit plane reference wave:

$$G(x, y) = |W(x, y, z_o, t_o) + \exp i(\eta x + \zeta y)|^2 = 1 + |W(x, y, z_o, t_o)|^2 + W^*(x, y, z_o, t_o) \exp i(\eta x + \zeta y) + W(x, y, z_o, t_o) \exp -i(\eta x + \zeta y) \ . \quad (9)$$

As well as in the previous case, it is necessary to choose the values η, ζ according to Eq. (7). The inclination of the reference wave front related to them, will transform into a displacement of the finite spatial spectrum of two last components in Eq. (9), which are the conjugate analytic signals of both x and y coordinates.

The Amplitude of the Analytic Signal and the Envelope of the Parametric Function Family

The solution of the wave equation for vacuum Eq. (5) can involve a multiplicative constant, for example, in the form $\exp i\varphi_o$. The change of the initial phase φ_o creates the parametric family of the functions $W(x, y, z, t)\exp i\varphi_o$. The functions of the family may have real envelope $a(x, y, z, t)$, which contacts them at some points and does not depend on the initial phase.

Let us find the relation between the envelope and the amplitude of the field.[8,19] When finding the envelope[11] it is necessary to solve the system of equations

$$\begin{cases} \mathrm{Re}\,[W(x, y, z, t)\exp i\varphi_o] = a \ , \\ \dfrac{\partial}{\partial\varphi_o}\mathrm{Re}\,[W(x, y, z, t)\exp i\varphi_o] = 0 \ . \end{cases}$$

Having substituted expression (5) in the second equation of the system and taking derivatives, we shall find

$$\begin{cases} \mathrm{Re}\,[W(x, y, z, t)\exp i\varphi_o] = a \ , \\ \mathrm{Im}\,[W(x, y, z, t)\exp i\varphi_o] = 0 \ . \end{cases}$$

After multiplication of the second equation by the imaginary unit and summing it with the first one we obtain the square of a modulus of a sum as the solution of the system

$$a^2(x, y, z, t) = |W(x, y, z, t)|^2.$$

For the monochromatic wave time t can be a parameter of the family. Having excluded from Eq. (5) the integration over the frequency ω, we can obtain the system of equations for the envelope

$$\begin{cases} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|S(\alpha, \beta)|\cos[\alpha x + \beta y + \gamma z - \omega t + \arg S(\alpha, \beta)]\,d\alpha\,d\beta = a \ , \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|S(\alpha, \beta)|\sin[\alpha x + \beta y + \gamma z - \omega t + \arg S(\alpha, \beta)]\,d\alpha\,d\beta = 0 \ . \end{cases}$$

Similarly to the previous case we obtain the solution $a^2(x, y, z) = |W(x, y, z)|^2$. The square of the parametric family envelope does not in this case depend on time as the intensity of the field.

Thus, the envelope and the amplitude of the AS are identical in the above considered cases, but the possibility of parametrization is limited by the physical properties of the problem, by the wave equation. In a particular case this property is monochromaticity, in a more general case it is the absence of passive medium and sources.

4. INTERRELATIONS OF CAUSALITY OF FOURIER TRANSFORM OF A COMPLEX FUNCTION AND MONOTONIC BEHAVIOR OF ITS PHASE

By analyzing a recorded interference pattern, for example, the expression (9), without the account for the process of its obtaining, one can hardly find out, whether the field $W(x, y)$ has the causal spatial spectrum, especially when the interference fringes are curvilinear and have variable width. Having in mind that the shift of a spectrum in the frequency region is equivalent to addition of a linear function to a signal phase, we connect the causality of a complex function spectrum to the monotonicity of its phase.[21] The monotonic property manifests itself in the experiments in the fact that the interference fringes have full profiles in the linear cross–sections of the interferogram.

Let us take the Bernstein inequality[13,21]:

$$\max\left|\frac{dW(x)}{dx}\right| \le \Delta\alpha \max|W(x)| \ , \qquad (10)$$

here the function $W(x) = a(x)\exp i\varphi(x)$ belongs to a class of functions with a finite spectrum, $\Delta\alpha$ is the halfwidth of the spectrum.

Let $a(x) = \mathrm{const}$, then we find $\max|\varphi'(x)| \le \Delta\alpha$. According to the theorem on the shift of a spectrum in the frequency region the function $\exp i[\varphi(x) + \Delta\alpha\, x]$ has the causal spectrum, and from the Bernstein inequality it follows that $\varphi(x) + \Delta\alpha\, x$ is obviously a monotonic function. That is causality is a sufficient condition of monotonicity. Therefore, cases are possible, when the monotonicity is present, but the causality is not present.

We consider the case, when $a(x) \ne \mathrm{const}$. From Bernstein inequality we obtain

$$\Delta\alpha^2 \ge \frac{\max[a'^2(x) + a^2(x)\,\varphi'^2(x)]}{\max a^2(x)} \ge$$

$$\ge \frac{\max[a^2(x)\,\varphi'^2(x)]}{\max a^2(x)} = r^2\max\varphi'^2(x) \ ,$$

where $r^2 \le 1$, then $\Delta\alpha \ge r\max|\varphi'(x)|$. The latter inequality allows three situations: the causality without monotonicity, the monotonicity without causality, and the presence of both properties.

The Dispersion Causality

Let us now elucidate which statement is true for the functions that, probably, have no finite spectrum. Let us define the half-width of the Fourier spectrum $S(\alpha)$ for the function $W(x)$ in a dispersion sense[23]

$$\Delta\alpha^2 = \int\limits_{-\infty}^{\infty}|S(\alpha)|^2(\alpha - \alpha_c)^2\,d\alpha =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\left\{a^2(x)\left[\varphi'(x) + \alpha_c\right]^2 + a'^2(x)\right\}dx \ .$$

We assume that $\alpha_c = 0$ and obtain the estimate $\Delta\alpha^2 \le A^2\max\varphi'^2(x) + B^2$, where

$$A^2 = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}a^2(x)\,dx \ ; \quad B^2 = \frac{1}{2p}\int\limits_{-\infty}^{\infty}a'^2(x)\,dx \ .$$

Let, without loss of generality, $A^2 = 1$, then the spectrum of the function:

$$a(x) \exp i \left\{\varphi(x) + x \sqrt{B^2 + \max \varphi'^2(x)}\right\}$$

is causal in the dispersion sense, and its phase is a monotonic function, as it follows from the inequality:

$$\Delta\alpha < \sqrt{B^2 + \max \varphi'^2(x)} > \max \varphi'(x).$$

Hence, at a reasonably large linear phase shift the monotonic behavior and the dispersion causality are observed in common.

Let $a(x) = \text{const}$, $\varphi'(x) \in L_2$. We find $\max \varphi'(x) > \Delta\alpha$, therefore, the monotony at the assumptions made is the sufficient condition of the dispersion causality, including the case when $\varphi(x) \in L_2(T)$ and $W(x) \in L_2(T)$, and they are the T-periodic functions.

Conditional Monotony

The cosine function, entering in the expression for interferogram, is even and periodic. Therefore, nonmonotonic phase can give the same interference pattern, as a monotonic one. Let x_o be an extreme point of the phase, then $\cos \Phi(x) = \cos \tilde{F}(x)$, where

$$\tilde{F}(x) = \begin{cases} \pm \Phi(x)\ , \ x \le x_o\ , \\ 2\pi n \pm \Phi(x)\ , \ x > x_o, \ \ n = 0, 1, 2, \dots \ . \end{cases} \quad (11)$$

Even if the phase $\tilde{F}(x)$ is discontinuous at the point x_o, its derivative remains continuous and limited in modulus by the same value, as $\Phi(x)$. The phase functions determined by Eq. (1) we call hereinafter as the conditionally monotonic.

Numerical Experiment

The conformity of causality and monotony was investigated numerically. Using a random number generator with the normal distribution we generated a periodic discrete samples of the analytic signal $W_k = \{U_k + iV_k\}\exp i\, \alpha_c k$ at $\alpha_c = +15$, $k \in [1, N]$, $N = 256$. Then the carrier frequency α_c changed in the range from + 15 up to −15 with the unit step. For each value of the carrier frequency the estimates of the degree of the phase monotony p_m and the causality degree p_c we then calculated their average values over thirty realizations of the random process W_k. To do this we used the following expressions:

$$p_m = \sqrt{\sum_{k=1}^{N-1} {}_+p_k^2 - {}_-p_k^2 \Big/ \sum_{k=1}^{N-1} {}_+p_k^2 + {}_-p_k^2}\ , \quad (12)$$

$${}_+p = \begin{cases} 0\ , & p_k < 0 \\ p_k, & p_k > 0, \end{cases} \qquad {}_-p = \begin{cases} 0\ , & p_k > 0 \\ p_k, & p_k < 0, \end{cases}$$

$$p_k = \arg W_{k+1} - \arg W_k\ ,$$

$$p_c = \sqrt{\left(\sum_{k=1}^{NN} |S_k|^2 - \sum_{k=NN+1}^{NN} |S_k|^2\right) \Big/ \sum_{k=1}^{N} |S_k|^2}\ ,$$

where $NN = N/2 + 1$ is the Nyquist frequency, and S_k is the discrete spatial spectrum.

Figure 2*a* shows the results of the experiment for the case, when the process W_k has undergone both the phase, and amplitude modulation. Its spectral density is constant and not equal to zero in the frequency interval $[\alpha_c - 4, \alpha_c + 4]$; outside this interval the spectral density is equal to zero. In Fig. 2*b* the results of studying of purely phase process, when $|W_k| = \text{const}$, are shown. The spectral density of the process visually is in the same interval, but its width and shape were not controlled.

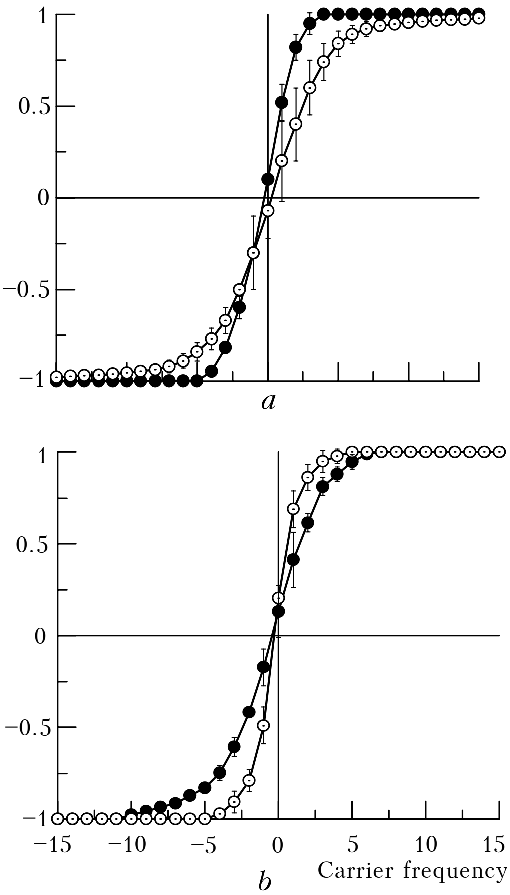

FIG. 2. The estimates of the causality degree p_c, see open circles (–○–) and the monotony degree p_m, see black circle (–●–). (a) is the signal with a finite spectrum and amplitude–phase modulation, (b) is the signal with phase modulation. As a confidence intervals we show the rms deviations of the corresponding functions.

The analysis made and the numerical experiment revealed four typical situations.

The first shows that the causality and monotony are observed simultaneously. It is true for the functions with only phase modulation, and for those modulation, as well with the amplitude–phase. We have in this case that $\alpha_c > \Delta\alpha$, $W(x)$ is an AS, $\alpha_c \in [-15,-5] \cup [7,11]$ in Fig. 2*a*.

In the second situation the complex function $W(x)$ has the causal spectrum, it is an AS, $\alpha_c > \Delta\alpha$, but its phase is not monotonic. It is observed with functions having the finite spectrum and amplitude–phase modulation, $\alpha_c \in [-15, -5] \cup [4, 15]$, Fig. 2*a*.

In the third situation $W(x)$ has not monotonic phase and it is not an AS, $\alpha_c < \Delta\alpha$. But, if the interference fringes have full profile, the function $W(x)$ has the conditionally monotonic phase.

And finally, the fourth situation is in the absence of the causal spectrum for the function $W(x)$, it is not an AS, but its phase is a monotonic one, $\alpha_c < \Delta\alpha$, $\alpha_c \in [-9, -5] \cup [5, 7]$, Fig. 2*b*.

The first and second situations are the most convenient for application of the algorithms, based on the Hilbert transform, and the experiment should be arranged so that these situations occur. But the cases, in which no causality spectrum occurs and there is the monotony or conditional monotony of the phase, are not lost. We pass to consideration of this question.

5. TRANSFORMATION OF THE COMPLEX FUNCTION WITH THE MONOTONIC PHASE IN THE ANALYTIC SIGNAL

The monotony enables one to increase the degree of causality of a spectrum and to improve the estimates of the phase. Let the function $W(x) = a(x) \exp i\, \Phi(x)$ which has a monotonic phase $\Phi(x)$, but is not an AS. If to transform the variable x so, that $\Phi[x(\tau)] = \alpha_c\tau$, then the spectrum of function $W(\tau)$ will become considerably narrower and located in the vicinity of the point $\alpha = \alpha_c$.

The width of this spectrum relative α_c, without the account of the transformation errors, is determined only by the amplitude $a(\tau)$, which changes a little bit as compared with $a(x)$, if the last is low–frequency. At $a(x) = \text{const}$, the function $W(x)$ is causal in the dispersion sense, and the transformed function $W(\tau)$ will be a harmonic fluctuation with the frequency α_c.

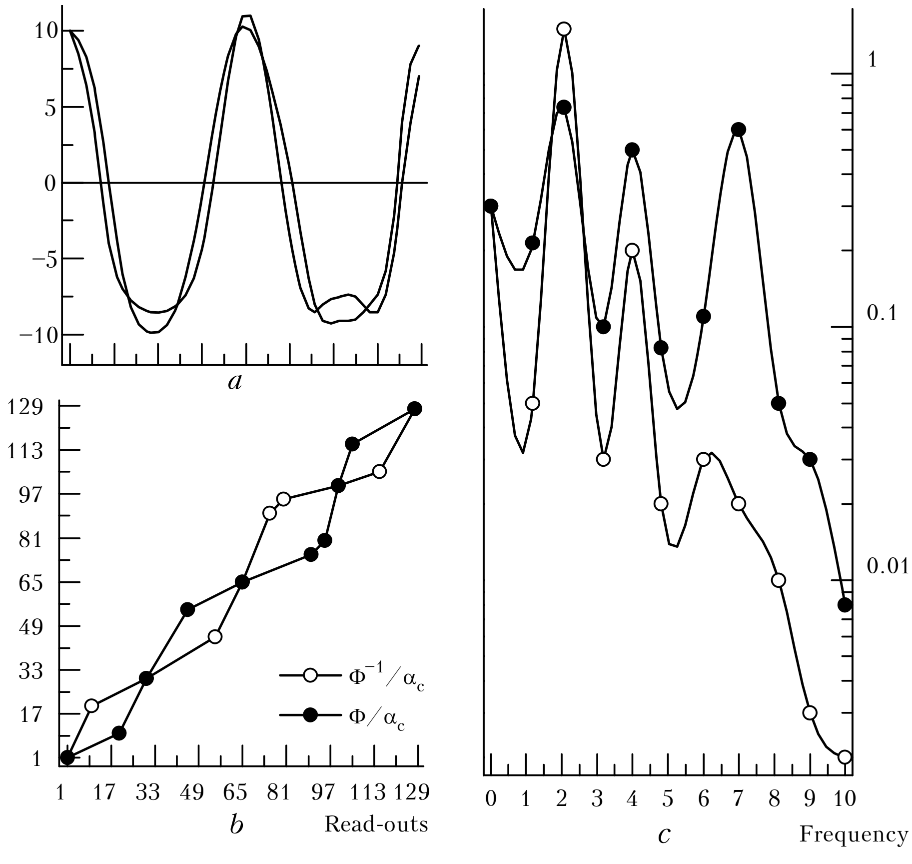

FIG. 3. The extension–compression transformation of a signal with a monotonic phase: the initial signal, (– – – –), and signal after transformation, (——) (a); normalized phase, (– –●– –), and normalized inverse phase of the initial signal, (–○–), (b); the spectrum of the initial signal, (–●–), and the spectrum of the transformed signal, (–○–) (c).

Such an extension–compression transform may be take place according to the expressions

$$\cos \Phi(x) \to \cos \Phi[\Phi^{-1}(\alpha_c\tau)] = \cos \alpha_c\tau ,$$
$$\underset{t}{\mathbf{H}} \cos \alpha_c\tau = \sin \alpha_c\tau \to \sin \Phi(x) , \quad (13)$$

where Φ^{-1} is the inverse function of Φ. The uniqueness of this inverse function is provided by the monotony of the direct function. If, in addition, the derivative $\Phi(x)$ is not equal to zero, the inverse function will not have a discontinuity, that it is especially important for its numerical calculations. Figure 3 illustrates this transformation.

There arises a question on the performance of the operations described. Really, to define the phase $\Phi(x)$, knowing only $U(x)$, this phase needs to be known for making transformations (13). But, it is possible to assume here, that function

$$\Phi_o(x) = \arctan \left[\underset{x}{\mathbf{H}} U(x) \,/\, U(x)\right] \quad (14)$$

is sufficient for performing the initial compression of the spectrum. Then the iterative process is executed following equation:

$$\Phi_{n+1}(x) = \arctan \left[\frac{\underset{t}{\mathbf{H}}\, U(\mathrm{F}_n^{-1}(\mathrm{a_c t}))}{U(\mathrm{F}_n^{-1}(\mathrm{a_c t}))}\right]_{\tau = \Phi_n(x)/\alpha_c} . \quad (15)$$

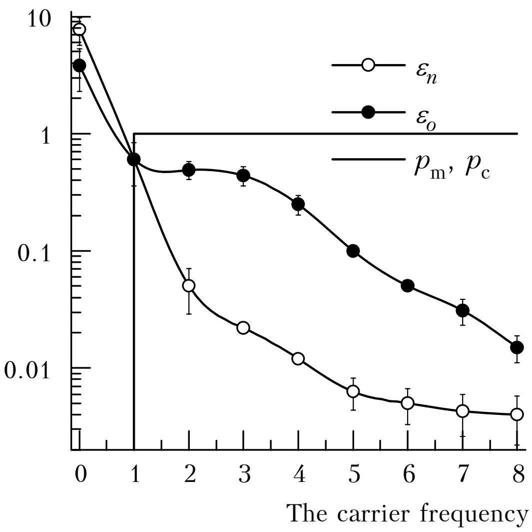

FIG. 4. Numerical simulation of the iterative equation for the estimation of a phase with the use of the extension–compression transformation. The estimation of the accuracy of a signal phase without iterations, (see – –●– –). The estimate of the accuracy of a signal phase at four iterations, (see –○–). The estimate of causality degree of a signal after its transformation and degree of monotony of a signal phase, (see - - -). As confidence intervals we show the rms deviations of the corresponding values.

For numerical simulations of the iteration transform we generated the realizations of the random phase φ_k, $k \in [1, N]$, $N = 256$ within the limit of one period, as in the previous experiment. The spectral density of these realizations is constant, when the frequency $\alpha \in [2, 5]$, $\sigma_\varphi = \pi/5$, and the average is equal to zero. Then a linear component was added to φ_k and this sum $\Phi_k = \varphi_k + \alpha_c k$ was obviously monotonic at $\alpha_c = 8$.

The complex process $W_k = \cos \Phi_k + i \sin \Phi_k$ was constructed and the estimates p_c were calculated for W_k,

$$\varepsilon_n = \sqrt{\sum_{k=1}^{N} [\Phi_k - \Phi_{nk}]^2 \Big/ \sum_{k=1}^{N} \varphi_k^2}$$

and p_m for Φ_k at $\alpha \in [0, 8]$. The value Φ_{nk} was determined from the iterative Eq. (15) at $n = 4$ and from the Eq. (14) at $n = 0$. The averaging of the estimates was conducted over thirty realizations of the random process for each value α_c independently. The results are shown in Figure 4. It is seen, that four iterations essentially improve the estimates of the phase when $p_m = 1$ and Φ_k is monotonic.

6. UNIQUENESS OF THE ANALYTIC SIGNAL IN THE MULTIDIMENSIONAL CASE

As is seen, the analytic signals, Eq. (6) and (8), are the cross–sections of the same function $W(x, y, z, t)$. Therefore, the amplitudes and phases in these expressions are also the cross–sections of the same four–dimensional functions.

The function $W(x, y, z_o, t_o)$, being a cross–section of the expression (5), is not an analytic signal of the variables x or y, being different from it only by the linear phase additive AS, is contained in the interferogram, Eq. (9). For functions with a monotonic phase the convertible transformation in the AS, Eq. (13), also exists.

These properties are a consequence of the limited bands of the temporal spectrum and the spectrum of spatial frequencies, connected with the wave propagation direction. The optical wave in the quasimonochromatic, geometrical, parabolic approximation has these properties. Besides the form of the wave equation solution, when the spatial variables and time are additive arguments of the exponential function (5), allows to use the spatial and the temporal carrier frequencies in the analysis jointly.

All this give a possibility to define the uniform four–dimensional analytic signal, its envelope and phase as follows:

$$W(x, y, z, t) \overset{\text{def}}{=} U(x, y, z, t) + iV(x, y, z, t) ,$$

$$V(x, y, z, t) = \underset{t}{\mathbf{H}} U(x, y, z, t) = \underset{z}{\mathbf{H}} U(x, y, z, t) ,$$

$$V[p(x, y, z), t] = \underset{x}{\mathbf{H}} U[p(x, y, z), t] =$$

$$= \underset{y}{\mathbf{H}} U[p(x, y, z),$$

$$V[x(\tau), y(\tau), z(\tau)] = \underset{r}{\mathbf{H}} U[\Phi^{-1}(a_c r)] \Big|_{r = \Phi(\tau)/\alpha_c} .$$

The practical use of these expressions is that they enable to design the algorithms of measurement of phase in various spatiotemporal, one–dimensional and multidimensional cross–sections and provide its coincidence with the uniform four–dimensional phase in these cross–sections. Determined by AS the amplitude and the phase are invariant relative to the replacement of the Hilbert transform argument and consequently are unique.

CONCLUSIONS

For the description of a wave field in the multidimensional case we introduced the analytic signal, invariant relative to the replacement of the Hilbert transform argument, determining unique multidimensional phase of the optical wave.

It is shown, that unique four–dimensional phase exists only as a consequence of the narrow–band temporal and the spatial spectra of the wave.

Conformities between the causality of the Fourier transform of a complex function and the monotony of its phase is established. A concept of the dispersion causality is introduced using the dispersion definition of width.

It is shown, that from the monotony of the phase follows the dispersion causality of the Fourier transform of the complex wave function with constant amplitude.

REFERENCES

1. M.V. Keldysh, Dokl. Akad. Nauk SSSR **47**, No. 4, 243–245 (1945).
2. D. Gabor, J. of IEE **93**, Pt. 3, 429–441 (1946).
3. E.C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Gostekhizdat, Moscow–Leningrad, 1948), 500 pp.
4. B.Ya. Levin, *Distribution of the Entire Function Roots* (Gostekhizdat, Moscow, 1956), 583 pp.
5. V.I. Tichonov, Radiotekhn. Elektron. **2** No. 4, 562–568 (1957).
6. G.B. Velker, Proc. IEEE **1**, No. 3, 5–20 (1966).
7. L. Mandel, J. Opt. Soc. Am. **57**, No. 5, 613–617 (1967).
8. V.I. Korjik, Radiotekhnika **23**, No. 4, 1–6 (1968).
9. V.I. Korjik, Problems of the information transfer **5**, No. 4, 3–18 (1969).
10. M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1959).
11. R. Courant, *Differential and integral calculus,* Vol. **2** (Nauka, Moscow, 1970), 672 pp.
12. A.M. Trachtman, V.A. Trachtman, Radiotekhnika **25**, No. 3, 85–89 (1970).
13. A. Papoulis, *Systems and Transforms with Applications in Optics* (McGrow Hill, New York, 1968).
14. L.M. Soroko, *Principles of Holography and Coherent Optics* (Nauka, Moscow, 1971), 616 pp.
15. Ya.I. Hurgin and V.P. Yakovlev, *The Finite Functions in Physics and Technology* (Nauka, Moscow, 1971), 408 pp.
16. E.C. Titchmarsh, *Introduction to the Generalized Spectral Theory of Signals* (Sov. Radio, Moscow, 1972), 352 pp.
17. L.M. Soroko, in: *Proceedings of the Fifth All–Union School on Holography* (Institute of Nuclear Physics, Leningrad, 1973) pp. 40–95.
18. V.A. Zverev, *Radiooptics* (Sov. Radio, Moscow, 1975), 304 pp.
19. V.A. Tartakovski, in: *Proceedings of the Third All–Union Symposium on Propagation of Laser Radiation through the Atmosphere* (Institute of Atmospheric Optics of the Siberian Branch of the Academy of Sciences of the USSR, Tomsk, 1975), pp. 230–231.
20. H.M. Nussenzveig, *Causality and Dispersion Relations* (Academic Press, New York, 1972).
21. Ya.I. Hurgin and V.P. Yakovlev, Proc. IEEE **65**, No. 7, 16–45 (1977).
22. A.G. Rekisha, Proc. IEEE **68**, No. 3, 5–30 (1980).
23. S.K. Pohlig, Proc. IEEE **68**, No. 5, 100–101 (1980).
24. G.N. Pihteev, *The Precise Methods of Calculation of Cauchy Type Integrals* (Nauka, Novosibirsk, 1980), 121 pp.
25. V.A. Tartakovski and V.V. Pokasov, *The Phase and the Envelope Determination in the Optics and Their Interrelation* (Institute of Atmospheric Optics of the Siberian Branch of the Academy of Sciences of the USSR, Tomsk, 1980) 13 pp., Dep. v VINITI, 15.01.80, Reg. No. 356–80, Moscow.
26. S.O. Rais, Proc. IEEE **70,** No. 7, 5–13 (1982).
27. L.A. Vainshtein and D.E. Vakman, *The Frequency Separation in the Theory of Oscillations and Waves* (Nauka, Moscow, 1983), 287 pp.
28. V.A. Tartakovski, Izv. Vyssh. Uchebn. Zaved. SSSR, Ser. Fizika, No. 11, 96–105 (1985).
29. I.D. Zolotarev, in: *Proceedings of the Second International Science–Technical Conference "Problems of the Electronic Instrument Making"* (Novosibirsk, 1994) pp. 17–19.
30. I.V. Kostuhina, in: *Proceedings of the Interinstitutes Science and Technology Conference "Microelectronics and Informatics"* (Moscow, Apr. 12-14, 1995) pp. 31–32.
31. V.A. Tartakovski, *Topical meeting on Adaptive Optics,* (Garching bei München, Oct. 2–6, 1995), ESO conference and workshop proceedings, No. 54, pp.141–147.