INTERFEROGRAM EVALUATION BY 4D ANALYTIC SIGNAL THEORY

Valeri A. Tartakovski
Institute of Atmospheric Optics of the Russian Academy of Sciences ❄ Tomsk

1. ABSTRACT

The Gabor analytic signal is constructed by means of the Hilbert transform. This linear operator gives the single-valued determination of the envelope and phase of one-dimensional signal with the support from infinity to infinity. But the wave field has four arguments and the interferogram is a two-dimensional function. This problem is discussed below. In addition, the error estimation of phase restoration and a consequence of phase dislocations are considered. The results of numerical experiments on the interferogram inversion are given.

2. ANALYTIC SIGNAL INTRODUCTION IN MULTIDIMENSIONAL CASE

This paper describes the wave field in multidimensional space by the analytic signal (AS). Let us consider the total solution of the scalar wave equation corresponding to the propagation of a quasi monochromatic wave through the homogeneous medium along the z-axis.

$$W(x,y,z,t)=\int_0^\infty d\omega \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\alpha,\beta)\exp\{-i(\alpha x+\beta y+\gamma z-\omega t)\}d\alpha\, d\beta. \qquad (1)$$

Here α, β and $\gamma=\sqrt{\omega^2/c^2-\alpha^2-\beta^2}$ are the spatial frequencies, ω is the temporal frequency, and c is the velocity of light. Such representation is sufficient to describe the process of the wave propagation through the optical system of the interferometer. The sign of the frequencies ω and γ is not changed, therefore the function $W(x,y,z,t)$ is the analytic signal of the z-variable as well as of the t one. However, the sign can be changed, for example, due to appearance of waves being reflected by the optical system surfaces and propagating toward the reference wave. But we assume that the amplitude of a counter wave is small, moreover, the optical diagram of the interferometer makes it possible to remove it from the direct wave in the range of spatial frequencies (α,β).

The variables z and t are fixed during the experiment and the interferogram is assigned as a function of x and y, therefore it is necessary to introduce the analytic signal depending on these variables. However, in the spatial frequencies plane the coordinate origin is inside the support of spectrum $S(\alpha,\beta)$, therefore $W(x,y,z_0,t_0)$ is not the analytic signal[1,2].

Let us rotate the recording plane (x,y) with respect to the x and the y axes at the certain angle. In this new plane the coordinate z will not be already fixed and changed as $p(x,y,z)$: $z_0-z+\eta x+\zeta y=0$. After these operations expression (1) becomes:

$$W\{p(x,y,z),t\}=\int_0^\infty d\omega \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\hat\alpha,\hat\beta)\exp\{-i(\hat\alpha x+\hat\beta y+\varphi_0-\omega t)\}d\hat\alpha\, d\hat\beta \qquad (2)$$

where $\hat\alpha=\alpha+\eta\gamma$, $\hat\beta=\beta+\zeta\gamma$, $\varphi_0=\gamma z_0$.

The maximal width of $S(\alpha,\beta)$, in a dispersion sense, is small compared to the quantity

$\omega_c / c \approx 10^4 \text{mm}^{-1}$, where ω_c is a central frequency of a temporal spectrum. It allows one to choose such quantities η and ζ that the spectrum $S(\hat{\alpha},\hat{\beta})$ is localised in one quadrant of the frequency plane $(\hat{\alpha},\hat{\beta})$ and owing to this the function $W(x,y,z,t)$ in the plane $p(x,y,z)$ is the AS depending on both x and y coordinates. Moreover, spectrum $S(\alpha,\beta)$ must be considered having a finite support. It is necessary for representation of the function $W(p,t)$ by a number of discrete readings. If even $|W(p,t)| < \infty$, then $W(p, t)$ is the entire exponential function (EEF) of A-type[3].

We estimate the quantities η and ζ and the values of a slope angle of a recording plane with respect to the z-axis. Let the light wavelength be $\lambda = 0.63\ 10^{-3}$ mm and the width of $S(\alpha,\beta)$ be $\eta_s = 200\pi\ \text{mm}^{-1}$, that corresponds to the spatial period 0.01 mm. The quantities η and ζ can be found from the condition $\eta > \eta_s / \sqrt{(2\pi/\lambda)^2 - \eta_s^2} \approx 0.063$. This corresponds to the slope of a recording plane with respect to the z-axis at an angle of $\theta = \pi/2 - \text{arctg}\ \eta = 86^\circ$.

The angle θ is experimentally established as such quantity that the light beam, incident on the recording plane, is found to be on one side from a certain normal to this plane. This condition of causality for the angular spectrum of the function $W\{p(x,y,z),t\}$ is provided by the above fact. The rotation of the recording plane results in the appearance of the spatial carrier frequency of the wave in this plane. It is an effect of the variables x, y, and z being additive arguments of the exponential function in Eq. (1). Similarly, the variable t is introduced into this function. Therefore, the phase and amplitude determined by the analytic signal, which is introduced by the time or the z-coordinate, are identical to the phase and amplitude determined by the analytic signal, which is introduced by the x or y coordinates when $p(x,y,z) = 0$.

This can be expressed in the form (3), where $W(x,y,z,t)$ is the 4D analytic signal, $U(\,)$ is the AS real part, $V(\,)$ is the AS imaginary part, H is the Hilbert transform operator.

$$
\begin{aligned}
W(x,y,z,t) &\overset{def}{=} U(x,y,z,t) + iV(x,y,z,t),\\
V(x,y,z,t) &= \underset{t}{\mathrm{H}}\, U(x,y,z,t) = \underset{z}{\mathrm{H}}\, U(x,y,z,t),\\
V\{p(x,y,z),t\} &= \underset{x}{\mathrm{H}}\, U\{p(x,y,z),t\} = \underset{y}{\mathrm{H}}\, U\{p(x,y,z),t\}.
\end{aligned}
\tag{3}
$$

The senses of the above expressions consists in the AS unicity, its amplitude and phase, in the AS invariance with respect to the Hilbert transform argument in the case of 4D space.

We are going to consider below the analytic signal $W(x,y,z,t)$ as a function of the parameter τ of some scanning line in the four-dimensional space. In the special case it may be the straight section of wave in x or y-directions in the recording plane, namely $W(\tau) = U(\tau) + iV(\tau) = |W(\tau)| \exp i\Phi(\tau)$, $\Phi(\tau) = \varphi(\tau) + \eta_c\tau$, here $\Phi(\tau)$ is the phase and $\varphi(\tau)$ is the object phase, η_c is some line-carrier frequency.

May the function $W(\tau)$ be the analytic signal as a function of the parameter τ when $W(x,y,z,t)$ is AS? We consider this question and relate causality of the complex function spectrum to monotony

of its phase $\Phi(\tau)$. Let us use the Bernstein inequality[4] for EEF: $\max|W'(\tau)| \le \eta_c \max|W(\tau)|$, here η_c is equal to the spectrum half-width. Let $|W(\tau)| = \text{const}$, then we can obtain $\max|\varphi'(\tau)| \le \eta_c$. According to the theorem on the shift within the frequency range, the function $\exp i\{\varphi(\tau) + \eta_c\tau\}$ has the causal spectrum and it follows from the Bernstein inequality that $\Phi(\tau) = \varphi(\tau) + \eta_c\tau$ is the certainly monotonic function. This means that causality is the sufficient condition of the presence of monotony. Therefore, it is obvious, that the cases when there is monotony and no causality are possible. Monotony is manifested in the experiment that the interference fringes are untruncated and characterised by the fill profile without ripples.

Let us consider the next consecutive transformations:

$$\cos\Phi(\tau) \rightarrow \cos\Phi\left(\Phi^{-1}(l)\right) = \cos l, \quad \underset{l}{\mathrm{H}}\cos l = \sin l \rightarrow \sin\Phi(\tau).$$

Such transformations are possible by the single-valued inverse function $\Phi^{-1}(l)$, it can be guaranteed by the monotony of the direct function $\Phi(\tau)$. In addition, we assume $\Phi'_\tau \neq 0$, then the inverse function will be continuous. Hence, we can expand equations (3) writing

$$V\{x(\tau), y(\tau), z(\tau)\} = \left[\underset{l}{\mathrm{H}} U\left\{\Phi^{-1}(l)\right\}\right]_{l=\Phi(\tau)}, \text{ if } U^2(\tau) + V^2(\tau) = \text{const}. \tag{4}$$

It gives a guarantee that phase of the four-dimensional wave function, determined by the AS, introduced as a function of the line parameter τ, on the scanning line, is the corresponding cross section of the unique four-dimensional phase function determined by the AS as a function of the time or propagation direction. As a result, it is possible now to measure the two or one-dimensional sections of the four-dimensional phase separately and apart from each other.

Naturally, the function $\Phi(\tau)$ could be known, if $V(\tau)$ had been defined before. From expression (4) follows, firstly, the existence of the problem solution and the solution in itself can be found by the method of successive approximation.

The matrix of readings of the interferogram is redundant with respect to the object phase, because the number of readings is chosen proceeding from the frequency of the space carrier, namely, the number of interference fringes, and on the frequency properties of the noise. Therefore the idea occurs to analyse the interferogram only in its individual sections from which the phase sections are restored. Their composition represents a two-dimensional function.

We consider a simple interferogram model on a line: $G(\tau) = |1 + W(\tau)|^2$. Let us denote a filtration operator by the symbol F_{bc}. The filter corresponding to this operator passes all the spectrum components at frequencies range $[b, c]$ without distortions and suppresses completely the components being out of this frequency range, therefore this filter makes the AS on its output. We assume that the spectrum of the analytical signal $W(\tau)$ does not overlap other components of the interferogram $G(\tau)$. This is sufficient in order to write the relation for direct demodulation of the interferogram:

$$\Phi(\tau) = \arg \mathrm{F}_{bc} G(\tau). \tag{5}$$

This relation is based on the known principle of heterodyning, which is extended here to the case of curvilinear scanning in the 3D space. The different modifications of this principle are used

in holography, e.g., the Leith-Upatnieks hologram. Its application to phase restoration from interferograms as the Fourier transform method was first described in Ref. 5 and as the technical solution for designing the interference wave-front sensor in Ref. 6.

3. RESTORED-PHASE ERROR ESTIMATION

It can be shown[7] that the estimate of the phase could be improved, if it is calculated as a mean value over the phase estimates of interferogram samples, which differ in the number of the interference fringes. Also the testing object is not modified or moved under registration of those samples. It is an obligatory condition.

he number of fringes on the interferogram is proportional to the carrier spatial frequency of the object wave in the interferogram, as well as it is connected with the tilt, e.g., of reference mirror in the interferometer. In consequence of the above-mentioned, the set of fringe numbers must be constructed knowing the a priori width of the spectrum of spatial frequencies of the object wave. If object waves with different carrier do not overlap in the plane of spatial frequencies, then the dispersion of the phase retrieval error decreases proportionally to the number of interferogram samples. It is necessary also that the noise in the interferogram should be additive, normal, and white.

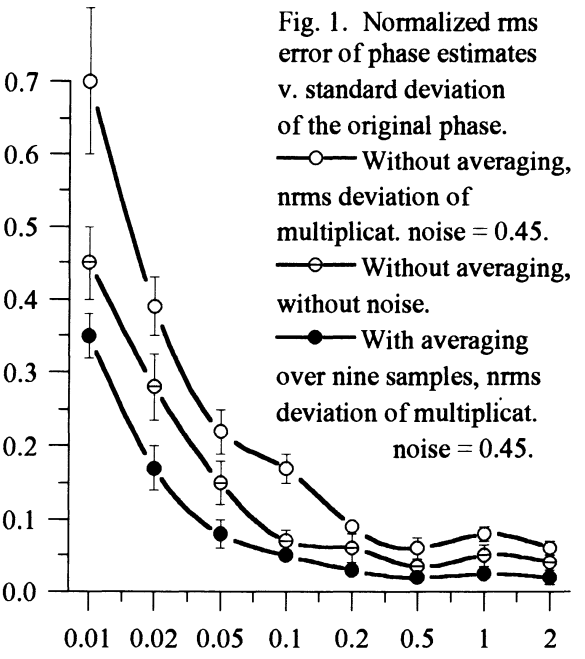

Fig. 1. Normalized rms error of phase estimates v. standard deviation of the original phase.

The numerical experiment was taken for the investigation of the phase estimation under other conditions, namely, by multiplicative, band limited noise in the interferogram. The half-width of the noise band was equal to 5% from the Nyquist frequency. The interference fringe number varied from 10 to 18, step 1. It can be seen in Fig. 1 the average of the phase estimates has always a good effect.

It is possible to estimate the restored phase error without using some statistical procedure, when only one interferogram is available. It is reached on the base of Eq. (3, 4) using some orthogonal scanning directions. Figures 2*a, b* demonstrate the connection between the real errors and the calculated estimates of the restored phase errors. They fitted well by the first order regression line. The correlation coefficient was equal 0.99 for the normalised rms error (Fig. 2*a*) and 0.94 for the peak-to-valley error (Fig. 2*b*). The standard deviation were equal 0.01 and 0.04 respectively. These graphical displays were given as a result of the closed numerical experiments with the following interferogram data: The interferogram matrix order was equal to 60, the interference fringe numbers is 12, the standard deviation of the original phase is 0.5, the nrms deviation of multiplicative noise was taken from {0.0, 0.1, 0.15, 0.2}, the interferogram quantization-level numbers were taken from {2, 4, 5, 6, 10, 256}, the mode of power spectrum of the original phase

for the Zernike set was located in "defocusing", the original phase was apodisated.

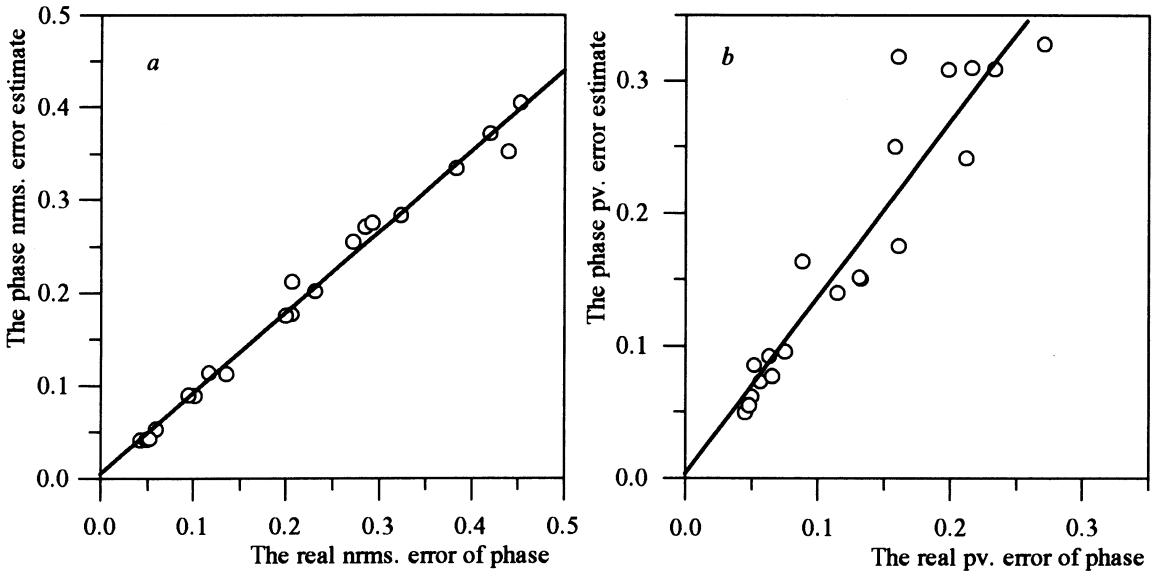

Fig. 2*a, b*. The relation between the real restored-phase error and its estimate.

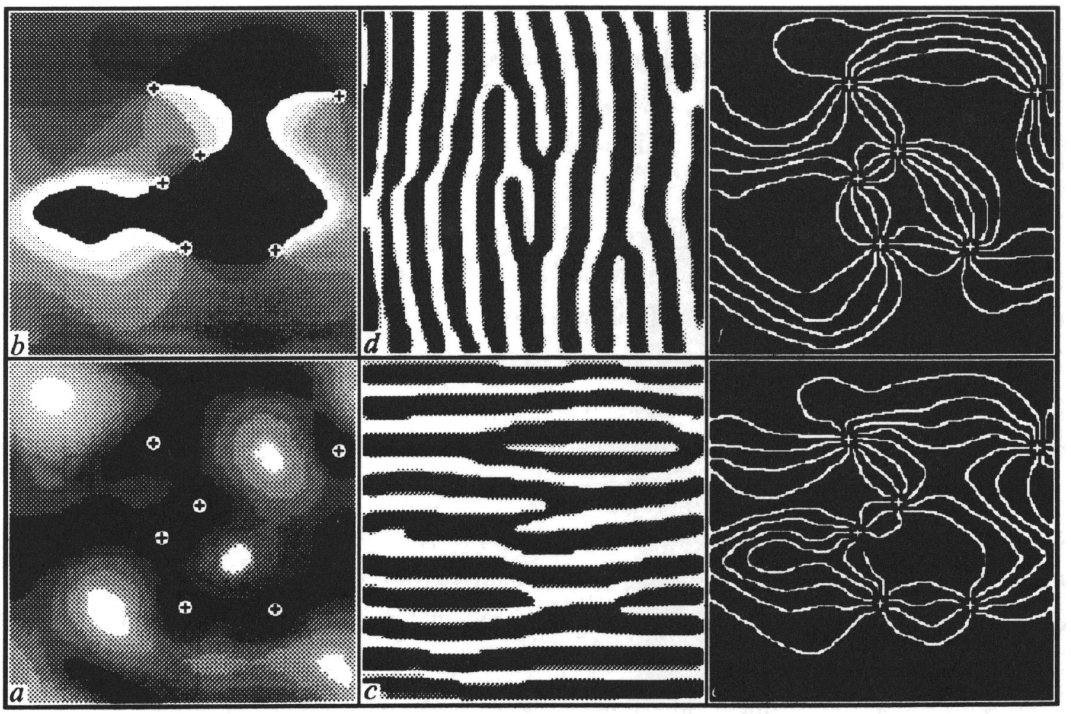

Fig. 3. Wave phase dislocations and structures created by them. Crosses denote zeros of the intensity and corresponding points of phase dislocations: wave intensity (*a*); wave phase (*b*); interference pattern for unit amplitude and carrier wave perpendicular to the horizontal coordinate axis (*c*); interference pattern for unit amplitude and carrier wave perpendicular to the vertical coordinate axis (*d*); contour lines of the phase sine (*e*); and, contour lines of the phase cosine (*f*).

4. CONSEQUENCE OF PHASE DISLOCATIONS

The appearance of phase dislocations of a light wave propagating through a randomly inhomogeneous medium was studied in quasi monochromatic and parabolic approximations[8]. To this end, the known numerical model was used[9]. In this model, the method of splitting and FFT by the Singleton algorithm were used for solving the wave equation. The wave and its angular spectrum were approximated by periodic functions entered in a computer in the form of two–dimensional matrices of their readings. Randomly inhomogeneous medium was modelled by the spectral power density of the field of refractive index of power–law type typical of the atmospheric turbulence. The propagation path was 6 km long, and the wavelength was 0.6328 μm. The magnitude of the wave fluctuations was characterised by Fried's coherence scale for both weak and strong intensity fluctuations.

As seen from Fig. 3*a*, the phase dislocations appear at the points where the intensity reaches its minimum. These points correspond to zeros of the wave function or AS. In the vicinity of these points the phase varies spirally. Along the whole length of boundaries between white and black areas in Fig. 3*b*, between two points of dislocation formation, the phase surface undergoes discontinuity of $\pm 2\pi$. Such a discontinuity cannot be removed with the use of translations of surface fragments. The dichotomy of maximum and minimum interference fringes on the pattern, and appearance and disappearance of interference fringes start at points of dislocations (Figs. 3*c, d*). Contour lines of phase cosine and sine form a radial structure in the vicinity of dislocation points and converge to them (Figs. 3*e, f*).

We have performed numerical experiment to compare the behaviour of wave–function scintillations and angular spectrum with the number of phase dislocations appearing with the increase in the turbulence intensity. The presence of dislocation was determined by calculating the phase gradient between neighbouring points arranged in a closed contour drawn around the point of the phase function under analysis. Dislocation occurred if the phase gradient was $\geq 2\pi$ or $\leq -2\pi$. The phase was calculated as inverse tangent of the ratio between imaginary and real components of the wave function. We normalised the number of dislocations to the ratio between the number of counts of calculational grid to the number of counts in a circle where a dislocation was determined. The wave scintillation index was calculated as a normalised variance of wave intensity, and the angular spectrum scintillation index was calculated as a normalised variance of the square of the modulus of its Fourier transform. We normalised the variances to the mean square of the corresponding parameter. Estimates of all three parameters under investigation were averaged over nine experiments.

Results of the experiment are shown in Fig. 4. of Ref. 10. In the region of large Fried's coherence scale that corresponds to weak turbulence, the wave scintillation index varies linearly, dislocations are absent, and the angular spectrum scintillation index reaches maximum values. Saturation of the wave scintillation index and normalised number of dislocations at unity level takes place with the increase in the turbulence intensity. The angular spectrum scintillation index saturates at unity level as well, but the dependence is reverse as compared with the two other plots. As one would expect, the dislocation number saturates since dislocation density cannot be greater

than unity. However, it is interesting that maximum density of dislocations is achieved together with saturation of scintillation indices of a wave and its angular spectrum. It should be pointed out that the phase dislocations and, correspondingly, the zeros of the wave function appear when the wave scintillation index approaches unity, that is at the origin of the region of strong fluctuations.

5. CONCLUSIONS

The facts are as follows, the wave-front dislocations are very probable for this model of the random inhomogenous medium, just for a vertical path through the turbulent atmosphere. Phase dislocations were to disturb the results of wave-front sensors if they are based on *a priori* continuity of a phase function. These disturbances will be even due to only one dislocation.

Flexible mirrors are incapable to reproduce some singular phase function too, and consequently the singular wave-front must be approximated by a continuous function. On the other side, wave-front dislocations are some onset of the wave disintegration on separate beams that will be further uncorrelated. Also there exists not so much energy in the vicinity of the dislocation appearance points. Therefore an adaptive mirror should not be bent as a singular form in these areas.

By our mind, the measurement of separate areas of wave-front and separately fitting these areas to adaptive mirrors is the right way designing a wave-front sensor and other adaptive parts as well. On this way 4D analytic signal allows us to go round singular points and to connect parts of the wave front to the unique phase function.

6. REFERENCES

1. V.A. Tartakovski, Abstracts of Reports at the Sixth Symposium on Propagation of Laser Radiation in the Atmosphere, Part 3, IOA SB RAS, Tomsk, 32-34 (1981).
2. V.A. Tartakovski, Abstracts of Reports at the IV All-Union Conference on Holography, Vol. 1, VNIIRI, Erevan, 723-727 (1981).
3. B.Ya. Levin, Distribution of Entire Function Roots, Gostechizdat, Moscow (1956).
4. Ya. I. Hurgin and V.P. Yakovlev, TIIER, Vol. 65, № 7, 16-45 (1977).
5. M. Takeda, H. Ina, and S. Kobayashi, J.Opt.Soc.Am. Vol. 72, № 1, 156-160 (1982).
6. E.A. Vitrichenko, L.A. Pushnoi, and V.A. Tartakovski, Author's Certificate, № 1024746, of 18.02.82, Bulletin, № 23, June (1983).
7. E.A. Vitrichenko, V.P. Lukin, L.A. Pushnoi, and V.A. Tartakovski, Problems of Optical Test, Nauka, Novosibirsk (1990).
8. N.N. Myer and V.A. Tartakovski, Atmospheric and Oceanic Optics, Vol. 8, № 3, 231-234 (1995).
9. V.P. Lukin, N.N. Myer, and B.V. Fortes, Atmospheric and Oceanic Optics, Vol. , № 12, 896-899 (1991).
10. V.A. Tartakovski & N.N. Myer. In this issue.