PHASE DISLOCATIONS AND MINIMAL–PHASE REPRESENTATION OF THE WAVE FUNCTION

N.N. Maier and V.A. Tartakovskii

Institute of Atmospheric Optics,
Siberian Branch of the Russian Academy of Sciences, Tomsk
Received September 16, 1994

Vanishing of the wave amplitude accompanied by the appearance of phase dislocations in these points violates the signal continuity in phase measuring and adaptive optical systems and makes them inoperative. The well–known descriptions of wave process either admit of appearance of singular phase points and become ineffective or ignore these points and have limited area of application. All this is a manifestation of complication of the same physical reality. A solution to this problem is likely to be found by representation of the wave function in terms of the components with regular phase. In this paper, we consider the minimal–phase components of an analytic signal.

1. INTRODUCTION

Let $U(r, t)$, $r\{x, y, z\}$ be the real function approximating a light wave, which propagates through an inhomogeneous medium along the z axis, and the following representation be necessary for this function: $U = A(r, t) \cos \Phi(r, t)$, where A is the amplitude, and Φ is the phase of the wave.

A salient feature of this function is that its phase may be written in the form $\Phi = \varphi(r, t) + kz + \omega t$, with the values of k and ω being such that the Fourier transforms of $\cos kz$ and $\cos \omega t$ oscillation modes do not intersect the Fourier transforms of the functions $A(r, t) \cos \varphi(r, t)$ and $A(r, t) \sin \varphi(r, t)$ in corresponding arguments and occur at much higher frequencies. Such properties of a light wave may be most naturally expressed if $U(r, t)$ is the entire exponential function of each variable. The theorem proved in Ref. 1 allows this approximation to be as accurate as is wished if

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

where

$$M(p) = \max |U'_x(r, t)| \text{ for } |x| < p .$$

In doing so the physical properties, namely, monochromaticity and parabolic nature, are extended to the wave approximation $U(r, t)$. Such an approximation is no longer the exact solution to the wave equation. However, there is no need for it, since the notions of the amplitude and phase do not follow from this equation, but exist only in connection with their measurements or definition.

The Hilbert transform is the sole linear operator allowing noncontradictory definition of the amplitude and phase of wave process to be given.[2] This transform exists for square integrable functions, limited functions, and functions satisfying the Hlöder boundary condition. It follows from the properties of Hilbert transform under the above–mentioned assumptions about a wave model that

$$\underset{t}{\mathrm{H}}\, U(r, t) = \underset{z}{\mathrm{H}}\, U(r, t) \overset{\text{def}}{=} V(r, t) = A(r, t) \sin\Phi(r, t) . \quad (1)$$

Here, the symbol H denotes the operators of Hilbert transform over the variables t and z. Let us introduce the complex wave function as an analytic signal (AS). Keeping in mind expression (1), we derive

$$W(r, t) \overset{\text{def}}{=} U(r, t) + i\, V(r, t) = A(r, t) \exp i[\varphi(r, t) + kz + \omega t]. \quad (2)$$

The analytic signal may be introduced for slant cross sections $z = z_0 + x \cos\theta$ or $t = t_0 + x/v$ as well. In this case Hilbert transform is taken over the variable x. The parabolic and monochromatic nature of a wave makes it possible to choose the angle θ between the z axis and the image plane close to $\pi/2$ and the rate of scanning v in the plane z = const large enough. Then $A(r, t)$ and $\varphi(r, t)$ in slant and normal cross sections will be indistinguishable. In this sense the amplitude and phase defined by AS are invariant under a change of coordinate over which the Hilbert transform is taken. Hence, they are unique.

The dispersion relations can be derived[3] for the logarithm of analytic signal $W(r, t)$. One of them has the form:

$$\varphi(x, \overline{x}) = \underset{x}{\mathrm{H}} \ln A(x, \overline{x}) + l(x, \overline{x}) + 2\sum_n \arctan \frac{x - x_n(\overline{x})}{\eta_n(\overline{x})} . \quad (3)$$

Here the symbol $\overline{x} \subset \{r, t\}$ denotes all variables except x, $l(x, \overline{x})$ is the arbitrary function of $\overline{x}$ linear in x, x_n and η_n are the real and imaginary coordinates of the zeros of the function $W(x, \overline{x})$ in the complex half–plane $x + i\eta$, $\eta > 0$.

In the domain where the Fermat principle in its weak statement is valid, out of a set of phases possible for a given intensity and defined by expression (3), only stationary phase realizes, that is, the first derivative of phase should vanish. A phase addition, dependent on the coordinates of the zeros of the analytic signal $W(x, \overline{x})$, may be variable. For stationarity of expression (3), it is

0235-6880/95/03 231-04 $02.00

essential that some nonidentical sets of the zeros of the function $W(x + i\eta, \overline{x}_0)$, $\eta > 0$, yield the same phase. However, it is impossible in view of the uniqueness of the entire function representation by its zeros. Hence, the phase given by expression (3) is stationary if and only if there are no zeros in the upper half–plane.

It follows from the Weierstrass preliminary theorem (Ref. 4, p. 113) that in the domain in which an analytic function of complex variables is holomorphic, its zeros move along continual trajectories. Hence, the zeros of an entire function, before their appearance inside a domain, should appear on its boundary. In this case the line $\overline{x} = \overline{x}_0$ is such a boundary; as it takes place, x may be changed for z or t.

What this means is vanishing of the field intensities precedes the appearance of runs–on of the phase that are different from linear functions and uncorrelated with the logarithm of the amplitude of the Hilbert transform. It is also well known that weak statement of the Fermat principle admits of presence of caustics (see Ref. 5, p. 809), with the domain of the first caustics being coincident with the origin of the region of strong fluctuations in the light wave intensity.[6]

It is evident from the foregoing that such a physical phenomenon as light wave propagation has fixed complexity threshold above which its representation contains singular points.

2. MODELING OF THE PHASE DISLOCATIONS OF THE WAVE FUNCTION

The appearance of phase dislocations of a light wave propagating through a randomly inhomogeneous medium was studied in quasimonochromatic and parabolic approximations. To this end, the known numerical model described in Refs. 7 and 8 was used. In this model, the method of splitting and fast Fourier transform by the Singleton algorithm were used for solving the wave equation. A 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 modeled by the spectral power density of the field of refractive index of power–law type that is 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 characterized by Fried's coherence radius for both weak and strong intensity fluctuations.

As seen from Fig. 1*a*, the phase dislocations appear at the points where the intensity reaches its maximum. These points correspond to zeros of the wave function or AS. Near these points, the phase varies spirally. Along the whole length of boundaries between white and black fragments in Fig. 1*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 contour lines of the interference pattern, and appearance and disappearance of interference bands occur at points of dislocations (Figs. 1*c* and *d*). Contour lines of phase cosine and sine form a radial structure in the vicinity of dislocation points and converge to them (Figs. 1*e* and *f*).

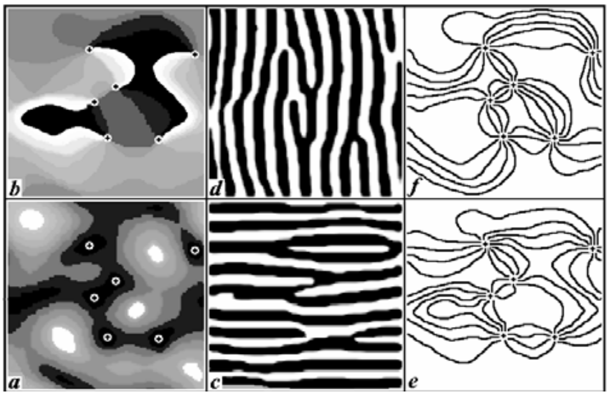

FIG. 1. Wave phase dislocations and structures created by them. Crosses denote zeros of the intensity and corresponding points of phase dislocations: wave intensity (a); wave phase (b); interference pattern for unit amplitude and carrier wave perpendicular to horizontal coordinate axis (c); interference pattern for unit amplitude and carrier wave perpendicular to vertical coordinate axis (d); contour lines of phase sine (e); and, contour lines of phase cosine (f).

We have performed numerical experiment to compare the behavior of wave–function scintillations and angular spectrum with the number of phase dislocations appearing with increase in turbulence intensity. The presence of dislocation was determined through calculation of phase gradient between neighbouring points arranged in a closed contour drawn around the point of 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 normalized 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 normalized variance of wave intensity, while the angular spectrum scintillation index was calculated as a normalized variance of the square of the modulus of its Fourier transform. We normalized the variances to the mean square of the corresponding parameter. Estimates of all three parameters under investigation were averaged over nine experiments.

Results of experiment are shown in Fig. 2. In the region of large Fried's coherence radii 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 normalized number of dislocations at unity level takes place with increase in the turbulence intensity. The angular spectrum scintillation index saturates at unity level as well, but the dependence is reverse as compared with 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.

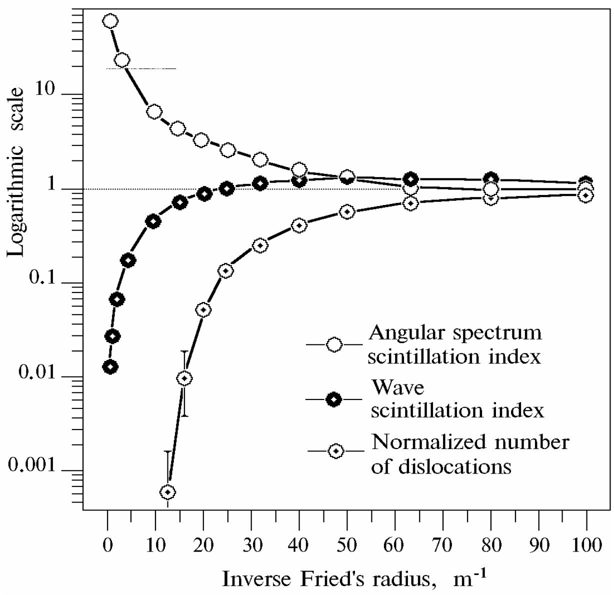

FIG. 2. Estimates of scintillation indices of wave and its angular spectrum and normalized number of wave phase dislocations. Standard deviations are indicated as confidence intervals. These deviations are not indicated if their values are less than point size.

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.

3. MINIMAL–PHASE COMPONENTS OF THE WAVE FUNCTION

Expression (3) relates the phase to the logarithm of the amplitude of the wave function or AS. However, the sum entering into the right–hand side of the expression hinders its use for determining the phase from the amplitude, since the location of zeros is usually unknown.

The principle question arises: has the function $W(x, \overline{x})$ zeros in the complex half–plane of any variable, when the rest of variables take real values? It follows from the theorem proved in Ref. 3 (see p. 323) that when analytic signal $W(x, \overline{x})$ is limited, all its zeros except, maybe, a zero–density set can lie within angles that are as small as is wished, near the real axis. Total absence of zeros in the complex half–plane is favorable to the Hilbert relation between the phase and logarithm of the amplitude, or, alternatively, zeros must be located far from the region of interest on the x axis so the sum in the right–hand side of expression (3) reduces to constant. This corresponds to the local validity of the Fermat principle, for example, in a paraxial region.

Rouche's theorem is most practical (see Ref. 4, p. 287). It follows from the theorem that if the wave function reduces to AS of both coordinates[9] in the image plane and its zero order has the amplitude which is greater than the sum of the rest of components, this complex function has no zeros in the complex half–planes of both coordinates of the image plane. Needless to say that it has no real zeros as well. In this case, in accordance with expression (3), the phase is minimal and the wave function is minimal–phase one.

If the modulus of the angular spectrum of the wave function has a pronounced global maximum, then this function may be represented in terms of four minimal–phase components. To this end, we must locate the global maximum at the origin of coordinates. Then each quadrant of spectral plane with the origin of coordinates will contain two–dimensional minimal–phase AS.

The amplitude of global maximum in the angular spectrum can be enhanced due to wave focusing and apodization as well as due to amplification of the zero order of the angular spectrum (Fig. 3) or suppression of its higher orders.

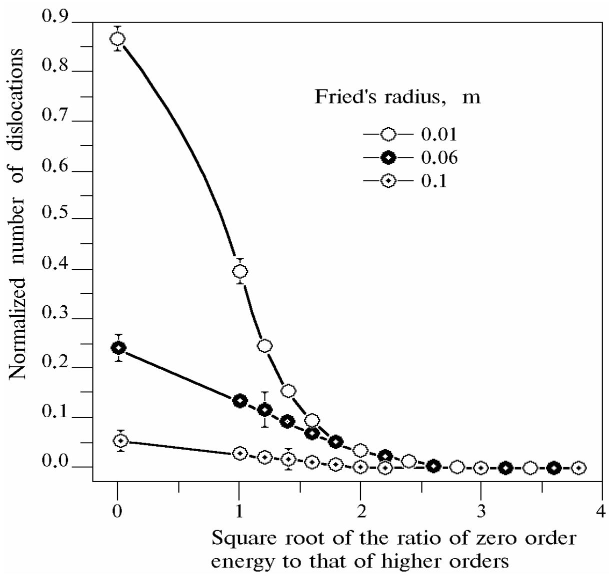

FIG. 3. Phase dislocation suppression due to amplification of the zero order of a wave. Standard deviations are indicated as confidence intervals. These deviations are not indicated if their values are less than point size. Zero abscissas correspond to the case in which the zero order was not changed.

Minimal–phase properties of the wave functions obtained by this way provide a possibility for inversion of transforms of the initial wave.

As is seen from behavior of curves in Fig. 3, the energy of zero order is bound to be specified times larger than the energy of the rest portion of wave, in order that the dislocations disappear and the wave becomes minimal–phase one. This amount of increase does not depend on the intensity of wave fluctuations and the initial number of dislocations.

ACKNOWLEDGMENTS

The authors would like to acknowledge N.N. Botygina and B.V. Fortes for their assistance and helpful discussions.

This work was supported in part by Russian Fundamental Research Foundation Grant No. 94002–03027–a.

REFERENCES

1. M.V. Keldysh, Dokl. Akad. Nauk SSSR **47**, No. 4, 243–245 (1945).

2. D.E. Vakman and L.A. Vainshtein, Usp. Fiz. Nauk **123**, No. 4, 657–682 (1977).

3. B.Ya. Levin, *Distribution of the Entire Function Roots* (Gostekhizdat, Moscow, 1956), 583 pp.
4. B.V. Shabat, *Introduction into an Analysis of Complex Functions. V. I. Functions of a Few Variables* (Nauka, Moscow, 1970), 400 pp.
5. M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1959).
6. Yu.A. Kravtsov, Zh. Eksp. Teor. Fiz. **55**, No. 3(9), 798–801 (1968).
7. P.A.Konyaev and V.P. Lukin, Izv. Vyssh. Uchebn. Zaved. SSSR, ser. Fiz., No. 2, 79–89 (1983).
8. V.P. Lukin, N.N. Maier, and B.V. Fortes, Atmos. Oceanic Opt. **4,** No. 12, 896–899 (1991).
9. V.A. Tartakovskii and V.V. Pokasov, *"Determination of wave envelope and phase as applied to optics and their interrelation",* VINITI, No. 356–80, Moscow, January 15, 1980, 13 pp.