PROPERTIES OF THREE-DIMENSIONAL HOLOGRAMS

V. V. ARISTOV and V. Sh. SHEKHTMAN

Institute of Solid State Physics, USSR Academy of Sciences

Usp. Fiz. Nauk 104, 51–76 (May, 1971)

CONTENTS

INTRODUCTION

HOLOGRAPHY methods, based on the principles developed in the pioneering papers of D. Gabor[1] have by now been extensively developed. The possibilities afforded by the use of holography for the nearest future are being extensively discussed in the scientific literature.

At the same time, the monographs and reviews (see, for example[2-6]) devoted to the physical principles and to the application of holography concern actually only holograms registered in the emulsion layers of ordinary photographic materials. It should be noted that such "planar" holography does not differ at all in principle, from the point of view of the optics of image-reproducing systems, from the traditionally known apparatus (devices) such as lenses, diaphragms, and flat diffraction gratings.

In this sense, a hologram recorded in a certain light-sensitive volume (three-dimensional hologram) is an example of a new physical instrument, which makes it possible to register all three components of the wave vector.

The first to note the important role played by the three-dimensional structure of the record in the shaping of the special properties of holograms was Yu. N. Denisyuk[7], who used an "opposing beam" scheme (the reference beam is directed at approximately 180° to the waves scattered by the object). It should be emphasized that the use of this scheme ensures the manifestation of new properties of holograms compared with the hologram of E. Leith and J. Upatnieks[8], even at a small thickness of the emulsion layer.

The first attempt at a theoretical analysis of certain properties of holograms obtained in a three-dimensional photosensitive medium was made by Van Heerden[9], who also touched upon the question of the information capacity of three-dimensional holograms.

By now, a definite volume of information has been accumulated pointing to the important features of this division of holography, which requires special theoretical treatment, and also the great prospects afforded by the use of three-dimensional holograms in a great variety of fields.

The purpose of the present review is to develop systematically the theory of the processes of image recording and reconstruction and the main features and possible practical applications of three-dimensional holograms. We consider also from a unified position the properties of planar holograms as the limiting case of three-dimensional ones.

The main content of the article is preceded by a brief review of modern three-dimensional photosensitive materials.

Materials for Obtaining Three-dimensional Holograms

For convenience in the exposition that follows, it is advantageous to use the following terminology: a hologram will be called three-dimensional if its thickness Z is comparable with its transverse dimensions X and Y and is much larger than the wavelength λ of the employed radiation; holograms for which $\lambda/4 \leq Z \leq X$ will be called intermediate and holograms with $Z < \lambda/4$ will be called planar. (A detailed justification for the "three-dimensionality" criteria will be given below.)

To obtain holograms of the three-dimensional or intermediate type, one uses either standard photographic materials, such as Lippmann emulsions, or photorecording media not customarily used in optics[10].

The development of holography has called for improvement in thick-layer emulsions, first developed by Lippmann (see, for example[11]). For holographic purposes, one uses emulsions with a resolution that reaches 5000 lines/mm. The thickness of the emulsion layer varies from several microns to several dozen microns. These photographic materials are used for the most part to obtain holograms by the method proposed in[7]. With the aid of these holograms it is possible to reconstruct an image in white light and to obtain color pictures.

Besides emulsions for the recording of intermediate holograms, photochromic and photopolymer films are coming into use[12,13]. These materials have definite advantages, such as the possibility of visual observation of the registration process, the absence of distortions connected with "wet" processes of photographic technology, etc.

Recording media that can be regarded more rigorously as three-dimensional ones are photochromic glasses, crystals with color centers, crystals having nonlinear properties, etc.

Let us stop to discuss these materials in somewhat greater detail. The use of photochromic glasses in holography was reported in[14-16]. These are substances that go over under the influence of light with a definite spectral composition from the ground state to an excited state that differs from the former one in its absorption spectrum. Under the influence of other radiation, the inverse transition takes place. The number of molecules executing the inverse transition is proportional to the light intensity. These glasses can therefore be used to record holograms. Photochromic glasses have so far relatively low resolution, for example 100 lines/mm[14]. The effective thickness of the holograms recorded in this photographic material reaches several millimeters.

The first to advance the idea of using crystals with color centers for the recording of holograms was Van Heerden[9]. A crystal in which a large number of color centers was produced beforehand is placed in the path of rays scattered by the object. This produces a definite emission structure that represents the interference pattern of the registered wave field. Colored crystals of KBr[17,18], KCl[18], CaF_2, and $SrTiO_3$ with different impurity contents[19] were used in a number of investigations. The KCl and KBr crystals are colored beforehand with γ rays at a dose on the order of 10^5—10^7 rad. An ultraviolet radiation source is necessary to color the crystals CaF_2 and $SrTiO_3$. A feature of these materials is that they are sensitive only in a spectral region within the limits of the light-absorption band of the color centers. Therefore the selection of crystals for producing holograms is dictated by the spectral characteristics of the available lasers. In particular, KCl and KBr crystals are suitable for producing holograms with helium-neon lasers. Another feature of colored crystals is the limited storage time of the holographic record. In addition, reconstruction of the image is accompanied by a rapid overall discoloring of the crystal, leading to erasure of the hologram. The resolution limit of these photographic materials is high; for KCl, for example, a resolution level on the order of 10^4 lines/mm can be attained. The effective recording thickness of the hologram exceeds 1 cm. The diffraction efficiency of such holograms amounts to a fraction of 1%.

Certain ferroelectric crystals, such as $LiNbO_3$, $Bi_4Ti_3O_{12}$, and others, make it possible to record phase holograms as a result of an optically-induced inhomogeneity of the refractive index[20-23]. The inhomogeneity of the refractive index depends on the intensity of the incident light and can be conserved for a long time—up to 100 hours in $LiNbO_3$, or else vanish after a short time interval following the turning off of the field, as for example with holograms are recorded in $Bi_4Ti_3O_{12}$. The diffraction efficiency of such holograms can exceed 40%, and the resolution is on the order of 1600 lines/mm[20].

All materials used for recording three-dimensional holograms have in common the possibility of erasure and repeated recording. These materials make it possible to monitor the quality of the image during the course of registration of the wave front and the time of information storage.

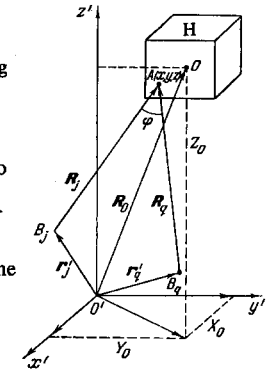

FIG. 1. Diagram showing the recording and reconstruction of three-dimensional holograms. H—hologram; B_j and B_q—two points of the object at recording of the hologram, at one of which it is possible to place a light source during the reconstruction of the hologram; O' and O—origins of coordinates for the object and hologram points. A(x, y, z)—arbitrary point inside the hologram.

On the whole, it can be stated that a limited number of materials for three-dimensional holograms is available at present. Moreover, the existing materials are more readily suitable for the study of three-dimensional holograms and the corresponding schemes than for extensive practical application.

Further research on photorecording materials is necessary to ensure utilization of the important properties of three-dimensional holograms.

I. ELEMENTS OF THE THEORY

1. Introductory Remarks. Kinematic Approximation

For concreteness in the exposition that follows, we start from the simplest scheme of recording and reconstruction of three-dimensional holograms that appears in Fig. 1. The image is reconstructed as a result of diffraction of the reconstructing beam by a certain three-dimensional structure of the hologram, which was produced during the recording process. We shall show that for holograms lying in the far field, this structure can be regarded as a superposition of harmonic distributions of the refractive index. Let the object be an arbitrary aggregate of coherently emitting points $U(x', y')$. The field $\Phi(x, y)$ registered on the hologram will be described by the formula

$$\Phi = \frac{i}{\lambda}\int \frac{U(x', y')}{|R|} e^{i\cdot\frac{2\pi}{\lambda}|R|}\,dx'\,dy'. \qquad (1)$$

The modulus of the vector R drawn from any point of the hologram to an arbitrary point of the object can be expanded in a series, since under the assumption made $R_0 = \sqrt{x_0^2 + y_0^2 + z_0^2}$ is much larger than the dimensions of the hologram and of the object. Therefore formula (1) is rewritten as follows:

$$\Phi = \int V(K_x, K_y)\, e^{i\cdot 2\pi(\mathbf{Kr})}\,dK_x\,dK_y, \qquad (2)$$

where r varies over the hologram and $V(K_X, K_Y)$, apart from coefficients, is equal to $U(K_X|R_0|\lambda, K_Y|R_0|\lambda)$. From (2) we see that

$$\Phi = F[V(K_x, K_y)\, e^{iKz\cdot 2\pi}], \qquad (3)$$

where $F[f]$ denotes the Fourier transform of the function f.

Let us assume that the hologram is a square-law detector. (No allowance is made for the nonlinearity,

since it is immaterial for the analysis of the main properties of three-dimensional holograms). In such a case the change Δn of the refractive index of the recording medium is proportional to the square of the wave field[7]:

$$\Delta n = \alpha |\Phi|^2, \tag{4}$$

where α is a proportionality coefficient and is usually very small, so that $\Delta n \ll n_0$. We now examine the structure of the recording in the hologram (Δn in accordance with formula (4)). Let the object consist of two points:

$$U(x', y') = f_1\delta(\mathbf{r}' - \mathbf{r}_1') + f_2\delta(\mathbf{r}' - \mathbf{r}_2'),$$

Then

$$|\Phi|^2 = |f_1|^2 + |f_2|^2 + f_1 f_2^* e^{i\cdot 2\pi r(K_1 - K_2)} + f_1^* f_2 e^{i\cdot 2\pi r(K_2 - K_1)}. \tag{5}$$

The constant-phase surfaces, and consequently the surfaces of constant Δn, are determined from the condition

$$c = (\mathbf{r}(\mathbf{K}_1 - \mathbf{K}_2)); \tag{6}$$

The maximum is reached at an integer value of c.

The structure of the recording in the hologram is a harmonic distribution of Δn. The surfaces $\Delta n = c$ are directed along the bisector of the angle φ of the convergence of the wave, and are arranged with a period $d = \lambda/2 \sin(\varphi/2)$ (Fig. 2). Such an analysis enables us to establish a deep analogy between the process of reconstruction of a three-dimensional hologram and diffraction of x-rays, electrons, or neutrons by crystal lattices*. It is therefore natural to use in the exposition that follows the main premises and terminology of the theory of x-ray scattering. It is known that there are two approximations of this theory[28]—kinematic for imperfect crystals and dynamic for crystals close to perfect. In the kinematic theory the following conditions are assumed to be satisfied: the intensity of the diffracted ray is negligibly small compared with the incident one; the secondary scattering of the diffracted rays can be neglected. Assuming that such conditions are fulfilled in the reconstruction of three-dimensional holograms, we can use integral transformations, just as in the case of planar holograms†. The kinematic

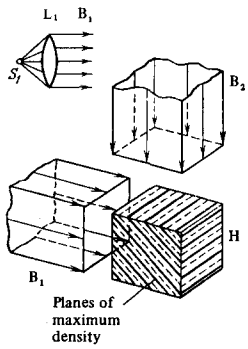

FIG. 2. Hologram from two collimated mutually-coherent monochromatic light beams B_1 and B_2 [4].

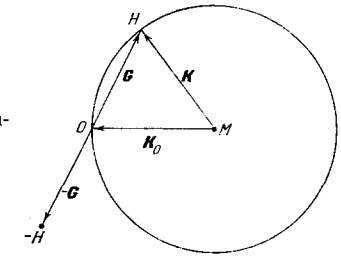

FIG. 3. Ewald's construction. O—Origin in Fourier space; the center of Ewald's sphere lies at the point M, with $\overrightarrow{OM} = -\mathbf{K}_0$.

approximation was used perferentially by a number of authors[7,18,29,30,31].

Thus, the result of recording the field from the object $U(x', y')$ is a variation of the refractive index of the hologram Δn. If the hologram is illuminated by a wave converging at the object point x_0', y_0', then diffraction results in an image that is described, apart from a constant, by the formula

$$U'(x', y') = \int |\Phi|^2 e^{i\cdot 2\pi(K - K_0)r}\, dx\, dy\, dz. \tag{7}$$

Comparing formulas (2), (3), and (7), we note that the hologram is recorded by a field corresponding to a Fourier transform with respect to x′ and y′, while the reconstructed image is obtained as a result of a three-dimensional Fourier transformation over the volume of the hologram. In other words, the Bragg condition should be satisfied in the reconstruction of the image. In the particular case of an object consisting of two points, this condition takes the form

$$\mathbf{K} - \mathbf{K}_0 = \pm \mathbf{G}; \tag{8}$$

here $\mathbf{G} = \mathbf{K}_1 - \mathbf{K}_2$ is the wave vector of one of the waves of the Fourier transformation of the distribution of the refractive index over the hologram (see (4)). We now use a construction in Fourier space, known as Ewald's construction (Fig. 3). The Fourier transform of a harmonic distribution, at sufficiently large dimensions of the hologram, will be the points H, which are mutually inverted relative to the origin. Satisfaction of the Bragg condition in this construction corresponds to location of the point H on a sphere of radius $1/\lambda$. We note that in the case under consideration, the reconstructed image is not a complete Fourier transform, since only one of the points H can be located on the Ewald sphere. It is clear that in reconstruction of an object consisting of many points, the Bragg condition should be satisfied simultaneously for a set of elementary periodics—$\mathbf{G}_q = \mathbf{K}_0 - \mathbf{K}_j$, or the points H_q—the nodes of the Fourier space—must fall simultaneously on the Ewald's sphere. The introduced concepts suffice to carry out geometrical analysis of the main experimental schemes. However, in the analysis of the intensities, the approximation in question cannot give satisfactory results, since the formulas are derived under the assumption of low diffraction efficiency. A more rigorous construction of the theory of three-dimensional holograms is possible in the dynamic approximation. As will be shown subsequently, this makes it possible to establish certain new features of three-dimensional holograms.

Inasmuch as the conclusions based on the kinematic theory are of qualitative character, they are valid for holograms with both phase and amplitude modulation.

*The formal aspect of the analogy, for the case of planar holograms, was traced in [24-27].

†The kinematic approximation employs in fact the principle of superposition of the operators of light scattering by the hologram, instead of the principle of superposition of electromagnetic waves.

In the dynamic theory, the absorption is taken into account by introducing the complex vectors K and G, which correspond to a complex dielectric constant of the hologram. We shall consider below principally phase holograms, which are of great practical interest, since they make it possible to obtain a high diffraction efficiency. A detailed solution of the problem of diffraction of light by a three-dimensional dielectric sinusoidal grating can be found in[52,32].

2. Dynamic Approximation

The principles of the dynamic theory of diffraction of electromagnetic waves by three-dimensional periodic structures were developed by Ewald and Laue[28]. At the present time, this theory is widely used to describe the diffraction of x-rays, electrons, and neutrons in nearly-perfect crystals. The need for developing a similar theory in holography was first indicated in in[32-35]. The approximations of the dynamic theory consist in the following; the dielectric constant $\eta = \eta_1 + \Delta\eta$ of the medium in which the hologram is recorded is a scalar quantity, the variation $\Delta\eta$ due to recording of the hologram is small, so that $\Delta\eta/\eta \ll 1$, and the medium is a nonmagnetic dielectric, $\mu = 1$. These assumptions are satisfied in experiment with a high degree of accuracy.

Under the foregoing assumptions, the induction D at any point inside the hologram should satisfy the equation

$$\nabla^2\mathbf{D} + \operatorname{rot}\operatorname{rot}\left(\mathbf{D}\frac{\Delta\eta}{\eta_1}\right) = \frac{\eta_1}{c^2}\frac{\partial^2\mathbf{D}}{\partial t^2}. \tag{9}$$

We seek a solution of this equation for the simplest hologram shown in Fig. 2. It follows from (4) and (5) that in this case

$$\frac{\Delta\eta}{\eta_1} = A_0 + A_{1,2}e^{2\pi i\mathbf{G}\mathbf{r}} + A_{2,1}e^{-2\pi i\mathbf{G}\mathbf{r}}. \tag{10}$$

Assume that a wave $\mathbf{D}_0\exp[2\pi i(\mathbf{K}_0\cdot\mathbf{r}) - i\omega t]$ propagates in a hologram with a dielectric-constant distribution (10). Then, accurate to the terms $A_{1,2}$, we can write

$$\mathbf{D} = \sum_{m=-1}^{+1}\mathbf{D}_m e^{2\pi i(\mathbf{K}_0+m\mathbf{G})\mathbf{r}-i\omega t} \tag{11}$$

If C is the polarization factor, we can substitute (11) and (10) in (9) to get

$$2\varepsilon_0 D_0 = D_0A_0 + C_1A_{21}D_1, \qquad 2\varepsilon_1 D_1 = D_1A_0 + C_1A_{12}D_0, \tag{12}$$

where

$$2\varepsilon_0 = \frac{K_0^2-k^2}{K_0^2}, \quad 2\varepsilon_1 = \frac{K_1^2-k^2}{K_1^2}, \quad \varepsilon \ll 1, \quad k^2 = \frac{\omega^2}{c^2}\eta_1\frac{1}{4\pi^2}.$$

Solving this system of equations, we obtain

$$\frac{D_1}{D_0} = \frac{C_1A_{1,2}}{2\varepsilon_1 - A_0} \tag{13}$$

and

$$(A_0 - 2\varepsilon_0)(A_0 - 2\varepsilon_1) = C_1^2|A_{1,2}|^2. \tag{14}$$

It is convenient to continue the analysis by constructing the dispersion surface, which is the geometric locus of the points satisfying Eq. (14), in Fourier space. We shall call these the wave points. The polarization factor is $C_1 = 1$ if $\mathbf{D}_0 \times \mathbf{K}_1 \times \mathbf{K}_0 = 0$, and $C_1 = \cos\varphi$ if $\mathbf{D}_0\cdot(\mathbf{K}_1 \times \mathbf{K}_0) = 0$. For one of these states of polarization, the form of the planar cross section of the dispersion surface is shown in Fig. 4. The asymptotes on this figure are parts of circular arcs with radii $|\mathbf{K}|(1 + A_0/2)$. Since $|\mathbf{K}| = 1/\lambda\sqrt{\eta_1}$, the arc segments for the light waves can be represented in the scale of the figure by straight lines. We note that the relative amplitudes D_0 and D_q depend on the position of the wave points relative to the diameter of the dispersion surface. Thus, if the wave point lies on the diameter, then $\epsilon_q = \epsilon_0$, $|\mathbf{K}_q| = |\mathbf{K}_0|$, and $D_q = D_0$. If the wave point is shifted, say, to the right, then the ratio D_q/D_0 tends to zero. At a sufficiently large distance, i.e., deviation from the Bragg condition, only one wave D_0 is retained (the single-wave case). We emphasize that at a specified D_0 the solution of (12) is satisfied by two wave points on different branches of the dispersion surface. This means that even for one of the noted polarization states there exist four waves D_0^1, D_q^1, D_0^2, and D_q^2 with different wave-vector lengths. This explains why the surface (14) is called the dispersion surface*.

The point L on Fig. 4 is defined by the wave vectors $|\mathbf{K}_0| = |\mathbf{K}_q| = 1/\lambda\sqrt{\eta_1}$ of the waves D_0 and D_q in the given medium prior to the recording of the hologram. The position of the wave points relative to L determines the refractive index for each wave. For example, for the waves corresponding to the lower branch of the dispersion surface, the refractive index will be smaller than the refractive index for the wave D_0 in the single-wave case. This difference leads to an interesting phenomenon when the refractive index of the medium becomes complex. Indeed, the absorption of the wave corresponding to the wave point lying on the lower branch of the dispersion surface will be smaller than the average absorption. Applying similar reasoning to the wave points on the upper branch, we shall see that under definite conditions one pair of waves will be almost completely absorbed, whereas the other propa-

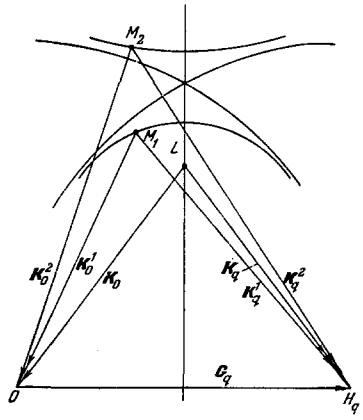

FIG. 4. Plane section of the dispersion surface of a hologram obtained from two plane monochromatic beams with wave vectors $\mathbf{K}_0$ and $\mathbf{K}_q$ ($= \mathbf{K}_1$ in formulas (11)–(14)). M_1 and M_2 are wave points.

*We point out the analogy between the behavior of an electromagnetic wave in a three-dimensional hologram and the behavior of electrons placed in a medium with a periodic potential. The energy-level splitting on the boundary of the Brillouin zone corresponds to splitting of the values of the wave vectors of the electromagnetic waves when the Bragg condition is satisfied.

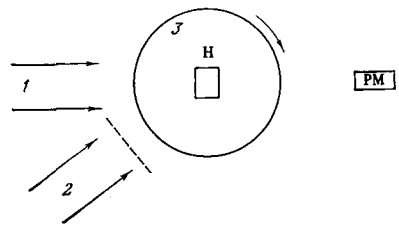

FIG. 5. Diagram of experiment for the observation of the Borrmann effect. 1, 2—plane coherent waves used to record the hologram; 3—rotating platform on which the hologram H is mounted. Wave 2 was masked out when the intensity of wave 1 was measured.

gates with anomalously low absorption. As a result, the wave emerging from the hologram will have a higher intensity than in the single-wave case. This effect was discovered by Borrmann[38] in a study of the diffraction of x-rays by a perfect crystal. Observation of the Borrmann effect in holography was reported in[31,35]. Figure 5 shows a diagram of the corresponding experiment[35]. The hologram was recorded in a colored KBr crystal from two plane waves ($\lambda = 0.63\ \mu$) incident on the surface of the crystal at an approximate angle 10° to each other. The intensity of beam 1, passing through the hologram, was measured as a function of the angle of rotation of the hologram on the platform. It turned out that the Bragg position corresponds to the maximum intensity of the transmitted beam (Fig. 6).

This result demonstrates convincingly the need for developing a dynamic approximation of diffraction theory for three-dimensional holograms.

We now consider a more complicated case, when the hologram is recorded from an object consisting of a discrete set of emitting points:

$$U(\mathbf{r}') = \sum_{j=0}^{N-1} f_j \delta(\mathbf{r}' - \mathbf{r}_j').$$

Let us find an expression for the wave field inside the hologram during the reconstruction. From (3), (4), and (10) it follows that in this case the relative variation of the dielectric constant $\Delta\eta/\eta_1$ is equal to

$$\sum_{j=0}^{N-1}\sum_{q=0}^{N-1} A_{j,q} e^{i\cdot 2\pi \mathbf{G}_{j,q}\mathbf{r}}; \qquad (15)$$

here

$$\mathbf{G}_{j,q} = \mathbf{K}_j - \mathbf{K}_q,\ \sum A_{jj} = A_0,\ A_{j,q} = f_j f_q^* \cdot \frac{2\alpha}{\sqrt{\eta_1}}.$$

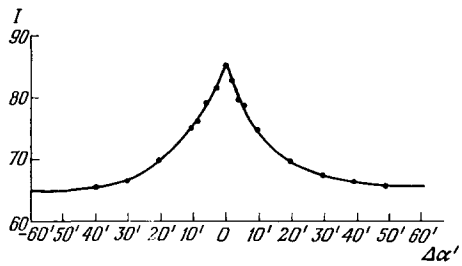

FIG. 6. The Borrmann effect. $\Delta\alpha'$—deviation of the angle between wave 1 and the plane of the hologram during the reconstruction, from the value of this angle during the recording of the hologram. I—intensity in arbitrary units.

If a wave $\mathbf{D}_0 \exp[2\pi i \mathbf{K}_0\mathbf{r} - i\omega t]$ propagates in such a hologram, then there should exist in the steady-state oscillation regime, besides this wave, also an infinite set of waves $\mathbf{D}_{j,q,m}$ with wave vectors $\mathbf{K}_{j,q,m} = \mathbf{K}_0 + m\mathbf{G}_{j,q}$. However, if $\Delta\eta/\eta_1 \ll 1$, just as in the case of the simplest holograms, all $\mathbf{D}_m$ for $|m| > 1$ will always be negligibly small compared with $\mathbf{D}_0$ and $\mathbf{D}_{j,q,1}$. We can therefore assume that there exists inside the hologram a wave that should be sought in the form

$$\mathbf{D} = \sum_{l=0}^{N-1} \mathbf{D}_l e^{i\cdot 2\pi(\mathbf{K}_l\mathbf{r}) - i\omega t}, \qquad (16)$$

where $\mathbf{K}_l = \mathbf{K}_0 + \mathbf{G}_l$ and $\mathbf{G}_l = \mathbf{G}_{j,q}$. Substituting (15) and (16) in (9) we obtain an equation for the field amplitudes

$$2\varepsilon_l D_l = \sum_{j=0}^{N-1} A_{l,j} D_j C_{l,j}; \qquad (17)$$

a particular case of this equation is (12). Starting from the same assumptions, we can obtain an equation for the wave field, when the hologram is produced by an object constituting a continuous aggregate of emitting points:

$$2\varepsilon D(\mathbf{K}) = \int A(\mathbf{K}, \mathbf{K}') D(\mathbf{K}') C(\mathbf{K}, \mathbf{K}')\, d\mathbf{K}'. \qquad (18)$$

From (17) it follows that in the Fourier space of the hologram there exists a many-sheeted dispersion surface (N sheets for N points of the object), which determines the wave vectors and amplitudes of all N^2 waves for each of the considered states of polarization. The transition from a discrete to a continuous aggregate of points on the object results in coalescence of part of the branches of the dispersion surface in the Fourier space of the corresponding hologram, and in formation of zones of allowed values of the wave vectors.

II. GEOMETRY OF RECONSTRUCTION AND INTENSITY OF RECONSTRUCTED IMAGE. ROLE OF REFERENCE BEAM

In this chapter we consider some consequences of the developed theory. We have seen (formula (15)) that for an object consisting of N points the structure of the hologram can be represented by a set of $(N^2 - N)/2$ elementary sinusoidal distributions of the refractive index Δn.

Let us examine the features of the distribution of the nodes in the Fourier space of the hologram. There are $N^2 - N$ nodes not lying in the origin, whose positions can be determined by the simple construction shown in Fig. 7 for the case N = 3. In Fourier space, the ends of all the vectors $\mathbf{G}_l = \mathbf{K}_j - \mathbf{K}_q (= \mathbf{G}_{j,q})$ lie on spheres of radius $1/\lambda$. Obviously, on each of these spheres there are N points $\mathbf{H}_{j,q}$, with q = const and j varying from 0 to N − 1, if the reconstructing wave has a wave vector $\mathbf{K}_q$. On the other hand, if the reconstruction is carried out by the wave $-\mathbf{K}_j$, then the Bragg condition is satisfied for the harmonics $\mathbf{G}_{j,q}$, j = const. In other words, when the hologram is illuminated by a wave from any point of the object, all the waves taking part in the recording of the hologram are reconstructed, and if they were constructed by

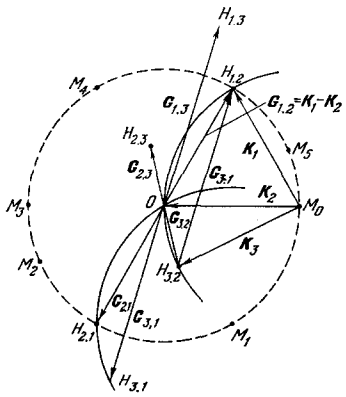

FIG. 7. Construction of nodes in Fourier space for a hologram of an object consisting of three points. Points M_q—centers of Ewald spheres, on each of which lie three points $H_{j,q}$; O—origin. The spheres are drawn from the points M_0 and M_1.

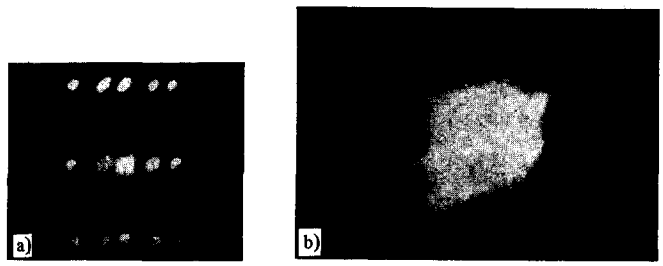

FIG. 9. a) Photograph of reconstructed image of points. The reconstruction was by waves from points that are more distinctly represented in the photograph. b) Photograph of image of triangular contour reconstructed by a wave from a point lying at the vertex of the triangle.

means of a wave converging towards one of the points of the object, a real image is reconstructed.

The character of the reasoning remains unchanged if the object is a continuous aggregate of luminous points. This is the case of a continuous distribution of nodes in Fourier space. It is described by the autocorrelation function of the distribution of the amplitudes and phases of the wave on the surface of the object (Fig. 8). To reconstruct the image of the object it is necessary that the reconstructing wave correspond to one of the waves that took part in the recording of the hologram. The position of the image of the object relative to the hologram does not depend on which wave was used for the reconstruction. Thus, it is immaterial how many waves take part in the reconstruction of the image.

To illustrate the foregoing, Fig. 9 shows photographs of images reconstructed from holograms by one and by several object points. In recording the holograms, no preferential reference beam was used.

We note that a three-dimensional hologram is an example of a square-law detector, which yields information not only on the wave amplitudes but also on the phases relative to any of them.

The dynamic approximation of the theory makes it possible to estimate the intensities of the waves in the reconstruction of the image. We shall use Eqs. (17) for the case of a hologram of an object consisting of N points. It is easily seen that in the general case the amplitudes of the reconstructed waves differ from the amplitudes of the corresponding waves used in recording the hologram. Only when the condition D_l/D_0 $\ll 1/N$ is satisfied for arbitrary l is it possible to write approximately

$$\frac{D_l}{D_0} = \frac{C_{0,l}A_{l0}}{2\varepsilon_l - A_0}, \tag{19}$$

i.e., the amplitude of the reconstructed wave is proportional to the amplitude of the corresponding wave used in recording the hologram if all the ϵ_l are equal (ϵ_l is determined from the boundary conditions and from the equation of the dispersion surface).

It is obvious that any wave from among those taking part in the recording of the hologram can be used as the reconstructing wave D_0. On the other hand, any object point can be regarded as the source of the reference recording wave for all the remaining waves emerging from the object. Thus, in the kinematic-theory approximation, there is no need for a special reference-beam for the reconstruction of the image of an object when a three-dimensional hologram is used.

Within the framework of the dynamic approximation, it is necessary to consider specially the case when one of the object points is outstanding in brightness. In this case one can assume that all the $A_{l,j} = 0$ if l and j do not pertain to the bright wave. For simplicity, we denote the index pertaining to the bright wave by zero, $A_{0,q}$ by A_{-q}, and $A_{q,0}$ by A_q. It is easy to show that to reconstruct the undistorted image it is necessary to have $\epsilon_l = \epsilon_{l'}$. In this case we obtain from (17)

$$\frac{D_l}{D_0} = \frac{A_l C_l}{\sum_j |C_j \cdot A_j|^2}(2\varepsilon_0 - A_0), \tag{20}$$

$$\sum_l \left|\frac{D_l}{D_0}\right|^2 = \frac{(2\varepsilon_0 - A_0)^2}{\sum_j |C_j \cdot A_j|^2}. \tag{21}$$

Thus, from the point of view of the dynamic theory, for a correct reproduction of the distribution of the

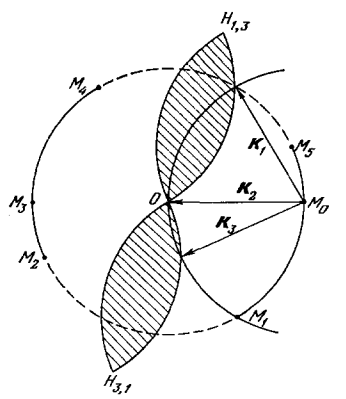

FIG. 8. Example of Fourier image of a hologram of a continuous object—a luminous band emitting coherent waves in the interval from K_1 to K_3 (shaded figure). Solid lines $M_1M_0M_5$ and $M_2M_3M_4$—centers of Ewald spheres corresponding to reconstructing waves that reconstruct virtual and real images of the object.

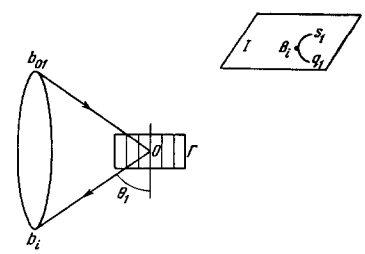

FIG. 10. Diagram illustrating the blurring of the image of an object point B_i when a three-dimensional hologram is reconstructed by an extended source. $\vec{b_{0i}O}$ and $\vec{b_iO}$ are parts of the vectors $\vec{B_0O}$ and $\vec{B_iO}$; 1—virtual image of the point B_i; θ_1—Bragg angle ($2\theta_1 = \varphi_1$).

intensities over the object it is useful to have one of the object points with preferential brightness.

This point can be regarded as the source of the reference wave, but, unlike planar holograms, it is not necessary to impose any conditions on the position of the preferential point.

III. SPECIAL CASES OF RECORDING AND RECONSTRUCTION OF THREE-DIMENSIONAL HOLOGRAMS

1. Use of Extended Sources

So far we have started from the assumption that the reconstruction of the image is carried out by a single wave. It is well known that in reconstructing an image from a planar hologram by means of an extended source it is necessary to distinguish between two cases:

a) The recording was made with the aid of a pointlike reference source; the reconstruction by an extended source leads to severe blurring of the image[3,37];

b) the recording was made with an extended source. Reconstruction of the image is possible if the reconstructing source coincides with the recording one. In this case the reconstruction of the image is equivalent to the integral operation

$$U(\mathbf{r}) * T(\mathbf{r}) * T^*(\mathbf{r}),$$

where $T(\mathbf{r})$ is the source function, and the symbol * denotes the autocorrelation operation. If $T(\mathbf{r})$ has a broad spectrum of spatial frequencies, then $T(\mathbf{r}) * T^*(\mathbf{r})$ is close to a δ function and the image of the object is reconstructed[3,38].

Let us consider now in the same sequence the reconstruction of a three-dimensional hologram by an extended source.

a) The hologram was recorded with a pointlike source. If none of the points of the reconstructing source coincides with the point B_0 that produced the reference wave during the recording of the hologram, then, as we have already seen, the image of the object is not reconstructed, and all that can appear are images of individual points outside the object. For convenience in the discussion that follows, let us analyze the process of production of the image for one of the object points B_i, if the reconstructing source contains the point B_0. The construction is shown in Fig. 10. Let K_0 be the wave vector of the wave emitted by the point of the reconstructing source which coincides with the object point B_0 during the recording of the hologram. Then K_1 is the wave vector of the reconstructed wave, forming the virtual image of the point B_1. It is seen from this scheme that there are also other directions K_0 (corresponding to the generatrices of a cone with aperture angle $180° - \varphi$), for which the appearance of a reconstructed wave of the type K_1 is possible.

We can now conclude that if the extended reconstructing source contains points that emit waves within the limits of the section of the diffraction cone, then the image of the object point B_i will be the line s_1q_1, analogous to the well known Kikuchi lines. As a result,

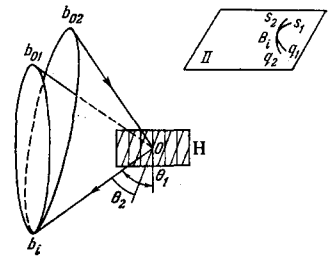

FIG. 11. Formation of the image of the object point B_i when hologram is recorded and reconstructed with an extended source. The notation is the same as in Fig. 10.

the image of the object in the reconstruction by the extended sources is blurred, as in the case of planar holograms.

b) The hologram was recorded with an extended source. We shall show that in this case the reconstruction of the image can be effected with the aid of a source of practically arbitrary form. Let B_i, as before, be one of the object points and B_{01} and B_{02} two arbitrary points of the recording extended source. It is obvious that the corresponding diffraction cones should have a common generatrix Ob_i (Fig. 11), since tangency or intersection of the cones fixes the direction of the wave that reconstructs the point B_i. Consequently, there is separated in the image plane a point of intersection of the lines s_1q_1 and s_2q_2. To each point B_{0j} of the recording source there corresponds a definite diffraction cone and a line s_jq_j in the image plane. It is clear that B_i is the point of intersection of all the lines s_jq_j, and is therefore of outstanding brightness. Thus, one should expect the reconstructed image of the object to be sharp if the number of lines taking part in the formation of each image point is sufficiently large. Obviously, this condition can be regarded as satisfied if the reconstructing and recording sources overlap at least in part. The contrast of the obtained image does not depend in this case on the total dimension and shape of the reconstructing source. Figure 12 shows a photograph of a reconstructed image of a triangle. The same extended source was used to record and reconstruct the hologram. Figure 13 shows an example of reconstruction of the image of the object (small triangle) when the recording source is a strip and is part of the reconstructing source (large triangle). The use of extended sources in the case of three-dimensional holograms makes it possible to obtain good-quality images, as is illustrated by the photograph of the reconstructed image of test pattern No. 4 on Fig. 14. We emphasize in connection with the foregoing that a three-dimensional hologram is capable of producing additionally a phantom image of the object from the part of the object contained in another object. The brightness of the obtained image is proportional to $|\int U(\mathbf{r})d\mathbf{r}|^2$, where the

FIG. 12. Photograph of reconstructed image of a triangle. The reconstruction was effected by the bright part of the triangle

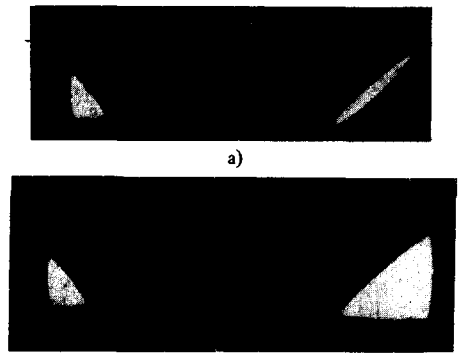

FIG. 13. a) Object (triangle) and extended reference source (strip). b) Photograph of reconstructed image of triangle (large triangle–reconstructing source).

integration is over the area that is common to the reconstructing and recording sources.

2. The Renninger Effect

In the preceding analysis we considered cases when the reconstructing wave is sent to the hologram from some external source. In this case there propagate inside the hologram waves that participate in the formation of the image of the object. The diffracted waves experience scattering in the volume of the hologram, which greatly complicates the process of reconstructing the image. Let us imagine that, for example, two holograms were recorded in a single photosensitive element by successive exposures, as shown in Fig. 15. We shall carry out the reconstruction with wave 1. The Bragg condition is then satisfied for one hologram. However, wave 3 is reconstructed with respect to the second hologram. Consequently, there will be reconstructed not only wave 3, but also an "extra" wave 2. The need for taking such a phenomenon into account when reconstructing holograms is confirmed by an experiment described in[39]. The corresponding effect is known in x-ray structure analysis as the "roundabout excitation" or the Renninger effect.

This effect can be considered only in the dynamic theory. Let us assume that we have first recorded a hologram of an object consisting of $N_1 - 1$ points with the aid of a reference wave of preferred brightness, denoted by the index 0. On the same photosensitive element, there is recorded a hologram of a second object, consisting of $N_2 - 1$ points, and one of the points of the first object, say the point with the index $q = N_1 - 1$, is the source of the reference wave. From (17) we find

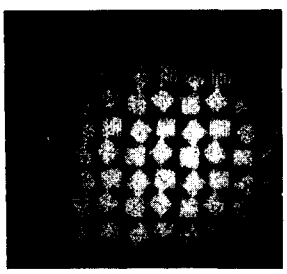

FIG. 14. Photograph of reconstructed image of pattern No. 4. The hologram was recorded with a round source of 0.2 mm diameter; the image was reconstructed with a source having a diameter of several millimeters; the centers of the two sources coincided.

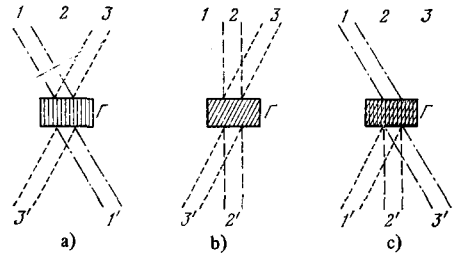

FIG. 15. Diagram of experiment for the observation of the Renninger effect. a) Recording of hologram with waves 1 and 3; b) recording of hologram with waves 2 and 3, which is reproduced after the recording of the hologram in the scheme a); c) reconstruction of image by wave 1.

expressions for the amplitudes of the reconstructed waves

$$\frac{D_q}{D_0} = \frac{A_q}{2\varepsilon_q - A_0}, \quad q \neq N_1 - 1, \qquad (22)$$

$$\frac{D_{N_1-1}}{D_0} = \frac{A_{N_1-1}C_{N_1-1,0}}{(2\varepsilon_{N_1-1} - A_0) - \sum_{m=0}^{N_2-1} \frac{|A_{N_1-1,m}C_{N_1-1,m}|^2}{2\varepsilon_m - A_0}},$$

$$\frac{D_m}{D_0} = \frac{A_{m,N_1-1}C_{m,N_1-1}}{2\varepsilon_m - A_0}\left(\frac{D_{N_1-1}}{D_0}\right);$$

here the indices q pertain to the image of the first object, and the indices m to the waves forming the image of the second. We see that the amplitudes of the waves D_q and D_m are determined by all the harmonic distributions of the refractive index, and that the amplitude D_m depends on the ratio of the amplitudes of the two reference waves.

The analyzed cases constitute an example of the extent to which the properties of three-dimensional holograms are unusual. It is interesting that the Renninger effect simulates, as it were, an associative memory similar, for example, to the linguistic memory of a man reading literature in a language with which he is not very familiar. He gains an idea of the subjects only after translating the words into his native tongue, i.e., in two stages, just as in the appearance of the "extra" wave in the reconstruction of the hologram[50]. The latter property of three-dimensional holograms can give a new impetus to discussion of the question as to whether the human memory is holographic[9,40-43].

IV. INFLUENCE OF LIMITED DIMENSIONS OF HOLOGRAM

1. Three-dimensional Hologram as an Optical Image-producing System

To estimate the possibilities of practical ultization of three-dimensional holograms, great importance attaches to the resolution during reconstruction, which depends on the real dimensions of the hologram.

The resolving power of planar holograms was considered in a number of papers (see, for example,[44]). It has been shown that the limiting resolution of points in the image of the object is limited by the diffraction of the wave by the planar aperture of the hologram. It will be shown below that registration of the wave field in three dimensions signals the creation of a qualitatively new optical image-producing system.

a) Resolution. We return to formula (7) and assume that the object is flat and has a field distribution $U(x', y')$. Since the hologram has finite dimensions, the integral (7) must be taken between certain limits. Let the hologram be a rectangular parallelepiped with dimensions X, Y, and Z. Then we have in place of (7) the function $U''(x', y')$, which describes the reconstructed image

$$U''(x', y') = U'(x', y') * L(\Delta \mathbf{K}), \tag{23}$$

where $L^2(\Delta \mathbf{K})$ is the Laue function,

$$L^2(\Delta \mathbf{K}) = \left(\frac{\sin \pi \Delta K_x X}{\pi \Delta K_x X}\right)^2 \left(\frac{\sin \pi \Delta K_y Y}{\pi \Delta K_y Y}\right)^2 \left(\frac{\sin \pi \Delta K_z Z}{\pi \Delta K_z Z}\right)^2, \tag{24}$$

and since $K_z = \sqrt{(1-\lambda^2) - K_x^2 - K_y^2}$,

$$-\Delta K_z \approx \frac{K_x}{K_z}\Delta K_x + \frac{K_y}{K_z}\Delta K_z + \frac{\Delta K_x^2}{2K_z} + \frac{\Delta K_y^2}{2K_z}.$$

$L(\Delta \mathbf{K})$ describes the image of the point as reconstructed from the hologram. At $Z = 0$, formula (23) gives the resolution of the object when the reconstruction is from a planar hologram. A three-dimensional hologram makes possible a higher resolution at given transverse dimensions X and Y. We shall illustrate this deduction with an Ewald construction. In the case of a bounded hologram, we have in the Fourier space, in place of the pointlike nodes, nodes with dimensions 2/S, 2/Y, and 2/Z, where 2/X is the distance between the first minima of $L(\Delta K_X)$. Figure 16 shows one of the nodes of the Fourier image of the hologram, which intersects the Ewald's sphere at a given direction of the reconstructing wave $\mathbf{K}_0$. The angle between the vectors $\mathbf{K} \pm \Delta \mathbf{K}_1$ and $\mathbf{K} \pm \Delta \mathbf{K}_2$ determines the resolution. It is seen from the figure that at the same dimension X, the increase of the thickness of the hologram ($Z_1 > Z_2$) leads to a decrease of the interval $\Delta \mathbf{K}$ (to a decrease in the dimensions of the image of the point).

The validity of the foregoing result was confirmed in an experiment whose scheme is shown in Fig. 17a[51]. The hologram was produced by two plane waves. A slit of width 200 μ was placed in front of the hologram during the reconstruction. Curves 1 and 2 of Fig. 17b show the intensity distribution in the reconstructed wave for holograms of thickness $Z_1 = 12$ mm and $Z_2 = 6$ mm, respectively. The positions of the minima of these curves coincide with those calculated by formula (24). The same figure shows for comparison the distribution of the intensity in diffraction of a plane wave by a slit of 200 μ (curve 3).

We call attention to the fact that the three-dimensional hologram registers values of the wave-vector components with accuracy governed by its dimensions.

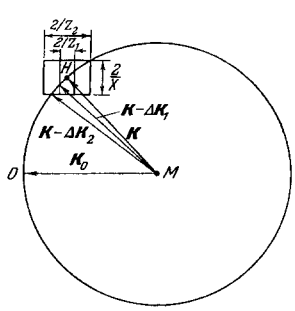

FIG. 16. Ewald construction for holograms having finite dimensions.

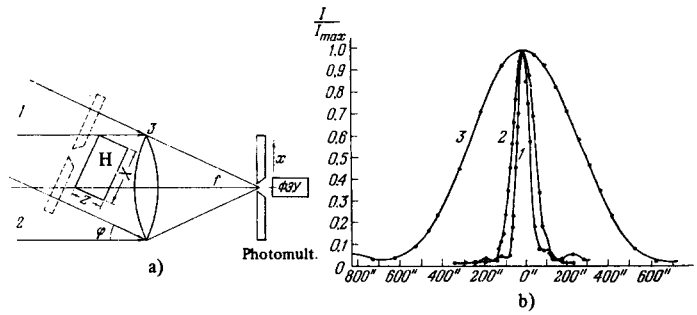

FIG. 17. a) Diagram of experiment that demonstrates the influence of the thickness of the hologram on the resolution with respect to ΔK_X (1, 2—plane waves incident at an angle $\varphi = 11.3°$ on the hologram H; 3—lens; the slit limiting the dimension of the hologram during the reconstruction is shown dashed. The intensity of the distribution was measured along the X axis). b) Experimental results (the intensity is given in relative units).

However, if the geometry in which the hologram has been obtained and hence reconstructed is known, the uncertainty in the value of the wave vector can be decreased compared with the value determined by the relation $\Delta K_X \Delta X = 1$.

For quantitative estimates it is important that when $K_X = K_Y = 0$ (for a point lying on the optical axis) the resolution increases insignificantly so long as $2X^2 > Z\lambda$.

At $X \approx Y \approx Z$, the resolution remains approximately constant as the point moves away from the optical axis. It is easy to show that a three-dimensional hologram gives an image with increased resolution also in the Z direction.

b) Spectral resolution. As noted many times in the literature[45-49], a three-dimensional hologram is a spectral instrument that makes it possible to reconstruct an image in white light. This important property of three-dimensional holograms can readily be explained by using Ewald's-construction. Unlike the preceding construction, we take the radius of the sphere to be unity, and then the modulus of the vector G is equal to λ/d. In this construction, the node in Fourier space has an additional blurring along the vector G when a nonmonochromatic light source is used for the reconstruction (Fig. 18). It follows from the construction that if the transverse dimensions of the hologram are sufficiently large, then the spectral resolution depends on the thickness Z and on the angle φ. For example, at the arrangement of the hologram and of

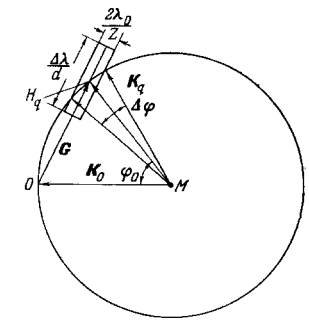

FIG. 18. Ewald's construction for the case when the reconstruction is by a polychromatic light source, and the hologram is symmetrically placed with respect to the waves $\mathbf{K}_0$ and $\mathbf{K}_q$.

the waves K_0 and K_q as shown in Fig. 18, the waves participating in the formation of the image lie in the interval $\Delta\lambda$ ($\Delta\lambda \ll \lambda_0$):

$$\Delta\lambda = \frac{2\lambda_0}{Z} \operatorname{ctg} \frac{\varphi}{2} \frac{\lambda_0}{2 \sin \frac{\varphi}{2}}.$$

Thus, a three-dimensional hologram recorded with light of wavelength λ_0 selects, when illuminated with a source of white light, a section of the spectrum in the vicinity of this wavelength. We see that the interval $\Delta\lambda$ decreases with increasing diffraction angle and becomes small at angles φ close to 180°, even at small values of Z. This explains the ability of holograms recorded on Lippmann emulsions in accordance with the Denisyuk scheme to reconstruct the image with a source of white light. It is obvious that when the hologram is recorded in complementary colors, it is possible to reconstruct a colored image of the object[49,53]. The use of an extended source during recording leads to an unexpected result. Then the spectral selectivity of the three-dimensional hologram becomes manifest only if the reconstructing source is also extended. Indeed, if the image is reconstructed with a pointlike polychromatic source, the Bragg condition is satisfied for a large spectral interval that includes λ_0. This results in a superposition of images produced by an entire set of wavelengths. If a "white" extended source is used for the reconstruction, the contrast of the image of each point of the object is determined by the intersection of a group of diffraction cones at a single generatrix (see Fig. 11). In this case the image of the object is sufficiently sharp and at the same time a narrow spectral interval is separated in the vicinity of λ_0.

In some experiments, the linear dimensions of the hologram can vary as a result, for example, of shrinkage of the film during the photographic processes, or of thermal expansion of photochromic material. Following a uniform change of the dimensions of the hologram, any vector G is replaced by G_β, where $\beta^3 = V_0/V$, and V_0 and V are the volumes of the hologram before and after the processing, respectively. The image of the object remains unchanged in this case. The reconstruction should be carried out with a wavelength $\lambda_1 = \lambda_0/\beta$. A nonisotropic change of the dimensions of the hologram can make it utterly impossible to reconstruct the image at any wavelength.

In concluding this section, let us estimate the resolving power of a three-dimensional hologram.[50] It is seen from Fig. 16 that if X = Y = Z, then the angular dimension of the image of a point remains constant at λ/X as the angle φ varies. From the Bragg equation we obtain $\Delta\lambda = d\cos(\varphi/2)\,\Delta\varphi$. If we define the resolving power R in the same manner as for plane diffraction gratings, then if N = X/d is the number of lines of the diffraction grating, $R = N/\cos(\varphi/2)$.

Thus, the resolving power of three-dimensional holograms is higher than that of plane diffraction gratings with an equal number N, and its dispersion region is no smaller. This can be of great practical importance when three-dimensional holograms are used in spectral instruments.

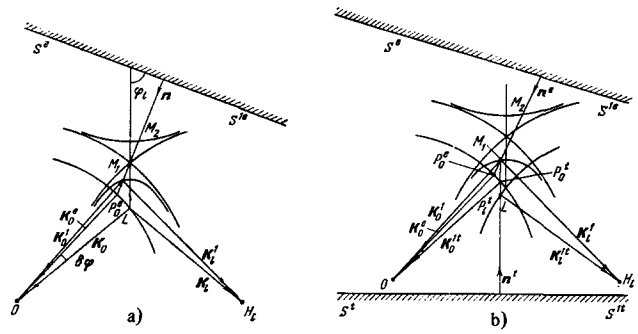

FIG. 19. Construction of the dispersion surface and determination of the wave points from the boundary conditions. M_1 and M_2—wave points chosen from the condition of equality of the wave-vector components parallel to the surface S^eS^{1e}; the points P_0^t and P_l^t on the circles are chosen from the same conditon on the surface SS^{1t}. For simplicity, the figure shows the construction only for the wave vectors connected with the point M_1. The wave vectors K_0^{2t} and K_l^{2t} of the waves emerging from the hologram are determined in the same manner. a) Semiinfinite hologram; b) hologram bounded on both sides. Since $|K^t| = |K^e| = |K|$, the points P lie on circles with radius equal to $|K| = LO = LH_l$.

2. Allowance for the Boundary Conditions

Allowance for the boundary conditions is essential for the calculation of the image intensities. The introduction of the boundary conditions is made simpler if account is taken of the fact that the hologram is surrounded by a medium having a dielectric constant η_1, whereas the dielectric constant of the hologram is $\eta_1 + \Delta\eta$, where, as before, $\Delta\eta/\eta_1 \ll 1$. We confine ourselves to the case when the hologram is recorded with a reference beam located on the same side as the waves from the object points.

If the reconstructing wave D_0 is incident on the interface between the hologram and the medium, then the condition that the phase jump be constant over the entire surface is satisfied. This means that the components of all the wave vectors parallel to the upper surface are equal to each other. This condition makes it possible to determine the wave points on the dispersion surface and, consequently, the amplitudes of the diffracted waves.

Let us perform the geometrical construction for specified boundary conditions in the two-wave case. In Fig. 19a, $\vec{LO}$ and $\vec{LH_l}$ are the wave vectors of the waves taking part in the recording of the hologram, $K_0^e = \vec{P_0^eO}$ is the vector of the reconstructing wave, and K_0^1 and K_l^1 are the vectors of waves propagating inside the hologram and specified by the positions of the points of intersection of the branches of the dispersion surface with the normal drawn from the point P_0^e to the trace of the interface S^eS^{1e}.

We denote $M_1P_0^e$ by $|K|q_1n^e$, where n^e is a unit vector normal to the surface S^eS^{1e}. Since by definition $\epsilon_l^1 = (|K_l^1| - |k|)/|k|$, it can readily be found from Fig. 19 that ϵ_l^1 is approximately equal to $g_1\sin(\psi_l - \varphi_l/2) + \delta\varphi\sin\varphi_l$. Substituting this value in (13), we obtain

$$\frac{D_l^1}{D_0} = \frac{C_l A_l}{2g_1 \sin\left(\psi_l - \frac{\varphi_l}{2}\right) + 2\delta\varphi \sin\varphi_l - A_0}. \quad (25)$$

If the hologram was recorded from a large number of points with a reference wave, then, as seen from formula (20), the solution of the wave equation does not differ formally from the solution for two waves. Therefore formula (25) gives the amplitudes of all waves inside the hologram. We denote the angles φ and ψ for the point lying in the center of the object by φ_0 and ψ_0. Then for all the remaining points $\psi_l = \psi_0 - \Delta\varphi_l/2$, $\varphi_l = \varphi_0 + \Delta\varphi_l$. We can thus write in place of (25)

$$\frac{D_l^1}{D_0} = \frac{C_l A_l}{2g_{l1} \sin\left(\psi_0 - \frac{\varphi_0}{2} - \Delta\varphi_l\right) - A_0 - 2\delta\varphi \sin(\varphi_0 + \Delta\varphi_l)}. \quad (26)$$

It is seen from (26) that the intensities of the points in the image of the object will always be distorted. This distortion is smaller for waves connected with that branch of the dispersion surface which lies closer to the point L. We note that if the hologram is obtained from a small object and $\varphi_0 = 0$, then at $\psi_0 = 90°$ the distortions are small.

We introduce analogously the boundary conditions on the output surface. Figure 19b shows the construction for the two-wave case. Here $\mathbf{K}_0^e$ is the wave vector of the reconstructing wave incident on the input surface, and $\mathbf{K}_0^{1t}$ and $\mathbf{K}_e^{1t}$ are the wave vectors of the waves emerging from the hologram.

We now consider the boundary conditions for the fields E and D. Under the assumption made above concerning the medium surrounding the hologram, neglecting quantities of the order $\Delta\eta/\eta_1$, we can write the boundary conditions on the upper surface for one reconstructing wave:

$$D_0 = \sum_{q=1}^{N+1} D_0^q, \qquad 0 = \sum_{q=1}^{N+1} D_l^q. \quad (27)$$

The waves D_l^q are specified by the wave vectors $\mathbf{K}_l^q$ for each branch of the dispersion surface. Equations (26) and (27) make it possible to determine uniquely the amplitudes of the field D_l^q in each experimental situation.

On the hologram output surface, the boundary conditions are written in the form

$$\sum_{q=0}^{N+1} M_l^q D_l^q = D_l^t M_l^{q,t}, \quad (28)$$

where $M_l^q = \exp(2\pi i \mathbf{K}_l^q \mathbf{r})$ are phase coefficients that depend on the thickness of the hologram.

We call attention to the fact that in the general case, $2N^2$ waves emerge from the hologram, and 2N different waves propagate in the direction of each point of the object. These waves form an interference pattern. The directions of all 2N waves coincide only in the case when the surfaces S^eS^{1e} and S^tS^{1t} are parallel. Therefore the condition that the hologram constitute a rectangular parallelepiped is necessary for reconstruction of the image. Assuming now that this condition is satisfied, we obtain the relative intensities of the waves $|D_l^t|^2$, if all $\epsilon_l = \epsilon_0$:

$$\left|\frac{D_l^t}{D_0^e}\right|^2 = T_l = \frac{|A_l|^2 |C_l|^2}{\sum_{q=1}^{N-1} |A_q C_q|^2} \sin^2\left(\frac{\pi k \sqrt{\sum_{q=1}^{N-1} |A_q C_q|^2}}{\cos\left(\frac{\varphi_l}{2}\right)} Z\right),$$
$$\left|\frac{D_0^t}{D_0^e}\right|^2 = 1 - \sum_{k=1}^{N-1} T_k. \quad (29)$$

Although formulas (29) were obtained for a particular case, they make it possible to draw a general conclusion: the intensities of the reconstructed waves oscillate with increasing thickness. The periods of the oscillations are proportional to $\overline{\Delta n}$, and the sum of the harmonic terms is preceded by a factor proportional to the intensity of the wave used to record the hologram. Obviously, the intensity of the reconstructed image, calculated with the aid of the dynamic theory, coincides with the intensity calculated from the formulas of the kinematic theory* if

$$Z \ll \frac{\lambda}{\Delta n}. \quad (30)$$

We call attention to the fact that the calculation made above pertains to phase holograms. For holograms recorded in a strongly absorbing medium ($\mu Z \gg 1$), the formulas become simpler, since it can be assumed that only waves possessing the smallest absorption emerge from the hologram. These waves, as already noted, correspond to the lower branch of the dispersion surface, that closest to the point L. In place of Eq. (28) we can write

$$\mathbf{D}_l^t = M_l \mathbf{D}_l.$$

In the general case, the absorption can be taken into account if it is assumed that the vectors $\mathbf{K}_{l,q}$ in (28) are complex quantities[52,53].

3. Information Capacity

As first shown by Van Heerden, a three-dimensional hologram has a large memory capacity[9]. Let us estimate the memory capacity within the framework of the kinematic approximation of diffraction theory. It was shown earlier that individual points of the image will be resolved if the corresponding nodes in the Fourier space do not overlap. If the information is recorded in the hologram with light of wavelength λ, then the periods of the harmonic distributions of Δn cannot be smaller than $\lambda/2$. This means that in Fourier space the nodes can be situated only within a sphere of radius $2/\lambda$, the center of which lies at the origin (Fig. 16). We shall assume that the hologram is a parallelepiped with dimensions X, Y, and Z. The dimensions of any node in Fourier space will be 2/X, 2/Y, and 2/Z. Then in the volume V_0 occupied by the limiting sphere we can place $n_0 = (4/3)\pi XYZ/\lambda^3$ nodes, i.e., independent memory elements. If the hologram is recorded from N points, then the Fourier space will contain $N^2 - N$ nodes and the maximum number N is equal to $\sqrt{n_0}$.

The most favorable case is when the recording is carried out with a reference wave of outstanding

*We note that the criterion (30) is simultaneously also the criterion for the smallness of the ratio D_l/D_0, when formulas (19) are valid.

brightness. Then it is necessary to take into account in Fourier space only 2N nodes, and $N = n_0/2 \approx 2.1\, V_h/\lambda^3$, where $V_h = XYZ$ is the volume of the hologram. We note that if the boundary sphere is densely filled, then the ratio of the signal power to the noise power is no smaller than the ratio of the area of all the diffraction maxima to the total area of the lateral maxima (see expression (24)). This makes it possible to obtain for the memory capacity a limiting value of approximately 5×10^2 bits for $\lambda = 1\ \mu$ and $V_h = 1\ cm^3$. This estimate actually shows the amount of information that can be recorded in a hologram by the interference method. In the holographic "reading" of this information, noise may arise due to the features of diffraction of a wave by a three-dimensional diffraction structure (for example, the Renninger effect). Therefore when $Z > \lambda/\Delta n$ the maximum number of independent points that can be recorded and reconstructed holographically will be approximately $n_0^{2/3}$; the maximum memory capacity decreases to 3×10^8 bits at $\lambda = 1\ \mu$ and $V_h = 1\ cm^3$ [50].

V. LIMITING TRANSITION FROM THREE-DIMENSIONAL HOLOGRAMS TO PLANAR ONES

1. Criteria of Transition

The separation of holograms into three-dimensional and "planar" ones is to some degree arbitrary. One can formulate only several methods of choosing the separation criteria. A criterion that meets with all the most important properties of three-dimensional holograms, which were considered in detail in the preceding chapters, can be written in the form

$$X \approx Y \approx Z, \quad V_r \gg \lambda^3. \tag{31}$$

It is frequently assumed that the main attribute of a three-dimensional hologram is the absence of a second image of the object upon reconstruction. This condition is satisfied if

$$Z \gg \frac{\lambda}{4 \sin^2(\varphi/2)}. \tag{32}$$

It is obvious that the condition (31) is more general, for its satisfaction also implies the satisfaction of the inequality (32) in a wide interval of angles φ up to values $(\lambda/Z)^{1/2}$. On the other hand, the inequality (32) can also be satisfied for thicknesses Z comparable with the wavelength λ. In this case the angle φ should be close to 180°. Such holograms, for which (32) is satisfied but not (31), possess only some of the properties of three-dimensional holograms under definite recording conditions. They must therefore be regarded as intermediate. If $Z < \lambda/4$, then the form of the wave field upon reconstruction does not depend significantly on the method used to record the hologram. The latter condition is a criterion for the limiting case of planar holograms.

The realization of the properties predicted for three-dimensional holograms by the dynamic theory is determined by formula (30). This formula gives a criterion for the applicability of the kinematic approximation of the theory. At large Δn, the effects of the dynamic theory can become manifest on holograms with small thickness Z. Therefore the thickness of the hologram must be estimated independently from the point of view of the criteria (31)—(32) as well as the criterion (30).

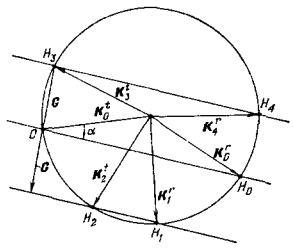

FIG. 20. Ewald's construction for a planar hologram. α—angle between the perpendicular to the plane of the hologram and the reconstructing wave $\mathbf{K}_0^t$; $\mathbf{K}_0^r$ is the specularly reflected beam. $\mathbf{K}_1^l$, $\mathbf{K}_2^t$, $\mathbf{K}_3^t$, and $\mathbf{K}_4^r$ are the diffracted waves.

2. Features of Planar Holograms

The theory developed in the preceding section can be used to describe the properties of holograms which were classified by us as planar and intermediate, and, in particular, permits a comparative analysis of planar and three-dimensional holograms. On going over to small thicknesses of the photosensitive layer, in practice it always suffices to use only the kinematic approximation. Let us consider the structure of Fourier space for a hologram. It was shown in Chap. IV that the dimension of the node in the Z direction is inversely proportional to the thickness of the hologram. Therefore all the nodes in the Fourier space are transformed into rods perpendicular to the surface of the hologram. The corresponding Ewald construction for one elementary harmonic recorded on the hologram is shown in Fig. 20. We see that the Fourier transform of one elementary harmonic can have four points of intersection with the Ewald sphere. These points correspond to diffracted waves with wave-vector directions K. The presence of the reflected wave is taken into account by constructing the rod for the zero harmonic; this rod passes through the origin. When the angle between the plane of the hologram and the reconstructing beam changes, a change takes place in the relative positions of the diffracted waves and in their intensities, and at a certain angle some waves vanish.

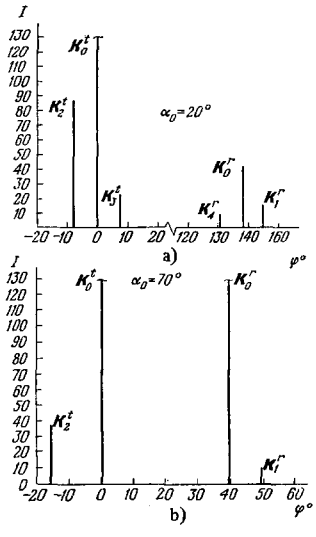

FIG. 21. Line diagram showing the dependence of the intensities and of the mutual placement of the waves K on the angle α_0 for the simplest hologram ($\alpha_0 = \alpha$ on Fig. 20). During the recording of the hologram, the angle φ_0 between the two plane waves was approximately 10°. The angle φ in the figure is reckoned from the direction of the wave $\mathbf{K}_0^t$.

An experimental confirmation of the reality of the described situation is shown in Fig. 21. A hologram recorded with two waves on an ordinary photographic emulsion was mounted on a goniometer. During the reconstruction, it was illuminated with a laser, and the intensities of the diffracted waves were registered with a photomultiplier[25]. Thus, to each node in Fourier space there correspond in the general case two diffracted rays on opposite sides of the hologram. It will suffice henceforth to consider only the rays passing through the hologram.

For any N points of the object, the Fourier space will contain $N^2 - N$ nodes, and the number of object-image points will correspond to the number of intersections of the nodes with the Ewald sphere. The locations of these points can be calculated in the following manner. Let the field distribution over the object be described by the function $U(x', y')$. Then the distribution recorded on the hologram will be the one described by formula (4) at $Z = 0$. If the hologram is reconstructed by a point source, we get

$$U'(x', y') \sim U(x', y') * U^*(x', y'). \quad (33)$$

Thus, what is reconstructed is not the field $U(x', y')$, but its auto-correlation function, i.e., the reconstructed image contains a set of displaced and overlapping images corresponding to recording by one of the points of the entire remainder of the object. In the special case when a preferential point producing a reference beam was used in the recording of the hologram, one of the images of the object is outstanding among the others in the reconstruction.

This image can be reconstructed in practice at an arbitrary angle between the reconstructing beam and the surface of the hologram. However, when this angle is varied, the locations and intensities of the image points vary. Therefore, to obtain a correct undistorted image of the object in the reconstruction, it is necessary to adhere rigorously to the geometry of the relative placement of the hologram and of the reference beam used during the recording. We note that in the case of three-dimensional holograms, satisfaction of the reconstruction condition corresponds to satisfaction of the condition of reconstruction of an undistorted image of the object. It thus turns out that it is preferable to use three-dimensional holograms in holographic interferometry, technology, and intrascopy, when it is necessary to obtain undistorted images of the objects.

A case of interest is when the reconstruction is with a nonmonochromatic light source with a spectral interval $\Delta\lambda = \lambda_2 - \lambda_1$. With the aid of the construction shown in Fig. 22, it can be shown that each elementary harmonic will produce the spectrum of the source in the angle interval

$$\Delta\varphi = \varphi_{\lambda_2} - \varphi_{\lambda_1} = \frac{\Delta\lambda}{\lambda} \operatorname{tg} \varphi. \quad (34)$$

We see that at small and large angles φ, the quantity $\Delta\varphi$ and, consequently, also the smearing of the object points during the reconstruction with a polychromatic source of light will be negligible, and all the points will turn out to be colored. This shows the possibility of reconstructing planar and intermediate holograms in "white" light. We note that for intermediate holograms it is better to use large angles φ, since at such angles the condition for separating the images is simultaneously satisfied for these angles (see (32)).

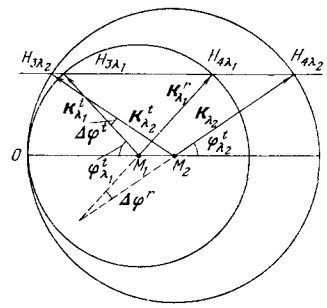

FIG. 22. Ewald's construction for the case of reconstruction of planar holograms by a source emitting electromagnetic waves in a spectral interval $\Delta\lambda = \lambda_2 - \lambda_1$. For simplicity, only one of the nodes in the Fourier space of the hologram is shown. $OM_1 = 1/\lambda_1$ and $OM_2 = 1/\lambda_2$.

In conclusion, we point out that the intermediate holograms have features in common with three-dimensional holograms not only in the simplest kinematic approximation. It has been shown experimentally that when ordinary emulsions are used one can observe the effects of the dynamic theory[31,35]. The dynamic theory for such holograms is more complicated than for three-dimensional ones, since it is necessary to deal here with extended nodes in Fourier space.

We have already mentioned that the effect of anomalously low absorption was discovered in the study of diffraction of x-rays by perfect crystals. The same effect, observed with thin holograms, is of definite interest, since it has been possible for the first time to observe anomalous passage for periodic structures whose thickness is only a few wavelengths. The performance of such an experiment in the x-ray band is difficult. The latter remark shows that three-dimensional and intermediate holograms can also be used for simulation of certain situations of x-ray and electron diffraction by perfect crystals[35,50].

VI. CONCLUSION

In spite of the technical difficulties determined by the need for creating and mastering the use of more perfect three-dimensional photosensitive materials, there are no grounds for doubting that three-dimensional holograms should make an essentially new contribution to the entire set of problems connected with the development of holography.

This pertains primarily to the theory of processes involved in the recording and reconstruction of the image, since, as shown above a three-dimensional hologram has more features in common with natural three-dimensional crystal lattices than with ordinary optical systems. The use of three-dimensional holograms assigns a much more modest role to the so-called "reference" beam, which in this case is no longer essential in principle for the production of the image of the object. On the whole, the registration of the wave field in a three-dimensional element is a more general case compared with its registration in a plane (for example, on a photographic plate), and this can lead to interesting consequences with respect to the possibilities of obtaining complete (amplitude and phase) information concerning the object.

It is presently difficult to speak of all the possible applications of three-dimensional holograms. In some cases they have considerable advantages over planar holograms, for example, when it is necessary to obtain an exact undistorted image of the object, in systems for storage and processing of information, etc. In different spectral instruments, three-dimensional holograms can ensure resolving powers close to those provided by Michelson or Fabry-Perot interferometers, at a dispersion width no smaller than that of the optical grating. One can assume that when holographic motion-picture and television systems are developed, "instantaneous" three-dimensional holograms will be used, for example, recorded with the aid of nonlinear effects, making it possible to record and read the frames in sequence.

Special emphasis must be placed on the possibilities arising in the experimental investigation of the phenomena of diffraction by three-dimensional holograms. Indeed, the holograms recorded with the aid of color centers or nonlinear optical effects make it possible to study certain properties of crystals, in view of the sensitivity to very small variations of the refractive index (or of the absorption).

On the other hand, recording of a hologram can ensure a specified structure of a three-dimensional diffraction grating. By the same token, a way is opened for simulation of different problems of structural analysis for the purpose of obtaining, in the optical band, experimental data that make it possible to evaluate complicated wave fields in the diffraction of electrons, neutrons, and x-rays. The undertaking of such investigations would be advantageous from the point of view of the dynamic theory of scattering, which requires further development.

[1] D. Gabor, Nature 161, 777 (1948); Proc. Roy. Soc. A197, 454 (1949); Proc. Roy. Soc. B64, 449 (1951).
[2] L. M. Soroko, Usp. Fiz. Nauk 90, 3 (1966) [Sov. Phys.-Usp. 9, 643 (1967)].
[3] G. W. Stroke, Introduction to Coherent Optics and Holography Academic, 1966.
[4] Ya. A. Smorodinskiĭ and L. M. Soroko, Fizika golografii (Physics of Holography), Dubna, 1968.
[5] A. L. Mikaélyan, Golografiya (Holography), Znanie, 1968.
[6] J. B. De Velis and G. O. Reynolds, Theory and Application of Holography, Addison-Wesley Publ. Co., Reading, Mass., 1967.
[7] Yu. N. Denisyuk, Dokl. Akad. Nauk SSSR 144, 1275 (1962) [Sov. Phys.-Doklady 7, 543 (1962)]; Opt. Spektrosk. 15, 522 (1963); 18, 275 (1965).
[8] E. N. Leith and J. Upatnieks, J. Opt. Soc. Am. 52, 1123 (1962).
[9] P. J. Van Heerden, Appl. Opt. 2, 393 (1963).
[10] V. N. Sintsov, 2-ya Vsesoyuznaya shkola po fizicheskim osnovam golografii (Second All-union School of the Physical Fundamentals of Holography), Moscow, 1970.
[11] Yu. N. Denisyuk and I. R. Protas, Opt. Spektrosk. 5, 721 (1963).
[12] D. H. Close and A. D. Jacobson, et al., Appl. Phys. Lett. 14, 159 (1969).
[13] M. Lescinsky and M. Miller, Opt. Comm. 1 (9), 417 (1970).
[14] J. P. Kirk, Appl. Opt. 5, 1684 (1968).
[15] V. A. Tsekhomskiĭ and I. V. Tunimanova, in "Golografiya i ee ispol'zovanie v optike" (Holography and its Use in Optics), Part II, Leningrad, (1970), p. 14.
[16] V. A. Tsekhomskiĭ, V. I. Sukhanov, D. N. Sitnik, and I. V. Tunimanova, ibid., p. 22.
[17] K. A. Garibashvili, V. V. Mumladze, and O. A. Chikhladze, Pribory i Tekhn. Éksperim. 5, 190 (1968).
[18] V. V. Aristov, V. G. Lysenko, V. B. Timofeev, and V. Sh. Shekhtman, Dokl. Akad. Nauk SSSR 183, 1039 (1969) [Sov. Phys.-Doklady 13, 1222 (1969)].
[19] D. R. Bosomworth and H. J. Gerritsen, Appl. Opt. 7, 95 (1968).
[20] F. S. Chen, J. T. LaMacechia, and D. B. Fraser, Appl. Phys. Lett. 13, 223 (1967).
[21] Bell Lab. Rec. 47, 1, 30 (1969).
[22] Bull. Am. Phys. Soc. 12, 670 (1967).
[23] L. H. Lin, JEEE 57, 2, 252 (1969).
[24] V. V. Aristov, V. L. Broude, et al., Dokl. Akad. Nauk SSSR 177, 65 (1967).
[25] V. V. Aristov, V. L. Broude, V. B. Timofeev, and V. Sh. Shekhtman, in the collection "Kvantovaya élektronika" (Quantum Electronics), No. 4, 1969.
[26] G. W. Stroke, Proc. Sym. Modern. Opt., Polytechnic Press, N. Y., 1967, p. 589.
[27] G. W. Stroke and R. C. Restrick, Nature 209, 5023, 603 (1966).
[28] P. P. Ewald, Revs. Modern. Phys., 37, 46 (1965).
[29] D. Gabor and G. W. Stroke, Proc. Roy. Soc. (London) A304, 1478, 279 (1968).
[30] E. G. Ramberg, RCA Rev. 27, 467 (1966).
[31] E. N. Leith, J. Upatnieks, A. Kozma, J. Marks, and N. Massey, Appl. Opt. 5, 1303 (1966).
[32] H. Kogelnik, Proc. Sum. Modern. Opt., Polytechnic Press, Brooklyn, 605 (1967).
[33] E. J. Saccocio, Appl. Phys. 38, 9, 10, 3995 (1967).
[34] C. B. Burchardt, JOSA 56, 11, 1502 (1966).
[35] V. V. Aristov, V. Sh. Shekchtman, and V. B. Timofeev, Phys. Lett. 28A, 10, 700 (1969).
[36] G. Borrmann, Z. Phys. 127, 297 (1950).
[37] Suzuki Takeonu and Hioki Pyulcki, Jap. J. Appl. Phys. 5, 9, 814 (1966).
[38] L. V. Koval'skiĭ and V. K. Polyanskiĭ, Opt. Spektrosk. 28 (2), 338 (1970).
[39] V. V. Aristov, V. G. Lysenko, V. Sh. Shekchtman, and V. B. Timofeev, Phys. Lett. A31 (4), 169 (1970).
[40] J. D. Valentine, Nature 220, 5166, 474 (1968).
[41] H. C. Longuet-Higgins, Nature 217, 5123, 104 (1968).
[42] P. Greguss, Nature 219, 5153, 482 (1968).
[43] N. S. Neidell, Nature 221, 5182, 755 (1969).
[44] B. P. Konstantinov, A. N. Zaĭdel', V. B. Konstantinov and Yu. I. Ostrovskiĭ, Zh. Tekh. Fiz. 36 (9), 1718 (1966) [Sov. Phys.-Tech. Phys. 11, 1279 (1967)].
[45] J. Upatnieks, J. Marks, and R. Fedorowicz, Appl. Phys. Lett. 8 (11), 287 (1966).
[46] F. A. Friesem and R. Fedorowicz, Appl. Opt. 5, 6, 1085 (1966).
[47] G. W. Stroke and A. Labeyrie, Phys. Lett. 20 (4), 368 (1966).
[48] L. Lin, K. Pennington, G. W. Stroke, and

A. Labeyrie, Bell Syst. Techn. J. 27 (4), 659, 660 (1966).
[49]G. W. Stroke, The Sci. Teacher 34, 7 (1967).
[50]V. V. Aristov, 2-ya Vsesoyuznaya shkola po fizicheskim osnovam golografii (Second All-union School of the Physical Fundamentals of Holography), Moscow, 1970.
[51]V. V. Aristov, V. G. Lysenko, V. B. Timofeev, and V. Sh. Shekhtman, Dokl. Akad, Nauk SSSR 191, 795 (1970) [Sov. Phys.-Doklady 15, 347 (1970)].
[52]H. Kogelnik, Bell. Syst. Techn. J. 48, 9 (1969).
[53]A. A. Friesem and J. L. Walker, Appl. Opt. 9, 201 (1970).

Translated by J. G. Adashko