OPTICAL METHODS OF LASER RADAR FIELD PROCESSING

N. D. Ustinov, I. N. Matveev, and V. V. Protopotov*

The book describes modern methods of receiving laser signals, viz., by interferometry, by nonlinear optics, by laser amplification, and by heterodyne mixing. The theory of these methods is considered as applied to the problem of laser radars. Investigated in detail are features such as the resolving power of interferometry and nonlinear-optics method, the efficiency of parametric conversion of laser signals, the sensitivity of laser amplifiers, and problems of optimal heterodyne detection. Methods of speckle interferometry and adaptive optics are described. The book is aimed at a large circle of scientists and engineers active in the problem of recording of laser signals, as well as to students in advanced courses of the corresponding specialties.

INTRODUCTION

The problems of laser ranging have much in common with those of radar. As for the methods of their solution in laser ranging, they frequently differ in principle from those in radar. The main reason is the difference in the wavelengths of the ranging radiation. Radar deals with electromagnetic radiation in the radio band, with wavelengths from tens of meters to millimeters. Typical laser-ranging wavelengths lie in the optical band and the wave sources are lasers of various types.

Ruby is the first subject in which lasing was realized, in 1960, at a wavelength 0.69 μm. The ruby laser is used to this day in laser ranging. More effective operating media were found subsequently. The possibility of lasing with neodymium ions at a wavelength 1.06 μm was first demonstrated in 1961. Widely used at present for lasers of this type are matrices based on yttrium–aluminum–garnet single crystals, as well as amorphous matrices of special glass.

Development of the carbon dioxide laser, operating at 10.6 μm, was reported in 1964. From among all the laser types used today in laser ranging, the carbon dioxide laser has the highest efficiency. The high output radiation powers, the ability to operate in the cw and in the pulsed regimes, as well as the high transparency of the atmosphere in this spectral region make the carbon dioxide laser most promising for the solution of ranging problems. The search for lasers of new types for laser ranging is continuing.

The wavelength difference is the foremost cause of the difference between the capabilities of laser ranging and radar. Above all, laser ranging has immeasurably greater capabilities of obtaining both coordinate and noncoordinate information on objects. At equal sizes of the receiving apertures, the potential resolving power in the optical band is larger by five orders of magnitude than, for example, in the centimeter radio band. Laser ranging, in contrast to radar, enables the operators to obtain the image of the object after one or several illumination cycles.

The reaching of the potential resolving power of laser-ranging devices is hindered, however, by such factors as the turbulence of the atmosphere and distortions (aberrations) of the optical receiving systems. To compensate for these phenomena there were developed in astronomy special methods of increasing the resolving power, such as intensity interferometry, speckle interferometry, etc. These methods, however, call for prolonged processing of a large number of measurement results.

The problems encountered in laser radars (lidars) called for the development of new identification methods that permit operation in real time. The results were the methods of the

*N. D. Ustinov — General Editor.

Translated from Metody Obrabotki Opticheskikh Polei v Lazernoi Lokatsii, pp. 3-269, 1983.

0270-2010/84/0504-0413$08.50

intensity holograms and of the modified Michelson interferometry; these are based on the use of the high monochromaticity of the lidar signal and make it possible to obtain under turbulent atmospheric conditions, within a single illumination pulse, a resolving power close to the diffractive value.

One of the most important problems in laser radars is that of raising the efficiency of reception of reflected laser signals. The reception efficiency and the associated signal/noise ratio determine the operating range of the lidar, as well as many other of its characteristics, including the accuracy.

Traditional methods of receiving laser signals, based on direct photodetection of the received radiation, have been used so far principally in the visible range. In the near and particularly middle infrared, the efficiency of the existing photoreceivers is drastically lower. At the same time, it is just these bands, particularly 1.06 and 10.6 μm, which attract the attention of the lidar developers, in view of the availability of sufficiently powerful lasers for these wavelengths, which furthermore are not greatly attenuated by the atmosphere. Recent revolutionary increase in the sensitivity of receivers for these wavelengths, on the other hand, were attained by using predetector laser amplification and parametric conversion of the frequency of the received radiation.

A widely used method of increasing the sensitivity of microwave radars is heterodyne detection. The use of this method has now become possible also in laser radars because of the substantial success reached in the required stabilization of laser transmitters. Investigations have shown that laser heterodyne receivers have much higher sensitivity than direct-detection receivers. Another important feature of the heterodyne-detection method is that in this case the electric signal at the output of the photoreceiver retains the information on the phase of the output field. A special role is therefore played by space–time fluctuations of the ranging field, as well as by aberrations of the receiver-channel optical elements. Heterodyne detection is already coming into use in modern lidars for various purposes and will certainly find ever-increasing applications.

Common to the foregoing methods is that they process, in real time, the laser ranging field ahead of the detector, and make use to one degree or another of the high monochromaticity of the laser signal. At the same time, some of the methods used in laser radars do not require the received radiation to be highly monochromatic. Speckle interferometry has already demonstrated its high effectiveness. As for adaptive optics, it still has to demonstrate its capabilities.

The problems mentioned are by far not all encountered in laser radars, but are typical of them and give an idea of their highly promising capabilities.

CHAPTER 1

BASIC CHARACTERISTICS OF LASER RADAR SIGNAL

For a correct understanding of any field-processing method used in laser radars, it is necessary to determine first the most important features of laser-radar signals and to formulate a mathematical model of these signals. The character of the information obtained from a laser radar depends to a considerable degree on such factors as the state of the atmosphere, the surface structure of the ranged object, and whether the laser operates in the pulsed or in the cw mode. The same factors determine the advantages of using a particular method of processing the radar field. Therefore, the mathematical model of the signal must adequately describe the spatial and temporal coherence of the field at the receiving aperture of the radar, the character of the reflection from the ranged object, the amplitude–phase distribution of the reflected field, and the influence of atmospheric turbulence.

The consideration of these problems must start out with a critical analysis of the forms of the laser signals used in laser radar.

1.1. Forms of Signals of Laser Radars

1. Basic Laser Types. The most widely used lasers used at present for laser radars are of three types: ruby laser (λ = 0.69 μm), neodymium–ion laser (λ = 1.06 μm), and carbon dioxide gas laser (λ = 10.6 μm).

Ruby (Al_2O_3 matrix doped with Cr^{3+}) was the first substance in which lasing was effected. This was done by Maiman and co-workers in 1960 (see [1]). They determined somewhat earlier the lifetimes and probabilities of the transitions between the levels [2] as well as the ruby absorption spectra [3].

A ruby laser is pumped in accord with a three-level scheme (Fig. 1.1). The ground band $^4F_2$ has a large absorption coefficient in the 0.55-μm region. The nonradiative transition to the upper laser level ^2E is much faster than to the ground level $^4A_2$. This leads to effective population of the ^2E level and produces the necessary population inversion of the levels $^4A_2$ and ^2E. Lasing takes place at the wavelength λ = 0.6943 μm between the sublevel $\bar{E}$ of the ^2E level and the ground level $^4A_2$. Owing to the difficulty of inverting the populations of the working levels in a three-level scheme, ruby lasers operate predominantly in the pulsed mode. A typical pulse duration in the case of Q switching is 30-50 nsec.

The main advantage of the ruby laser, from the point of view of laser radar, is that it operates in the visible region of the spectrum. This permits the use of highly sensitive photoreceivers, such as photomultipliers, for the radiation reception.

The feasibility of lasing on neodymium ions Nd^{3+} was first shown by Jackson and Nassu [4]. The matrix used was $CaWO_4$. The matrices most widely used at present are based on yttrium–aluminum–garnet (YAG) single crystals and amorphous matrices of special glass.

Pumping and lasing in neodymium lasers is in accord with a four-level scheme (Fig. 1.2). Since the lower laser level $^4I_{11/2}$ is approximately $2 \cdot 10^3$ cm^{-1} away from ground level $^4I_{9/2}$, its population is quite low. This is the cause of the considerably higher efficiency of the neodymium laser compared with the ruby laser. Just as in the ruby laser, a neodymium laser is pumped optically, mainly through an absorption band in the range 0.5-0.8 μm. The lasing corresponds to transition from the metastable level $^4F_{3/2}$ to the levels $^4I_{9/2}$, $^4I_{11/2}$, $^4I_{13/2}$ with respective wavelengths 0.9, 1.06, and 1.33 μm. The strongest luminescence line corresponds to the $^4F_{3/2}$–$^4I_{11/2}$ transition with wavelength 1.06 μm. The width of this line at room temperature is 20 nm for a glass laser and 0.67 nm for an YAG laser. The wavelength changes insignificantly with temperature.

A neodymium laser can operate in either the pulsed or the cw mode. In lidars, however, the pulsed regime is used in most cases, since it is capable of ranging over an appreciably larger distance.

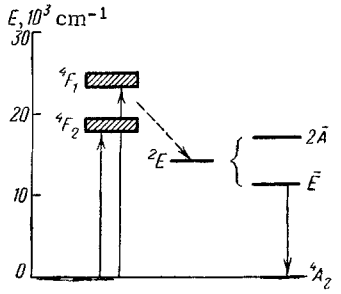


Fig. 1.1

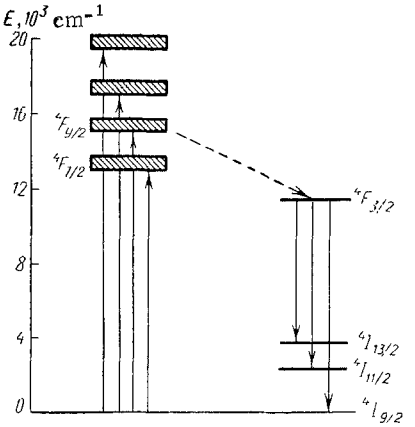


Fig. 1.2

An YAG laser operates stably at pulse repetition frequencies up to 100 Hz. It is easy to obtain from a Q-switched standard crystal 60 mm long and 3 mm in diameter a power on the order of 10 MW at a pulse duration 10-20 nsec.

The emission of a neodymium-ion laser can be effectively converted into $\lambda = 0.53$ μm visible radiation by using second-harmonic generation in a nonlinear crystal. The energy conversion coefficient can reach 80%. Neodymium lasers with second-harmonic generation are quite promising for lidar use, since the generated wavelength is in the region of maximum sensitivity of photoreceivers.

At the present time, of the three types of lasers used for lidars, the highest efficiency and largest average output power are those of gas lasers using the mixture CO_2:N_2:He [15]. A laser of this type was developed in 1964 by Patel [5, 6]. Population inversion is produced in a CO_2 laser between vibrational levels (more accurately, between vibrational—rotational bands) of the electronic ground state of the CO_2 molecule. The energy needed for population inversion is imparted to the CO_2 molecules by inelastic collisions with electrons and by resonant transfer from vibrationally excited N_2 molecules. The vibrational-level scheme of the CO_2 and N_2 molecules is shown in Fig. 1.3a.

The upper laser level of CO_2 (00°1) is only 18.6 cm^{-1} higher than the vibrational level of N_2 (v = 1). A very strong resonance interaction takes place, therefore, between these levels and leads to excitation of the upper laser level of CO_2. The lower laser levels are CO_2 (10°0) and (02°0). The lasing transition (00°1) → (10°0) occurs at the wavelength 10.6 μm, and the (00°1) → (02°0) transition at the wavelength 9.6 μm. The laser levels are split into rotational sublevels corresponding to different values of the rotational quantum number j. Therefore, the laser transition consists of R modes ($\Delta j = -1$) and P modes ($\Delta j = +1$). The corresponding transitions are designated R(j) and P(j) (Fig. 1.3b).

The high output emission powers, the possibility of both cw and pulsed operation, as well as the existence of a "transparency window" in the atmosphere at 10.6 μm make the CO_2 laser one of the most promising devices for lidars.

Besides those considered, mention should be made also of dye, excimer, and photodissociation lasers. Although not yet extensively used in lidars, one can hope them to become suitable for the solution of radar problems in the near future.

2. Analytic Signal. We shall describe hereafter the optical fields of a laser signal by using the complex representation of the electric field intensity $\mathbf{E}(\mathbf{r}, t)$, where $\mathbf{r}$ is the radius vector in the recording plane and t is the time. In most cases we shall confine ourselves to do consideration of plane-parallel radiation and use, without further stipulation, the scalar notation $E(\mathbf{r}, t)$. Cases when the radiation polarization plays a special role will be considered separately.

The central point in the description of the laser radiation field by a complex function is occupied by the concept of the analytic signal. The foundations for a correct description of physical quantities by complex functions were laid by Gabor [7]. The term "analytic signal" was first introduced by Ville [8]. The method for describing the fields consists of introducing, in lieu of a real physical quantity $E_r(\mathbf{r}, t)$, a complex quantity $E(\mathbf{r}, t)$ called the analytic signal:

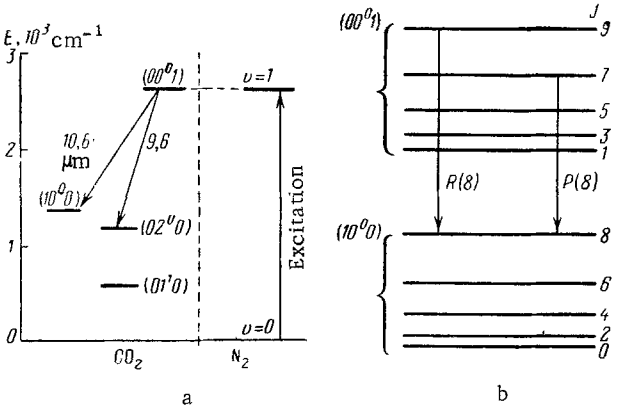


Fig. 1.3

$$E(\boldsymbol{r}, t) = E_r(\boldsymbol{r}, t) + i E_i(\boldsymbol{r}, t). \tag{1.1}$$

Since the second term in the right-hand side of (1.1) is introduced artificially, for an adequate mathematical description of $E_r(\mathbf{r}, t)$ we must impose definite conditions on the function $E_i(\mathbf{r}, t)$ [and hence on $E(\mathbf{r}, t)$]. These conditions were formulated by Gabor [7] and reduce to the following.

We first stipulate that the function $E(\mathbf{r}, t)$, on going into the complex plane of the argument, be regular and analytic in the upper half plane. Another requirement imposed on $E(\mathbf{r}, t)$ is that on going to the complex plane of the argument the modulus of the function $E(\mathbf{r}, t)$ tend rapidly to zero as the imaginary part of the argument tends to infinity in the upper half plane. It can be proved (we shall not dwell on this in detail) that $E_r(\mathbf{r}, t)$ and $E_i(\mathbf{r}, t)$ turn out to be related by a Hilbert transformation [9]. This satisfies the condition that a representation in the form (1.1) be unique.

By way of example of an analytic signal we can cite the function

$$\exp(i\omega t - ikz) = \cos(\omega t - kz) + i \sin(\omega t - kz). \tag{1.2}$$

It can be shown that $\cos(\omega t - kz)$ and $\sin(\omega t - kz)$ are a pair of Hilbert transforms [10]. The choice of the signs of the arguments in (1.2) does not play a special role and its result is that the functions turn out to be analytic either in the upper half plane when the sign of the argument is positive, or in the lower half plane in the opposite case.

The main advantage of the analytic signal is the convenience of operating with complex values, including with the complex values, including with the complex exponential (1.2). An analytic signal guarantees uniqueness of the representation of a real physical quantity, particularly of the electric field intensity $E_r(\mathbf{r}, t)$, in the form of a complex function. This eliminates automatically the ambiguity in the choice of the amplitude and phase of the field when modulated optical signals are considered: the amplitude of the field is equal to the modulus of the analytic signal, and the phase of the field is equal to the argument of the analytic signal.

In many cases, however, a rigorous expression for the analytic signal is difficult to write because of mathematical complications encountered in the calculation of Hilbert transforms. Moreover, even in those cases when such calculations can be carried through to conclusion, the complicated form of the obtained analytic signal frequently brings to naught the advantages afforded by the transition to the complex domain. One frequently uses therefore for the electric-field intensity expressions that represent only approximately an analytic signal.

One such approximation is the representation of a pulsed signal in the form

$$E(t) = A(t) \exp[i\Omega t + i\varphi(t)], \tag{1.3}$$

where $A(t)$ and $\varphi(t)$ are arbitrary real functions. The real and imaginary parts of (1.3), equal, respectively, to

$$E_r(t) = A(t) \cos[\Omega t + \varphi(t)], \quad E_i(t) = A(t) \sin[\Omega t + \varphi(t)], \tag{1.4}$$

are, generally speaking, not Hilbert transforms. Consequently, expression (1.3) is not an analytic signal, and the amplitude and phase of the field can be chosen to some degree arbitrarily.

We elucidate in this connection the constraints that must be imposed on functions A(t) and $\varphi(t)$ in order that (1.3) be approximately regarded as an analytic signal. To make the calculations that follow easily understood, we change to the frequency domain by taking the Fourier transforms of the expression in (1.4):

$$V_r(\omega) = \int_{-\infty}^{+\infty} E_r(t) \exp(-i\omega t)\, dt, \tag{1.5}$$

$$V_i(\omega) = \int_{-\infty}^{+\infty} E_i(t) \exp(-i\omega t)\, dt.$$

In the case when $E_r(t)$ and $E_i(t)$ are Hilbert transforms, the complex functions $V_r(\omega)$ and $V_i(\omega)$ are connected by the relation [11]

$$V_i(\omega) = \begin{cases} -iV_r(\omega), & \omega > 0, \\ 0, & \omega = 0, \\ iV_r(\omega), & \omega < 0. \end{cases} \tag{1.6}$$

The Fourier transform of the initial signal (1.3) is written in the form

$$V(\omega) = \int_{-\infty}^{+\infty} E(t) \exp(-i\omega t)\, dt = V_r(\omega) + iV_i(\omega). \tag{1.7}$$

Using property (1.6), we find that in the case when E(t) is an analytic signal it is necessary to satisfy the condition

$$V(\omega) = \begin{cases} 2V_r(\omega), & \omega > 0, \\ V_r(\omega), & \omega = 0, \\ 0, & \omega < 0. \end{cases} \tag{1.8}$$

It can be shown that (1.8) is the necessary and sufficient condition for the function E(t) to be an analytic signal. Thus, for the function E(t) to be an analytic signal it is necessary that its spectrum V(ω) vanish at ω < 0.

Expression $A(t) \exp[i\varphi(t)]$ in (1.3) plays the role of a modulating function. We designate the spectrum of this function in the form

$$V_M(\omega) = \int_{-\infty}^{+\infty} A(t) \exp[i\varphi(t) - i\omega t]\, dt. \tag{1.9}$$

The spectrum of the initial signal (1.3) is then expressed as

$$V(\omega) = \int_{-\infty}^{+\infty} A(t) \exp[i\varphi(t) - i(\omega - \Omega)t]\, dt = V_M(\omega - \Omega). \tag{1.10}$$

We define now the effective half-width $\Delta\omega_M$ of the modulation spectrum as that value of ω at which the spectrum $V_M(\omega)$ is approximately zero:

$$V_M(\omega) \approx 0 \quad \text{at} \quad |\omega| > \Delta\omega_M.$$

In this case the necessary condition that the initial signal (1.3) be analytic takes the form

$$V_M(\omega - \Omega) = 0 \quad \text{at} \quad \omega < 0$$

or, with allowance for the given definition,

$$\Omega \geqslant \Delta\omega_M. \tag{1.11}$$

The modulated function (1.3) can therefore be approximately regarded as an analytic signal in the case of narrow-band modulation. The stronger the inequality (1.11), the better the approximation of an analytic signal by the function E(t).

The lasers used for ranging operate in either the pulsed or in the cw regime. In the pulsed regime the width of the laser emission spectrum is as a rule of the order of 10^{11} Hz, whereas the emission frequencies of the ruby, neodymium, and CO_2 lasers are, respectively, $4.3 \cdot 10^{14}$, $2.8 \cdot 10^{14}$, and $2.8 \cdot 10^{13}$ Hz.

Thus, in the overwhelming majority of the cases the pulse modulation of the laser emission can be regarded as narrow-band, and the representation of the electric field intensity in the form (1.3) as an analytic signal. Some of the lidars employed, both cw and pulsed, mostly CO_2, have an even narrower emission band, on the order of 10^4-10^5 Hz. In this case the analytic-signal approximation is even more accurate.

3. Correlation Function of Single-Mode Laser Emission. We consider now quasimonochromatic single-mode cw laser emission. Experiments show that the electric field intensity of this emission be represented in the form

$$E(t) = a(t) \exp[i\omega t + i\varphi(t)], \tag{1.12}$$

where $a(t)$ is the real envelope of the process and $\varphi(t)$ is the phase of the field. We are not considering here the dependence on the spatial coordinates. In practice, the envelope $a(t)$ and the phase $\varphi(t)$ are random functions. As a result, the function E(t) is also a random quantity.

The fluctuations of the field amplitude are due mainly to fluctuations of the gain of the active medium. The field-phase changes are due to many factors: change of the refractive index of the medium, smooth variation of the cavity length, or vibrations of the cavity mirrors. As a rule the amplitude fluctuations of the output emission are easier to suppress than the phase fluctuations. In many cases it is possible to lower the relative fluctuation amplitudes to a level less than 10%. The main contribution to the decrease of the field coherence is then made by the phase fluctuations, and the field itself can approximately be represented as

$$E(t) = a_0 \exp[i\omega t + i\varphi(t)]. \tag{1.13}$$

Since the phase $\varphi(t)$ of the field depends on many independent parameters, it obeys Gaussian statistics. By virtue of the arbitrary choice of the initial phase, it can be assumed that the mean value of $\varphi(t)$ is zero. Let us obtain in this approximation the correlation function of the field (1.13). The normalized correlation function of single-mode laser emission is of the form

$$R(t_1, t_2) = \langle E(t_1) E^*(t_2) \rangle / a_0^2. \tag{1.14}$$

Here and below the angle brackets denote averaging over the ensemble, while an asterisk denotes a complex conjugate.

For an ergodic process such as the time-dependent laser-emission field, averaging over the ensemble is equivalent to time-averaging. Assume that the process E(t) is stationary. Then, as is well known,

$$R(t_1, t_2) = R(t_1 - t_2) = R(t_2 - t_1) = R(\tau), \quad \tau = |t_1 - t_2|. \tag{1.15}$$

Substituting (1.13) in (1.14) we obtain

$$R(\tau) = \exp(i\omega\tau) \langle \exp[i\varphi(t+\tau) - i\varphi(t)] \rangle. \tag{1.16}$$

Since $\varphi(t)$ is a normal random quantity with zero mean value, the difference $\varphi(t+\tau) - \varphi(t)$ is also a normal random quantity with a distribution given by

$$w(x)\,dx = (2\pi D)^{-1/2} \exp(-x^2/2D)\,dx, \tag{1.17}$$

where

$$D = \langle [\varphi(t+\tau) - \varphi(t)]^2 \rangle = \langle x^2 \rangle, \quad x = \varphi(t+\tau) - \varphi(t).$$

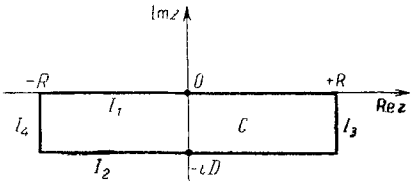


Fig. 1.4

The averaging in (1.16) can be carried out directly with the aid of (1.17). Recognizing that we shall need such mean values in the future, we shall carry out the calculation in detail. To this end we consider the integral

$$I=(2\pi D)^{-1/2}\int_{-\infty}^{+\infty}\exp(ix-x^2/2D)\,dx. \tag{1.18}$$

After simple transformations of the argument of the exponential we get

$$I=(2\pi D)^{-1/2}\exp(-D/2)\int_{-\infty}^{+\infty}\exp[-(x-iD)^2/2D]\,dx. \tag{1.19}$$

Besides the integral (1.19) we consider the integral along the contour C (Fig. 1.4) in the complex plane of the argument z:

$$I_C=(2\pi D)^{-1/2}\int_C\exp(-z^2/2D)\,dz. \tag{1.20}$$

Since the integrand in (1.20) is analytic within the integration contour, the integral (1.20) is equal to zero. The integrals along the lateral parts of the contour C are equal to

$$I_3=(2\pi D)^{-1/2}\exp(-R^2/2D)\int_0^{-iD}\exp[(2Rix+x^2)/2D]\,dx, \tag{1.21}$$

$$I_4=(2\pi D)^{-1/2}\exp(-R^2/2D)\int_{-iD}^{0}\exp[(2Rix+x^2)/2D]\,dx. \tag{1.22}$$

In the limit as $R\to\infty$ both integrals tend to zero, and the integrals along the horizontal sections of the contour C tend to the integrals

$$I_1=(2\pi D)^{-1/2}\int_{-\infty}^{+\infty}\exp(-x^2/2D)\,dx, \tag{1.23}$$

$$I_2=-(2\pi D)^{-1/2}\int_{-\infty}^{+\infty}\exp[-(x-iD)^2/2D]\,dx. \tag{1.24}$$

Since $I_C = I_1 + I_2 + I_3 + I_4 = 0$, we obtain the equality

$$(2\pi D)^{-1/2}\int_{-\infty}^{+\infty}\exp[-(x-iD)^2/2D]\,dx=(2\pi D)^{-1/2}\int_{-\infty}^{+\infty}\exp(-x^2/2D)\,dx=1. \tag{1.25}$$

Thus, returning to (1.19), we see that the correlation function of the considered emission field of a single-mode laser is equal to

$$R(\tau)=\exp(i\omega\tau)\exp\{-\langle[\varphi(t+\tau)-\varphi(t)]^2\rangle/2\}. \tag{1.26}$$

The averaging in (1.26) is carried out in the usual manner. As a result we have

$$\langle[\varphi(t+\tau)-\varphi(t)]^2\rangle=2\sigma_\varphi^2[1-C_\varphi(\tau)], \tag{1.27}$$

where σ_φ^2 is the variance of the phase and $C_\varphi(\tau)$ is the normalized correlation function of the phase. Substituting (1.27) in (1.26) we obtain ultimately

$$R(\tau) = \exp(i\omega\tau)\exp\{-\sigma_\varphi^2[1 - C_\varphi(\tau)]\}. \tag{1.28}$$

4. Correlation Function of the Field of Multimode Laser Radiation. We assume now that the laser emission field contains N modes with different frequencies ω_m. In this case the model (1.13) is valid for each mode. The total radiation field then takes the form

$$E(t) = \sum_{m=-N}^{N} a_m \exp[i\omega_m t + i\varphi_m(t)]. \tag{1.29}$$

Let us obtain its correlation function. We note for this purpose that the phases of each of the modes are statistically independent. In addition, the phase of the field and its amplitude are two independent parameters. On this basis, we obtain

$$\langle E(t)E^*(t+\tau)\rangle = \sum_{m=-N}^{N}\sum_{k=-N}^{N}\langle a_m a_k\rangle\langle\exp[-i\varphi_m(t+\tau)+i\varphi_k(t)]\rangle\exp[i(\omega_k-\omega_m)t - i\omega_m\tau]. \tag{1.30}$$

Since we regard the amplitude and the phase of the field as statistically independent quantities, the averaging over these parameters is carried out independently. In addition, by virtue of the independence of the phases of the different modes, we have

$$\langle\exp\{-i[\varphi_m(t+\tau)-\varphi_k(t)]\}\rangle = \langle\exp\{-i[\varphi_m(t+\tau)-\varphi_m(t)]\}\rangle\delta_{mk}, \tag{1.31}$$

where δ_{mk} is the Kronecker delta. Averaging the exponential in (1.31) yields the slowly varying part of the correlation function of the single-mode radiation field (1.28). Summing in (1.30) with δ_{mk}, we obtain

$$\langle E(t)E^*(t+\tau)\rangle = \sum_{m=-N}^{N}\langle a_m^2\rangle\exp\{-i\omega_m\tau - \sigma_{\varphi m}^2[1 - C_{\varphi m}(\tau)]\}. \tag{1.32}$$

To analyze in detail the behavior of (1.32), we make the simplifying assumption that all the $\sigma_{\varphi m}$ and $C_{\varphi m}$ are the same, and that the ω_m are equidistant on the frequency axis:

$$\omega_m = \omega_0 + \Delta\omega\, m. \tag{1.33}$$

Then

$$\langle E(t)E^*(t+\tau)\rangle = R(\tau)\sum_{m=-N}^{N}\exp(-i\,\Delta\omega\tau m)\langle a_m^2\rangle/\langle a_0^2\rangle, \tag{1.34}$$

where $R(\tau)$ is the correlation function of the field of one emission mode. In the case when N is so large that

$$\langle a_N^2\rangle/\langle a_0^2\rangle \ll 1, \tag{1.35}$$

the summation limits in (1.34) can be extended to infinity. In this case the series in (1.34) is a Fourier series of a periodic function of argument τ with a period $2\pi/\Delta\omega$. We denote this function by $F(\tau)$. The correlation function of the multimode-radiation field can then be written in the form

$$\langle E(t)E^*(t+\tau)\rangle = R(\tau)F(\tau). \tag{1.36}$$

If the characteristic width of the spectrum of the multimode radiation is $\Delta\Omega$ (Fig. 1.5), the characteristic decay time of the function $F(\tau)$ is $\Delta\Omega^{-1}$. The characteristic decay time of the function $R(\tau)$ (i.e., the coherence time of one mode) is equal in order of magnitude to the reciprocal of the spectral width of one mode. Since the total spectral width in multimode laser emission is as a rule much larger than the spectral width of one mode, the coherence function of multimode radiation is determined mainly by the function $F(\tau)$.

We consider now multimode pulsed laser emission. Its field can be represented in the form

$$E(t) = a(t)\sum_{m=-N}^{N} a_m \exp(i\omega_m t + i\varphi_m). \tag{1.37}$$

In this model, $a(t)$ is the real envelope of the radiation pulse, ω_m and a_m are the frequency and amplitude of the m-th radiation mode excited during the pulse-generation time. Since the pulse is as a rule very short, the phase of an individual mode has no time to change substantially. We shall therefore not consider the time dependence of the phase in this model.

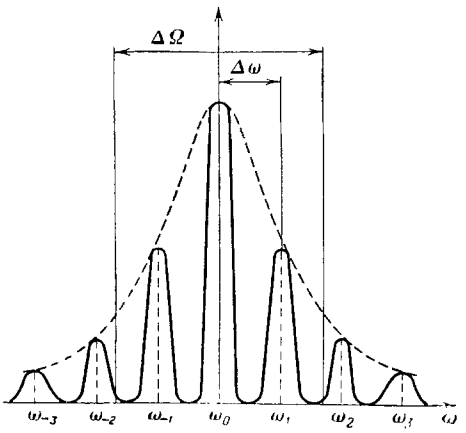


Fig. 1.5

We obtain now the correlation function of a pulsed radiation field in the form (1.37). Proceeding as in the case of cw radiation before, we obtain

$$\langle E(t)E^*(t+\tau)\rangle = a(t)a(t+\tau)\sum_{m=-N}^{N}\langle a_m^2\rangle \exp(-i\omega_m\tau). \tag{1.38}$$

We note first that in this case we are dealing with a complex correlation function of a nonstationary process. This circumstance calls for additional clarification. In the analysis of correlation functions of stationary random fields the fact that the expressions obtained were complex did not introduce additional complications.

In fact, the complex correlation function of a stationary analytic signal is uniquely connected with the real correlation function of a real physical field. This connection is expressed in the form [9]

$$R_E(\tau) = 2[R_r(\tau) + iR_{ri}(\tau)]. \tag{1.39}$$

Here $R_E(\tau)$ is the correlation function of the analytic signal E(t), $R_r(\tau)$ is the correlation function of the real part of the analytic signal $E_r(t)$, and $R_{ri}(\tau)$ is the mutual correlation function of the real and imaginary parts of the analytic signal. Expression (1.39) is a consequence of the fact that $E_r(t)$ and $E_i(t)$ are Hilbert transforms and are furthermore stationary random functions.

When pulsed signals are considered, $E_r(t)$ and $E_i(t)$ are not, as indicated, exact Hilbert transforms and are not stationary random processes. Therefore relation (1.39) is generally speaking not applicable in this case. If, however, the pulse duration is long compared with the coherence time of its carrier, relation (1.39) can be used approximately for normalized correlation functions that turn out to be approximately stationary during a time interval equal to the pulse duration.

5. Propagation of Radiation in Free Space. We consider now the laws of the spatial distribution of the field of a sounding signal in the target region. We disregard the time dependence of the field and deal only with the dependence of the field on the spatial coordinates.

To find the field distribution in the target region we must solve the problem of the diffraction of a light field of known form by the radiating aperture. To gain a clear idea of the restrictions inherent in the results that follow, we derive in detail the fundamental diffraction equation that will be used subsequently throughout the book. We use as the basis the Green-function formalism [12].

Our task is to find a function that would connect the amplitude E of the light-wave electric field intensity at an arbitrary point P_0 of a space Ω, on the one hand, with the amplitude E of the electric field intensity on the surface Σ of the given region, on the other (Fig. 1.6).

The amplitude E of the light-wave field in free space is known to satisfy the Helmholtz equation

$$\Delta E + k^2 E = 0, \tag{1.40}$$

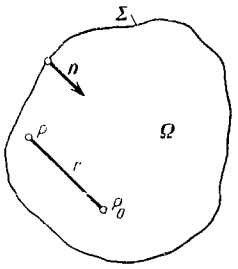


Fig. 1.6

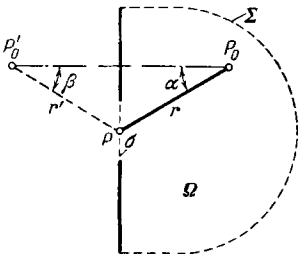


Fig. 1.7

where $k = 2\pi/\lambda$ is the wave number and Δ is the Laplace operator. For brevity, we shall denote hereafter the operator $\Delta + k^2$ by the single letter L. Equation (1.40) then takes the form

$$LE = 0 \tag{1.41}$$

with boundary conditions

$$E_{P \in \Sigma} = E(P), \tag{1.42}$$

where E(P) is an arbitrary function specified beforehand.

The Green function for the considered problem is defined as a function $G(P, P_0)$ of two variable points P and P_0, and satisfies the following conditions:

$$\left.\begin{aligned} &L_P G(P, P_0) = 0 \text{ for the entire region } \Omega, \\ &\qquad\qquad \text{excluding the point } P_0; \\ &G(P, P_0)_{P \in \Sigma} = 0; \\ &G(P, P_0) = \exp(ikr)/4\pi r + g(P, P_0). \end{aligned}\right\} \tag{1.43}$$

Here L_P denotes the action of the operator L over the variable P; $g(P, P_0)$ is a function that is regular in the region Ω and satisfies Eq. (1.41) at each point of the region Ω; the function $\exp(ikr)/4\pi r$ is called the fundamental solution of the Helmholtz equation; r is the distance between the points P and P_0.

With the aid of the Green function constructed in this manner we can relate the field $E(P_0)$ at the point P_0 with the field E(P) on the boundary of the region Ω. This connection is expressed in the form

$$E(P_0) = \int_\Sigma E(P) \frac{dG(P, P_0)}{dn} d\Sigma, \tag{1.44}$$

where n is the inward normal to the surface Σ. Without stopping to prove (1.44), we note only that it is based on the use of the known Green formula in the region Ω [12].

The problem is now to find a suitable function $g(P, P_0)$. Obviously, the form of this function depends on the concrete form of the surface Σ, since the condition (1.43b) must be satisfied. We, therefore, use a more specific formulation of the problem and consider diffraction by an opening in a flat screen. Sommerfield [13] proposed for this case the following convenient form of the function $g(P, P_0)$:

$$g(P, P_0) = -\exp(ikr')/4\pi r', \tag{1.45}$$

where r' is the distance between the point P of the surface Σ and a point P_0' symmetric to the flat part of the surface Σ (Fig. 1.7). For this choice of the function $g(P, P_0)$ by virtue of the construction, the Green function $G(P, P_0)$ is exactly equal to zero on the flat part of the surface Σ. On the remaining part of Σ the Green function is generally speaking different from zero. Therefore, a function based on (1.43) and (1.45) is approximate.

Substituting (1.45) in (1.43c) we obtain

$$G(P, P_0) = \exp(ikr)/4\pi r - \exp(ikr')/4\pi r'. \tag{1.46}$$

Calculating the derivative $dG(P, P_0)/dn$ contained in (1.44) we get

$$\frac{dG(P,\ P_0)}{dn} = -\frac{\cos\alpha}{4\pi r}\exp(ikr)\left(ik-\frac{1}{r}\right) - \frac{\cos\beta}{4\pi r'}\exp(ikr')\left(ik-\frac{1}{r'}\right). \tag{1.47}$$

The meaning of the angles α and β is clear from Fig. 1.7. On the flat part of the surface Σ this derivative is obviously

$$\left(\frac{dG(P,\ P_0)}{dn}\right)_{P\in\Sigma} = -2\frac{\cos\alpha}{4\pi r}\exp(ikr)\left(ik-\frac{1}{r}\right). \tag{1.48}$$

Neglecting the integral over the remaining part of the surface Σ, we obtain

$$E(P_0) = -\int_{\Sigma} E(P)\frac{\cos\alpha}{2\pi r}\exp(ikr)\left(ik-\frac{1}{r}\right)d\Sigma. \tag{1.49}$$

We make an additional approximation. Inasmuch as in the optical band the following inequality is obviously valid

$$k = 2\pi/\lambda \gg 1/r, \tag{1.50}$$

we can neglect the term 1/r compared with ik in the parentheses of (1.49). In addition, we confine ourselves to problems in which the angle α is small enough, so that cos α in (1.49) can be replaced by unity:

$$\cos\alpha \approx 1. \tag{1.51}$$

In this approximation we obtain

$$E(P_0) = \frac{1}{i\lambda}\int_{\Sigma} E(P)\frac{1}{r}\exp(ikr)\,d\Sigma. \tag{1.52}$$

From the practical point of view, however, Eq. (1.52) is useless, inasmuch as to solve the diffraction problem with its aid and find the field $E(P_0)$ we must know beforehand the field distribution E(P) on the right-hand side of the screen, i.e., know beforehand the solution of the diffraction problem. The usual way out of the resultant difficult situation is to use the Kirchhoff approximation. In this approximation the field E(P) on the right-hand side of the screen is assumed equal to zero, and the field E(P) in the plane of the aperture σ is assumed equal to the field that would be present in this plane in the absence of the screen. Under all these approximations we obtain ultimately

$$E(P_0) = \frac{1}{i\lambda}\int_{\sigma} E(P)\frac{1}{r}\exp(ikr)\,d\sigma. \tag{1.53}$$

6. Fresnel and Fraunhofer Diffraction Approximations. Using Eq. (1.53) we obtain the distribution of the laser-radiation field near the target (Fig. 1.8). We choose on a planar aperture σ an arbitrary origin O_1. From this point we draw along the normal **n** an axis and choose on it an arbitrary origin O_2 in the region of the target. The distance between the points O_1 and O_2 will be designated L. According to (1.53) the field at an arbitary point **r** on the target surface is

$$E(\boldsymbol{r}) = \frac{1}{i\lambda}\int_{\sigma} E(\boldsymbol{\rho})\frac{1}{R(\boldsymbol{r},\ \boldsymbol{\rho})}\exp[ikR(\boldsymbol{r},\ \boldsymbol{\rho})]\,d\sigma. \tag{1.54}$$

Usually (we consider everywhere from now on just this case) the distance L to the target is much larger than the sizes of the aperture and of the target. We can, therefore, put approximately in the denominator $R(\mathbf{r}, \boldsymbol{\rho}) \approx L$ and take this quantity outside the integral sign. We then obtain the following expression for the field at an arbitrary point **r** of the target region:

$$E(\boldsymbol{r}) = \frac{1}{i\lambda L}\int_{\sigma} E(\boldsymbol{\rho})\exp[ikR(\boldsymbol{r},\ \boldsymbol{\rho})]\,d\sigma. \tag{1.55}$$

We consider now in greater detail the dependence of R on the coordinates **r** and **ρ**. From geometric considerations we have the obvious equality

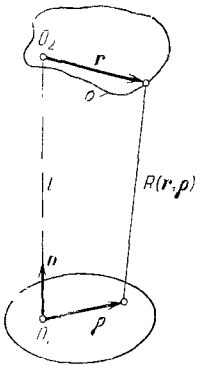


Fig. 1.8

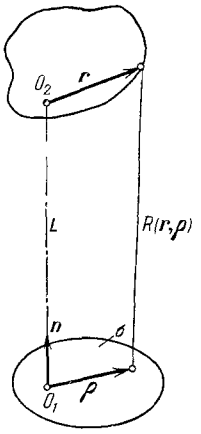


Fig. 1.9

$$R(\boldsymbol{r}, \boldsymbol{\rho}) = |L\boldsymbol{n} + \boldsymbol{r} - \boldsymbol{\rho}| \tag{1.56}$$

or

$$R(\boldsymbol{r}, \boldsymbol{\rho}) = \sqrt{(L\boldsymbol{n} + \boldsymbol{r} - \boldsymbol{\rho})^2}. \tag{1.57}$$

By simple transformations, recognizing that **n** and ρ are perpendicular,

$$\boldsymbol{n}\boldsymbol{\rho} = 0, \tag{1.58}$$

we obtain

$$R(\boldsymbol{r}, \boldsymbol{\rho}) = L\left(1 + \frac{r^2 + \rho^2}{L^2} + 2\frac{\boldsymbol{n}\boldsymbol{r}}{L} - 2\frac{\boldsymbol{r}\boldsymbol{\rho}}{L^2}\right)^{1/2}. \tag{1.59}$$

Recognizing that

$$r, \rho \ll L, \tag{1.60}$$

we can expand (1.59) in powers of the small parameters **r**/L and ρ/L and confine ourselves to the first terms of the expansion. Then

$$R(\boldsymbol{r}, \boldsymbol{\rho}) = L\left(1 + \frac{r^2 + \rho^2}{2L^2} + \frac{\boldsymbol{n}\boldsymbol{r}}{L} - \frac{\boldsymbol{r}\boldsymbol{\rho}}{L^2}\right). \tag{1.61}$$

Since the dimensions of the target are comparable with those of the aperture, all the terms in the expansion (1.61) are of the same order. Therefore, (1.55) takes the form

$$E(\boldsymbol{r}) = \frac{1}{i\lambda L}\exp\left[ikL + ik\left(\frac{r^2}{2L} + \boldsymbol{n}\boldsymbol{r}\right)\right] \int_\sigma E(\boldsymbol{\rho})\exp\left[ik\left(\frac{\rho^2}{2L} - \frac{\boldsymbol{r}\boldsymbol{\rho}}{L}\right)\right] d\sigma. \tag{1.62}$$

Even though condition (1.60) is always satisfied for laser radars, the exponential under the integral sign in (1.62) cannot be set equal to unity, and its variation with changing ρ must be taken into account in the integration. In fact, substituting the numerical values $|\boldsymbol{\rho}|$ = 1 m, L = 100 km, λ = 0.69 μm, we obtain

$$k\rho^2/2L \approx 150\,\pi,$$

i.e., the integration interval spans approximately 70 periods of the exponential.

The amplitude distribution of the field in the region of the target is determined by the modulus of the complex function

$$A(\boldsymbol{r}) = \int_\sigma E(\boldsymbol{\rho})\exp\left[ik\left(\frac{\rho^2}{2L} - \frac{\boldsymbol{r}\boldsymbol{\rho}}{L}\right)\right] d\sigma. \tag{1.63}$$

The expression in the right-hand side of (1.63) is called the Fresnel transform of the function E(ρ) [14], and it is stated in this case that the object is located in the Fresnel diffraction zone.

If the dimensions of the radiating aperature are so small that the condition

$$k\rho^2/2L \ll \pi \tag{1.64}$$

is satisfied, we can neglect on the entire surface of the aperture the corresponding term in (1.63). The amplitude distribution function is then

$$A(\boldsymbol{r}) = \int_{\sigma} E(\boldsymbol{\rho}) \exp\left(-i\, 2\pi \frac{\boldsymbol{r}\boldsymbol{\rho}}{\lambda L}\right) d\sigma, \tag{1.65}$$

i.e., when condition (1.64) is satisfied the function $A(\mathbf{r})$ and the field at the transmitting aperture are connected by a Fourier transformation. It is then said that the target is in the Fraunhofer diffraction zone.

Let the field at the radiating aperture consitute an assembly of fields of several laser modes:

$$E(\boldsymbol{\rho}) = \sum_{k=0}^{N} E_k(\boldsymbol{\rho}). \tag{1.66}$$

The laser modes are mutually orthogonal in the plane of the exit pupil, i.e., the following relation is satisfied:

$$\int_{\sigma} E_k(\boldsymbol{\rho}) E_m^*(\boldsymbol{\rho})\, d\sigma = \delta_{km} P_k, \tag{1.67}$$

where δ_{km} is the Kronecker delta and P_k is the total power of the k-th mode at the exit aperture. We shall not dwell in detail on questions connected with the theory of laser modes. We note only that relation (1.67) is a consequence of the fact that $E_k(\boldsymbol{\rho})$ is an eigenfunction of the equation for the field in the laser cavity.

We shall show that the field $E_k(\mathbf{r})$ produced in the target region by the k-th mode of the laser radiation is orthogonal to the field of the $E_m(\mathbf{r})$ of the m-th mode on a flat surface parallel to the aperture plane. To this end we calculate directly an integral of the form

$$J_{km} = \int_{S} E_k(\boldsymbol{r}) E_m^*(\boldsymbol{r})\, d^2r, \tag{1.68}$$

where S is the area of the plane over which the integral is calculated. Substituting (1.62) in (1.68) we obtain

$$J_{km} = \int_{S} \frac{1}{(\lambda L)^2} \int_{\sigma_1}\int_{\sigma_2} E_k(\boldsymbol{\rho}_1) E_m^*(\boldsymbol{\rho}_2) \exp\left[-i\frac{2\pi}{\lambda L}\boldsymbol{r}(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2) + i\frac{2\pi}{\lambda L}(\rho_1^2 - \rho_2^2)\right] d\sigma_1 d\sigma_2 d^2r. \tag{1.69}$$

In this expression we calculate first the integral

$$\frac{1}{(\lambda L)^2}\int_{S} \exp\left\{-i\frac{2\pi}{\lambda L}[r_x(\rho_{1x} - \rho_{2x}) + r_y(\rho_{1y} - \rho_{2y})]\right\} dr_x dr_y, \tag{1.70}$$

where ρ_x, ρ_y are the projections of the vector $\boldsymbol{\rho}$ in Cartesian coordinates. We use the fact that

$$\int_{-\infty}^{+\infty} \exp\left[-i\frac{2\pi}{\lambda L} r_x(\rho_{1x} - \rho_{2x})\right] dr_x = \lambda L\, \delta(\rho_{1x} - \rho_{2x}), \tag{1.71}$$

where $\delta(\rho_{1x} - \rho_{2x})$ is the Dirac delta function. Then, extending the integration limits in (1.70) to infinity we obtain after substituting (1.71) in (1.69)

$$J_{km} = \int_{\sigma_1}\int_{\sigma_2} E_k(\boldsymbol{\rho}_1) E_m^*(\boldsymbol{\rho}_2) \exp\left[i\frac{2\pi}{\lambda L}(\rho_1^2 - \rho_2^2)\right] \delta^2(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)\, d\sigma_1 d\sigma_2, \tag{1.72}$$

where we have introduced the notation

$$\delta^2(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2) = \delta(\rho_{1x} - \rho_{2x})\,\delta(\rho_{1y} - \rho_{2y}). \tag{1.73}$$

The integral (1.72) can be easily integrated with respect to $d\sigma_1$ (or $d\sigma_2$). Then, using the selectivity property of the delta function, we obtain

$$J_{km} = \int_{\sigma_2} E_k(\rho_2) E_m^*(\rho_2)\, d\sigma_2 = \delta_{km} P_k. \tag{1.74}$$

Thus, the fields $E_k(\mathbf{r})$ and $E_m(\mathbf{r})$, which belong to different modes, turn out to be orthogonal. Obviously, if we put k = m in (1.74) we obtain an expression that follows from the energy conservation law:

$$J_{kk} = P_k. \tag{1.75}$$

This equality means that the total radiation power P_k that left the transmitting aperture passed completely through the plane S over which the integral J_{kk} was calculated.

1.2. Types of Targets and Character of Reflected Signal

1. Model of Target and Description of Reflected Field. The great variety of target types encountered in lidar installations calls for finding for the target a mathematical model (or several models) describing correctly the main properties of the target. Such a model may possibly not affect the secondary parameters individually — target parameters that do not alter the general character of the signal reflected from the target.

Experimental data show that to facilitate the mathematical description all the possible target types can be subdivided into two basic classes, regardless of their sizes and of the distances to the receiving aperture, namely diffusely and specularly reflecting.

Diffusely reflecting targets have very rough surfaces, and light reflection from such a surface is described by Lambert's law. Experiments have shown that a diffusely reflecting target is equivalent to an aggregate of a large number of radiating point sources randomly disposed over the target surface. The amplitude of the light wave formed by an individual point source is proportional to the reflection coefficient of the given region of the target surface. The phases of the elementary sources can be regarded as independent. In such a model, the field reflected from the target surface is delta-correlated:

$$\langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2)\rangle = I(\boldsymbol{r}_1)\,\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2), \tag{1.76}$$

where $I(\mathbf{r}_1)$ is the average intensity of the laser radiation on the target surface at the point $\mathbf{r}_1$. Property (1.76) is valid for a diffusely reflecting target and does not depend on the correlation properties of the sounding radiation.

The radiation field reflected from the target and considered at the receiving aperture can be represented as a superposition of the fields from the individual point sources on the target surface. Introducing on the target surface a coordinate $\mathbf{r}$, and on the receiving aperture a coordinate ρ (Fig. 1.9), the field $E_k(\rho, \mathbf{r}_k)$ radiated by the k-th source located at the point $\mathbf{r}_k$ of the target surface is written at the point ρ of the receiving aperture in the form

$$E_k(\rho,\ \boldsymbol{r}_k) = a_k \frac{1}{R(\rho,\ \boldsymbol{r}_k)} \exp\left[i\varphi_k + i\frac{2\pi}{\lambda} R(\rho,\ \boldsymbol{r}_k)\right]. \tag{1.77}$$

Expression (1.77) is, as indicated in Sec. 1.1, a fundamental solution of the Helmholtz equation; it describes a spherical wave propagating from the point $\mathbf{r}_k$.

The total field $E(\rho)$ from all the point sources on the target surface is equal to the sum of all the fields, namely

$$E(\rho) = \sum_k E_k(\rho,\ \boldsymbol{r}_k). \tag{1.78}$$

Representing each k-th source as a delta function, we can formally express the right-hand side of this equation with the aid of an integral over the object surface σ:

$$E(\rho) = c \int_\sigma E(\boldsymbol{r}) \frac{1}{R(\rho,\ \boldsymbol{r})} \exp[ikR(\rho,\ \boldsymbol{r})]\, d\sigma, \tag{1.79}$$

where

$$E(\mathbf{r}) = \sum_k a_k e^{i\varphi_k} c^{-1} \delta(\mathbf{r} - \mathbf{r}_k).$$

This expression should, on going in the limit to a plane surface σ, coincide with Eq. (1.54). From this we get

$$c = \frac{1}{i\lambda}. \tag{1.80}$$

If the object contains one or several regions that produce specular reflection in the direction of the receiving aperture, it can be assumed that the object is "specular." The justification for this is the obvious argument that the radiation power incident on the receiving aperture from specular regions exceeds by many orders of magnitude the radiation power incident on the receiving aperture from diffusely reflecting regions, so that it is technically impossible to record the diffuse component of the reflecting signal. Such objects are, therefore, seen by the observer as an assembly of a small number of specularly reflecting regions.

If the specular-reflection regions are not pointlike but have finite areas, it is obvious that the object-surface sections corresponding to them are flat (or almost flat) and parallel to the receiving aperture. In this case the field $E(\rho)$ at the receiving aperture is a superposition of fields reflected from each specular region σ_m:

$$E(\rho) = \sum_m E_m(\rho). \tag{1.81}$$

The field $E_m(\rho)$ reflected from a flat region σ_m can in turn be calculated by Eq. (1.54) if the distribution of the field $E_m(\mathbf{r})$ in the region σ_m is known:

$$E_m(\rho) = \frac{1}{i\lambda} \int_{\sigma_m} E_m(\mathbf{r}) \frac{1}{R(\mathbf{r}, \rho)} \exp[ikR(\mathbf{r}, \rho)]\, d\sigma_m. \tag{1.82}$$

Substituting (1.82) in (1.81) we obtain

$$E(\rho) = \frac{1}{i\lambda} \int E(\mathbf{r}) \frac{1}{R(\mathbf{r}, \rho)} \exp[ikR(\mathbf{r}, \rho)]\, d\sigma, \tag{1.83}$$

where the integration is over all the specular regions of the target, and the field $E(\mathbf{r})$ is the total field on the target surface:

$$E(\mathbf{r}) = \sum_m E_m(\mathbf{r}). \tag{1.84}$$

Obviously, the field $E(\mathbf{r})$ is here not delta-correlated. The correlation properties of $E(\mathbf{r})$ are determined by the correlation properties of the sounding radiation and by the sizes σ_m of the specular regions. Therefore, even though (1.83) and (1.79) are outwardly identical in form, the signals recorded at the receiving aperture in the case of diffuse and specular targets are different in character. The processed signals reflected from the diffuse and specular targets turn therefore out to be different.

The character of the reflection from the target influences also the statistics of the field (or the intensity) of one mode of the reflected signal, i.e., inside the region of spatial coherence of the field at the receiving aperture. In diffuse reflection from the target, many reflection points participate in the formation of the field of one reflected-signal mode. The number of this point is very large, and the amplitude of each is random. Therefore, by virtue of the central limit theorem, the field of one radiation mode reflected from a diffuse target obeys Gaussian statistics. The statistical properties of the recorded signal of the photoreceiver on whose input acts a field of the described type have been well investigated [16].

In reflection from a specular target, the contributions made to the field by one reflected-signal mode comes mainly from a limited number of specular regions. In this case the field of one mode of the reflected signal cannot, generally speaking, be regarded as Gaussian. The statistics of the amplitude and intensity of such a field was investigated in a number of papers, e.g., in [17, 18]. It was established there that if the total number of specular regions is less than six, the statistics of the field differ greatly from Gaussian, and in the case

when the total number of reflecting elements exceeds six, the statistics of the received fiedl can be regarded as approximately Gaussian.

2. Spatial Coherence of Signal Field at the Aperture. We consider the case of a diffusely reflecting target. The signal $E(\rho)$ on the surface of the receiving aperture is expressed in terms of the field $E(\mathbf{r})$ on the surface of the target in the form of the integral (1.79) on the target surface:

$$E(\rho)=\frac{1}{i\lambda}\int_{\sigma} E(\boldsymbol{r})\frac{1}{R(\boldsymbol{r},\ \rho)}\exp[ikR(\boldsymbol{r},\ \rho)]\,d\sigma. \tag{1.85}$$

We have already encountered (in Sec. 1.1) such an expression in the description of the field formed by a radiating aperture in the target region [formula (1.54)]. Now we must solve the inverse problem.

We introduce a coordinate frame similar to that assumed in Fig. 1.9. We carry out the same transformations as in Sec. 1.1. We obtain then a final formula similar to (1.62), but with the vector $\mathbf{r}$ replaced by the vector ρ:

$$E(\rho)=\frac{1}{i\lambda L}\exp\left[ikL+ik\frac{\rho^2}{2L}\right]\int_{\sigma} E(\boldsymbol{r})\exp\left[ik\left(\frac{r^2}{2L}-\frac{\boldsymbol{r}\rho}{L}+\boldsymbol{n}\boldsymbol{r}\right)\right]d\sigma. \tag{1.86}$$

The integration in (1.86) is over the surface of the target.

Using (1.86), we obtain the correlation function of the field $E(\rho)$ on the surface of the receiving aperture

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle.$$

Substituting (1.86) we get

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=\frac{1}{\lambda^2L^2}\exp\left[ik\frac{\rho_1^2-\rho_2^2}{2L}\right]\int_{\sigma_1}\int_{\sigma_2}\langle E(\boldsymbol{r}_1)\,E^*(\boldsymbol{r}_2)\rangle\exp\left[ik\left(\frac{r_1^2-r_2^2}{2L}-\frac{\boldsymbol{r}_1\rho_1-\boldsymbol{r}_2\rho_2}{L}+\boldsymbol{n}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right)\right]d\sigma_1\,d\sigma_2. \tag{1.87}$$

The field on the surface of a diffusely reflecting target is delta-correlated:

$$\langle E(\boldsymbol{r}_1)\,E^*(\boldsymbol{r}_2)\rangle=I(\boldsymbol{r}_1)\,\delta(\boldsymbol{r}_1-\boldsymbol{r}_2). \tag{1.88}$$

Substituting (1.88) in (1.87) and integrating with respect to $d\sigma_1$ (or $d\sigma_2$) we obtain

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=\frac{1}{\lambda^2L^2}\exp\left[ik\frac{\rho_1^2-\rho_2^2}{2L}\right]\int_{\sigma} I(\boldsymbol{r})\exp\left[-ik\frac{\boldsymbol{r}(\rho_1-\rho_2)}{L}\right]d\sigma. \tag{1.89}$$

Equation (1.89) is the van Zittert–Zernike theorem, which establishes the connection between the correlation function of a field at an aperture and the distribution of the intensity $I(\mathbf{r})$ on the surface of a diffusely radiating body.

If the object is pointlike, i.e.,

$$I(\boldsymbol{r})=I_0\,\delta(\boldsymbol{r}_0-\boldsymbol{r}), \tag{1.90}$$

the correlation function is found to be

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=\frac{I_0}{\lambda^2L^2}\exp\left[ik\frac{\rho_1^2-\rho_2^2}{2L}-ik\frac{\boldsymbol{r}_0(\rho_1-\rho_2)}{L}\right]. \tag{1.91}$$

Thus, in this case the correlation function oscillates with change of the coordinates of the points ρ_1 and ρ_2. This behavior of the correlation function is due to the fact that the wavefront incident on the receiving aperture from the point $\mathbf{r}_0$ is spherical and "inclined" to the aperture plane at an angle $\theta \approx |\mathbf{r}_0|/L$. As a result, the phases of the waves at the points ρ_1 and ρ_2 are different, and it is this which leads to a relation of the form (1.91).

Introducing the field correlation function for a point source placed at the point $\mathbf{r} = 0$ and denoting this correlation function by $B(\rho_1,\ \rho_2)$, we can rewrite (1.89) in the form

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=B(\rho_1,\ \rho_2)\int_{\sigma} I_{\text{п}}(\boldsymbol{r})\exp\left[-ik\frac{\boldsymbol{r}(\rho_1-\rho_2)}{L}\right]d\sigma, \tag{1.92}$$

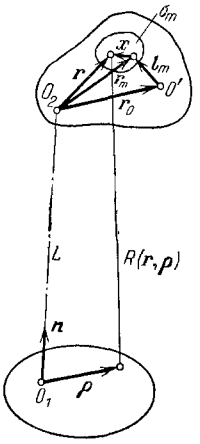


Fig. 1.10

where $I_n(\mathbf{r})$ is the normalized intensity distribution on the target surface and is equal to

$$I_n(\boldsymbol{r}) = I(\boldsymbol{r})/I_0, \quad \int_\sigma I_n(\boldsymbol{r})\, d\sigma = 1. \tag{1.93}$$

Expression (1.92) shows that an increase of the target size leads, at fixed wavelength and range, to a decrease of the coherence region of the field at the receiving aperture. Since a pointlike source produces on the aperture a spatially coherent field, the decrease of the coherence in the case of an extended source was determined by the modulus of the integral in the right-hand side of (1.92). The function

$$F(\Delta\rho) = \left| \int_\sigma I_n(\boldsymbol{r}) \exp\left[-ik\frac{\boldsymbol{r}\Delta\rho}{L} \right] d\sigma \right| \tag{1.94}$$

determines the dimension d_c of the coherence region. For example, if the target is a uniformly illuminated circle of radius r_0, it can be easily shown that the function $F(\Delta\rho)$ first vanishes at

$$|\Delta\rho| = 1.22L\lambda/2r_0. \tag{1.95}$$

The size of the coherence region can, therefore, be estimated by means of the quantity

$$d_c = L\lambda/r_0. \tag{1.96}$$

We rewrite (1.96) in a somewhat different form

$$2r_0 = 2L\lambda/d_c. \tag{1.97}$$

On the other hand, it is known that the minimum target size $2r_0$, resolved in accord with the Rayleigh criterion by a telescopic system with a receiving aperture diameter D, is (19)

$$d_t = 2r_0 \approx 2L\lambda/D. \tag{1.98}$$

Analyzing relations (1.97) and (1.98) we arrive at the following important conclusion: if the target is not resolved by the receiving aperture, the field reflected from the target is spatially coherent on this aperture. If, however, the receiving aperture resolves N elements on a diffusely reflecting target, the same aperture will span N coherence regions of the reflected field.

3. Spatial Coherence of the Field of a Signal Reflected from a Specular Target. We now obtain the correlation function of a field produced at a receiving aperture by an assembly of specular regions present on the target surface. In this case the field $E(\rho)$ on the surface of the receiving aperture is expressed in terms of the field $E_m(\mathbf{r})$ of the specular region on the target surface by means of Eq. (1.83):

$$E(\rho) = \frac{1}{i\lambda} \sum_m \int_{\sigma_m} E_m(\boldsymbol{r}) \frac{1}{R(\boldsymbol{r}, \rho)} \exp[ikR(\boldsymbol{r}, \rho)]\, d\sigma_m, \tag{1.99}$$

where the integral is calculated over the surface σ_m of each specular region.

Before we proceed, we introduce some additional notation (Fig. 1.10). As before, we choose on the aperture the origin O_1. On the continuation of the normal **n** to the plane of the aperture we choose the origin O_2 at a distance L from the aperture. In the target region, at the center of the cluster of specular regions, we choose a point O' such that it is equidistant on the average from all the specular regions σ_m. The characteristic dimension of the m-th specular region we denote by f_m. The vector joining the point O_2 with the center of the m-th specular region will be denoted $\mathbf{r}_m$. The following equality holds in this case (Fig. 1.10):

$$\boldsymbol{r}_m = \boldsymbol{r}_0 + \boldsymbol{l}_m. \tag{1.100}$$

It follows from (1.99) that the field at the receiving aperture is a superposition of fields from individual specular regions, each of which produces at the aperture a field

$$E_m(\boldsymbol{\rho}) = \frac{1}{i\lambda}\int\limits_{\sigma_m} E_m(\boldsymbol{r})\frac{1}{R(\boldsymbol{r},\ \boldsymbol{\rho})}\exp[ikR(\boldsymbol{r},\ \boldsymbol{\rho})]\,d\sigma_m. \tag{1.101}$$

Carrying out the same calculations as in the derivation of (1.86), we obtain

$$E_m(\boldsymbol{\rho}) = \frac{1}{i\lambda L}\exp\left[ik\left(L+\frac{\rho^2}{2L}\right)\right]\int\limits_{\sigma_m} E_m(\boldsymbol{r})\exp\left[ik\left(\frac{r^2}{2L}-\frac{\boldsymbol{r}\boldsymbol{\rho}}{L}+\boldsymbol{n}\boldsymbol{r}\right)\right]d\sigma_m. \tag{1.102}$$

We introduce now additionally a vector **x** located in the plane of the m-th specular region and equal to

$$\boldsymbol{x} = \boldsymbol{r} - \boldsymbol{r}_m. \tag{1.103}$$

Expression (1.102) then takes the form

$$E_m(\boldsymbol{\rho}) = \frac{1}{i\lambda L}\exp\left[ik\left(L+\frac{r_m^2}{2L}+\boldsymbol{n}\boldsymbol{r}_m+\frac{\rho^2}{2L}\right)\right]\int\limits_{\sigma_m} E_m(\boldsymbol{r}_m+\boldsymbol{x})\exp\left[ik\left(\frac{\boldsymbol{r}_m\boldsymbol{x}}{L}+\frac{x^2}{2L}+\boldsymbol{n}\boldsymbol{x}-\frac{\boldsymbol{r}_m\boldsymbol{\rho}}{L}-\frac{\boldsymbol{\rho}\boldsymbol{x}}{L}\right)\right]d^2x. \tag{1.104}$$

We note now that the surface of the specular region is approximately perpendicular to the normal **n**. We can therefore put approximately in (1.104)

$$\boldsymbol{n}\boldsymbol{x} = 0. \tag{1.105}$$

In addition, it follows from the experimental data that the characteristic dimension f_m of the specular regions is small enough and as a rule does not exceed 0.1 m. Under these conditions it can be assumed that

$$kf_m^2/2L \ll \pi. \tag{1.106}$$

For example, at f_m = 0.1 m, λ = 10.6 μm, and L = 10 km we obtain

$$kf_m^2/2L \approx 0.1\pi. \tag{1.107}$$

Thus, we can neglect in (1.104) the term $k\mathbf{x}^2/2L$ in the argument of the exponential. As a result we get

$$E_m(\boldsymbol{\rho}) = \frac{1}{i\lambda L}\exp\left[ik\left(L+\frac{\rho^2}{2L}+\frac{r_m^2}{2L}+\boldsymbol{n}\boldsymbol{r}_m-\frac{\boldsymbol{r}_m\boldsymbol{\rho}}{L}\right)\right]\int\limits_{\sigma_m} E_m(\boldsymbol{r}_m+\boldsymbol{x})\exp\left[ik\left(\frac{\boldsymbol{r}_m}{L}-\frac{\boldsymbol{\rho}}{L}\right)\boldsymbol{x}\right]d^2x. \tag{1.108}$$

Assume that the target is uniformly illuminated over the entire surface. In this case we can put

$$E_m(\boldsymbol{r}_m+\boldsymbol{x}) = E_0\exp(ik\theta\boldsymbol{x}), \tag{1.109}$$

where θ is the angle at which reflection takes place from the specular region. The integral in (1.108), which describes the amplitude distribution of the field on the aperture surface, can be written in the form of a certain function $A_m(\boldsymbol{\rho} - \mathbf{r}_m - L\theta)$, and the field $E(\boldsymbol{\rho})$ can be represented as the sum of the fields $E_m(\boldsymbol{\rho})$:

$$E(\rho)=\sum_m A_m(\rho-r_m-L\theta)\exp\left[ik\frac{\rho^2}{2L}+i\varphi_m-ik\frac{r_m\rho}{L}\right]. \tag{1.110}$$

In (1.110) we have introduced the notation

$$\varphi_m=k\left(\frac{r_m^2}{2L}+nr_m\right), \tag{1.111}$$

and the factor exp (ikL)/iλL, from (1.108) is included in $A_m(\rho - r_m - L\theta)$.

In the case when each individual specular region σ_m is not resolved by the receiving aperture, i.e., when the condition is

$$\lambda L/D\gg f_m, \tag{1.112}$$

where D is the diameter of the receiving aperture, we can assume approximately that within the limits of the receiving aperture

$$A_m(\rho-r_m-L\theta)\approx A_m. \tag{1.113}$$

In fact, the function $A_m(\rho - r_m - L\theta)$ is the Fourier transform of the distribution of the field on the area σ_m. Since we have assumed a uniform distribution of the field on σ_m, the dimension f_m of this area is connected with the characteristic width $\Delta\rho$ of the function A_m by a relation known from the theory of Fourier transforms:

$$\Delta\rho\, f_m\approx\lambda L. \tag{1.114}$$

Inasmuch as by assumption an individual region σ_m is not resolved by the receiving aperture, i.e., relation (1.112) holds, it follows from (1.114) and (1.112) that

$$\Delta\rho\gg D. \tag{1.115}$$

Consequently, the function A_m changes little within the limits of the aperture, and relation (1.113) is valid.

Thus, by using (1.100), we can finally write

$$E(\rho)=\exp\left[ik\left(\frac{\rho^2}{2L}-\frac{r_0\rho}{L}\right)\right]\sum_m A_m\exp\left(i\varphi_m-ik\frac{l_m\rho}{L}\right). \tag{1.116}$$

We now obtain the correlation function of the field $E(\rho)$. From (1.116) we have directly

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=B(\rho_1,\rho_2)\left\langle\sum_{m,n}A_m A_n^*\exp\left[i(\varphi_m-\varphi_n)-ik\left(\frac{l_m\rho_1}{L}-\frac{l_n\rho_2}{L}\right)\right]\right\rangle\left[\sum_m\langle|A_m|^2\rangle\right]^{-1}, \tag{1.117}$$

where we have introduced, by definition, the correlation function of a point source

$$B(\rho_1,\rho_2)=\exp\left\{ik\left[\frac{\rho_1^2-\rho_2^2}{2L}-\frac{r_0(\rho_1-\rho_2)}{L}\right]\right\}\sum_m\langle|A_m|^2\rangle. \tag{1.118}$$

The correlation function (1.118) is completely analogous to the previously introduced correlation function (1.91) of a point source. The only difference is that in (1.118) the role of the total intensity is played by the sum of the intensities of all the specular components.

Assume that the fields reflected from each specular region are statistically independent. In addition, it is natural to assume that the dimension of the specular region and its relative position on the surface of the target are statistically independent. Then the mean values of all the cross products in (1.117) vanish because the m-th and n-th fields are independent, and the averages over the phases and over the amplitudes of the diagonal terms are separated. As a result we get

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=B(\rho_1,\rho_2)\sum_m\langle|A_m|^2\rangle\left\langle\exp\left[-i\frac{k}{L}l_m(\rho_1-\rho_2)\right]\right\rangle\left[\sum_m\langle|A_m|^2\rangle\right]^{-1}. \tag{1.119}$$

If the average statistical properties of all the specular regions are the same, Eq. (1.119) simplifies and reduces to the form

$$\langle E(\rho_1)\,E^*(\rho_2)\rangle=B(\rho_1,\rho_2)\left\langle\exp\left[-i\frac{k}{L}l(\rho_1-\rho_2)\right]\right\rangle, \tag{1.120}$$

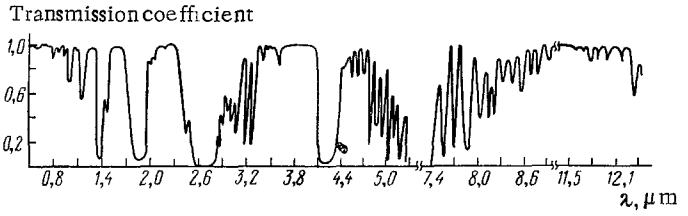


Fig. 1.11

where the vector $\boldsymbol{l}$ pertains to an arbitrary specular region. The problem has thus been reduced to finding the mean value of the second factor in (1.120). The averaging should be carried out over the angle α between the vectors $\boldsymbol{l}$ and $\boldsymbol{\rho}_1-\boldsymbol{\rho}_2$ and over the modulus of the vector $\boldsymbol{l}$.

In accordance with the remarks made earlier concerning the choice of the point O' in the target region (Fig. 1.10), the probability densities for the angle α and for the modulus of the vector $\boldsymbol{l}$ can be assumed to be uniform:

$$P(\alpha)=1/2\pi, \quad P(|\boldsymbol{l}|)=1/R, \tag{1.121}$$

where R is the mean radius of the target. The averaging over the angle is carried out using the integral representation of Bessel functions [20]:

$$\left\langle \exp\left(-i\frac{k}{L}|\boldsymbol{l}|\,|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\cos\alpha\right)\right\rangle=\frac{1}{2\pi}\int_0^{2\pi}\exp\left(-i\frac{k}{L}|\boldsymbol{l}|\,|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\cos\alpha\right)d\alpha=J_0\left(\frac{k}{L}|\boldsymbol{l}|\,|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right), \tag{1.122}$$

where J_0 is a Bessel function of zero order. We now average over the amplitude:

$$\left\langle J_0\left(\frac{k}{L}|\boldsymbol{l}|\,|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right)\right\rangle=\frac{1}{R}\int_0^R J_0\left(\frac{k}{L}|\boldsymbol{l}|\,|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right)d|\boldsymbol{l}|=H\left(\frac{k}{L}R|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right)\left(\frac{k}{L}R|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right)^{-1}. \tag{1.123}$$

The function H(x) introduced in this expression is equal to

$$H(x)=\int_0^x J_0(y)\,dy. \tag{1.124}$$

We obtain ultimately the following expression for the correlation function of the field reflected from a target with specular regions that are not resolved by the aperture:

$$\langle E(\boldsymbol{\rho}_1)E^*(\boldsymbol{\rho}_2)\rangle=B(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)H\left(\frac{k}{L}R|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right)\left(\frac{k}{L}R|\boldsymbol{\rho}_1-\boldsymbol{\rho}_2|\right)^{-1}. \tag{1.125}$$

In the limiting case $R\to 0$ the correlation function of the field goes over into the correlation function for a point source. Appendix I contains a detailed table of the values of the functions H(x) and H(x)/x.

1.3. Effect of Atmosphere on Laser-Signal Propagation

1. General Description of the Atmosphere. The action of the atmosphere on laser radiation propagating in it reduces mainly to three factors: attenuation of the radiation by scattering and absorption, refraction, and turbulent distortion of the wave front.

Attenuation in an atmosphere free of clouds and fog consists of scattering of the light by the gas and water-vapor molecules, and of selective absorption. In the visible part of the spectrum, at wavelengths 0.4-0.75 μm, the principal attenuation is due to scattering of the radiation by molecules and solid particles. The coefficient of molecular scattering is proportional to λ^{-4}, where λ is the radiation wavelength. The empirically obtained law governing the variation of the coefficient of scattering by solid particles is of the form $\lambda^{-1.75}$. The role of the selective absorption in the wavelength range 0.4-0.75 μm is negligible. At wavelengths above 1 μm the principal role in the attenuation of the radiation is played by selective absorption (Fig. 1.11). The spectrum sections with high transmission coefficients are called the "transparency windows" of the atmosphere. The "transparency windows" are largest in the ranges 0.95-1.06, 1.2-1.3, 1.5-1.8, 2.1-2.4, 3.3-4.0, 8.0-12.0 μm.

Astronomic refraction is the refraction of rays coming from a celestial luminary or another source (at high altitude) to the observer. When the ray passes through a layer of the atmosphere whose refractive index increases as the earth's surface is approached, the ray trajectory is convex towards the zenith. Depending on the zenith angle of observation, the refraction error can range from zero (at the zenith) to several minutes of degree. The refraction error must be calculated beforehand with sufficient accuracy.

2. Radiation Propagation in a Turbulent Atmosphere. The atmosphere's refractive index is not constant. Wind and local temperature inhomogeneities cause the atmosphere's refractive index to become dependent on the time and the coordinates. These inhomogeneities, called turbulences, lead to random changes of the amplitude and phase of laser radiation propagating through the turbulence region. Random amplitude–phase modulation of laser radiation from the target leads to loss of coherence of the propagating radiation, to random displacements of the laser beam, and to an overall deterioration of the lidar-signal processing conditions.

The propagation of laser radiation in a turbulent atmosphere is the subject of many studies (see, e.g., [21, 22]). Without undertaking to summarize all these investigations or to single out the most general approach to this problem we shall attempt, following the character of exposition used earlier, to demonstrate in illustrative form one of the possible solutions of this problem. In so doing we shall follow [23] and [24].

The mathematical formulation of the problem is the following: it is required to find the value of the complex field E(P) at an arbitrary point P in space, given the distribution of the refractive index n(P) of the medium in the considered space and the field distribution $E(P_1)$ at an arbitrary point P_1 on a finite surface σ of the target.

We consider next the scalar equations for the field, assuming them to be the indentical and independent for each orthogonal component of the field polarization. Under these conditions the equation for the electric field intnesity of a monochromatic light wave is of the form

$$[\Delta_P + k^2 n^2(P)]\, E(P) = 0, \tag{1.126}$$

where ΔP is the Laplace operator acting on the coordinates of the point P; $k = 2\pi/\lambda$ is the wave number. In (1.126) it was assumed that the distribution of the refractive index n(P) does not manage to change while the field E propagates in the considered volume of space. This condition is satisfied in practically all cases.

The experimental data show that the time interval during which the distribution of the refractive index does not manage to change is about 1 msec. During this time the laser beam covers an approximate distance 300 km. Recognizing that the length of the turbulence region that must be taken into account does not exceed as a rule several dozen kilometers, it becomes clear that the condition indicated above is satisfied with a large margin. We shall assume that an average refractive index n(P) = 1 in the entire considered space, and that the fluctuations of the refractive index are themselves small:

$$n(P) = 1 + n_1(P), \quad \langle n_1(P) \rangle = 0, \quad |n_1(P)| \ll 1. \tag{1.127}$$

Our problem is to find the solution of (1.126) in a closed volume V, given the boundary conditions on the boundary σ of the volume V (Fig. 1.12). A consistent rigorous application of the Green-function formalism to the solution of this problem turns out to be impossible, since great difficulties are encountered when it comes to ensuring zero boundary conditions when the Green functions are chosen (Eq. (1.43)). We shall therefore use another solution method, but one similar to the Green-function method.

We note first of all that at $n_1(P) = 0$ the fundamental solution of Eq. (1.126) is of the form

$$v_0(P_1, P) = \frac{1}{r(P_1, P)} \exp[ikr(P_1, P)], \tag{1.128}$$

where $r(P_1, P)$ is the distance between the points P_1 and P. The basic assumption that we shall use throughout is that Eq. (1.126) has in the general case a singular solution with a pole in the form $1/r(P_1, P)$ and that this solution can be represented as

$$v(P_1, P) = v_0(P_1, P) \exp[\psi(P_1, P)], \tag{1.129}$$

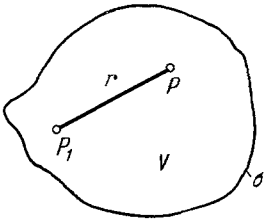


Fig. 1.12

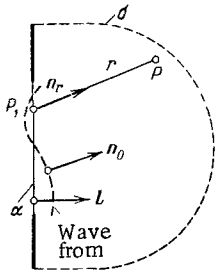


Fig. 1.13

where $\psi(P_1, P)$ is an analytic function. The basis for this statement is that in the limiting case of geometric optics a solution in the form (1.129) exists, with

$$\psi(P_1, P) = ik \int n_1(P)\, ds. \tag{1.130}$$

The curvilinear integral in (1.130) is calculated along the path of the light beam propagating from the point P_1 to the point P (or back).

When Eq. (1.126) has a solution that satisfies all the foregoing conditions, any other solution E(P) of (1.126) is expressed in terms of $v(P_1, P)$ [25]:

$$E(P) = \frac{1}{4\pi} \int_\sigma \left[-v(P_1, P) \frac{\partial E(P_1)}{\partial l} + E(P_1) \frac{\partial v(P_1, P)}{\partial l} \right] d\sigma, \tag{1.131}$$

where l is the inward normal to the surface σ, while the point P_1 is located on σ (Fig. 1.13).

We now stipulate more specifically the surface σ and assume that the field E differs from zero only within an area that is part of an infinite plane. Since the functions E and v satisfy the Sommerfield radiation condition, the integral over an infinite closed part of the surface σ is zero. In addition, we assume that the wave-front curvature radius on the surface α is much larger than the wavelength. This condition permits considerable simplification of (1.131). In fact, the normal derivative of the function E, which is contained in (1.131), is none other than the scalar product of the gradient of the function and the inward normal to the surface α (Fig. 1.13):

$$\frac{\partial E}{\partial l} = (\nabla E \cdot \boldsymbol{l}). \tag{1.132}$$

In accord with the assumption made that the wavelength is short compared with the curvature radius of the wave front on the surface α, a small-enough section of the wave front can be regarded as a segment of a plane wave propagating in a direction of the normal $\mathbf{n}_0$ to the wave front. We can thus write approximately

$$\frac{\partial E}{\partial l} \approx (ik\boldsymbol{n}_0 \cdot \boldsymbol{l})\, E. \tag{1.133}$$

As a result, the first term under the integral sign in (1.131) is written in the form

$$-v(P_1, P) \frac{\partial E(P_1)}{\partial l} \approx -E(P_1) \frac{1}{r(P_1, P)} \exp[ikr(P_1, P) + \psi(P_1, P)]\, ik\boldsymbol{n}_0\boldsymbol{l}. \tag{1.134}$$

We consider now the second term. Substituting the expression for the function $v(P_1, P)$ we obtain

$$E(P_1) \frac{\partial v(P_1, P)}{\partial l} = E(P_1) \left(\nabla \left\{ \frac{1}{r(P_1, P)} \exp\left[ikr(P_1, P) + \psi(P_1, P) \right] \right\} \cdot \boldsymbol{l} \right). \tag{1.135}$$

The gradient $\nabla\{\ldots\}$, contained in (1.135), is equal to

$$\exp[ikr(P_1, P) + \psi(P_1, P)] \left\{ \nabla \left(\frac{1}{r(P_1, P)} \right) + \frac{1}{r(P_1, P)} [-ik\boldsymbol{n}_r + \nabla\psi(P_1, P)] \right\}, \tag{1.136}$$

where $\mathbf{n}_r$ is a unit vector in the direction from the point P_1 to the point P. Recognizing that

$$\nabla \frac{1}{r(P_1, P)} = \boldsymbol{n}_r \frac{1}{r^2(P_1, P)}, \tag{1.137}$$

we obtain

$$\nabla\left\{\frac{1}{r(P_1, P)} \exp[ikr(P_1, P) + \psi(P_1, P)]\right\} = \exp[ikr(P_1, P) + \psi(P_1 P)]\left\{\frac{\boldsymbol{n}_r}{r^2(P_1, P)} - \frac{ik\boldsymbol{n}_r}{r(P_1, P)} + \frac{\nabla\psi(P_1, P)}{r(P_1, P)}\right\}. \tag{1.138}$$

Since we are considering a case when the distance $r(P_1, P)$ is larger than the wavelength by many orders, the first term on the right in the curly brackets of (1.138) can be neglected in comparison with the second. Next, recognizing that in the geometric-optics approximation the function $\psi(P_1, P)$ is determined by Eq. (1.130), we obtain

$$\nabla\psi(P_1, P) \approx ikn_1(P_1)\boldsymbol{n}_0, \tag{1.139}$$

where $n_1(P_1)$ is the fluctuating part of the refractive index. Since the absolute value of these fluctuations is small ($|n_1| = 10^{-6}$), the third term in the curly brackets in (1.138) can also be neglected compared with the second.

Taking the foregoing remarks into account, we can write Eq. (1.131) in simpler form:

$$E(P) = -\frac{ik}{4\pi}\int_\alpha E(P_1)\exp[ikr(P_1, P) + \psi(P_1, P)]\ \frac{1}{r(P_1, P)}[(\boldsymbol{n}_0 \cdot \boldsymbol{l}) + (\boldsymbol{n}_r \cdot \boldsymbol{l})]\, d\sigma. \tag{1.140}$$

Usually the paraxial approximation is sufficient when solving radar problems. The reason is that the dimensions of the target and the dimensions of the receiving aperture are much shorter than the distance that separates the target and the receiving aperture. The relative changes of the vector $\mathbf{n_r}$ on going from point to point on the surface of the target or of the aperture are quite small. The plane α can be chosen to be perpendicular to the direction to the receiving aperture. The directions of the vectors $\mathbf{n_r}$ and $\boldsymbol{l}$ are the approximately the same and we can put

$$(\boldsymbol{n}_r \cdot \boldsymbol{l}) = 1.$$

Taking this into account, Eq. (1.140) can be written in the form

$$E(P) = -\frac{ik}{4\pi}\int_\alpha E(P_1)\exp[ikr(P_1, P) + \psi(P_1, \boldsymbol{P})]\ |1 + (\boldsymbol{n} \cdot \boldsymbol{l})|\frac{1}{r(P_1, P)}\, d\sigma. \tag{1.141}$$

In this expression the vector $\mathbf{n_0}$ is a function of the coordinates of the point P_1.

We introduce one more restriction. Assume that the normal $\mathbf{n_0}$ to the wave front of the field changes little within the limits of the plane surface α. This assumption corresponds in laser radars to the case when the reflection from the target is specular. Let, according to this assumption, the normal to the wave front change so little that we can write approximately

$$(\boldsymbol{n}_0 \cdot \boldsymbol{l}) = 1. \tag{1.142}$$

Equation (1.141) takes then a still simpler form

$$E(P) = \frac{1}{i\lambda}\int_\alpha E(P_1)\exp[ikr(P_1, P) + \psi(P_1, P)]\frac{1}{r(P_1, P)}\, d\sigma. \tag{1.143}$$

This equation is in fact a generalization of Eq. (1.53) to the case of field propagation in a turbulent medium. Besides the analyticity requirement referred to above, it is necessary to impose on the function $\psi(P_1, P)$ one more condition that ensures satisfaction of the reciprocity (reversibility) principle. The reciprocity principle consists in the following. Assume that a source located at the point P_1 produces at the point P a field with complex amplitude A. Then, when placed at the point P, the same source will produce at the point P_1 a field whose complex amplitude is also A.

For the reciprocity principle to be satisfied when (1.143) is used, we must have

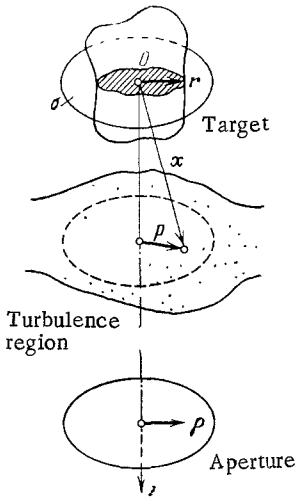


Fig. 1.14

$$\frac{1}{r(P_1, P)}\exp[ikr(P_1, P)+\psi(P_1, P)] \quad \frac{1}{r(P, P_1)}\exp[ikr(P, P_1)+\psi(P, P_1)]. \tag{1.144}$$

In the absence of turbulence, i.e., when $\psi(P_1, P) = \psi(P, P_1) = 0$, Eq. (1.144) is obviously satisfied, since the distance r between the points P_1 and P does not depend on their permutation. It turns out that in the presence of turbulence Eq. (1.144) is also satisfied, and Eq. (1.143) satisfies the reciprocity principle if, as assumed above, the function $v(P_1, P)$ (Eq. (1.129)) is a solution of the wave equation (1.126). A proof of this statement is given in Appendix II.

If we wished to used Eq. (1.143) to calculate the field reflected from the target and passing through a turbulent atmosphere, we could formally use this equation only in the case of a specularly reflecting target. The assumptions made to derive this equation are valid, as indicated, only in this case. Nonetheless, by the same reasoning as in Subsec. 1 of Sec. 1.2, we can generalize (1.143) to the case of a diffusely reflecting target. The integral is then calculated not over a plane surface, but over the target surface, and the field $E(P_1)$ is delta-correlated.

3. Effect of Atmospheric Turbulence on Spatial Coherence of the Signal Field. We consider now the correlation properties of a field passing through a layer of turbulent atmosphere. When solving this problem we follow the approach developed in [26].

Assume that the target is illuminated by laser radiation. The reflected radiation, after passing through the layer of turbulent atmosphere, produces at the receiving aperture a random field $E(\rho)$. This field is random for two reasons: the random character of the reflection from the target surface and the random fluctuations of the turbulent atmosphere. The influence of the character of the reflection from the target surface on the correlation properties of the field at the aperture was investiagated by us earlier, and we shall therefore not consider this question further.

We introduce a coordinate system that relates the target with the receiving aperture (Fig. 1.14). We choose the origin O in the target region, and direct the z axis from the origin perpendicular to the aperture plane. We specify the coordinates of the receiving-aperture points by a vector ρ that starts out from the point of intersection of the z axis with the aperture plane. The points on the target surface are specified by a vector $\mathbf{r}$ located in the σ plane (perpendicular to the z axis and passing through the origin O), and by the coordinate z of the given point. In other words, the vector $\mathbf{r}$ is the projection, on the σ plane, of the radius vector of a point of the target surface. An arbitrary point of space in the turbulence region, defined by a radius vector $\mathbf{x}$, will be characterized by the coordinate z and by a vector $\mathbf{p}$ that is a projection of the vector $\mathbf{x}$ on a plane passing through the given point and perpendicular to the z axis.

Under these conditions the field $E(\rho)$ at the receiving aperture is expressed in terms of the field $E(\mathbf{r})$ at the target surface by Eq. (1.143). In the new notation this equation takes the form

$$E(\rho)=\frac{1}{i\lambda}\int_{\sigma} E(r)\,v(\rho,\ r)\,d\sigma, \tag{1.145}$$

where the integration is on the target surface, and

$$v(\rho,\ r)=\frac{1}{R(\rho,\ r)}\exp[ikR(\rho,\ r)+\psi(\rho,\ r)]. \tag{1.146}$$

Here $R(\rho, \mathbf{r})$ is the distance between the point ρ of the receiving aperture and the point $\mathbf{r}$ of the target surface.

To obviate the need for retaining in the subsequent calculations cumbersome terms of little significance, we make the following simplifying assumption. Namely, we assume that the target is planar and its surface is perpendicular to the z axis. The integration over the target surface is automatically transformed into integration on the flat area σ.

The correlation function of the field at the receiving aperture is equal to

$$\langle E(\rho_1)E^*(\rho_2)\rangle \quad \frac{1}{\lambda^2}\int_{\sigma_1}\int_{\sigma_2}\langle E(r_1)E^*(r_2)\rangle\langle v(\rho_1,r_1)v^*(\rho_2,r_2)\rangle\,d\sigma_1 d\sigma_2. \tag{1.147}$$

It was naturally assumed in (1.147) that the field $E(\mathbf{r})$ at the target surface and the atmospheric fluctuations are statistically independent quantities. Substituting for the function $v(\rho, \mathbf{r})$ its explicit expression (1.146), we obtain

$$\langle v(\rho_1,\ r_1)v^*(\rho_2,\ r_2)\rangle=[R(\rho_1,\ r_1)R(\rho_2,\ r_2)]^{-1}\quad \exp[ikR(\rho_1,\ r_1)-ikR(\rho_2,\ r_2)]F(\rho_1,\ r_1,\ \rho_2,\ r_2), \tag{1.148}$$

where we have introduced the notation

$$F(\rho_1,\ r_1,\ \rho_2,\ r_2)=\langle\exp[\psi(\rho_1,\ r_1)+\psi^*(\rho_2,\ r_2)]\rangle. \tag{1.149}$$

Thus, to find the correlation function of the field at the receiving aperture from a specified correlation function of the field at the target surface we must know the function $F(\rho_1, \mathbf{r}_1, \rho_2, \mathbf{r}_2)$. It is easy to explain its physical meaning. Assume that in the target plane σ are located two mutually coherent point sources at the points $\mathbf{r}_1$ and $\mathbf{r}_2$. The coherence function of the combined field, considered at the points ρ_1 and ρ_2 of the aperture, is the function $F(\rho_1, \mathbf{r}_1, \rho_2, \mathbf{r}_2)$. The light field propagating through the turbulent medium passes through a large number of elementrary turbulence sections. There are therefore grounds for assuming that $\psi(\rho, \mathbf{r})$ is a Gaussian random function. In this case, taking (1.28) into account, we have [26]

$$F(\rho_1,\ r_1,\ \rho_2,\ r_2)=\exp\{-[\langle|\psi|^2\rangle-\langle\psi(\rho_1,\ r_1)\psi^*(\rho_2,\ r_2)\rangle]\}. \tag{1.150}$$

Thus, the problem of calculating the correlation function of the field at the aperture was reduced to the problem of calculating the correlator

$$\langle\psi(\rho_1,\ r_1)\psi^*(\rho_2,\ r_2)\rangle.$$

Calculating the value of this correlator at $\rho_1 = \rho_2$, $\mathbf{r}_1 = \mathbf{r}_2$, we obtain the second moment $\langle|\psi|^2\rangle$ contained in Eq. (1.150). To calculate the correlator we must find the explicit form of the function $\psi(\rho, \mathbf{r})$.

Assume that we know the field $E_T(\rho, \mathbf{r})$ produced at a point ρ of the receiving aperture by a point source placed at the point $\mathbf{r}$ of the target. The field $E_T(\rho, \mathbf{r})$ takes into account the turbulence of the atmosphere. Let, in addition, the field $E_0(\rho, \mathbf{r})$ from the same source in the absence of turbulence be known. Then, obviously, the function $\psi(\rho, \mathbf{r})$ satisfies the equation

$$\exp[\psi(\rho,\ r)]=E_T(\rho,\ r)/E_0(\rho,\ r). \tag{1.151}$$

It is natural now to expand the left- and right-hand sides of (1.151) in terms of a small parameter, such as the fluctuating part $n_1(\mathbf{x})$ of the refractive index. Equating the terms of like powers of $n_1(\mathbf{x})$ in the left- and right-hand sides of (1.151), we obtain an expression for the first approximation of the function $\psi(\rho, \mathbf{r})$:

$$\psi(\rho, r) = E_1(\rho, r)/E_0(\rho, r), \tag{1.152}$$

where the first term of the expansion is [21]

$$E_1(\rho, r) = \frac{k^2}{2\pi}\int_V E_0(x, r)\, n_1(x)\, \frac{1}{R(x, \rho)} \exp[ikR(x, \rho)]\, d^3x, \tag{1.153}$$

$$E_0(x, r) = \frac{1}{R(x, r)} \exp[ikR(x, r)]. \tag{1.154}$$

The integration in (1.153) is over the entire volume in which the radiation propagates.

With the aid of (1.152)-(1.154) we obtain

$$\langle \psi(\rho_1, r_1)\psi^*(\rho_2, r_2)\rangle = \frac{k^4}{(2\pi)^2}[E_0(\rho_1, r_1)E_0^*(\rho_2, r_2)]^{-1}\int_{V_1}\int_{V_2} B_n(x_1, x_2)\times$$
$$\times E_0(x_1, r_1)E_0^*(x_2, r_2)[R(x_1, \rho_1)R(x_2, \rho_2)]^{-1} \exp\{ik[R(x_1, \rho_1) - R(x_2, \rho_2)]\}\, d^3x_1 d^3x_2, \tag{1.155}$$

$$B_n(x_1, x_2) = \langle n_1(x_1)\, n_1(x_2)\rangle. \tag{1.156}$$

In the case when the turbulent medium is stationary and isotropic, the function B_n can be written in the form

$$B_n(x_1, x_2) = B_n(x_1 - x_2, \eta), \tag{1.157}$$

where $\eta = (z_1 + z_2)/2$, and z_1 and z_2 are the coordinates of the points $\mathbf{x}_1$ and $\mathbf{x}_2$ on the z axis. This form of the function B_n takes into account the slow variation of the properties of the atmosphere along the route.

In the calculation of $R(\mathbf{x}, \rho)$ we use the same approximation as before [Eq. (1.161)]. Namely, we replace the function $R(\mathbf{x}, \rho)$ in the denominator by the z-axis segment corresponding to the points $\mathbf{x}$ and ρ. As a result we get

$$\langle \psi(\rho_1, r_1)\psi^*(\rho_2, r_2)\rangle = \left(\frac{k^2 z}{2\pi}\right)^2 \exp\left\{-\frac{ik}{2z}[(r_1 - \rho_1)^2 - (r_2 - \rho_2)^2]\right\}\times$$
$$\times \int_{V_1}\int_{V_2} \frac{1}{z_1 z_2 (z - z_1)(z - z_2)} B_n(x_1 - x_2, \eta) \exp\left\{ik\left[\frac{(p_1 - r_1)^2}{2z_1} - \right.\right. \tag{1.158}$$
$$\left.\left. - \frac{(p_2 - r_2)^2}{2z_2} + \frac{(\rho_1 - p_1)^2}{2(z - z_1)} - \frac{(\rho_2 - p_2)^2}{2(z - z_2)}\right]\right\} d^3x_1 d^3x_2.$$

We now express the function B_n in the form of a Fourier integral of the two-dimensional spectral density of the refractive-index fluctuations:

$$B_n(x_1 - x_2, \eta) = \int F_n(\beta, \eta, \xi) \exp[-i\beta(p_1 - p_2)]\, d^2\beta, \tag{1.159}$$

where $\xi = z_1 - z_2$, and $F_n(\beta, \eta, \xi)$ is the two-dimensional (in β) spectral density of the refractive-index fluctuations. The Fourier expansion in (1.159) is with respect to two coordinates in a plane perpendicular to the z axis. The dependence of $z_1 - z_2$ remains in parametric form. Substitution of (1.159) in (1.158) yields

$$\langle \psi(\rho_1, r_1)\psi^*(\rho_2, r_2)\rangle = \left(\frac{k^2 z}{2\pi}\right)^2 \exp\left\{-\frac{ik}{2z}[(r_1 - \rho_1)^2 - (r_2 - \rho_2)^2]\right\}\int Q(\beta)\, d^2\beta, \tag{1.160}$$

$$Q(\beta) = \int_{V_1}\int_{V_2} F_n(\beta, \eta, \xi)\, \frac{1}{z_1 z_2 (z - z_1)(z - z_2)}\times \tag{1.161}$$
$$\times \exp\left\{ik\left[\frac{(p_1 - r_1)^2}{2z_1} - \frac{(p_2 - r_2)^2}{2z_2} + \frac{(\rho_1 - p_1)^2}{2(z - z_1)} - \frac{(\rho_2 - p_2)^2}{2(z - z_2)}\right] - i\beta(p_1 - p_2)\right\} d^3x_1 d^3x_2.$$

In (1.161) we can integrate directly with respect to $d^2\mathbf{p}_1$ and $d^2\mathbf{p}_2$, omitting the integration with respect to dz_1 and dz_2. These calculations are similar to those used in the integration of (1.18). As a result we get

$$Q(\beta) = \left(\frac{2\pi}{kz}\right)^2 \exp\left\{\frac{ik}{2z}\left[(r_1 - \rho_1)^2 - (r_2 - \rho_2)^2\right]\right\} \int_0^z\int_0^z F_n(\beta, \eta, \xi) \exp\left\{-i\left[\beta\left(\frac{z_1}{z}\rho_1 - \frac{z_2}{z}\rho_2\right) + \right.\right. \tag{1.162}$$

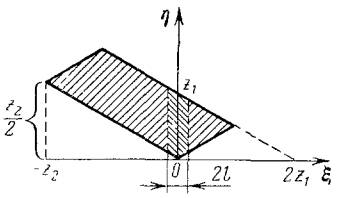


Fig. 1.15

$$+\beta\left(\frac{z-z_1}{z}\boldsymbol{r}_1-\frac{z-z_2}{z}\boldsymbol{r}_2\right)+\frac{\beta^2}{2kz}\left(z_1(z-z_1)-z_2(z-z_2)\right)\Big]\Big\}\,dz_1dz_2. \tag{1.162}$$

We transform from the variables z_1 and z_2 to the variables

$$\xi=z_1-z_2,\quad \eta=(z_1+z_2)/2. \tag{1.163}$$

This is accompanied by a corresponding change of the integration region (Fig. 1.15). The integration changes from a square of side z to a parallelogram. Since with respect to the coordinate ξ the function $F_n(\beta, \eta, \xi)$ differs from zero only in the small interval $(-l, +l)$, in which the correlation function B_n differs from zero, the integration with respect to the coordinate ξ can be extended to infinity.

Next, with respect to the coordinate β the function $F_n(\beta, \eta, \xi)$ differs from zero only within the region $|\beta| \leq l^{-1}$. This allows us to neglect the last term in the exponential under the integral sign in (1.162). In fact, we have

$$\frac{\beta^2}{2kz}[z_1(z-z_1)-z_2(z-z_2)]=\frac{\beta^2}{2k}\xi\left(1-\frac{\eta}{z}\right). \tag{1.164}$$

Taking the foregoing into account, we find that this term is of the order of $\lambda/l \ll 1$; it can therefore be neglected.

Finally, since the function F_n is even in ξ, the integration with respect to ξ can be expressed in terms of an integral only over the positive ξ axis. As a result we get

$$Q(\beta)=2\left(\frac{2\pi}{kz}\right)^2\exp\left\{\frac{ik}{2z}[(\boldsymbol{r}_1-\boldsymbol{\rho}_1)^2-(\boldsymbol{r}_2-\boldsymbol{\rho}_2)^2]\right\}\int_0^z\int_0^\infty F_n(\beta,\ \eta,\ \xi)\exp\left\{-i\left[\frac{\eta}{z}\boldsymbol{\rho}\beta+\left(1-\frac{\eta}{z}\right)\boldsymbol{r}\beta\right]\right\}d\xi d\eta, \tag{1.165}$$

where the notation $\boldsymbol{\rho}=\boldsymbol{\rho}_1-\boldsymbol{\rho}_2$, $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ was introduced.

We now call attention to the following circumstance. The three-dimensional spectral density Φ_n of the fluctuations of the refractive index can be obtained from F_n by an additional Fourier transformation with respect to the coordinate ξ. The three-dimensional vector γ of the spatial frequency is connected with the two-dimensional vector β by the relation

$$\gamma=\beta+\delta,$$

where δ is a certain constant orthogonal to β:

$$\beta\delta=0.$$

Thus,

$$\Phi_n(\gamma,\ \eta)=\frac{1}{\pi}\int_0^\infty F_n(\beta,\ \eta,\ \xi)\exp(-i\delta\xi)\,d\xi. \tag{1.166}$$

Whence it follows directly at $\delta = 0$ that

$$\Phi_n(\beta,\ \eta)=\frac{1}{\pi}\int_0^\infty F_n(\beta,\ \eta,\ \xi)\,d\xi. \tag{1.167}$$

With allowance for this equality, Eq. (1.165) is written in the form

$$Q(\beta)=2\pi\left(\frac{2\pi}{kz}\right)^2\exp\left\{\frac{ik}{2z}[(\boldsymbol{r}_1-\boldsymbol{\rho}_1)^2-(\boldsymbol{r}_2-\boldsymbol{\rho}_2)^2]\right\}\int_0^z\Phi_n(\beta,\ \eta)\exp\left\{-i\left[\frac{\eta}{z}\boldsymbol{\rho}\beta+\left(1-\frac{\eta}{z}\right)\boldsymbol{r}\beta\right]\right\}d\eta. \tag{1.168}$$

Let now the turbulence region, which is homogeneous in its properties, be bounded by a distance region (z_a, z_b). The function $\Phi_n(\beta, \eta)$ can then be represented in parametric form:

$$\Phi_n(\beta, \eta) = \Phi_{n0}(\beta) S(\eta), \tag{1.169}$$

where

$$S(\eta) = \begin{cases} 1, & z_a \leqslant \eta \leqslant z_b, \\ 0, & \eta < z_a \text{ and } \eta > z_b. \end{cases} \tag{1.170}$$

From (1.160) and (1.168) we get

$$\langle \psi(\rho_1, r_1) \psi^*(\rho_2, r_2) \rangle = 2\pi k^2 \int \Phi_{n0}(\beta) \int_0^z S(\eta) \exp\left\{-i\left[\frac{\eta}{z}\rho\beta + \left(1 - \frac{\eta}{z}\right) r\beta\right]\right\} d\eta\, d^2\beta. \tag{1.171}$$

We transform in the inner integrals to the dimensionless variable

$$t = \frac{\eta}{z}$$

and integrate with respect to the angle coordinate in the outer integral. We then obtain

$$\langle \psi(\rho_1, r_1) \psi^*(\rho_2, r_2) \rangle = 4\pi^2 k^2 z \int_0^\infty \Phi_{n0}(|\beta|) |\beta| \int_{t_a}^{t_b} J_0[|t\rho + (1-t)r||\beta|] dt\, d|\beta|, \tag{1.172}$$

where $t_\alpha = z_\alpha/z$, $t_b = z_b/z$, and J_0 is a Bessel function of zero order. Substituting (1.172) in (1.150) we get

$$F(\rho, r) = F(\rho_1, r_1, \rho_2, r_2) = \exp\left[-4\pi^2 k^2 \int_{t_a}^{t_b} \int_0^\infty \Phi_{n0}(|\beta|) |\beta| \{1 - J_0[|t\rho + (1-t) r||\beta|]\} d|\beta| dt\right]. \tag{1.173}$$

In most cases of practical importance the spectrum $\Phi_{n0}(|\beta|)$ can be approximated by the power law [21]:

$$\Phi_{n0}(|\beta|) = 0.132 C_n^2 |\beta|^{-11/3}, \tag{1.174}$$

where C_n^2 is called the structure constant of the refractive-index fluctuations. Substitution of (1.174) in (1.173) and the corresponding calculations yield [27]

$$F(\rho, r) = \exp\left[-1{,}45 C_n^2 k^2 z \int_{t_a}^{t_b} |t\rho + (1-t) r|^{5/3} dt\right]. \tag{1.175}$$

We consider a few particular cases. We find first the correlation function of the field produced by a point source located in target plane z = 0. Equations (1.147) and (1.175) yield

$$\langle E(\rho_1) E^*(\rho_1 + \rho) \rangle = B(\rho) F(\rho, 0), \tag{1.176}$$

where B(ρ) is the correlation function of a point source in the absence of turbulence and is already known to us. Calculating F(ρ, 0) directly, we obtain

$$F(\rho, 0) = \exp\left[-0.54 C_n^2 k^2 z |\rho|^{5/3} \left(t_b^{8/3} - t_a^{8/3}\right)\right]. \tag{1.177}$$

If the turbulence occupies the entire space between the target and the aperture, then $t_b = 1$, $t_\alpha = 0$, and we have

$$F(\rho, 0) = \exp\left[-0{,}54 C_n^2 k^2 z |\rho|^{5/3}\right]. \tag{1.178}$$

Such a relation is a feature of routes of short length near the ground.

Introducing, by definition, the correlation radius of the field at an aperture near the ground

$$\rho_0 = (0.54 C_n^2 k^2 z)^{-3/5}, \tag{1.179}$$

we can rewrite (1.178) in the form

$$F(\rho, 0) = \exp\left[-(|\rho|/\rho_0)^{5/3}\right]. \tag{1.180}$$

TABLE 1. Dependence of Structure Constant of Refractive-Index Fluctuations on the Height

Altitude, km	C_n^2, $m^{-2/3}$		
	weak turbulence	medium turbulence	strong turbulence
0,01	$52 \cdot 10^{-17}$	$75 \cdot 10^{-16}$	$10 \cdot 10^{-14}$
2,0	1,0	2,5	0,65
4,0	0,38	1,5	0,59
6,0	0,36	1,1	0,35
8,0	0,31	0,75	0,18

When ranging an object in outer space (say an artificial satellite), the atmosphere occupies a very small part of the route over which the radiation propagates. In that case $t_b = 1$ and $t_a \approx t_b$. By taking the limit as $z \to \infty$ we obtain from (1.177)

$$z\left(t_b^{8/3} - t_a^{8/3}\right) = z\left[1 - (1 - L/z)^{8/3}\right] \approx \frac{8}{3} L,$$

where $L = z_b - z_a$ is the length of the atmospheric section. Accordingly,

$$F(\rho, 0) = \exp\left[-1.45 C_n^2 k^2 L |\rho|^{5/3}\right]. \tag{1.181}$$

If the field correlation radius is assumed to be

$$\rho_0 = (1.45 C_n^2 k^2 L)^{-3/5}, \tag{1.182}$$

we arrive again at (1.180).

The foregoing analysis does not take into account the dependence of the structure constant of the refractive-index fluctuations on the altitude. Such a dependence actually exists and is characterized by the data listed in Table 1 [28]. Nevertheless, the results can be used for estimates.

We consider one example. Assume that the object range is an artificial satellite with orbit altitude $z = 300$ km. For the structure constant we assume the value corresponding to medium altitude at 2 km altitude, namely $C_n^2 = 2.5 \cdot 10^{-16}$ $m^{-2/3}$. We obtain the field correlation radius ρ_0 for the wavelengths $\lambda_1 = 0.69$ μm, $\lambda_2 = 1.06$ μm, and $\lambda_3 = 10.6$ μm, assuming that the atmospheric section of the route has an effective length $L = 10$ km. From (1.182) we obtain

$$\rho_{01} = 3.2 \text{ cm}, \quad \rho_{02} = 5.4 \text{ cm}, \quad \rho_{03} = 0.86 \text{ m},$$

which agrees quite well with the experimental data.

CHAPTER 2

HOLOGRAPHIC AND INTERFEROMETRIC METHODS OF PULSED-SIGNAL RECEPTION

One of the principal problems of laser radar is to obtain information on the geometric dimensions and shape of an object. The main difficulty in solving this problem is that the atmospheric turbulence distorts the target information contained in the wave front of the light wave reflected from the target. The simplest method of recognizing objects is to form their image with a telescopic system. However, the turbulence of the atmospheric layer through which the radiation reflected from the target propagates impairs the resolving power of the system.

Experiments show that when an image is formed by an ordinary telescopic system the atmospheric turbulence makes impossible a resolving power better than 1-5 sec of angle in the visible band [1, 2]. Yet the tasks of laser radar require a resolving power better than 0.5 sec of angle. It is therefore obvious that special methods must be found for processing the laser signals if the influence of the air turbulence is to be eliminated.

In this chapter we consider three basic methods used in laser radar to solve the foregoing problem: the intensity-hologram method, the modified-Michelson-interferometer method, and the Goodman method.

2.1. Intensity-Hologram Method

1. General Description of Method. The main reason why the atmosphere impairs the resolving power of optical systems is that the atmospheric turbulence distorts the wavefront of the wave reflected from the object. The problem would be solvable if there existed some method of determining the wavefront distortions due to the atmosphere from the wavefront distortions due to the object itself (more accurately, due to its not being pointlike). It would only be necessary to separate the "atmospheric" distortions and remove them, after which usual methods could be used to form the now undistorted image of the object.

Under real conditions, however, we have no a priori information whatever on the object. It is therefore impossible in principle to separate the "atmospheric" distortions of the wavefront from the "object" ones. We must consequently adopt an extreme viewpoint and assume that the "phase" information on the object is completely "spoiled" and must be discarded. We are left then only with amplitude information from which we must deduce the shape and size of the investigated object. It is precisely this discard of the phase information which is the basis of the intensity-hologram method proposed in 1968 by Kuriksha [3].

To elucidate the gist of the intensity-hologram method, we confine ourselves first to a simplified analysis. Let the remote object be illuminated by a short laser-radiation pulse. The radiation reflected from the object is recorded on a photographic plate (or film) located at the exit pupil of a telescope (Fig. 2.1a). If the illuminating pulse satisfies definite monochromaticity conditions (these conditions will be discussed in Subsec. 4) and its duration is too short for a noticeable displacement of the object (i.e., spatial coherence of the reflected field is preserved), an interference pattern will be recorded on the photographic plate. This photograph is called the intensity hologram.

We assume throughout that the photographic density of the plate emulsion is proportional to the field intensity. The transmission coefficient $T(\rho)$ of the photographic plate can then be expressed in the form

$$T(\rho) = 1 - \varepsilon |E(\rho)|^2, \tag{2.1}$$

where ε is a coefficient that relates the photographic density of the plate emulsion with the intensity of the incident light; $E(\rho)$ is the field in the recording plane.

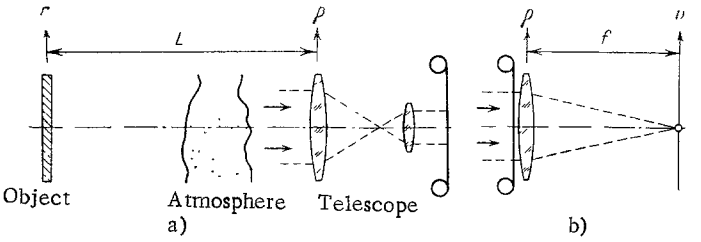


Fig. 2.1

We assume next that the target is in the Fraunhofer-diffraction zone relative to the receiving aperture. In addition, we assume for simplicity that the target is planar and parallel to the receiving aperture. In the absence of atmospheric turbulence, as shown in Chap. 1, the field $E(\rho)$ at the aperture surface is then connected with the field $E(\mathbf{r})$ on the target surface by the Fourier transformation:

$$E(\rho) = \alpha \exp\left(i \frac{k}{2L} \rho^2\right) \int_{\sigma} E(\boldsymbol{r}) \exp\left(-i \frac{k}{L} \rho \boldsymbol{r}\right) d^2 r, \tag{2.2}$$

where α is a constant coefficient of no importance in this particular case, and L is the distance between the origins O_1 and O_2.

In the presence of atmospheric turbulence the phase distribution of the field $E(\rho)$ at the aperture is changed. Since the illuminating pulse is very short, it can be stated that, within the time that it is recorded, the state of the atmosphere remained unchanged. As a result, we have

$$E(\rho) = \alpha \exp\left[i\varphi(\rho) + i \frac{k}{2L} \rho^2\right] \int_{\sigma} E(\boldsymbol{r}) \exp\left(-i \frac{k}{L} \rho \boldsymbol{r}\right) d^2 r, \tag{2.3}$$

where the function $\varphi(\rho)$ describes the phase distortions introduced by the atmosphere. Substituting (2.3) in (2.1), we obtain for the transmission coefficient of the photographic plate the expression*

$$T(\rho) = 1 - \varepsilon_1 \int E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) \exp\left[-i \frac{k}{L} \rho (\boldsymbol{r}_1 - \boldsymbol{r}_2)\right] d^2 r_1 d^2 r_2, \tag{2.4}$$

where we introduced the abbreviated notation $\varepsilon_1 = \varepsilon |\alpha|^2$.

The reduction of the recorded intensity hologram consists in the following. The developed plate is placed in a parallel-ray beam. Directly behind the photographic plate is placed a lens and the image produced in its focal plane is recorded (Fig. 2.1b). It is known [4] that the field $E(\mathbf{v})$ in the focal plane of the lens is connected with the field $E(\rho)$ in the plane of this lens by the Fourier transformation.

$$E(\boldsymbol{v}) = c \exp\left(i \frac{k}{2f} \boldsymbol{v}^2\right) \int E(\rho) \exp\left(-i \frac{k}{f} \rho \boldsymbol{v}\right) d^2 \rho, \tag{2.5}$$

where c is a certain coefficient of no importance in the analysis that follows, and f is the focal length of the lens. The field passing through the photographic plate is equal to

$$E(\rho) = AT(\rho), \tag{2.6}$$

where A is the amplitude of the incident plane wave.

Using (2.4)-(2.6), we obtain the field in the focal plane of the lens:

$$E(\boldsymbol{v}) = cA \exp\left(i \frac{k}{2f} \boldsymbol{v}^2\right) \int \exp\left(-i \frac{k}{f} \rho \boldsymbol{v}\right) \left[1 - \varepsilon_1 \int E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) \exp\left[-i \frac{k}{L} \rho (\boldsymbol{r}_1 - \boldsymbol{r}_2)\right] d^2 r_1 d^2 r_2\right] d^2 \rho. \tag{2.7}$$

*Strictly speaking, when making this substitution, account must be taken of the change of the coordinate scale, due to the fact that the telescope magnification is not equal to unity. This scale change, however, is so trivial that it will be omitted hereafter for brevity.

The first term in (2.7) can be approximately written in the form

$$cA\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int\exp\left(-i\frac{k}{f}\boldsymbol{\rho}\boldsymbol{v}\right)d^2\rho\approx cA\lambda^2f^2\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\delta(\boldsymbol{v}), \tag{2.8}$$

where $\delta(\mathbf{v})$ is the Dirac delta function. Thus, the first term in (2.7) describes the field distribution, which is a narrow maximum of the focused radiation on the lens axis in its focal plane. This part of the field $\mathbf{E}(\mathbf{v})$ contains no information on the object.

The information on the object is contained in the second term of (2.7). This term can be approximately written in the form

$$cA\varepsilon_1\lambda^2f^2\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int\delta\left[\boldsymbol{v}+\frac{f}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right]E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)\,d^2r_1d^2r_2. \tag{2.9}$$

Integrating next with respect to the coordinate $\mathbf{r}_2$, we obtain

$$cA\varepsilon_1\lambda^2f^2\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E(\boldsymbol{r})E^*\left(\boldsymbol{r}+\frac{L}{f}\boldsymbol{v}\right)d^2r. \tag{2.10}$$

Substituting (2.8) and (2.10) in (2.7), we obtain the following expression for the field in the focal plane of the lens:

$$E(\boldsymbol{v})=cA\lambda^2f^2\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\left[\delta(\boldsymbol{v})-\varepsilon_1\int E(\boldsymbol{r})E^*\left(\boldsymbol{r}+\frac{L}{f}\boldsymbol{v}\right)d^2r\right]. \tag{2.11}$$

When recording this field, the central maximum determined by the first term in the square brackets can be filtered out by a suitable screen. The area of this screen must be small enough not to obscure the basic information-carrying part determined by the second term in the square brackets of (2.11). We introduce for this term a special designation

$$E_i(\boldsymbol{v})=\varepsilon_2\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E(\boldsymbol{r})E^*\left(\boldsymbol{r}+\frac{L}{f}\boldsymbol{v}\right)d^2r, \tag{2.12}$$

where ε_2 is a constant coefficient.

The function $\mathbf{F}(\mathbf{x})$ defined by an integral operation on the function $\varphi(\boldsymbol{y})$, of the form

$$F(\boldsymbol{x})=\int\varphi(\boldsymbol{y})\varphi^*(\boldsymbol{y}+\boldsymbol{x})\,d^2y, \tag{2.13}$$

is called the autocorrelation function [5]. Thus, if the trivial phase factor preceding the integral in (2.12) is discarded, the informative part of the field in the focal plane of the lens is the autocorrelation function of the target-image field.

The photographic plate placed in the focal plane of the lens that affects the Fourier transformation of the image in the reduction of the intensity hologram responds to the field intensity. It will therefore record an emulsion photographic-density distribution proportional not to the field autocorrelation function $\mathbf{E_i(v)}$, but the square of its modulus (intensity),

$$I_i(\boldsymbol{v})=|\varepsilon_2|^2\int E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)E^*\left(\boldsymbol{r}_1+\frac{L}{f}\boldsymbol{v}\right)E\left(\boldsymbol{r}_2+\frac{L}{f}\boldsymbol{v}\right)d^2r_1d^2r_2. \tag{2.14}$$

Turning again to the expression for $\mathbf{E_i(v)}$, we see that the field formed as a result of processing the intensity hologram does not depend on the turbulent state of the atmosphere. This was accomplished, however, at the cost of losing part of the information on the object. Generally speaking, it is impossible to reconstruct unambiguously the image of the object by means of the autocorrelation function of the image field. Yet some information on the size of the object remains: the dimension of the region occupied by the autocorrelation function along any direction is equal approximately to double the dimension of the object in the same direction.

Thus, by measuring the width of the field autocorrelation function [more accurately, of the intensity $\mathbf{I_i(v)}$], we can estimate the dimensions of the object itself. Finally, it is possible to prepare beforehand an assembly of standard autocorrelation functions corresponding to objects of known shape, and compare the unknown correlation functions with the standard ones. Problems connected with the effectiveness of such comparisons are already in the realm of image-recognition theory and will not be considered here.

2. Averaging over Realizations of a Gaussian Field. We return now to Eq. (2.14). The intensity distribution $I_i(\mathbf{v})$ recorded on the photographic plate was the result of reduction of one sounding laser pulse. The distribution of the field $E(\mathbf{r})$ corresponding to one pulse can differ strongly from the field distribution corresponding to another pulse. As a result, the recorded picture of the intensity distribution $I_i(\mathbf{v})$ is different for different realizations.

Assume that the target is immobile. In this case the general relative positions of the target and aperture do not change from pulse to pulse. However, owing to the different types of the external actions on the target and on the receiving apertures, there are always small shifts of the target relative to the aperture. As a result of these microscopic shifts (of the order of the wavelength) the amplitude–phase distribution of the field $E(\mathbf{r})$ is different for different pulses even if the target is immobile.

This raises the following question: What is the average intensity distribution in the plane of the photographic plate when several intensity holograms are processed?

We consider first the case of a diffusely reflecting target. According to (2.13) we have

$$\langle I_i(\boldsymbol{v})\rangle = |\varepsilon_2|^2 \int \left\langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) E^*\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right) E\left(\boldsymbol{r}_2 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle d^2r_1 d^2r_2. \tag{2.15}$$

Consequently, to solve the problem, it is necessary to calculate a correlation function of the form

$$\langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) E(\boldsymbol{r}_3) E^*(\boldsymbol{r}_4)\rangle. \tag{2.16}$$

This can be done by assuming that the field reflected from the target surface is Gaussian. This assumption is valid only for a diffusely reflecting target. In this case the field is the result of the action of a large number of random parameters determined by the diffuse character of the reflection. By virtue of the central limit theorem, such a field can be regarded as Gaussian. It must be emphasized that a field reflected from a specular target cannot be regarded as Gaussian. Therefore, the result obtained below pertains only to the case of a diffusely reflecting target.

The following equality holds for the random Gaussian quantity $E(\mathbf{r})$ [6]:

$$\langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) E(\boldsymbol{r}_3) E^*(\boldsymbol{r}_4)\rangle = \langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2)\rangle \langle E(\boldsymbol{r}_3) E^*(\boldsymbol{r}_4)\rangle + \langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_4)\rangle \langle E^*(\boldsymbol{r}_2) E(\boldsymbol{r}_3)\rangle, \tag{2.17}$$

and is a particular case of a more general property of Gaussian fields [6]:

$$\langle E_1 \dots E_n E^*_{n+1} \dots E^*_{2n}\rangle = \sum_a \langle E_1 E^*_{n+a(1)}\rangle \dots \langle E_n E^*_{n+a(n)}\rangle, \tag{2.18}$$

where $a(j)$ denotes all the possible permutations of the numbers 1, 2, ..., n, and the summation is over all these permutations. Equation (2.18) is valid for a Gaussian quantity E with a zero mean value.

We shall use (2.17) to calculate the integral (2.15). The integrand is here

$$\left\langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) E^*\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right) E\left(\boldsymbol{r}_2 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle = \langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2)\rangle \left\langle E^*\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right) E\left(\boldsymbol{r}_2 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle + \tag{2.19}$$
$$+ \left\langle E(\boldsymbol{r}_1) E^*\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle \left\langle E^*(\boldsymbol{r}_2) E\left(\boldsymbol{r}_2 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle.$$

The field at the surface of a diffusely reflecting target is delta-correlated. We therefore obtain ultimately for the integrand

$$\left\langle E(\boldsymbol{r}_1) E^*(\boldsymbol{r}_2) E^*\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right) E\left(\boldsymbol{r}_2 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle = \langle I(\boldsymbol{r}_1)\rangle \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2) \left\langle I\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2) + \tag{2.2}$$
$$+ \langle I(\boldsymbol{r}_1)\rangle \delta\left(\frac{L}{f}\boldsymbol{v}\right) \langle I(\boldsymbol{r}_2)\rangle \delta\left(\frac{L}{f}\boldsymbol{v}\right),$$

where $I(\mathbf{r})$ is the intensity of the field at the target surface. Substituting (2.20) in (2.15), we obtain

$$\langle I(\boldsymbol{v})\rangle = |\varepsilon_2|^2 \int \langle I(\boldsymbol{r}_1)\rangle \left\langle I\left(\boldsymbol{r}_1 + \frac{L}{f}\boldsymbol{v}\right)\right\rangle d^2r_1 + |\varepsilon_2|^2 \delta\left(\frac{L}{f}\boldsymbol{v}\right)\left[\int \langle I(\boldsymbol{r}_1)\rangle d^2r_1\right]^2. \tag{2.21}$$

It can be seen that the mean value of the intensity in the focal plane of the lens that affects the Fourier transformation in the reduction of the intensity hologram consists of two terms. The first term of (2.21) carries information on the object in the form of the autocorrelation function of the average value of the intensity. The second term in (2.21) describes a narrow intensity maximum in the center of the focal plane, with an amplitude that depends on the total energy reflected from the target. This term contains no information on the shape of the target, and can be easily filtered out by a suitable screen.

Taking this into account, we are justified in writing down the informative part of the mean value of the intensity in the form

$$\langle I_{\text{в}}(\boldsymbol{v})\rangle = |\varepsilon_2|^2 \int \langle I(\boldsymbol{r})\rangle \left\langle I\left(\boldsymbol{r} + \frac{L}{f}\boldsymbol{v}\right)\right\rangle d^2r. \tag{2.22}$$

The intensity values in (2.22) are averages over many realizations. This means that the distribution $\langle I(\mathbf{r})\rangle$ in the integrand corresponds to the usual and not to the atmosphere-distorted target image obtained in white light. The spotty structure typical of images obtained in coherent light is smoothed out by the averaging. The average distribution of the informative part of the signal recorded by reducing the intensity hologram is thus the autocorrelation function of the object image obtained in white light and not distorted by the atmosphere.

We consider by way of example the autocorrelation functions typical of some simplest objects. An illustrative method of forming the autocorrelation function of images with uniform intensity distributions is to shift the object image relative to itself and calculate the area of the common section of the two shifted figures (Fig. 2.2). Figure 2.3 shows the autocorrelation functions of certain objects. The ordinates are the amplitudes of the autocorrelation functions.

It is important to note that the main fraction of the information obtained from the hologram-intensity method is information on the target size. These data are contained within the boundaries of the autocorrelation function. These boundaries, however, are smeared out since, as can be seen from Fig. 2.3, the amplitude of the autocorrelation function on the boundaries tends to zero. The last circumstance complicates greatly the analysis and the reduction of the intensity holograms.

3. Inclusion of the Fresnel Approximation in the Hologram-Intensity Method. The preceding analysis is valid when the target is in the Fraunhofer-diffraction zone relative to the aperture and is planar. The field at the target surface and the field at the aperture are connected by Fourier transformation only in this case. In most cases it is necessary, when determining the propagation of the field from the target to the aperture, to take into account in the integrand terms quadratic in the coordinates (see Chap. 1). It is necessary, in addition, to take the longitudinal dimensions of the target into account. In place of Eq. (2.2) we then have the following expression for the field at the aperture:

$$E(\boldsymbol{\rho}) = \alpha \exp\left(i\frac{k}{2L}\boldsymbol{\rho}^2\right)\int_{\sigma} E(\boldsymbol{r}) \exp\left[ik\left(\frac{\boldsymbol{r}^2}{2L} + n\boldsymbol{r} - \frac{\boldsymbol{r}\boldsymbol{\rho}}{L}\right)\right] d^2r, \tag{2.23}$$

where α is a factor inessential for the analysis that follows.

In the presence of a turbulent atmosphere we should write, in accord with (1.143),

$$E(\boldsymbol{\rho}) = \alpha \exp\left(i\frac{k}{2L}\boldsymbol{\rho}^2\right)\int_{\sigma} E(\boldsymbol{r}) \exp\left[ik\left(\frac{\boldsymbol{r}^2}{2L} + n\boldsymbol{r} - \frac{\boldsymbol{r}\boldsymbol{\rho}}{L}\right) + \psi(\boldsymbol{r}, \boldsymbol{\rho})\right] d^2r, \tag{2.24}$$

where the function $\psi(\mathbf{r}, \boldsymbol{\rho})$ takes into account the turbulence of the atmosphere. We note that, in a rigorous analysis to which we now proceed, it is necessary to take into account in the function ψ the dependences on both $\boldsymbol{\rho}$ and $\mathbf{r}$. Therefore, Eq. (2.3), in which only the dependence of the function ψ on $\boldsymbol{\rho}$ is considered, is, generally speaking, approximate. The validity of such an approximation will be discussed later.

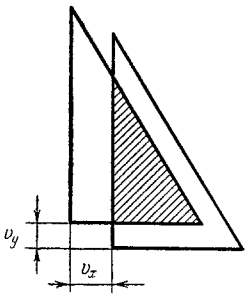


Fig. 2.2

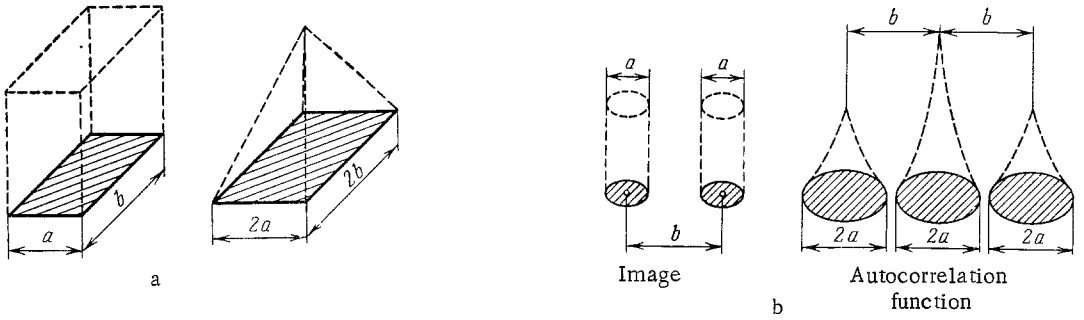


Fig. 2.3

Using (2.24) for the field at the receiving aperture, we obtain an expression for the transmission coefficient of the photographic plate that records this field:

$$T(\rho)=1-\varepsilon_1\int E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)\exp\left\{ik\left[\frac{\boldsymbol{r}_1^2-\boldsymbol{r}_2^2}{2L}+\left(\boldsymbol{n}-\frac{\rho}{L}\right)(\boldsymbol{r}_1-\boldsymbol{r}_2)\right]+\psi(\boldsymbol{r}_1,\ \rho)+\psi^*(\boldsymbol{r}_2,\ \rho)\right\}d^2r_1d^2r_2, \quad (2.25)$$

where, as before, $\varepsilon_1 = \varepsilon|\alpha|^2$. Double integration is carried out in (2.25) over the target surface. After processing the recorded intensity hologram, the field at the focal plane of the lens is equal to [in accord with (2.5) and (2.6)]

$$E(\boldsymbol{v})=cA\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int\exp\left(-i\frac{k}{f}\rho\boldsymbol{v}\right)\times\left[1-\varepsilon_1\int E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)\exp\left\{ik\left[\frac{\boldsymbol{r}_1^2-\boldsymbol{r}_2^2}{2L}+\left(\boldsymbol{n}-\frac{\rho}{L}\right)(\boldsymbol{r}_1-\boldsymbol{r}_2)\right]+\psi(\boldsymbol{r}_1,\ \rho)+\psi^*(\boldsymbol{r}_2,\ \rho)\right\}d^2r_1d^2r_2\right]d^2\rho. \quad (2.26)$$

The first integral in (2.26) is calculated in the plane of the lens that affects the Fourier transformation. This integral breaks up into two terms. The first is of the form

$$E_1(\boldsymbol{v})=cA\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int_0^{2\pi}\int_0^R\exp\left(-i\frac{k}{f}|\rho||\boldsymbol{v}|\cos\theta\right)|\rho|\,d|\rho|\,d\theta, \quad (2.27)$$

where θ is the angle between the directions of the vectors ρ and $\mathbf{v}$, and R is the radius of the lens. It is easily seen that the distribution of the field $E_1(\mathbf{v})$ is the well-known diffraction distribution in the focal plane of the lens:

$$E_1(\boldsymbol{v})=cA\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\pi R^2\cdot 2J_1\left(\frac{k}{f}|\boldsymbol{v}|R\right)\Big/\frac{k}{f}|\boldsymbol{v}|R. \quad (2.28)$$

In the limit as $R\to\infty$ this distribution goes over into the distribution (2.8).

The second term in (2.26), the informative part of the field in the focal plane of the lens, is

$$E_{\text{i}}(\boldsymbol{v})=\varepsilon_2\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int\exp\left(-i\frac{k}{f}\rho\boldsymbol{v}\right)E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)\times$$
$$\times\exp\left\{ik\left[\frac{\boldsymbol{r}_1^2-\boldsymbol{r}_2^2}{2L}+\left(\boldsymbol{n}-\frac{\rho}{L}\right)(\boldsymbol{r}_1-\boldsymbol{r}_2)\right]+\psi(\boldsymbol{r}_1,\ \rho)+\psi^*(\boldsymbol{r}_2,\ \rho)\right\}d^2r_1d^2r_2d^2\rho. \quad (2.29)$$

It can be seen that, in the general case, the informative part of the field obtained as a result of the data reduction is not equal to the autocorrelation function of the image field

$E(\mathbf{r})$. In addition, the field $E_i(\mathbf{v})$ is found to depend on the turbulent state of the atmosphere. Obviously, total compensation for the atmospheric turbulence is possible only under the condition

$$\psi(\boldsymbol{r}_1, \boldsymbol{\rho}) = -\psi^*(\boldsymbol{r}_2, \boldsymbol{\rho}). \tag{2.30}$$

The function $\psi(\mathbf{r}, \boldsymbol{\rho})$ describes both the phase and the amplitude distortions of the field at the receiving aperture. In 1 of Sec. 2.1 we used a phase approximation for the function ψ and assumed furthermore that the function φ which describes the phase fluctuations of the atmosphere depends only on the coordinate $\boldsymbol{\rho}$ on the receiving aperture:

$$\psi(\boldsymbol{r}, \boldsymbol{\rho}) \approx i\varphi(\boldsymbol{\rho}), \tag{2.31}$$

where $\varphi(\boldsymbol{\rho})$ is a real function. In this approximation the intensity-hologram method cancels out completely the distortions due to the atmospheric turbulence.

Equation (2.31) imposes simultaneously two conditions on the function $\psi(\mathbf{r}, \boldsymbol{\rho})$. The first is that account be taken of only the dependence on the coordinate $\boldsymbol{\rho}$. In the geometric-optics approximation this means that all the rays that go from different points $\mathbf{r}$ of the object surface to one point $\boldsymbol{\rho}$ of the receiving aperture pass through the same sections of the turbulent layer of the atmosphere. Obviously, the smaller the object, i.e., the closer to one another the outermost points $\mathbf{r}$ on its surface, the more accurately is the equality $\psi(\mathbf{r}, \boldsymbol{\rho}) = \psi(\boldsymbol{\rho})$ satisfied. The range of coordinates $\mathbf{r}$ within which this equality holds is called the isoplanatism region or zone. When objects in outer space are viewed from the earth's surface, the principal role is played not by the linear but by the angular dimension of the isoplanatism zone. For observation angles close to that of the zenith, this dimension is equal to several seconds of angle at a wavelength $\lambda = 0.5$ μm.

The second condition is that the real part of $\psi(\mathbf{r}, \boldsymbol{\rho})$ be discarded as insignificant. By the same token, account is taken of the phase and not of the amplitude field fluctuations due to the passage of the radiation through the turbulent atmosphere. If the atmospheric-turbulence region were an infinitely thin phase transparency placed in front of the receiving aperture, condition (2.31) would be satisfied for points $\mathbf{r}$ separated by arbitrary distances, and the isoplanatism region would be infinitely large. Equation (2.31) is therefore frequently called the phase-transparency approximation.

Let us assume that condition (2.31) is satisfied, and let us examine how the terms quadratic in $\mathbf{r}$ of the expansion in the exponential under the integral sign of (2.29) influence the result. In accord with the foregoing we have

$$E_u(\boldsymbol{v}) = \varepsilon_2 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)\times$$
$$\times \exp\left\{ik\left[\frac{\boldsymbol{r}_1^2-\boldsymbol{r}_2^2}{2L} + \boldsymbol{n}(\boldsymbol{r}_1-\boldsymbol{r}_2) - \left(\frac{\boldsymbol{v}}{f}+\frac{\boldsymbol{r}_1-\boldsymbol{r}_2}{L}\right)\boldsymbol{\rho}\right]\right\} d^2r_1 d^2r_2 d^2\rho. \tag{2.32}$$

We separate in this expression the integral with respect to the coordinate $\boldsymbol{\rho}$, calculated in the plane of the Fourier-transforming lens:

$$\int \exp\left[-ik\left(\frac{\boldsymbol{v}}{f}+\frac{\boldsymbol{r}_1-\boldsymbol{r}_2}{L}\right)\boldsymbol{\rho}\right] d^2\rho. \tag{2.33}$$

If the dimensions of the integration region are allowed to tend to infinity, the integral in (2.33) can be assumed approximately equal to

$$L^2\lambda^2\delta\left(\boldsymbol{r}_{1\rho}-\boldsymbol{r}_{2\rho}+\frac{L}{f}\boldsymbol{v}\right), \tag{2.34}$$

where $\delta\left(\boldsymbol{r}_{1\rho}-\boldsymbol{r}_{2\rho}+\frac{L}{f}\boldsymbol{v}\right)$ is the Dirac delta function and the subscript ρ of the vectors $\mathbf{r}_{1\rho}$ and $\mathbf{r}_{2\rho}$ denotes the projections of these vectors on the aperture plane. Since the integration will hereafter be carried out everywhere on the target surface visible from the receiving aperture, a one-to-one correspondence exists between the vector $\mathbf{r}$ and the vector $\mathbf{r}_\rho$. The vector $\mathbf{r}$ is then represented as the sum $\mathbf{r}_\rho + \mathbf{r}_n$, where $\mathbf{r}_n$ is the projection of the vector $\mathbf{r}$ on the vertical axis of the receiving aperture, whose direction is specified by the unit vector $\mathbf{n}$. The scalar products $\mathbf{r}\boldsymbol{\rho}$ is then equal to $\mathbf{r}_\rho\boldsymbol{\rho}$. In addition, owing to the one-to-one correspondence between the vector $\mathbf{r}$ and the vector $\mathbf{r}_\rho$, a delta function of type (2.34)

has the same "cut-out" action on the functions that depend on $\mathbf{r}_\rho$ and on the functions that depend on $\mathbf{r}$.

An approximation such as (2.34) for the integral (2.33) is valid when the minimum dimension d of the region we wish to distinguish on the target surface satisfies the relation

$$dD/L\lambda \gg 1, \tag{2.35}$$

where D is the diameter of the aperture. The physical meaning of this inequality is that the size of the diffraction resolution element on the target should be many time smaller than the minimum size of the region we wish to resolve on the target.

With allowance for these values, expression (2.32) reduces to

$$E_i(\boldsymbol{v}) = \varepsilon_3 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E(\boldsymbol{r}_1)E^*(\boldsymbol{r}_2)\exp\left[ik\left(\frac{\boldsymbol{r}_1^2-\boldsymbol{r}_2^2}{2L}+\boldsymbol{n}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right)\right]\delta\left(\boldsymbol{r}_{1\rho}-\boldsymbol{r}_{2\rho}+\frac{L}{f}\boldsymbol{v}\right)d^2r_1d^2r_2, \tag{2.36}$$

where $\varepsilon_3 = \varepsilon_3 L^2\lambda^2$. In view of the one-to-one correspondence between $\mathbf{r}$ and $\mathbf{r}_\rho$, the field E can be written as a function of the argument $\mathbf{r}_\rho$. We integrate in (2.36) with respect to the argument $\mathbf{r}_2$. This calls for putting $\mathbf{r}_{2\rho} = \mathbf{r}_{1\rho} + [(L/f)\mathbf{v}]$ everywhere in the integrand. This value of the projection on the aperture plane is possessed by the vector $\mathbf{r}_2$, whose projection on the $\mathbf{n}$ direction is equal to $\mathbf{r}_{2n}$. In this notation, we obtain

$$E_i(\boldsymbol{v}) = \varepsilon_3 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E(\boldsymbol{r}_{1\rho})E^*\left(\boldsymbol{r}_{1\rho}+\frac{L}{f}\boldsymbol{v}\right)\exp\left[ik\left(-\frac{\boldsymbol{r}_{1\rho}\boldsymbol{v}}{f}-\frac{L}{2f^2}\boldsymbol{v}^2+\frac{\boldsymbol{r}_{1n}^2-\boldsymbol{r}_{2n}^2}{2L}+\boldsymbol{r}_{1n}-\boldsymbol{r}_{2n}\right)\right]d^2r_1, \tag{2.37}$$

where $\mathbf{r}_{1n}$ is the projection of the vector $\mathbf{r}_1$ on the $\mathbf{n}$ direction. It must be emphasized that $\mathbf{r}_{2n}$, unlike $\mathbf{r}_{1n}$, is in this expression the $\mathbf{n}$-projection of a vector $\mathbf{r}_2$ such that its projection on the aperture plane is $\boldsymbol{r}_{2\rho} = \boldsymbol{r}_{1\rho} + \frac{L}{f}\boldsymbol{v}$. Thus, $\mathbf{r}_{2n}$ is a function of $\mathbf{r}_{1\rho}$.

It follows from (2.37) that allowance for the quadratic terms of the expansion in (2.29) connects the informative part of the field in the focal plane of the transforming lens with the field on the target surface by a more complicated integral transformation than the autocorrelation transformation. It must also be emphasized that the exponential $\exp\left(-i\frac{k}{f}\boldsymbol{r}_{1\rho}\boldsymbol{v}\right)$ in (2.37) is rapidly oscillating and cannot be replaced by unity. In fact, the maximum value of the argument of the exponential is of the order of

$$2\pi D_t|\boldsymbol{v}|/\lambda f, \tag{2.38}$$

where D_t is the transverse dimension of the target. From elementary geometric relations we have*

$$|\boldsymbol{v}/f| = |\boldsymbol{r}_\rho/L| \sim D_t/L. \tag{2.39}$$

As a result we find that the maximum value of the argument of the exponential is

$$2\pi D_t^2/\lambda L. \tag{2.40}$$

As a rule $D_t \approx D$. It follows, therefore, from (2.35) that

$$2\pi D_t^2/\lambda L \gg 2\pi \tag{2.41}$$

and, consequently, the exponential in (2.37) is rapidly oscillating. Thus, in each individual realization the result of the intensity-hologram reduction differs from the autocorrelation function of the field at the target surface.

Let us see to what the reduction result averaged over all the realization corresponds. From (2.37) we obtain the sought expression for the resultant intensity distribution

$$\langle I_i(\boldsymbol{v})\rangle = |\varepsilon_3|^2\int\left\langle E(\boldsymbol{r}_{1\rho})E^*(\boldsymbol{r}'_{1\rho})E^*\left(\boldsymbol{r}_{1\rho}+\frac{L}{f}\boldsymbol{v}\right)E\left(\boldsymbol{r}'_{1\rho}+\frac{L}{f}\boldsymbol{v}\right)\right\rangle \times$$
$$\times \exp\left\{ik\left[\frac{\boldsymbol{r}_{1n}^2-\boldsymbol{r}_{2n}^2-\boldsymbol{r}_{1n}'^2+\boldsymbol{r}_{2n}'^2}{2L}+\boldsymbol{r}_{1n}-\boldsymbol{r}_{2n}-\boldsymbol{r}'_{1n}+\boldsymbol{r}'_{2n}-f^{-1}(\boldsymbol{r}_{1\rho}-\boldsymbol{r}'_{1\rho})\boldsymbol{v}\right]\right\}d^2r_1d^2r'_1. \tag{2.42}$$

*We disregard here the scale change due to the magnification of the telescope (see footnote on p. 444).

We consider only the case of a diffusely reflecting target. We then obtain, on the basis of (2.20),

$$\langle I_{\mathrm{i}}(\boldsymbol{v})\rangle = |\varepsilon_3|^2 \int \langle I(\boldsymbol{r}_{1\rho})\rangle \left\langle I\left(\boldsymbol{r}_{1\rho}+\frac{L}{f}\boldsymbol{v}\right)\right\rangle \delta(\boldsymbol{r}_{1\rho}-\boldsymbol{r}'_{1\rho}) \times \tag{2.43}$$

$$\times \exp\left\{ik\left[\frac{\boldsymbol{r}_{1n}^2-\boldsymbol{r}_{2n}^2-\boldsymbol{r}_{1n}'^2+\boldsymbol{r}_{2n}'^2}{2L}+\boldsymbol{r}_{1n}-\boldsymbol{r}_{2n}-\boldsymbol{r}'_{1n}+\boldsymbol{r}'_{2n}-f^{-1}(\boldsymbol{r}_{1\rho}-\boldsymbol{r}'_{1\rho})\boldsymbol{v}\right]\right\} d^2r_1 d^2r'_1 + h\delta\left(\frac{L}{f}\boldsymbol{v}\right),$$

where h is a certain constant whose exact value is immaterial to us. The second term in this equation is a narrow intensity maximum at the center of the pattern and contains no information on the shape of the object. Of interest is only the first term, which after integration with respect to the coordinate $\mathbf{r}'_1$ is transformed ultimately into

$$\langle I_{\mathrm{i}}(\boldsymbol{v})\rangle = |\varepsilon_3|^2 \int \langle I(\boldsymbol{r}_{1\rho})\rangle \left\langle I\left(\boldsymbol{r}_{1\rho}+\frac{L}{f}\boldsymbol{v}\right)\right\rangle d^2r_1.$$

All the terms containing $\mathbf{r}_n$ cancel each other, since $\mathbf{r}_{1n} = \mathbf{r}'_{1n}$ at $\mathbf{r}_{1\rho} = \mathbf{r}'_{1\rho}$. Thus, the intensity averaged over all the realizations is the autocorrelation function of the average intensity distribution over the object.

It must be emphasized once more that this result pertains only to the case of a diffusely reflecting target. In the case of a specular target, as indicated earlier, the result of the reduction is not the intensity autocorrelation function. This fact is confirmed by experiments. We note also that the intensity hologram is not sensitive to changes of the reflected-field polarization. The last circumstance is quite obvious. In fact, if we introduce the unit vector $\mathbf{e}(\rho)$ of the field polarization at the aperture, the field takes the form

$$\boldsymbol{E}(\rho) = \boldsymbol{e}(\rho)E(\rho).$$

Calculating the field intensity, we get

$$I(\rho) = \boldsymbol{e}(\rho)\boldsymbol{e}(\rho)|E(\rho)|^2 = |E(\rho)|^2,$$

i.e., the intensity hologram does not depend on the polarization of the incoming radiation.

4. Resolving Power of the Intensity-Hologram Method. It is clear from the foregoing that the intensity-hologram method does not always, generally speaking, eliminate completely the influence of the atmospheric turbulence. In the general case, therefore, the possibility of target recognition is determined both by intrinsic (instrumental) properties of the apparatus and by the turbulence of the atmosphere.

In some cases, as shown, the intensity-hologram method cancels out completely the atmospheric turbulence. The character of the recorded pattern is then determined only by the properties of the recording apparatus itself. In contrast, however, to the usual systems that produce a simple image of the target, the intensity-hologram does not make it possible to establish a one-to-one correspondence between the object and the recorded pattern, which is the autocorrelation function of the image of this object. Therefore, the standard concept of resolving power, based on the Rayleigh criterion, cannot be applied directly in this case. Nonetheless, in practice it is necessary to assess the suitability of some particular optical system for use in the intensity-hologram method. We must therefore establish some criterion for the quality of optical systems used in the intensity-hologram method.

To this end, we turn again to Eq. (2.32), which expresses the informative part of the field at the plane where the autocorrelation function is recorded in terms of the field distribution on the target surface. An exponential factor in the form

$$\exp\left[ik\left(\frac{r^2}{2L}+\boldsymbol{n}\boldsymbol{r}\right)\right]$$

is contained in (2.32) and in the entire preceding algebra, in the form of a multiplier of the distribution of the field $E(\mathbf{r})$ on the target surface. It is therefore natural to introduce a generalized amplitude–phase distribution of the field $E_1(\mathbf{r})$ defined as

$$E_1(\boldsymbol{r}) = E(\boldsymbol{r})\exp\left[ik\left(\frac{r^2}{2L}+\boldsymbol{n}\boldsymbol{r}\right)\right]. \tag{2.44}$$

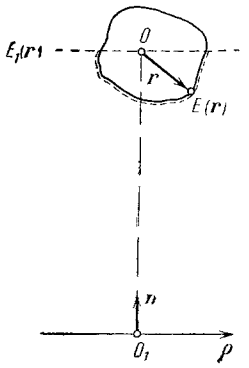


Fig. 2.4

The part

$$E(\boldsymbol{r})\exp(ik\boldsymbol{n}\boldsymbol{r})$$

of (2.44) constitutes the field reflected from the target and considered in a plane perpendicular to the axis of the receiving aperture (the **n** direction) and passing through the origin of the vector **r** (Fig. 2.4).

We set the generalized field distribution $E_1(\mathbf{r})$ in correspondence with the correlation function

$$F(\boldsymbol{x}) = \int E_1(\boldsymbol{r})\, E_1^*(\boldsymbol{r}+\boldsymbol{x})\, d^2r. \tag{2.45}$$

The integration in (2.45) is over the target surface. Since a one-to-one correspondence exists between the vector **r** and its projection $\mathbf{r}_\rho$ on the aperture plane, the vector $\mathbf{r}_\rho$ can be chosen as the argument of the function E_1. The integration in (2.45) can also be transformed to a plane parallel to the receiving aperture, replacing d^2r by Jd^2r_ρ, where J is the Jacobian of the transformation.

In the case of a specular target, its surface is parallel to the aperture plane (1 of Sec. 1.2) and the Jacobian is equal to unity. If, however, the target is diffuse, the integration over the target surface is, in accord with what was said in 1 of Sec. 1.2, in fact summation over individual independent point sources, and the function $E_1(\mathbf{r})$ is proportional to the density of placement of these sources. The number of summed points does not depend on the Jacobian that transforms the elementary integration areas. Therefore, Eq. (2.45) can be formally written in the form

$$F(\boldsymbol{x}) = \int E_1(\boldsymbol{r}_\rho)\, E_1^*(\boldsymbol{r}_\rho+\boldsymbol{x})\, d\sigma_\rho, \tag{2.46}$$

where the vector **x** lies in the same plane as the vector $\mathbf{r}_\rho$, and the symbol $d\sigma_\rho$ denotes integration over this plane.

The intensity-hologram method forms in each realization an analog of distribution (2.46); this analog is somewhat "spoiled" by the optical data-reduction system. It can thus be assumed that, in the intensity-hologram method, the primary autocorrelation function $F(\mathbf{x})$ is transformed into the recorded autocorrelation function $F_1(\mathbf{x})$. The more accurately $F_1(\mathbf{x})$ duplicates $F(\mathbf{x})$, the better the data-reduction system.

We show now that $F_1(\mathbf{x})$ is the result of a linear transformation of the function $F(\mathbf{x})$. In fact, it follows from (2.32) and (2.33) that the informative part of the recorded field is

$$E_1(\boldsymbol{v}) = \varepsilon_2 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E_1(\boldsymbol{r}_{1\rho})\, E_1^*(\boldsymbol{r}_{2\rho})\quad G\left(\boldsymbol{r}_{1\rho}-\boldsymbol{r}_{2\rho}+\frac{L}{f}\boldsymbol{v}\right)d\sigma_{\rho 1}d\sigma_{\rho 2}, \tag{2.47}$$

where

$$G\left(\boldsymbol{r}_{1\rho}-\boldsymbol{r}_{2\rho}+\frac{L}{f}\boldsymbol{v}\right) = \int_S \exp\left[-ik\left(\boldsymbol{r}_{1\rho}-\boldsymbol{r}_{2\rho}+\frac{L}{f}\boldsymbol{v}\right)\frac{\rho}{L}\right]d^2\rho. \tag{2.48}$$

The integration in (2.47) is over a surface parallel to the receiving aperture and in (2.48) over the surface of the transforming lens.

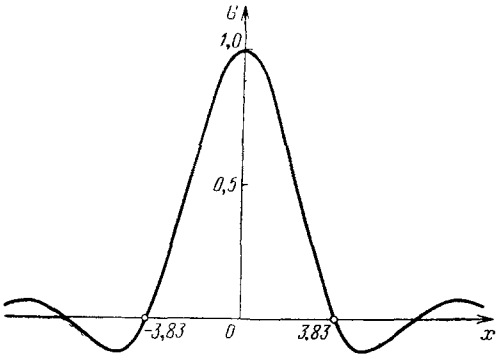


Fig. 2.5

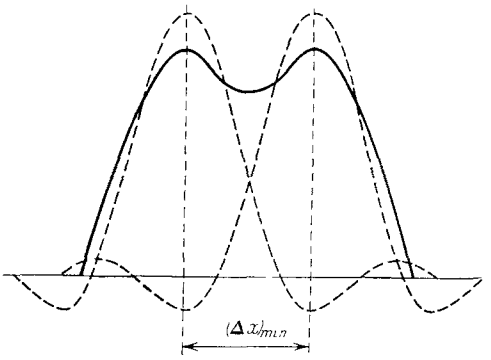


Fig. 2.6

We introduce a new system of variables

$$\boldsymbol{u} = \boldsymbol{r}_{1\rho} - \boldsymbol{r}_{2\rho} + \frac{L}{f}\boldsymbol{v}, \quad \boldsymbol{r}_{1\rho} = \boldsymbol{r}_{1\rho}. \tag{2.49}$$

Expression (2.47) can thus be written in the form

$$E_i(\boldsymbol{v}) = -\varepsilon_2 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E_1(\boldsymbol{r}_{1\rho})\, E_1^*\left(\boldsymbol{r}_{1\rho} + \frac{L}{f}\boldsymbol{v} - \boldsymbol{u}\right) G(\boldsymbol{u})\, d\sigma_{\rho 1} d\sigma_u, \tag{2.50}$$

where $d\sigma_u$ denotes integration over a new region defined by the transformation (2.49). Using (2.46), we obtain

$$E_i(\boldsymbol{v}) = -\varepsilon_2 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int F\left(\frac{L}{f}\boldsymbol{v} - \boldsymbol{u}\right) G(\boldsymbol{u})\, d\sigma_u. \tag{2.51}$$

As indicated in 2 of Sec. 2.1, for an ideal optical system with infinite resolving power we have

$$G(\boldsymbol{u}) = \lambda^2 L^2 \delta(\boldsymbol{u}). \tag{2.52}$$

Substituting this expression in (2.51) we have

$$E_i(\boldsymbol{v}) = c\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right) F\left(\frac{L}{f}\boldsymbol{v}\right), \tag{2.53}$$

where $c = \lambda^2 L^2 \varepsilon_2$.

Actually, any real optical system has a finite resolving power. The recorded field is found to be proportional to the function $F_1[(L/f)\mathbf{v}]$:

$$E_i(\boldsymbol{v}) = c\exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right) F_1\left(\frac{L}{f}\boldsymbol{v}\right), \tag{2.54}$$

where F_1 is a linear transformation of the function F:

$$F_1(\boldsymbol{x}) = \int F(\boldsymbol{x} - \boldsymbol{u})\, G(\boldsymbol{u})\, d\sigma_u. \tag{2.55}$$

Equation (2.55) is a perfect analog of a formula, known from the theory of optical systems, which connects the image field $\mathbf{E}(\mathbf{x})$ at the entry to the optical system with the image field $\mathbf{E}_1(\mathbf{y})$ at the exit from the optical system [4]:

$$E_1(\boldsymbol{y}) = \int E(\boldsymbol{y} - \boldsymbol{x})\, G(\boldsymbol{x})\, d^2x, \tag{2.56}$$

where the function $\mathbf{G}(\mathbf{x})$ is the so-called scattering function of the system.

The scattering function is usually introduced for image transmission systems. Its physical meaning is then the distribution of the field produced in the recording plane by a point source at the system entry. The width of the scattering function determines the resolution of the system for image transmission. For an aberration-free optical system, the function $\mathbf{G}(\mathbf{x})$ normalized to unity at zero is equal to

$$G(\boldsymbol{x}) = 2J_1\left(\frac{2\pi}{\lambda f}|\boldsymbol{x}|R\right)\Big/\frac{2\pi}{\lambda f}|\boldsymbol{x}|R, \tag{2.57}$$

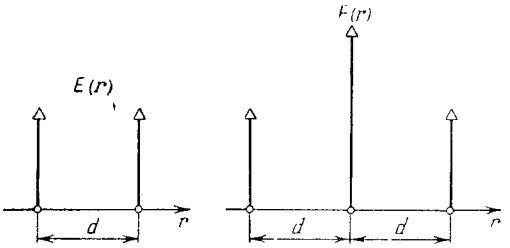


Fig. 2.7

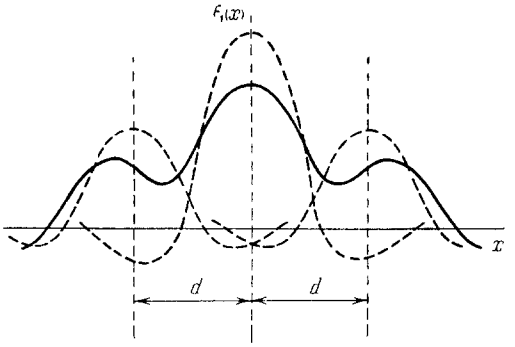


Fig. 2.8

where R is the radius of the exit pupil of the optical system, f is the focal length of the lens, and J_1 is a Bessel function of first order. The form of the function G(**x**) is shown in Fig. 2.5.

In optics, the resolving power of incoherent optical systems is determined by using the Rayleigh criterion [7]. According to this criterion, the measure of the resolving power of an ideal incoherent optical system is taken to be the distance between the origin and the first zero of the function G(**x**). In accordance with this definition, for an incoherent optical system the minimum resolvable size of an element in the target plane is equal to

$$(\Delta x)_{\min} = \frac{3.83\lambda}{2\pi R} f = 1.22 \frac{\lambda}{D} f,$$

where D is the diameter of the exit pupil.

The basis for the Rayleigh criterion is the experimental fact that the human eye is capable of distinguishing between the images of two incoherent point sources if they are separated by a distance not less than $(\Delta x)_{\min}$. In this case, the intensity in the minimum between two images is about 70% of the intensity at the maximum. When an image is formed in coherent light, this relation is satisfied at

$$(\Delta x)_{\min} = \frac{5.3\lambda}{2\pi R} f = 1.7 \frac{\lambda}{D} f. \tag{2.58}$$

The distribution of the total field, obtained in this case, is shown in Fig. 2.6.

Using the complete analogy between (2.55) and (2.56), we introduce, starting with (2.48), the scattering function for the intensity-hologram method:

$$G(\boldsymbol{u}) = \int_S \exp\left(-i \frac{k}{L} \boldsymbol{u}\boldsymbol{\rho}\right) d^2\rho. \tag{2.59}$$

It can be seen that the introduced scattering function coincides with the scattering function for an ideal optical image-formation system if we replace in the latter the focal length f of the lens by the distance L to the target.

Assume now that the data were reduced using the intensity hologram of the field produced by two point sources:

$$E(r) = \delta(r + d/2) + \delta(r - d/2), \tag{2.60}$$

where d is the distance between the sources. The undistorted autocorrelation function is (Fig. 2.7)

$$F(r) = \delta(r + d) + \delta(r - d) + 2\delta(r). \tag{2.61}$$

The recorded autocorrelation function is obtained by substituting (2.61) in (2.55):

$$F_1(x) = G(x + d) + G(x - d) + 2G(x), \tag{2.62}$$

where x = Lv/f.

If the optical system used to reduce the intensity hologram is aberration-free, we have for a round aperture, according to the definition (2.59),

$$G(x)=2J_1\left(\frac{2\pi}{\lambda L}xR\right)\Big/\frac{2\pi}{\lambda L}xR. \tag{2.63}$$

Hence, according to the Rayleigh criterion, a distinction can be made between the individual components contained in (2.62) if $d > (\Delta x)_{min}$, where the minimum resolvable size of an element in the target plane is

$$(\Delta x)_{\min}=1.7\frac{\lambda}{D}L. \tag{2.64}$$

The picture, corresponding to this case, of the distribution $F_1(x)$ in the registration plane is shown in Fig. 2.8.

The angular resolution of the intensity-hologram method is thus

$$(\Delta\theta)_{\min}=\frac{(\Delta x)_{\min}}{L}=1.7\frac{\lambda}{D}. \tag{2.65}$$

Comparison of (2.65) with (2.58) shows that the intensity-hologram method has the same angular resolving power with respect to the autocorrelation function as are possessed, relative to an object image in the absence of turbulence, by diffraction-limited optical image-formation systems.

5. Radiation Monochromaticity Requirements. One of the features of the intensity-hologram method is the need for ensuring high monochromaticity of the sounding laser radiation. The point is that the interference pattern on the surface of the receiving aperture must be maintained stationary during the entire length of the sounding pulse. This means, in turn, that the field on the entire surface of the target should be spatially coherent at each instant of time.

The field of the sounding radiation is in first-order approximation on assembly of trains within which the field is coherent in time, while the individual trains are not coherent with one another. As a result, when such a wave is incident on an object, surface regions are produced on it with fields that are not mutually coherent. These regions move over the object together with the sounding wave, causing blurring of the interference pattern on the receiving aperture. Let us consider this phenomenon in greater detail.

Let the field on the aperture be $E(\boldsymbol{\rho}, t)$. The spatial coherence of the field is determined by the interference term $E(\boldsymbol{\rho}_1, t)E^*(\boldsymbol{\rho}_2, t)$. The intensity hologram calls for conservation of the spatial coherence over the entire duration T of the sounding pulse, i.e., the function

$$S(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)=\int_0^T E(\boldsymbol{\rho}_1, t)E^*(\boldsymbol{\rho}_2, t)\,dt \tag{2.66}$$

must be close to its maximum value for all $\boldsymbol{\rho}_1$, $\boldsymbol{\rho}_2$ on the receiving aperture. In each realization, the fields $E(\boldsymbol{\rho}, t)$ are different. It is therefore necessary to require that the mean value of the function $S(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$, which is equal to

$$\langle S(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)\rangle=\int_0^T\langle E(\boldsymbol{\rho}_1, t)E^*(\boldsymbol{\rho}_2, t)\rangle\,dt, \tag{2.67}$$

be close to the maximum for all $\boldsymbol{\rho}_1$, $\boldsymbol{\rho}_2$.

We express the field $E(\boldsymbol{\rho}, t)$ of the receiving aperture in terms of the complex reflection coefficient $C(\mathbf{r})$ from the target surface. Assume that the sounding field is a plane wave propagating perpendicular to the aperture in the direction toward the target:

$$A\left(t-\frac{\mathbf{nr}}{c}\right)\exp(ik\mathbf{nr}), \tag{2.68}$$

where $\mathbf{n}$ is a unit vector in the direction of the outward normal to the aperture, and $A(t-\mathbf{nr}/c)$ is the complex amplitude of the field. Then the field reflected from the target surface is equal to

$$E(\boldsymbol{r}, t)=C(\boldsymbol{r})A\left(t-\frac{\boldsymbol{nr}}{c}\right)\exp(ik\boldsymbol{nr}). \tag{2.69}$$

Taking (2.68) and (2.69) into account, the field on the receiving aperture is

$$E(\rho, t)=\frac{1}{i\lambda L}\int A\left(t-\frac{nr}{c}-\frac{R(\rho, r)}{c}\right)C(r)\exp[ikR(\rho, r)+iknr]\,d^2r, \tag{2.70}$$

where L is the distance to the target and $R(\rho, \mathbf{r})$ is the distance between the points ρ and $\mathbf{r}$. The integration is over the surface of the target. Carrying out the usual expansion of the function $R(\rho, \mathbf{r})$ (Chap. 1), we confine ourselves to the principal term of this expansion:

$$R(\rho, r)=nr+L. \tag{2.71}$$

Substituting (2.71) and (2.70) in (2.67), we obtain

$$\langle S(\rho_1, \rho_2)\rangle=\frac{1}{\lambda^2L^2}\int_0^T\int\left\langle A\left(t-2\frac{nr_1}{c}-\frac{L}{c}\right)A^*\left(t-2\frac{nr_2}{c}-\frac{L}{c}\right)\right\rangle C(r_1)C^*(r_2)\times \tag{2.72}$$
$$\times\exp[ik(R(\rho_1, r_1)-R(\rho_2, r_2)+n(r_1-r_2))]\,d^2r_1d^2r_2dt.$$

Since we are interested only in the temporal coherence of the sounding radiation, only the function A in (2.72) need be averaged. By definition, the temporal correlation function of the sounding field is equal to

$$B(\tau)=\langle A(t)A^*(t+\tau)\rangle. \tag{2.73}$$

Expression (2.72) is then reduced to the form

$$\langle S(\rho_1, \rho_2)\rangle=\frac{T}{\lambda^2L^2}\int B\left(2\frac{|n(r_1-r_2)|}{c}\right)C(r_1)C^*(r_2)\exp[ik(R(\rho_1, r_1)-R(\rho_2, r_2)+n(r_1-r_2))]\,d^2r_1d^2r_2. \tag{2.74}$$

The distribution of the recorded intensity hologram is obtained from (2.74) by putting $\rho_1=\rho_2$. In fact, when the intensity hologram is formed, the photographic density of the emulsion is proportional (within the limits of the linearity region) to the total radiation energy absorbed at a given point ρ of the photorecorder, i.e., proportional to the quantity

$$\int_0^T|E(\rho, t)|^2\,dt=S(\rho, \rho).$$

The mean value of this quantity is expressed, obviously, by Eq. (2.74) at $\rho_1=\rho_2$.

If the coherence of the sounding radiation is so small that the function B in (2.74) changes much more rapidly than the exponential in the same formula, the correlation function B can be approximately represented in the form

$$B\left(2\frac{|n(r_1-r_2)|}{c}\right)=\alpha\langle|A|^2\rangle\delta(r_1-r_2), \tag{2.75}$$

where α is a normalization coefficient. We carry out the normalization in such a way that the mean value of the power incident on the receiving aperture corresponds to uniform scattering into a solid angle of 2π sr. Such a normalization corresponds to the considered case of a diffusely reflecting target.

Substituting (2.75) in (2.74) we obtain the following expression for the average power $P(\rho)$ at the point ρ of the receiving aperture:

$$\frac{\langle S(\rho, \rho)\rangle}{T}=P(\rho)=\frac{1}{\lambda^2L^2}\alpha\langle|A|^2\rangle\int|C(r)|^2\,d^2r. \tag{2.76}$$

Since this expression is independent of ρ, the total power incident on the entire aperture is obtained from (2.76) by simple multiplication by the area S_a of the receiving aperture. On the other hand, calculating the average power that propagates after reflection from the target into the solid angle subtended by the receiving aperture, we find that the total power incident on the aperture is equal to

$$P=\frac{S_a}{2\pi L^2}\langle|A|^2\rangle\int|C(r)|^2\,d^2r. \tag{2.77}$$

Comparing (2.77) and (2.76), we arrive at the conclusion that the normalization coefficient is

$$\alpha = \lambda^2/2\pi. \tag{2.78}$$

Returning to (2.74), we find that, under condition (2.75), the distribution of the recorded intensity hologram is of the form

$$\langle S(\rho, \rho)\rangle = \frac{T}{2\pi L^2}\langle |A|^2\rangle \int |C(r)|^2 d^2r \tag{2.79}$$

and does not depend on ρ. Consequently, its contrast is zero in the considered case, and the intensity hologram itself contains no information whatever on the shape of the object.

In order for the contrast of the recorded hologram to be close to the maximum, it is necessary to satisfy the condition

$$\left\{2\frac{|n(r_1 - r_2)|}{c}\right\}_{\max} \ll T_c, \tag{2.80}$$

where T_c is the coherence time of the sounded radiation. As an estimate of the left-hand side of (2.80) we can assume

$$\left\{2\frac{|n(r_1 - r_2)|}{c}\right\}_{\max} \approx 2\frac{D_t}{c}, \tag{2.81}$$

where D_t is the longitudinal dimension of the target along the vector **n**. Consequently, the condition imposed on the sounding-radiation coherence time is of the form

$$T_c \gg 2\frac{D_t}{c}. \tag{2.82}$$

The spectral width $\Delta\lambda$ of the sounding radiation is therefore obtained from the formula

$$\Delta\lambda = \frac{\lambda^2}{cT_c}.$$

Let us cite some estimates. Let λ = 0.69 μm and D_t = 1 m. We then obtain from (2.82)

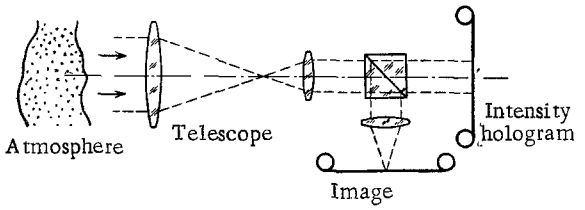
$$\Delta\lambda \ll \frac{\lambda^2}{2D_t} = 2 \cdot 10^{-4} \text{ nm}.$$

The formulated requirement on the coherence of the sounding radiation is very stringent. The physical explanation of such stringent requirements is that, in the intensity-hologram method, one has interference of waves from different points of the target surface. It is therefore necessary that the coherence length of the sounding radiation be much greater than the longitudinal dimensions of the target. The development of lasers with sufficient output pulse energy and with approximate spectrum width 10^{-4} nm is a complicated problem. It is precisely the stringent requirements on the coherence of the laser radiation which limit the use of the intensity-hologram method, despite the relatively simple apparatus needed for its realization.

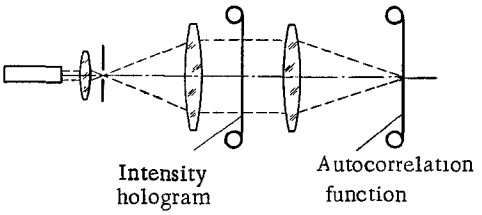
6. Experimental Results. An experimental setup for the intensity-hologram method is illustrated in Fig. 2.9. Figure 2.9a shows only the receiving part of the installation. A pulsed ruby laser (λ = 0.69 μm) operates in the free lasing regime. An optical system guides the radiation to the object located about 1 km away. The reflected radiation is received with a telescope. A special optical system transforms the focused radiation at the output of the telescope into a parallel beam and directs it to the recording photographic film. Thus, the image of the input pupil of the receiving telescope is changed in scale and is transferred to the photographic film. Part of the parallel beam is diverted to an objective that forms in its focal plane the object image recorded on the film.

The recorded hologram is processed in accordance with the scheme shown in Fig. 2.9b. The reconstruction is by a laser operating at 0.63-μm wavelength in the fundamental transverse mode. A diaphragm placed in the focal plane of the first lens affects additional mode selection. The photographic film with the recorded hologram is placed in the parallel beam ahead of a second focusing objective. The autocorrelation function is recorded in the focal plane of the objective.

Figure 2.10 shows successive films of originals of different objects (a), their images (b) distorted by turbulent atmosphere, the intensity holograms (c), and the autocorrelation functions reconstructed from them (d).



a



b

Fig. 2.9

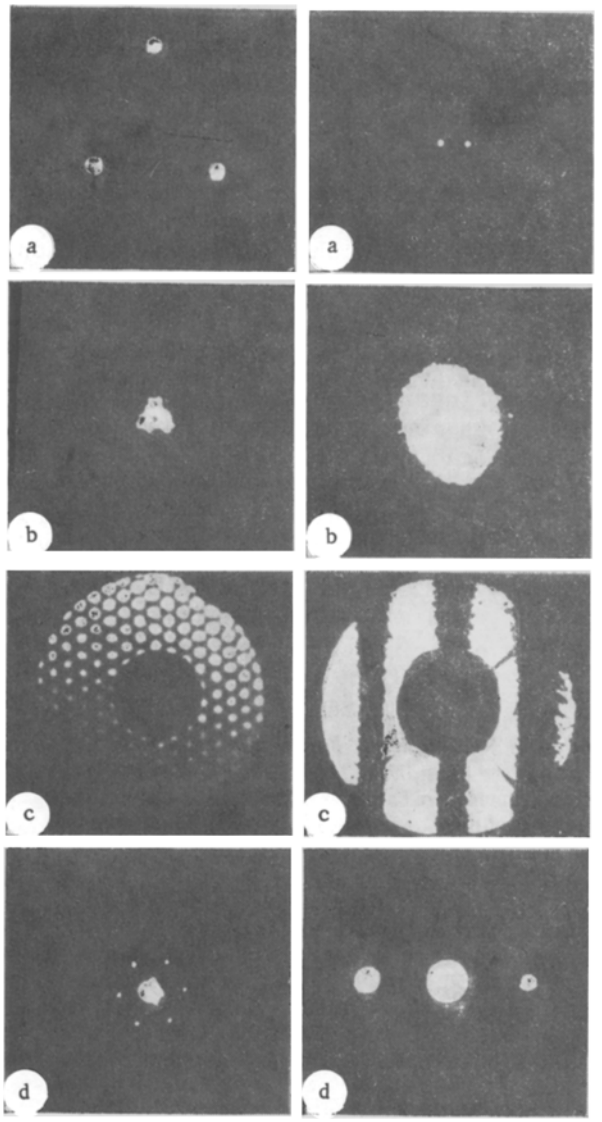


Fig. 2.10

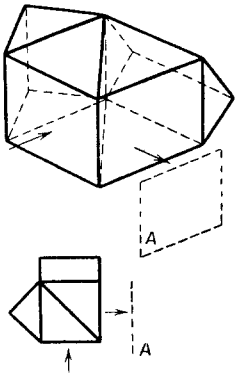


Fig. 2.11

The first object comprises three pointlike sources. The angular dimension of the object is ≈6 sec of angle. In the direct image channel this object can hardly be resolved. The images of the individual points are strongly blurred by the amplitude and phase fluctuations of the atmosphere, and also by the aberrations of the receiving optical system. The intensity hologram of such an object has a regular structure not distorted by the atmosphere. Against the background of the ordered structure of the intensity hologram are visible the amplitude fluctuations. Six points at the vertices of a regular hexagon are located in the reconstructed autocorrelation function around the zeroth order. An estimate of the quality of the obtained autocorrelation function has shown that the resolving power of the method was, in this case, ≈1 sec of angle. The diffractive resolving power of the employed telescope in coherent light at a wavelength 0.69 μm was 0.51 sec of angle.

The second object comprises two point sources. The angle distance between the two points is 0.3 sec of angle. In the direct-image channel this object cannot be resolved at all, but the obtained autocorrelation function made it possible to resolve distinctly two points symmetric about the center. In this experiment, the diffractive resolving power of the receiving telescope in coherent light of 0.69-μm wavelength was 0.15 sec of angle.

2.2. Method of Modified Michelson Interferometer

1. General Description of Method. The main feature of the intensity-hologram method considered in the preceding section is the need for using lasers having a very high degree of radiation monochromaticity. The modified-Michelson-interferometer method, hereafter called for brevity simply the interferometric method, relaxes substantially the requirements imposed on the laser-emission coherence [8].

The interferometric method is based on a modification of the Michelson interferometer (Fig. 2.11) [9, 10]. Two cube-prisms mounted in the interferometer arms are skewed 90° relative to each other. In the recording plane A the interferring fields are those reflected from each cube-prism. If a field E(x, y) enters the interferometer, the field interferring in the recording plane A are $E_1 = E(-x, y)$ and $E_2 = E(x, -y)$. By making the change of variable $x_1 = -x$, we find that $E_1 = E(\rho)$, $E_2 = E(-\rho)$, where $\rho = (x_1, y)$. Thus, one of the two interferring fields is equal to the other, but is turned through 180°.

It will be shown below that for a practical realization of the interferometric method it is necessary to tilt the cube prisms in the interferometer arms. If the edge of one of the cube prisms is deflected from the vertical axis by an angle α_1, and the edge of the other is deflected from the horizontal by an angle α_2, the fields interferring in the recording plane A are

$$E_1 = E(\rho)\exp(ik\alpha_1 y), \quad E_2 = E(-\rho)\exp(ik\alpha_2 x).$$

As a result, the interference term is equal to

$$E_1 E_2^* = E(\rho)E^*(-\rho)\exp(ik\boldsymbol{\alpha}\rho), \tag{2.83}$$

where $\boldsymbol{\alpha} = (-\alpha_2, \alpha_1)$, $k = 2\pi/\lambda$ is the wave number. By the same token, an additional spatial frequency $k\boldsymbol{\alpha}$ is introduced into the interference term.

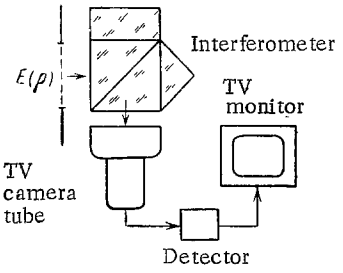


Fig. 2.12

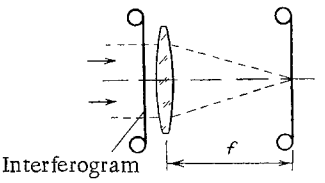


Fig. 2.13

The fields interferring in the modified Michelson interferometer are produced by each individual target-surface point in the two interferometer arms. To produce the interference pattern, it is not at all necessary to have, during the recording interval, total coherence of the fields produced by the different points of the target surface. On the contrary, as will be made clear in the following, the interferometric method calls for absence of coherence during the recording interval between the individual radiating points on the target surface.

At the same time, the sounding radiation must not be excessively broadband, since mutual coherence of the fields in the interferometer arms is essential. Namely, the coherence length of the sounding radiation should be larger than the path difference of the rays in the interferometer arms. In practice, it is not particularly difficult to achieve a ray-path difference not larger than 1 mm. At a wavelength $\lambda = 0.69$ μm such a coherence length corresponds to a spectrum width $\Delta\lambda = 0.48$ nm. It is therefore usually required that the width of the sounding-radiation spectrum be equal to several millimeters, so that the field on the target surface can be regarded as spatially incoherent during the recording interval.

We consider now the gist of the interferometric method (Fig. 2.12). The field reflected from the target passes through a layer of turbulent atmosphere and is incident on the plane of the receiving aperture, thus producing a certain field distribution $\mathbf{E}(\boldsymbol{\rho})$. This field passes further through the modified Michelson interferometer and produces an interference pattern in the recording plane. The recorder is a TV camera tube. The video signal and the output of this tube is proportional to the field intensity in the recording plane.

When working in the pulsed regime, the TV tube memorizes the pulsed intensity distribution during one frame. The video signal picked off the TV tube is square-law-detected. The detector output signal is proportional to the square of the field intensity in the recording plane. The detected signal is extracted in the form of a picture on the screen of a video monitor. The image on this screen is the photographed and Fourier-transformed as described in the treatment of the intensity hologram in Sec. 2.1 (Fig. 2.13).

We proceed now to a detailed description of the interferometric method. When account is taken of the turbulence of the atmosphere, the field at the receiving aperture is expressed in the form [Eq. (2.24)]

$$E(\boldsymbol{\rho}) = a\exp\left(ik\frac{\rho^2}{2L}\right)\int_{\sigma} E(\boldsymbol{r})\exp\left[ik\left(\frac{r^2}{2L}+\boldsymbol{nr}-\frac{\boldsymbol{r\rho}}{L}\right)+\psi(\boldsymbol{r},\ \boldsymbol{\rho})\right]d^2r, \tag{2.84}$$

where a is a constant factor of no significance in what follows, $\mathbf{E}(\mathbf{r})$ is the field on the target surface, $\mathbf{n}$ is a unit vector normal to the aperture, ψ is a function that takes into account the turbulence of the atmosphere, and L is the distance to the target.

Since the sounding radiation is not coherent during the recording interval, Eq. (2.84) must be supplemented by the time dependence of the field. We then obtain an expression similar to (2.70):

$$E(\boldsymbol{\rho},\ t) = a\exp\left(ik\frac{\rho^2}{2L}\right)\int_{\sigma} A\left(t-2\frac{\boldsymbol{nr}}{c}-\frac{L}{c}\right)C(\boldsymbol{r})\ \exp\left[ik\left(\frac{r^2}{2L}+\boldsymbol{nr}-\frac{\boldsymbol{r\rho}}{L}\right)+\psi(\boldsymbol{r},\ \boldsymbol{\rho})\right]d^2r. \tag{2.85}$$

Here A(t) is a random function of the time and describes the time dependence of the sounding field, $\mathbf{C}(\mathbf{r})$ is the complex reflection coefficient of the target reflection; the integration is over the target surface.

The field intensity at the interferometer output is

$$I(\rho,\ t)=|E(\rho,\ t)+E(-\rho,\ t)|^2, \tag{2.86}$$

and the TV transmitting tube responds to the total field energy received during the recording time:

$$I_{\text{т}}(\rho)=\int_0^T I(\rho,\ t)\,dt, \tag{2.87}$$

where T is the recording interval. Substituting (2.86) in (2.87), we obtain

$$I_{\text{т}}(\rho)=\int_0^T |E(\rho,\ t)|^2\,dt+\int_0^T |E(-\rho,\ t)|^2\,dt\ \int_0^T E(\rho,\ t)\,E^*(-\rho,\ t)\,dt+\int_0^T E^*(\rho,\ t)\,E(-\rho,\ t)\,dt. \tag{2.88}$$

We introduce the skew of the interferometer cube-prisms. We then obtain from (2.83)

$$I_{\text{т}}(\rho)=\int_0^T |E(\rho,\ t)|^2\,dt+\int_0^T |E(-\rho,\ t)|^2\,dt+\exp(ik\alpha\rho)\int_0^T E(\rho,\ t)\,E^*(-\rho,\ t)\,dt+ \tag{2.89}$$

$$+\exp(-ik\alpha\rho)\int_0^T E^*(\rho,\ t)\,E(-\rho,\ t)\,dt.$$

The integration with respect to time in (2.89) means, in fact, averaging over time. Substituting (2.85) in (2.89), we obtain, apart from an inessential constant, the following expression for the first term in (2.89):

$$I_{\text{т1}}(\rho)=\int\left\langle A\left(t-2\frac{n\boldsymbol{r}_1}{c}-\frac{L}{c}\right)A^*\left(t-2\frac{n\boldsymbol{r}_2}{c}-\frac{L}{c}\right)\right\rangle C(\boldsymbol{r}_1)\,C^*(\boldsymbol{r}_2)\exp\left[ik\left(\frac{1}{2L}(\boldsymbol{r}_1^2-\boldsymbol{r}_2^2)+\boldsymbol{n}(\boldsymbol{r}_1-\boldsymbol{r}_2)-\right.\right. \tag{2.90}$$

$$\left.\left.-\frac{1}{L}\rho(\boldsymbol{r}_1-\boldsymbol{r}_2)\right)+\psi(\boldsymbol{r}_1,\ \rho)+\psi^*(\boldsymbol{r}_2,\ \rho)\right]d^2r_1d^2r_2.$$

The angle brackets in this expression mean time averaging.

We now make two important assumptions. First, that the atmosphere can be represented in the form of a phase transparency. We then have, according to (2.31),

$$\psi(\boldsymbol{r},\ \rho)\approx i\varphi(\rho). \tag{2.91}$$

The second assumption is that the time dependence of the field can be regarded as delta-function correlated. The validity of this assumption is based on the fact that the coherence length of the sounding field is very small and does not exceed several millimeters at $\Delta\lambda = 0.3$ nm. Taking this into account, we obtain

$$I_{\text{т1}}(\rho)=\langle|A|^2\rangle\int|C(\boldsymbol{r})|^2\,d^2r. \tag{2.92}$$

An analogous expression is obtained also for the second term in (2.89):

$$I_{\text{т2}}(\rho)=\langle|A|^2\rangle\int|C(\boldsymbol{r})|^2\,d^2r. \tag{2.93}$$

The first two terms describe uniform illumination on the TV camera tube. For the sum of the third and fourth terms of (2.89) we obtain

$$I_{\text{т3}}(\rho)+I_{\text{т4}}(\rho)=2\int\langle|A|^2\rangle|C(\boldsymbol{r})|^2\ \cos\left[-2\frac{k}{L}\boldsymbol{r}\rho+\varphi(\rho)-\varphi(-\rho)+k\alpha\rho\right]d^2r. \tag{2.94}$$

The uniform illumination described by (2.92) and (2.93) is not sensed at the output of the TV camera tube. As a result, the video signal at the input of the square-law detector is describable by Eq. (2.94). The product $\langle|A|^2\rangle|C(\mathbf{r})|^2$ in this expression can be interpreted as the average intensity of the reflected field $I(\mathbf{r})$.

The intensity of the image on the screen of the video monitor is proportional to the square of expression (2.94):

$$I_{\mathrm{v}}(\rho)=4\int I(r_1)\,I(r_2)\,\cos\left[-2\frac{k}{L}r_1\rho+\varphi(\rho)-\varphi(-\rho)+k\alpha\rho\right]\times$$
$$\times\cos\left[-2\frac{k}{L}r_2\rho+\varphi(\rho)-\varphi(-\rho)+k\alpha\rho\right]d^2r_1d^2r_2. \tag{2.95}$$

An intensity distribution of the form (2.95) is called an interferogram. The interferogram is photographed on a film whose transmission coefficient $T(\rho)$ is determined by Eq. (2.1):

$$T(\rho)=1-\varepsilon I_{\mathrm{v}}(\rho). \tag{2.96}$$

The subsequent reduction of the interferogram consists of placing the photographic film on a lens, transilluminating it with coherent laser emission, and recording the intensity at the focal plane of the lens. The unity term in (2.96) gives an intensity distribution in the focal plane of the lens in the form of a delta function that contains no information on the shape of the object.

We now consider the informative part, determined by the second term of (2.96), of the field in the focal plane of the lens. The field $E(\mathbf{v})$ in the focal plane of the lens is expressed in terms of the field $E(\rho)$ in the plane of the lens by Eq. (2.5):

$$E(v)=a\int E(\rho)\exp\left(i\frac{k}{2f}v^2-i\frac{k}{f}\rho v\right)d^2\rho. \tag{2.97}$$

The informative part of this field is obviously equal to

$$E_{\mathrm{i}}(v)=a_1\int I_{\mathrm{v}}(\rho)\exp\left(i\frac{k}{2f}v^2-i\frac{k}{f}\rho v\right)d^2\rho, \tag{2.98}$$

where we have introduced the notation $a_1 = -a\varepsilon$.

We represent the intensity $I_{\mathrm{v}}(\rho)$ as a sum of two terms:

$$I_{\mathrm{v}}(\rho)=2\int I(r_1)\,I(r_2)\cos\left[-2\frac{k}{L}\rho(r_1-r_2)\right]d^2r_1d^2r_2+$$
$$+2\int I(r_1)\,I(r_2)\cos\left[-2\frac{k}{L}\rho(r_1+r_2)+2(\varphi(\rho)-\varphi(-\rho))+2k\alpha\rho\right]d^2r_1d^2r_2. \tag{2.99}$$

It can be seen that the first term does not depend on the fluctuations of the atmosphere. The information concerning the object after separation of this term is not distorted by the atmosphere. It is found that this term is separated automatically in the focal plane of the lens because the second term of (2.99) contains an additional spatial frequency $2k\alpha$. As a result, the second term is shifted in the focal plane of the lens relative to the first by a distance $\Delta v = 2f\alpha$.

It now becomes understandable why a skew must be introduced in the arms of the interferometer: this enables us to separate that part of the field which is independent of the atmosphere's fluctuations. This part, in accord with (2.98) and (2.99), is equal to

$$E_{\mathrm{i}}(v)=2a_1\exp\left(i\frac{k}{2f}v^2\right)\int I(r_1)\,I(r_2)\,\cos\left[-2\frac{k}{L}\rho(r_1-r_2)\right]\exp\left(-i\frac{k}{f}\rho v\right)d^2r_1d^2r_2d^2\rho. \tag{2.100}$$

In this expression the integration with respect to $d^2\rho$ is in the plane of the transforming lens.

If the resolving power of the lens is high enough, we can write approximately

$$\int\cos\left[-2\frac{k}{L}\rho(r_1-r_2)\right]\exp\left(-i\frac{k}{f}\rho v\right)d^2\rho=\frac{1}{2}\delta(v/\lambda f+2(r_1-r_2)/\lambda L)+\frac{1}{2}\delta(v/\lambda f-2(r_1-r_2)/\lambda L). \tag{2.101}$$

Substituting the last expression in (2.100), we get

$$E_{\mathrm{i}}(v)=a_2\exp\left(i\frac{k}{2f}v^2\right)\left[\int I(r)\,I\left(r+\frac{L}{2f}v\right)d^2r+\int I(r)\,I\left(r-\frac{L}{2f}v\right)d^2r\right], \tag{2.102}$$

where a_2 is a new constant. The two terms in the right-hand side of this expression are equal, as can be easily verified by making, e.g., the change of variable $\mathbf{r}' = \mathbf{r} - \mathbf{L}\mathbf{v}/2\mathbf{f}$ in the second term. We ultimately have

$$E_{\mathrm{i}}(v)=2a_2\exp\left(i\frac{k}{2f}v^2\right)\int I(r)\,I\left(r+\frac{L}{2f}v\right)d^2r \tag{2.103}$$

Thus, the modulus of the informative part of the field produced by the interferometric method is proportional to the autocorrelation function of the field intensity at the target. The recorded intensity in the focal plane of the transforming lens is proportional to the square of the autocorrelation function of the field intensity at the target:

$$I_{\text{i}}(\boldsymbol{v}) = 4|a_2|^2\left[\int I(\boldsymbol{r})\, I\left(\boldsymbol{r} + \frac{L}{2f}\boldsymbol{v}\right) d^2r\right]^2. \tag{2.104}$$

Since the autocorrelation function of the intensity is real and positive, the contours of the recorded picture are determined by the first power of this function. The characteristic form of the autocorrelation function was considered earlier for a number of objects in the investigation of the intensity-hologram method.

2. Influence of Polarization of Received Field. Worthy of a separate consideration is the role played in the interferometric method by the field polarization. For the intensity-hologram method, the field polarization is immaterial, as is quite obvious. In the interferometric method the interferring fields come from different sections of the aperture, so that the existing difference between the field polarizations at different points of the aperture can influence the interferogram.

Actually, however, this is not the case. In fact, of principal importance for the interferograms are the last two terms of (2.89):

$$I_{\text{т3}}(\boldsymbol{\rho}) + I_{\text{т4}}(\boldsymbol{\rho}) = \exp(ik\boldsymbol{\alpha}\boldsymbol{\rho}) \int_0^T E(\boldsymbol{\rho},\ t)\, E^*(-\boldsymbol{\rho},\ t)\, dt + \exp(-ik\boldsymbol{\alpha}\boldsymbol{\rho}) \int_0^T E^*(\boldsymbol{\rho},\ t)\, E(-\boldsymbol{\rho},\ t)\, dt. \tag{2.105}$$

We introduce the vector $\mathbf{e}(\mathbf{r})$ of the field polarization on the target surface. The vector field $\mathbf{E}(\boldsymbol{\rho}, t)$ on the aperture surface is then [Eq. (2.85)]

$$\boldsymbol{E}(\boldsymbol{\rho},\ t) = a \exp\left(i\frac{k}{2L}\boldsymbol{\rho}^2\right) \int_\sigma A\left(t - 2\frac{\boldsymbol{nr}}{c} - \frac{L}{c}\right) \boldsymbol{e}(\boldsymbol{r})\, C(\boldsymbol{r}) \exp\left[ik\left(\frac{\boldsymbol{r}^2}{2L} + \boldsymbol{nr} - \frac{\boldsymbol{r}\boldsymbol{\rho}}{L}\right) + \psi(\boldsymbol{r},\ \boldsymbol{\rho})\right] d^2r. \tag{2.106}$$

We write down (2.105) with allowance for the vector character of the field $\mathbf{E}(\boldsymbol{\rho}, t)$. We then obtain, for the first term of this formula,

$$I_{\text{т3}}(\boldsymbol{\rho}) = a \exp(ik\boldsymbol{\alpha}\boldsymbol{\rho}) \int_\sigma \int_\sigma \int_0^T A\left(t - 2\frac{\boldsymbol{nr}_1}{c} - \frac{L}{c}\right) A^*\left(t - 2\frac{\boldsymbol{nr}_2}{c} - \frac{L}{c}\right) dt\, \boldsymbol{e}(\boldsymbol{r}_1)\, \boldsymbol{e}(\boldsymbol{r}_2)\, C(\boldsymbol{r}_1)\, C^*(\boldsymbol{r}_2) \times \tag{2.107}$$

$$\times \exp\left\{ik\left[\frac{1}{2L}(\boldsymbol{r}_1^2 - \boldsymbol{r}_2^2) + \boldsymbol{n}(\boldsymbol{r}_1 - \boldsymbol{r}_2) - \frac{1}{L}(\boldsymbol{r}_1 + \boldsymbol{r}_2)\boldsymbol{\rho}\right] + \psi(\boldsymbol{r}_1,\ \boldsymbol{\rho}) + \psi^*(\boldsymbol{r}_2,\ -\boldsymbol{\rho})\right\} d^2r_1 d^2r_2.$$

Integration with respect to time is equivalent to averaging over the recording time T. If A(t) is a delta-correlated random function, then

$$\int_0^T A\left(t - 2\frac{\boldsymbol{nr}_1}{c} - \frac{L}{c}\right) A^*\left(t - 2\frac{\boldsymbol{nr}_2}{c} - \frac{L}{c}\right) dt = a_2 \langle |A|^2 \rangle \delta\left(2\frac{\boldsymbol{n}(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{c}\right), \tag{2.108}$$

where a_2 is an inessential constant that depends on the assumed normalization. The argument of the delta function has the dimension of time, and the delta function (2.108) in expression (2.107) effects "cut-out" with respect to the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$. Transforming to the argument $\mathbf{r}_1 - \mathbf{r}_2$ in the delta function in (2.108), we get

$$a_2\delta\left(2\frac{\boldsymbol{n}(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{c}\right) = a_3\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2), \tag{2.109}$$

where a_3 is a new constant.

Assuming, as before, that function A(t) is delta-correlated, and using the phase-transparency approximation for the function $\psi(\mathbf{r}, \boldsymbol{\rho})$, we get, apart from an inessential constant,

$$I_{\text{т3}}(\boldsymbol{\rho}) = \exp(ik\boldsymbol{\alpha}\boldsymbol{\rho}) \int_\sigma \int_\sigma \langle |A|^2 \rangle \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, \boldsymbol{e}(\boldsymbol{r}_1)\, \boldsymbol{e}(\boldsymbol{r}_2) \times \tag{2.110}$$

$$\times C(\boldsymbol{r}_1)\, C^*(\boldsymbol{r}_2) \exp\left\{ik\left[\frac{1}{2L}(\boldsymbol{r}_1^2 - \boldsymbol{r}_2^2) + \boldsymbol{n}(\boldsymbol{r}_1 - \boldsymbol{r}_2) - \frac{1}{L}(\boldsymbol{r}_1 + \boldsymbol{r}_2)\boldsymbol{\rho}\right] + i\varphi(\boldsymbol{\rho}) - i\varphi(-\boldsymbol{\rho})\right\} d^2r_1 d^2r_2.$$

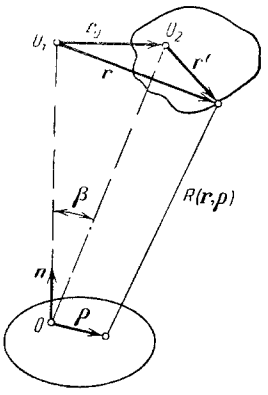


Fig. 2.14

In the integration with respect to d^2r_2 we are left in the integrand with a scalar product $\mathbf{e}(\mathbf{r}_1)\mathbf{e}(\mathbf{r}_1)$ equal to unity. As a result, we get

$$I_{r3}(\rho) = \langle |A|^2 \rangle \exp(ik\alpha\rho) \int |C(r)|^2 \exp\left[-i\frac{2k}{L} r\rho + i\varphi(\rho) - i\varphi(-\rho)\right] d^2r. \qquad (2.111)$$

Carrying out similar transformations for the second term in (2.105) and adding the result, we obtain (2.94).

3. Analysis of the Effect of the Tracking Error. A distinguishing feature of the interferometric method, compared with the intensity-hologram method, is the clearly pronounced dependence of the recorded interferogram on the angle β between the normal $\mathbf{n}$ to the receiving aperture and the target-sighting line OO_2 (Fig. 2.14).

As a rule, laser radar systems are designed to keep the target at the center of field of view of the receiving telescope. In this case, the vector $\mathbf{n}$ is directed toward the target and the angle β is zero. In practice, however, owing to all the possible perturbing actions as well as because of the target motion, it is impossible to keep the target at the center of the field of view at all times. In this case, the angle β, which we call the tracking error, differs from zero.

The main result of the tracking errors is that the field at the receiving aperture is additionally phase-modulated in space. This spatial phase modulation is due to the inclination, by an angle β, of the wavefront of the wave reflected from the target. In the intensity-hologram method the quantity recorded is the field intensity at the aperture. Therefore, the tracking errors, which cause only phase modulation of the field at the aperture, are of no importance in the intensity-hologram method.

The situation is different in the interferometric method. Interference takes place here between values of the field at different points of the aperture. Therefore, all sorts of phase distortions influence the complete form of the recorded interferogram. Further reduction can filter out the phase distortions that find their way into the interferogram. We have already shown, however, that this calls for artificially producing skews in the interferometer arms and for introducing spatial amplitude modulation into the interferogram. Without this modulation we cannot filter out the phase-distorted interferogram components. The tracking errors are therefore important in the interferometric method. Let us examine this question in greater detail.

To obviate the need of performing all the calculations for the case $\beta \neq 0$, we use the formal possibility of taking the tracking error into account directly in the final formula (2.99). We note for this purpose that, in the derivation of (2.99), no restrictions whatever were imposed on the relative scales of the vectors ρ and $\mathbf{r}$. All we assumed was satisfaction of the condition

$$|r|/L, \quad |\rho|/L \ll 1, \qquad (2.112)$$

with the vector $\mathbf{r}$ originating on the point O_1 on the receiving-aperture axis (Fig. 2.14).

In accord with the notation of Fig. 2.14, we introduce the vector $\mathbf{r}_0$ defined as

$$r_0 = r - r', \tag{2.113}$$

where **r'** is the radius vector of a point on the target vector and is reckoned from a point O_2 in the target region. At a sufficiently small angle **β**, we then have

$$\beta \approx r_0/L \tag{2.114}$$

or

$$r = \beta L + r'. \tag{2.115}$$

The term βL can be regarded as constant during the recording time. We can thus transform in Eq. (2.99), which describes the intensity on the screen of the video monitor, from the argument **r** to the argument **r'**. All the changes reduce then only to redefinition of **r** as **r'** and introduction of an additional term proportional to **β**L in the argument of the cosine.

To simplify the notation we shall omit the prime symbol of the vector **r'**. We then obtain the following expression for the intensity distribution on the screen of the video monitor:

$$I_v(\rho) = 2\int I(r_1) I(r_2) \cos\left[-2\frac{k}{L}\rho(r_1 - r_2)\right] d^2r_1 d^2r_2 + \\ + 2\int I(r_1) I(r_2) \cos\left[-2\frac{k}{L}\rho(r_1 + r_2) - 2k\beta\rho + 2k\alpha\rho + 2(\varphi(\rho) - \varphi(-\rho))\right] d^2r_1 d^2r_2. \tag{2.116}$$

Analyzing this expression, we conclude that the tracking error **β** is equivalent to introducing an addition to the spatial frequency of that interferogram term which must be filtered out.

Let us examine the consequences of the uncontrollable spatial-frequency change that takes place in the interferogram as a result of random changes of the tracking error.

First, the summary spatial frequency **α** − **β** may turn out to be equal or close to zero. In this case it may turn out to be impossible to isolate the first (undistorted by the atmosphere) term of the interference pattern from the second (which depends on the turbulence of the atmosphere). Since **β** is a random quantity, it is impossible to compensate for the change of **β** by a proper skew in the interferometer arms (i.e., by choosing **α**).

Second, it may turn out that the modulus of the spatial frequency $|\alpha - \beta|$ is so large that the interferogram modulation frequency is smaller than the resolution element of the recording systems. The interferogram modulation is then smoothed out, and the effectiveness of separating the two terms in (2.116) is decreased. It is therefore necessary to ensure small tracking errors in the interferometric method.

4. Resolving Power of the Modified-Michelson-Interferometer Method. We consider now the question of the resolving power of the interferometric method. We shall use the same approach as for the determination of the resolving power of the hologram-intensity method. Let us recall the gist of that approach. We introduce formally the undistorted intensity autocorrelation function on the target surface

$$F(x) = \int_\sigma I(r) I(r + x)\, d^2r. \tag{2.117}$$

We find next for the autocorrelation function the Green's function that converts formally, via a linear integral transformation, the formally introduced undistorted autocorrelation function into the true autocorrelation function obtainable in this method. The resolving power is then determined by the obtained Green's function. The specific formula for the angular resolving power is determined by the Rayleigh criterion.

To find the Green's function for the autocorrelation function we turn to Eq. (2.100). This equation describes the informative part, of interest to us, of the field in the focal plane of the transforming lens:

$$E_i(v) = 2a_1 \exp\left(i\frac{k}{2f}v^2\right)\int I(r_1) I(r_2) \cos\left[-2\frac{k}{L}\rho(r_1 - r_2)\right]\exp\left(-i\frac{k}{f}\rho v\right) d^2r_1 d^2r_2 d^2\rho. \tag{2.118}$$

The limits of integration with respect to the coordinate ρ are actually set in this equation by the dimensions of the receiving aperture on which the interferogram was recorded.

We integrate now with respect to the coordinate $\boldsymbol{\rho}$. For the case of a round aperture we have

$$\int \cos\left[-2\frac{k}{L}\boldsymbol{\rho}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right]\exp\left(i\frac{k}{f}\boldsymbol{\rho}\boldsymbol{v}\right)d^2\rho=\frac{1}{2}S\cdot 2J_1\left\{\left|\frac{k}{f}\boldsymbol{v}+2\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right|R\right\}\times$$
$$\times\left\{\left|\frac{k}{f}\boldsymbol{v}+2\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right|R\right\}^{-1}+\frac{1}{2}S\cdot 2J_1\left\{\left|\frac{k}{f}\boldsymbol{v}-2\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right|R\right\}\left\{\left|\frac{k}{f}\boldsymbol{v}-2\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\right|R\right\}^{-1}. \tag{2.119}$$

Here J_1 is a Bessel function of first order, R is the radius of the aperture, and $S = \pi R^2$ is the area of the aperture.

Expression (2.119) contains the well-known Airy function:

$$G(\boldsymbol{x})=2J_1(|\boldsymbol{x}|)/|\boldsymbol{x}|. \tag{2.120}$$

We substitute (2.119) in (2.118). Taking the notation (2.120) into account, we have

$$E_i(\boldsymbol{v})=a_1S\int I(\boldsymbol{r}_1)I(\boldsymbol{r}_2)G\left[2\frac{kR}{L}\left(\frac{L}{2f}\boldsymbol{v}+\boldsymbol{r}_1-\boldsymbol{r}_2\right)\right]d^2r_1d^2r_2$$
$$+a_1S\int I(\boldsymbol{r}_1)I(\boldsymbol{r}_2)G\left[2\frac{kR}{L}\left(\frac{L}{2f}\boldsymbol{v}-\boldsymbol{r}_1+\boldsymbol{r}_2\right)\right]d^2r_1d^2r_2. \tag{2.121}$$

We note that the first and second terms of (2.121) are equal. To check on this it suffices to replace $\mathbf{r}_1$ by $\mathbf{r}_2$, and vice versa. The result is

$$E_i(\boldsymbol{v})=2a_1S\int I(\boldsymbol{r}_1)I(\boldsymbol{r}_2)G\left[2\frac{kR}{L}\left(\frac{L}{2f}\boldsymbol{v}+\boldsymbol{r}_1-\boldsymbol{r}_2\right)\right]d^2r_1d^2r_2. \tag{2.122}$$

We introduce a new system of variables

$$\boldsymbol{r}=\boldsymbol{r}_1,\quad \boldsymbol{u}=\boldsymbol{r}_2-\boldsymbol{r}_1. \tag{2.123}$$

Substituting (2.123) in (2.122), we obtain

$$E_i(\boldsymbol{v})=2a_1S\int\int I(\boldsymbol{r})I(\boldsymbol{r}+\boldsymbol{u})\,d^2rG\left[2\frac{kR}{L}\left(\frac{L}{2f}\boldsymbol{v}-\boldsymbol{u}\right)\right]d^2u. \tag{2.124}$$

The inner integral in this equation is equal to the previously introduced undistorted autocorrelation function (2.117). We can thus write

$$E_i(\boldsymbol{v})=2a_1S\int F(\boldsymbol{u})G\left[2\frac{kR}{L}\left(\frac{L}{2f}\boldsymbol{v}-\boldsymbol{u}\right)\right]d^2u. \tag{2.125}$$

This expression shows that the informative part of the field is a linear integral transformation of the undistorted autocorrelation function, and in this transformation $G(\mathbf{x})$ assumes the role of the Green's function.

Before we proceed to the question of the resolving power of the interferometric method, we call attention to one substantial circumstance. Namely, in Eq. (2.125), the role of the distance to the target is played by the quantity $L' = L/2$. In fact, if f is the focal length of the converting lens, the coordinate $\mathbf{v}$ in its focal plane is transformed in the plane $\mathbf{u}$ of the undistorted autocorrelation function in accord with the equation (Fig. 2.15)

$$\boldsymbol{u}_1=\frac{L}{2f}\boldsymbol{v}. \tag{2.126}$$

Consequently, treatment of the interferogram alters the scale of the obtained autocorrelation function; this is equivalent to decreasing the distance to the object plane by one-half. This effect is due to the square-law detection of the video signal picked off the TV transmitting tube.

Using the "equivalent" distance L', we rewrite (2.125) in the form

$$E_i(\boldsymbol{u}_1)=2a_1S\int F(\boldsymbol{u})G\left[\frac{kR}{L'}(\boldsymbol{u}_1-\boldsymbol{u})\right]d^2u, \tag{2.127}$$

where, in accord with (2.126), $\mathbf{u}_1$ is the value of the coordinate $\mathbf{v}$ recalculated into the plane $\mathbf{u}$ of the undistorted autocorrelation function.

We shall use the Rayleigh criterion to determine the resolving power. The basis for this approach in the case of coherent optical systems was cited earlier in the analysis of the intensity-hologram method. From the criterion established in item 3 of Sec. 2.1, the

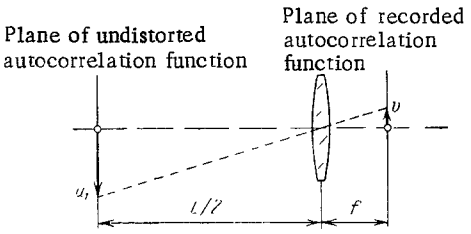


Fig. 2.15

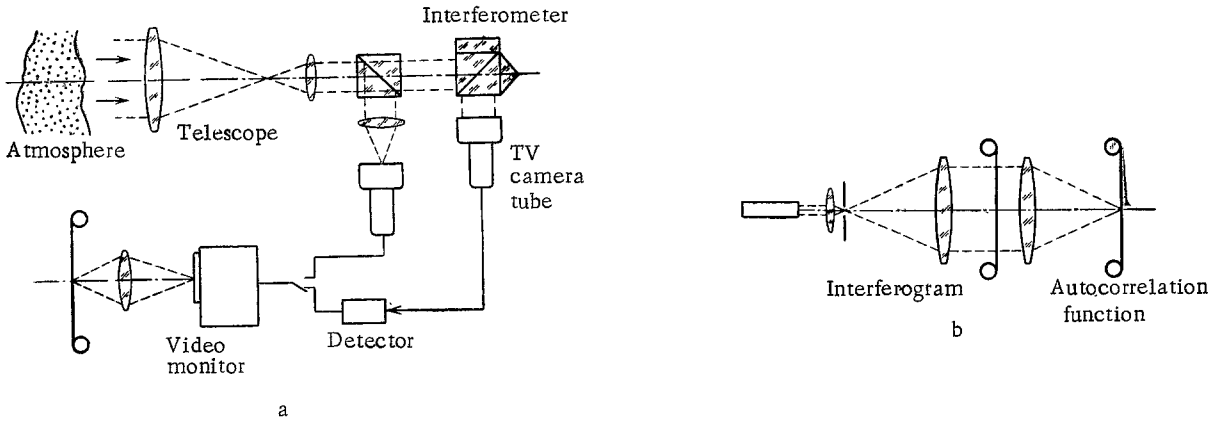


Fig. 2.16

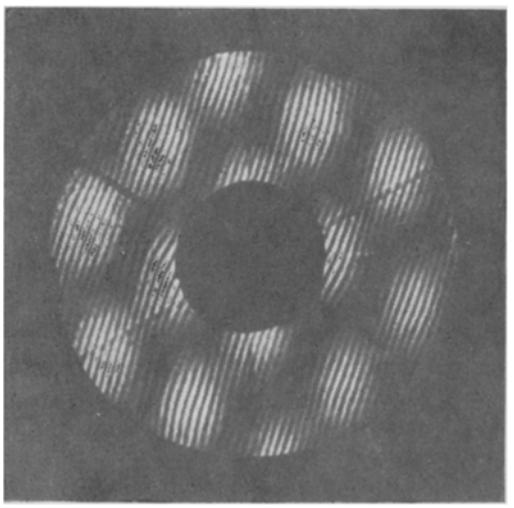


Fig. 2.17

linear dimension of the smallest object resolvable in the plane of the undistorted autocorrelation function is equal to

$$(\Delta u)_{\min} = 1.7 \frac{\lambda}{D} L', \tag{2.128}$$

where D is the diameter of the receiving aperture. From this we easily obtain the angular resolution of the interferometric method:

$$(\Delta\theta)_{\min} = \frac{(\Delta u)_{\min}}{L'} = 1.7 \frac{\lambda}{D}. \tag{2.129}$$

Thus, the limiting resolution angle of the interferometric method is equal to the angular resolving power of a coherent diffraction-limited optical system for image formation in the absence of a turbulent atmosphere. Consequently, the resolving power of the interferometric method and the resolving power of the intensity-hologram are identical.

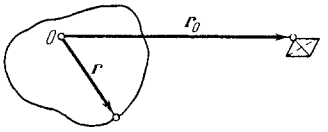


Fig. 2.18

5. Experimental Results. We consider some results of the use of the interferometric method in laser radars. Figure 2.16a shows schematically one of the possible practical realizations of this method. A laser transmitter illuminates the target through a collimating optical system. The reflected radiation passes through the layer of turbulent atmosphere, is gathered by a receiving telescope, and is incident on the interferometer. The interferogram at the exit from the interferometer is recorded with a TV transmitting tube whose output video signal is square-law-detected and fed to the screen of the video monitor.

The laser transmitter can operate in either the cw or the pulsed regime. In the latter, special synchronization is necessary to ensure that the interferogram is read from the tube during one frame. Otherwise, the screen might be photographed at instants at which the interferogram had not yet been read, or was already read and its image no longer on the tube.

The reduction procedure for the interferograms is practically the same as for the intensity holograms (Fig. 2.16b).

Provision was made in the experiment for obtaining the direct image of the object. To this end, part of the received signal was diverted from the output of the receiving telescope to a TV camera tube of the image channel. By switching the outputs of the TV camera tubes, it was possible to display on the screen either the image of the target distorted by the atmosphere or the interferogram of the reflected field. This made it possible to compare the resolving powers of the image and interferogram channels.

The investigations have shown that the interferometric method, just as the intensity-hologram method, permits a substantial increase of the resolving power of a laser radar system under turbulent-atmosphere conditions. Figure 2.17 shows as an example a photograph of an interferogram recorded from an object in the form of three points located at the vertices of an equilateral triangle. The resolving power of the receiving telescope at 0.69-μm wavelength in coherent light was 0.5 sec of angle.

2.3. Goodman's Method

1. General Description of Method. The methods considered in the preceding sections, notwithstanding their high resolving power, do not yield an object image undistorted by the atmosphere. Both in the intensity-hologram method and in the interferometric method, the object must be discerned by using the autocorrelation function of the field (or of the intensity) on the surface of the object. There is, however, no one-to-one correspondence between the autocorrelation function and the object image.

Goodman (see [11]) proposed a method that makes it possible to obtain, in principle, a target image undistorted by the atmosphere. This method was named Goodman's method. It consists of placing in the immediate vicinity of the object a pointlike reflector and recording on the receiving aperture a hologram of the radiation reflected from the target. The reference is the radiation reflected from the pointlike reflector. Since the reference radiation from the pointlike reflector and the radiation reflected from the target pass through the same region of the atmosphere, the distortions introduced by the atmosphere cancel one another in the recorded hologram.

Goodman's method can be regarded as a particular case of the hologram-intensity method. In fact, the target, together with the reflector placed alongside, can be treated as a single composite target (Fig. 2.18). In this case, the field of the composite target can be represented as

$$E_s(\boldsymbol{r}) = E(\boldsymbol{r}) + \delta(\boldsymbol{r} - \boldsymbol{r}_0), \tag{2.130}$$

where $E(\mathbf{r})$ is the field at the target itself, and $\mathbf{r}_0$ is the coordinate of the reflector. On the receiving aperture is recorded a hologram of the intensity of the field reflected from the entire composite target. Assume that the atmosphere can be represented by a phase trans-

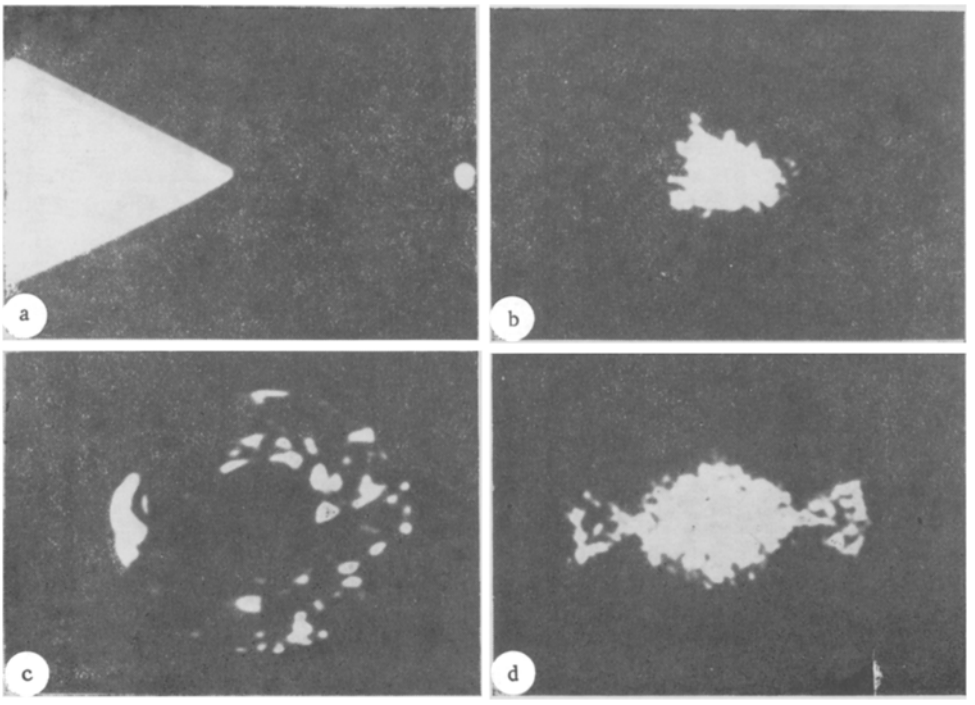


Fig. 2.19

parency. Reducing the intensity hologram, we obtain for the field in the focal plane of the converting lens [see Eq. (2.12)]

$$E_i(\boldsymbol{v}) = \varepsilon_2 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\int E_c(\boldsymbol{r})\,E_c^*\left(\boldsymbol{r}+\frac{L}{f}\boldsymbol{v}\right)d^2r. \tag{2.131}$$

Substituting (2.130) in (2.131), we obtain

$$E_i(\boldsymbol{v}) = \varepsilon_2 \exp\left(i\frac{k}{2f}\boldsymbol{v}^2\right)\left[\int E(\boldsymbol{r})\,E^*\left(\boldsymbol{r}+\frac{L}{f}\boldsymbol{v}\right)d^2r + \delta\left(\frac{L}{f}\boldsymbol{v}\right) + E\left(\boldsymbol{r}_0+\frac{L}{f}\boldsymbol{v}\right) + E\left(\boldsymbol{r}_0-\frac{L}{f}\boldsymbol{v}\right)\right]. \tag{2.132}$$

The first two terms in this expression describe the autocorrelation function of the field on the target surface and a narrow deltalike maximum located at the center of the obtained picture.

The basic information is contained in the third and fourth terms of (2.132). They constitute two distributions proportional to the field distribution on the target surface, symmetrical about the center of the field pattern obtained, and at a distance from the center

$$\Delta\boldsymbol{v} = \frac{f}{L}\boldsymbol{r}_0.$$

If the modulus of the vector $\mathbf{r}_0$ is larger than double the target size, the third and fourth terms of (2.132) are separated from the central autocorrelation function and do not overlap it. Since the quantity recorded is not the field $\mathbf{E}_i(\mathbf{v})$ but its intensity $\mathbf{I}_i(\mathbf{v})$, we get on the photographic film two symmetrically placed images of the target $\mathbf{I}(\mathbf{r})$:

$$I_i(\boldsymbol{v}) = |\varepsilon_2|^2\left[I\left(\boldsymbol{r}_0+\frac{L}{f}\boldsymbol{v}\right) + I\left(\boldsymbol{r}_0-\frac{L}{f}\boldsymbol{v}\right)\right] + F\left(\frac{L}{f}\boldsymbol{v}\right), \tag{2.133}$$

where $F[(L/f)\mathbf{v}]$ is the intensity of the autocorrelation function together with the central deltalike maximum. It can be seen that the two symmetric images of the target $I[\mathbf{r}_0 + (L/f)\mathbf{v}]$ and $I[\mathbf{r}_0 - (L/f)\mathbf{v}]$ do not depend on the turbulence of the atmosphere.

We shall not investigate Goodman's method further from the viewpoint of the resolving power, the required coherence of sounding radiation, and similar questions. All the results obtained for the intensity hologram remain in force also for Goodman's method. We note only one circumstance.

To prevent overlap of the target image and the autocorrelation function, it is desirable to choose as large a vector $\mathbf{r}_0$ as possible. On the other hand, the larger the modulus of the vector $\mathbf{r}_0$, the larger the differences between the conditions for the propagation through the atmosphere for the wave from the target and for the reference wave from the reflector. The region of target-space points for which these conditions are identical is called the isoplanatism region or zone. The vector $\mathbf{r}_0$ must thus lie in the isoplanatism zone around the target. In this case, the phase-transparency approximation can be used to take the atmosphere into account.

2. Experimental Results. Interest attaches to certain experimental results of Goodman's method. The experimental setup is similar to that used in the intensity-hologram method. Alongside the object, whose angular dimension is 4 sec of angle, is placed a pointlike mirror reflector. Figure 2.19 shows photographs of the composite object (a), its direct image obtained by photography through the turbulent atmosphere (b), the recorded hologram (c), and the result of using Goodman's method (d). Two images of the object are distinctly seen and are symmetric about a central spot, whereas on the photograph obtained by the direct method the object is not resolved. The resolving power of the employed telescope in coherent light was 0.5 sec of angle at 0.69-μm wavelength.

Thus, Goodman's method actually improves the resolving power when an image of an object is produced under turbulent-atmosphere conditions. Among all the experimental investigations performed to date by Goodman's method, notice should be taken of experiments under natural conditions [12, 13], which demonstrated the feasibility of obtaining undistorted images of cosmic objects with the spaceship "Gemini" as the example.

CHAPTER 3

METHODS OF AMPLIFYING AND TRANSFORMING LASER SIGNALS

The development of laser radar has led to the development of a number of promising methods of predetector processing of radar signals. The gist of these methods is that, to increase the effectiveness of further recording, the received radar signal is processed before it reaches the photodetector. We consider in this chapter two such methods: predetector amplification of the signal with the aid of a laser amplifier, and parametric conversion of the signal frequency in a nonlinear crystal. Both methods are intended principally for use in infrared laser radars (lidars) at long wavelengths, where the efficiency of receivers for direct detection is insufficient for radar applications.

3.1. Laser Amplifiers

1. Equations for the Density Matrix. We consider the principles of the amplification of optical signals on the basis of the quasiclassical model first proposed by Lamb. In this model the field is described by Maxwell's classical equations, while the active medium, which consists of a set of active atoms, is treated quantum-mechanically. The interaction of the field with the atoms is taken into account by perturbation theory. Only two working levels of the active atoms, a and b, are taken into consideration (Fig. 3.1). The upper and lower working levels are described by the wave functions ψ_a and ψ_b, respectively.

As a result of the interaction of the atom with the external field, the wave function of the atom becomes equal to the superposition of the wave functions ψ_a and ψ_b:

$$\psi = a(t)\psi_a + b(t)\psi_b, \tag{3.1}$$

where a(t) and b(t) are time-dependent coefficients, while the atom acquires a dipole moment

$$\Pi = \boldsymbol{p}(a^*b + ab^*). \tag{3.2}$$

Here **p** is the matrix element of the dipole-moment operator between the states with wave functions ψ_a and ψ_b. Expression (3.2) is obtained by direct calculation of average of the dipole-moment operator. Since the dipole moment is certainly different from zero in this case (these are the only transitions considered), the wave functions ψ_a and ψ_b are wave functions with different parity [2].

The terms a^*b and ab* in (3.2) are elements of the density matrix ρ(t, {i}), where {i} denotes the entire aggregate of parameters that characterize the given atom, viz., coordinates, velocity, transition frequency ω, etc. Besides their position in space, the atoms can differ in the rate of translational motion (for a gaseous active medium) and in the transition frequency (for atoms in a liquid or in a solid). The result can be the so-called inhomogeneous broadening of the active laser line.

In the model considered, the atom, together with the field acting on it, form a closed system with a definite wave function. The density matrix for a system having a definite wave function ψ is written in the form (see Chap. 2 of [2]):

$$\rho = \psi \cdot \psi.$$

Taking the foregoing into account, the density matrix of the system is given by

$$\rho(t, \{i\}) = \begin{pmatrix} \rho_{aa}(t, \{i\}) & \rho_{ab}(t, \{i\}) \\ \rho_{ba}(t, \{i\}) & \rho_{bb}(t, \{i\}) \end{pmatrix} = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}, \tag{3.3}$$

where the coefficients a and b also depend on the set of parameters {i} and on the time t. Using (3.3) and (3.2), we write the dipole moment of the atom in the form

$$\Pi = \boldsymbol{p}[\rho_{ab}(t, \{i\}) + \rho_{ba}(t, \{i\})]. \tag{3.4}$$

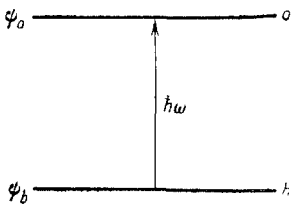


Fig. 3.1

The macroscopic polarization per unit volume of the active medium can be obtained by summing the dipole moments of all the atoms located in the given volume

$$\boldsymbol{P}(\boldsymbol{r}, t)=\sum_{\{i\}} \boldsymbol{p}[\rho_{ab}(t, \{i\})+\rho_{ba}(t, \{i\})], \tag{3.5}$$

where **r** is the radius vector of the considered point of space. Assume that we know the macroscopic polarization **P(r, t)** of the active medium. The passage of the field through the active medium can then be calculated by using Maxwell's classical equations

$$\begin{aligned} &\operatorname{div} \boldsymbol{D}=0, \quad \operatorname{rot} \boldsymbol{E}=-\frac{\partial \boldsymbol{B}}{\partial t}, \\ &\operatorname{div} \boldsymbol{B}=0, \quad \operatorname{rot} \boldsymbol{H}=\boldsymbol{J}+\frac{\partial \boldsymbol{D}}{\partial t}. \end{aligned} \tag{3.6}$$

Here $\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P}$ is the electric induction (bias); $\mathbf{B} = \mu_0\mathbf{H}$ is the magnetic induction; $\mathbf{P} = \mathbf{P}(\mathbf{r}, t)$ is the macroscopic polarization of the medium; $\mathbf{J} = \sigma\mathbf{E}$ is the current density in the medium and is introduced formally to take the radiation absorption into account; ε_0 and μ_0 are the dielectric constant and the magnetic permeability. Simple transformations yield

$$\operatorname{rot} \operatorname{rot} \boldsymbol{E}+\mu_0 \sigma \frac{\partial \boldsymbol{E}}{\partial t}+\mu_0 \varepsilon_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2}=-\mu_0 \frac{\partial^2 \boldsymbol{P}}{\partial t^2}. \tag{3.7}$$

We confine ourselves hereafter to analysis of plane-parallel radiation. We can then put

$$\boldsymbol{E}(\boldsymbol{r}, t)=\boldsymbol{e}E(\boldsymbol{r}, t), \quad \boldsymbol{P}(\boldsymbol{r}, t)=\boldsymbol{e}P(\boldsymbol{r}, t),$$

where **e** is the polarization unit vector. Recognizing that the speed of light in vacuum is $c_0 = (\varepsilon_0\mu_0)^{-1/2}$, we obtain from (3.7)

$$\operatorname{rot} \operatorname{rot} \boldsymbol{E}+\mu_0 \sigma \frac{\partial \boldsymbol{E}}{\partial t}+c_0^{-2} \frac{\partial^2 \boldsymbol{E}}{\partial t^2}=-\mu_0 \frac{\partial^2 \boldsymbol{P}}{\partial t^2}. \tag{3.8}$$

The polarization **P(r, t)** in this expression can be represented by two terms: the "active" polarization $\mathbf{P}_a$, which describes the amplification of the signal, and the "passive" one $\mathbf{P}_p$ due to the passive medium:

$$\boldsymbol{P}=\boldsymbol{P}_{\mathrm{a}}+\boldsymbol{P}_{\mathrm{p}}.$$

In this case, $\mathbf{P}_p = \chi\mathbf{E}$, where χ is the dielectric susceptibility of the passive medium, and $\mathbf{P}_a$ is determined by Eq. (3.5). As a result, we have the following equation for the field:

$$\operatorname{rot} \operatorname{rot} \boldsymbol{E}+\mu_0 \sigma \frac{\partial \boldsymbol{E}}{\partial t}+c^{-2} \frac{\partial^2 \boldsymbol{E}}{\partial t^2}=-\mu_0 \frac{\partial^2 \boldsymbol{P}_{\mathrm{a}}}{\partial t^2}, \tag{3.9}$$

where c is the speed of light in the medium.

To solve the problem of field propagation in an active medium, we must supplement (3.9) with equations that relate the polarization with the field. We turn again to Eq. (3.5). The equation of motion for the density matrix is of the form [1]

$$\frac{d\rho}{dt}=-\frac{i}{\hbar}[\mathscr{H}, \rho]-\frac{1}{2}\{\Gamma, \rho\}, \tag{3.10}$$

where Γ is a phenomenologically introduced diagonal matrix that accounts for the relaxation of levels a and b with relaxation rates γ_a and γ_b, respectively:

$$\Gamma=\begin{pmatrix}\gamma_a & 0\\ 0 & \gamma_b\end{pmatrix}, \tag{3.11}$$

$\mathcal{H}$ is the Hamiltonian of the system, the square brackets denote the commutation operation, and the round brackets denote anticommutation.

The Hamiltonian of the system is written, in the framework of perturbation theory, in the following form:

$$\mathcal{H}=\begin{pmatrix} W_a & -pE \\ -pE & W_b \end{pmatrix}, \tag{3.12}$$

where W_a and W_b are the energies of levels a and b, respectively. As a result, we obtain the following system of equations for the density-matrix elements:

$$\begin{aligned} \frac{d\rho_{aa}}{dt}&=-\gamma_a\rho_{aa}-i\frac{p}{\hbar}E(\rho_{ab}-\rho_{ba}),\\ \frac{d\rho_{bb}}{dt}&=-\gamma_b\rho_{bb}+i\frac{p}{\hbar}E(\rho_{ab}-\rho_{ba}),\\ \frac{d\rho_{ab}}{dt}&=-(\gamma_{ab}+i\omega)\rho_{ab}-i\frac{p}{\hbar}E(\rho_{aa}-\rho_{bb}),\\ \rho_{ba}&=\rho_{ab}^{*}, \end{aligned} \tag{3.13}$$

where we used the relation $W_a-W_b=\hbar\omega$ and introduced the notation $\gamma_{ab} = (\gamma_a + \gamma_b)/2$. Not withstanding such a definition of the parameter γ_{ab}, it is an independent parameter, not connected solely with γ_a and γ_b. It will be shown below that γ_{ab} determines the width of the gain band of the active medium. Experiments show that the parameter γ_{ab} exceeds as a rule the quantity $(\gamma_a + \gamma_b)/2$:

$$\gamma_{ab}>(\gamma_a+\gamma_b)/2.$$

The point is that, besides radiative decay which, as a matter of fact is taken into account by Eqs. (3.13), there exist various nonradiative relaxation mechanism. It is their presence which leads indeed to the increase of the parameter γ_{ab}.

Equations (3.13) were written for one atom. Using (3.4), we can obtain the following system of equations:

$$\begin{aligned} \frac{\partial^2\Pi}{\partial t^2}+2\gamma_{ab}\frac{\partial\Pi}{\partial t}+(\gamma_{ab}^2+\omega^2)\Pi&=-2\omega\frac{p^2}{\hbar}\boldsymbol{E}(\rho_{aa}-\rho_{bb}),\\ \frac{\partial\rho_{aa}}{\partial t}+\gamma_a\rho_{aa}&=(\hbar\omega)^{-1}\boldsymbol{E}\left(\frac{\partial\Pi}{\partial t}+\gamma_{ab}\Pi\right),\\ \frac{\partial\rho_{bb}}{\partial t}+\gamma_b\rho_{bb}&=-(\hbar\omega)^{-1}\boldsymbol{E}\left(\frac{\partial\Pi}{\partial t}+\gamma_{ab}\Pi\right). \end{aligned} \tag{3.14}$$

2. Approximation of Homogeneous Broadening of Spectral Line. We confine ourselves hereafter to the case of identical atoms, i.e., to the so-called homogeneous broadening of the spectral line. We assume in addition that $\gamma_a = \gamma_b = \gamma$. This assumption does not affect the matter in question, but makes it possible to arrive in final analysis at perspicuous analytic results. The assumption that the gain line is homogeneous is the most significant restriction. The point is that the amplifying media such as Nd^{3+} glass, high pressure CO_2, or dyes have inhomogeneously broadened gain lines. Nonetheless, a consideration of the case of homogeneous broadening is useful, since it can reveal the basic regularities that take place when signals are amplified by laser amplifiers. The more general case of inhomogeneous broadening is considered in detail in [3, 4].

In the homogeneous-broadening approximation we have

$$\boldsymbol{P}(r,\ t)=N\boldsymbol{p}(\rho_{ab}+\rho_{ba})=N\Pi, \tag{3.15}$$

where N is the number of active atoms per unit volume of the working medium. Recognizing that

$$\rho_{aa}N=N_a,\quad \rho_{bb}N=N_b, \tag{3.16}$$

where N_a and N_b are the populations of levels a and b, we obtain from (3.14) and (3.15)

$$\begin{aligned} \frac{\partial^2\boldsymbol{P}}{\partial t^2}+2\gamma_{ab}\frac{\partial\boldsymbol{P}}{\partial t}+(\gamma_{ab}^2+\omega^2)\boldsymbol{P}&=-2\omega\frac{p^2}{\hbar}\boldsymbol{E}(N_a-N_b),\\ \frac{\partial(N_a-N_b)}{\partial t}+\gamma(N_a-N_b)&=2(\hbar\omega)^{-1}\boldsymbol{E}\left(\frac{\partial\boldsymbol{P}}{\partial t}+\gamma_{ab}\boldsymbol{P}\right). \end{aligned} \tag{3.17}$$

Equations (3.9) and (3.17) form a closed system of equations:

$$\begin{gathered}\operatorname{rot}\operatorname{rot}\boldsymbol{E}+\mu_0\sigma\frac{\partial\boldsymbol{E}}{\partial t}+c^{-2}\frac{\partial^2\boldsymbol{E}}{\partial t^2}=-\mu_0\frac{\partial^2\boldsymbol{P}}{\partial t^2},\\ \frac{\partial^2\boldsymbol{P}}{\partial t^2}+2\gamma_{ab}\frac{\partial\boldsymbol{P}}{\partial t}+(\gamma_{ab}^2+\omega^2)\boldsymbol{P}=-2\omega\frac{p^2}{\hbar}\boldsymbol{E}(N_a-N_b),\\ \frac{\partial(N_a-N_b)}{\partial t}+\gamma(N_a-N_b)=2(\hbar\omega)^{-1}\boldsymbol{E}\left(\frac{\partial\boldsymbol{P}}{\partial t}+\gamma_{ab}\boldsymbol{P}\right).\end{gathered}\tag{3.18}$$

We consider next a plane wave propagating along the z axis:

$$E(\boldsymbol{r},\ t)=A(z,\ t)\cos(ft-kz)=E(z,\ t),\tag{3.19}$$

where k = f/c is the wave number and f the frequency of the amplified signal. The amplitude A(z, t) in this expression describes the pulse envelope of the wave. The function A(z, t) depends in fact on one parameter $\tau = t - z/c$:

$$A(z,\ t)=A(\tau).\tag{3.20}$$

Substitution of (3.19) in the first equation of the system (3.18) yields

$$-\frac{\partial^2E}{\partial z^2}+\mu_0\sigma\frac{\partial E}{\partial t}+c^{-2}\frac{\partial^2E}{\partial t^2}=-\mu_0\frac{\partial^2P}{\partial t^2}.\tag{3.21}$$

It is next natural to assume that the polarization of the active atoms varies in accord with the same harmonic law as the field (3.19). In the general case we can write

$$P(z,\ t)=C(z,\ t)\cos(ft-kz)+S(z,\ t)\sin(ft-kz),\tag{3.22}$$

where C and S are slowly varying functions. We substitute this expression in the second equation of (3.18). Equating the coefficients of cos (ft − kz) and sin (ft − kz), and recognizing also that $\omega^2 >> \gamma^2_{ab}$, we get

$$\begin{aligned}\frac{\partial C}{\partial t}&=-\gamma_{ab}C+(\omega-f)S,\\ \frac{\partial S}{\partial t}&=-\gamma_{ab}S-(\omega-f)C-\frac{p^2}{\hbar}A(N_a-N_b).\end{aligned}\tag{3.23}$$

In the derivation of expressions (3.23) we have neglected the derivatives, which vary slowly compared with the optical frequencies ω and f:

$$\begin{aligned}\frac{\partial C}{\partial t}&\ll\omega C, \quad \frac{\partial C}{\partial t}\ll fC,\\ \frac{\partial S}{\partial t}&\ll\omega S, \quad \frac{\partial S}{\partial t}\ll fS,\end{aligned}\tag{3.24}$$

and used, in addition, the approximate equality

$$\omega^2-f^2\approx(\omega-f)2f.\tag{3.25}$$

Indeed, the width of the gain band is of the order of γ_{ab}. It is natural to consider only a case when the external-field frequency lands in the gain band, i.e.,

$$|\omega-f|<\gamma_{ab}.\tag{3.26}$$

Since $\omega >> \gamma_{ab}$, we have

$$\omega+f\approx 2f,\tag{3.27}$$

the approximate formula (3.25) is therefore correct.

We substitute (3.22) and (3.19) in (3.21). Retaining only the principal-order terms, we obtain

$$\frac{\partial A}{\partial z}+c^{-1}\frac{\partial A}{\partial t}=-\beta A-\frac{k}{2\varepsilon_0}S,\tag{3.28}$$

where $\beta = \sigma/2\varepsilon_0 c$ plays the role of the absorption coefficient of the medium. Finally, we substitute (3.22) and (3.19) in the third equation of the system (3.18). Retaining in the right-hand side of this equation only the terms that vary slowly compared with ω and f, we get

$$\frac{\partial (N_a - N_b)}{\partial t} = -\gamma (N_a - N_b) + \hbar^{-1} AS. \tag{3.29}$$

Equations (3.28), (3.23), and (3.29) comprise a closed system:

$$\begin{aligned} \frac{\partial A}{\partial z} + c^{-1} \frac{\partial A}{\partial t} &= -\beta A - \frac{k}{2\varepsilon_0} S, \\ \frac{\partial C}{\partial t} &= -\gamma_{ab} C + (\omega - f) S, \\ \frac{\partial S}{\partial t} &= -\gamma_{ab} S - (\omega - f) C - \frac{p^2}{\hbar} A \Delta N, \\ \frac{\partial \Delta N}{\partial t} &= -\gamma \Delta N + \hbar^{-1} AS, \end{aligned} \tag{3.30}$$

where we have introduced the following notation for the population difference between levels a and b:

$$\Delta N = N_a - N_b.$$

Interest attaches usually to the temporal evolution of the amplified pulse. It is convenient here to transform to the "running" coordinate τ via the change of variables

$$\tau = t - z/c, \qquad z = z.$$

The system (3.30) then takes the form

$$\begin{aligned} \frac{\partial A}{\partial z} &= -\beta A - \frac{k}{2\varepsilon_0} S, \\ \frac{\partial C}{\partial \tau} &= -\gamma_{ab} C + (\omega - f) S, \\ \frac{\partial S}{\partial \tau} &= -\gamma_{ab} S - (\omega - f) C - \frac{p^2}{\hbar} A \Delta N, \\ \frac{\partial \Delta N}{\partial \tau} &= -\gamma \Delta N + \hbar^{-1} AS. \end{aligned} \tag{3.31}$$

3. Long-Pulse Approximation. We now note the following. When laser amplifiers are used in practice in actual devices, the duration of the amplified pulse exceeds greatly the quantity γ^{-1}_{ab}. The point is that when shorter pulses are amplified the output pulse is "stretched." We have here a rather good analog with radio engineering, where it is known that transmission of a short pulse through a narrow-band filter broadens the pulse and causes loss of part of its energy. In our case, the role of the narrow-band filter is played by the active medium, whose gain band width is proportional to γ_{ab}. When designing laser amplifiers the tendency, therefore, is to have the duration T of the amplified pulse satisfy the relation

$$T \gg \gamma_{ab}^{-1}. \tag{3.32}$$

The value of γ^{-1}_{ab} is different for different amplifying media. For example, for an Nd^{3+} laser amplifier the value of the parameter γ^{-1}_{ab} is 0.2-0.5 psec, while for a low-pressure CO_2 amplifier the value is about 20 psec. It follows, therefore, that for each specific amplifying medium the validity of condition (3.32) must be separately verified. When this condition is satisfied we can neglect the left-hand sides of the second and third equations in (3.31). We then obtain

$$S = -\frac{p^2}{\hbar \gamma_{ab}} AL (\omega - f) \Delta N, \tag{3.33}$$

where

$$L (\omega - f) = \frac{\gamma_{ab}^2}{(\omega - f)^2 + \gamma_{ab}^2} \tag{3.34}$$

is a function that describes the Lorentz shape of the spectral line.

Substituting the expression for S in the first and last equations of the system (3.31), we obtain a closed system of equations:

$$\begin{aligned}\frac{\partial A}{\partial z} &= -\beta A + \frac{kp^2}{2\varepsilon_0 \hbar \gamma_{ab}} AL(\omega - f)\,\Delta N,\\ \frac{\partial \Delta N}{\partial \tau} &= -\gamma\,\Delta N - \frac{p^2}{\hbar^2 \gamma_{ab}} A^2 L(\omega - f)\,\Delta N.\end{aligned} \tag{3.35}$$

We note that the first equation of this system can be represented in the form

$$\frac{\partial A}{\partial z} = (\alpha - \beta)\,A, \tag{3.36}$$

where

$$\alpha = \frac{kp^2}{2\varepsilon_0 \hbar \gamma_{ab}} L(\omega - f)\,\Delta N. \tag{3.37}$$

At constants α and β the solution of (3.36) is

$$A(z,\ t) = A_0(t) \exp[(\alpha - \beta)\,z]. \tag{3.38}$$

Thus, the parameter α, introduced in Eq. (3.37), plays the role of the gain of the active medium, the parameter β plays the role of the attenuation factor, while $\exp[(\alpha - \beta)z]$ is the overall gain of the amplifier.

It is of interest to analyze expression (3.37) for the gain. First of all, the gain depends on the frequency mismatch between the resonant frequency ω of the active atoms and the frequency f of the amplified signal. With increasing frequency difference $\omega - f$ the gain of the active medium decreases in accord with Eq. (3.34). The role of the gain bandwidth is played by $2\gamma_{ab}$. It is now clear that condition (3.32) means that the duration of the amplified pulse must be longer than the reciprocal gain bandwidth. In other words, the width of the input-signal spectrum should be less than the gain bandwidth.

The ratio of the gain α to the inverted population ΔN per unit volume is called the interaction cross section. This parameter has the dimension of area and is equal to

$$\varkappa = \frac{\alpha}{\Delta N} = \frac{kp^2}{2\varepsilon_0 \hbar \gamma_{ab}} L(\omega - f). \tag{3.39}$$

The second equation of the system (3.35) describes the decrease of the inverted population when the amplified field interacts with the active medium. A feature of the considered approximation is that the inverted population ΔN can only decrease. In fact, in the signal-amplification regime the right-hand side of the considered equation cannot be larger than zero, since the gain α must be positive. Consequently, the derivative of the inverted population is negative. It must be noted that this behavior of the inverted population is a consequence of the approximation (3.32). When short pulses are amplified and the reverse of (3.32) holds, the population can be increased by the passage of a light pulse [3].

The system (3.35) can be written also for the field intensity $I = A^2$:

$$\begin{aligned}\frac{\partial I}{\partial z} &= -2\beta I + \frac{kp^2}{\varepsilon_0 \hbar \gamma_{ab}} I L(\omega - f)\,\Delta N,\\ \frac{\partial \Delta N}{\partial \tau} &= -\gamma\,\Delta N - \frac{p^2}{\hbar^2 \gamma_{ab}} I L(\omega - f)\,\Delta N.\end{aligned} \tag{3.40}$$

It is of interest to consider one particular case. Assume that the following inequality holds ($\gamma = \gamma_a = \gamma_b$):

$$\gamma_{ab}^{-1} \ll T \ll \gamma^{-1}. \tag{3.41}$$

Such an inequality is usually satisfied in a neodymium-glass laser amplifier. In this case typical values of the parameters are

$$\gamma_{ab}^{-1} \approx 0.5 \text{ psec}, \quad T \approx 50 \text{ nsec}, \quad \gamma^{-1} \approx 0.5 \text{ msec}.$$

If condition (3.41) holds, we can neglect the quantity $\gamma\Delta N$ in the second equation of (3.40) and put $\gamma = 0$. Then a joint solution of Eqs. (3.40) yields the following integrodifferential equation:

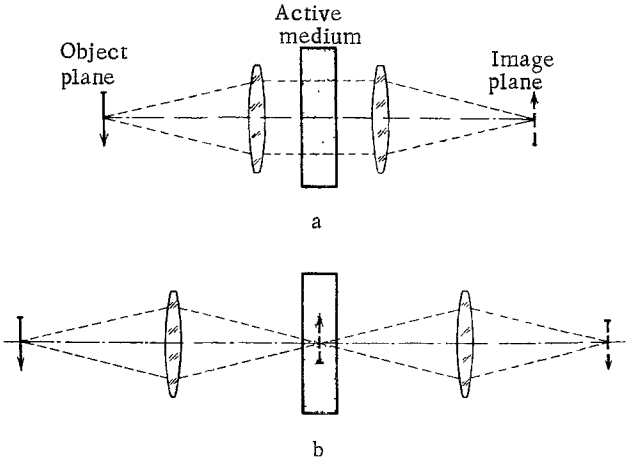


Fig. 3.2

$$\frac{\partial I(\tau, z)}{\partial z} = \left\{2\alpha_0 \exp\left[-\frac{2\varepsilon_0}{k} \varkappa \int_{-\infty}^{\tau} I(\tau', z)\, d\tau'\right] - 2\beta\right\} I(\tau, z), \tag{3.42}$$

where α_0 is the gain of the active medium prior to the arrival of the amplified pulse, and is equal to

$$\alpha_0 = \frac{kp^2}{2\varepsilon_0 \hbar \gamma_{ab}} L(\omega - f)\, \Delta N(0). \tag{3.43}$$

Equation (3.42) shows that the gain saturates when the input pulse is amplified, and this saturation is defined by the parameter

$$\eta = \exp\left[-\frac{2\varepsilon_0}{k} \varkappa \int_{-\infty}^{\tau} I(\tau', z)\, d\tau'\right]. \tag{3.44}$$

A characteristic feature is that the saturation parameter η in a given cross section of the amplifier z and at a given instant τ depends on the total energy of the signal

$$W = \int_{-\infty}^{\tau} I(\tau', z)\, d\tau',$$

that has passed through this cross section by the instant τ.

When laser amplifiers are used in laser-radar receivers, one deals as a rule with very weak light fluxes. In practice, therefore, gain saturation does not play a noticeable role. In this simplest case the field propagates through the amplifier in exactly the same way as in a passive medium. The only difference is that the amplitude of each plane wave is amplified in accord with the law (3.38), where z is the distance traversed in the amplifier by the given plane wave.

In amplification of a multimode field, which is a superposition of many plane waves, each plane wave is amplified independently of the others. In first-order approximation the phase of each plane wave is also conserved. As a result, the field spatial distribution formed at the output of the amplifier will be the same as in the case of a passive medium, but amplified by many times. Thus, it becomes possible to amplify the field that carries information on the image of the object.

Figure 3.2 shows two possible constructions of laser image amplifiers, with the amplification in parallel (a) and converging (b) beams, respectively. When the amplification is in a parallel beam, the active medium is used uniformly and is less subject to saturation than in the case of amplification in convergent beams. Amplification with convergent beams, however, is convenient, in that inhomogeneities of the refractive index of the medium have less effect on the quality of the obtained image. This result can be obtained on the basis

of the theory of linear optical systems [5]. In the parallel-beam amplification system the object and image planes coincide with the focal planes of the corresponding lenses.

The field at the amplifier output is the enhanced and somewhat distorted Fourier transform of the undistorted image field. The distortions introduced can be described by the amplitude–phase factor

$$a(\boldsymbol{r})\exp[i\varphi(\boldsymbol{r})],$$

where $a(\mathbf{r})$ are the amplitude distortions and $\varphi(\boldsymbol{r})$ the phase distortions due to the inhomogeneity of the active-medium refractive index. If there are no distortions, we have $a(\mathbf{r})$ = const, $\varphi(\boldsymbol{r})=0$ within the limits of the exit aperture of the amplifier. The resolving power of the entire system is determined by the system scattering function

$$G(\boldsymbol{x}) = \int a(\boldsymbol{r})\exp[i\varphi(\boldsymbol{r}) + i\boldsymbol{x}\boldsymbol{r}]\,d^2r,$$

where the integral is calculated over the surface of the output energy of the amplifier. The phase and amplitude distortions lead in this case to a broadening of the principal maximum of $G(\mathbf{x})$ and hence to a deterioration of the resolving power.

In the amplification system with converging beams it can be assumed that the field at the amplifier output is the undistorted image field multiplied by the earlier amplitude–phase factor. Since only the intensity of the image is of importance when the latter is recorded, the phase distortions play no role in this case. The amplitude distortions cause neighboring sections of the image to differ in brightness, but the resolving power does not become worse in first-order approximation.

We note also that more detailed investigations reveal certain differences between laser amplifiers and passive image-transmission systems. It turns out that laser amplifiers introduce in the produced image certain distortions that are peculiar only to them [6]. These questions, however, will not be considered here.

4. Spontaneous Emission in Laser Amplifiers. A very important feature of laser amplifiers is the spontaneous-emission noise. The presence of spontaneous emission is essentially a quantum effect and cannot be explained by classical theory. In quantum field theory one introduces the photon creation and annihilation operators C^+ and C, which act in one linear space containing N photons [7]. The linear photon space is, in the language of optics, one emission mode. It can be shown that the matrix elements of the operator C differ from zero only for transitions that take place with the total number of photons in the mode increased by unity. The probability of such a transition is proportional to the number of photons present in the mode prior to the instant of the transition. The probability of such a transition, i.e., the probability of absorbing a photon, is consequently equal to zero if the initial number of photons is zero.

It can be shown that the matrix elements of the operator C^+ differ from zero only for transitions in which the total number of photons in the mode is increased by unity. The probability of such a transition is found to be proportional to N + 1, where N is the number of photons present in the mode prior to the transitions. Thus, even if the initial number of photons is zero, the probability of emitting a photon is not zero. This probability is due to spontaneous emission. An important consequence is that the probability of spontaneous emission into one mode of the field is equal to the probability of emission of one photon into the same mode under the action of one external photon.

The higher the gain of the active medium, the higher (other conditions being equal) the spontaneous-emission power. The spontaneous emission determines in fact the sensitivity of a receiver with a laser amplifier. We thus have two contradictory requirements in the design of laser amplifiers: it is necessary to increase the gain and to decrease the spontaneous-emission power at the amplifier output. Taking into account the direct connection between the probabilities of the induced and spontaneous emission, an unambiguous relation can be established between the amplifier gain and the power of the spontaneous noise at its output.

Consider a volume element ΔV of the active medium. The spontaneous-emission power from this volume into a solid angle $d\Omega$ in the frequency interval (f, f + df) is

$$dP = \hbar f t_s^{-1} N_a L(f)\frac{d\Omega}{4\pi}\,df\,dV, \tag{3.45}$$

where N_α is the number of active atoms excited to the upper energy level α per unit volume of the medium; L(f) is the spectral-line shape function normalized in such a way that

$$\int_{-\infty}^{+\infty} L(f)\,df = 1, \tag{3.46}$$

and t_s is the average lifetime of the excited state relative to spontaneous transitions.

The spontaneous emission is amplified by the active medium just as the useful radiation of the signal. If the spontaneous emission arose in a volume element dV located at a distance z from the output, its power will be amplified at the output by a factor exp [2α(f)z] where the gain α(f) is determined by Eq. (3.37). Transforming from the function L(ω − f) in (3.37) to the function L(f) defined in (3.46), we get

$$2\alpha(f) = \frac{\pi}{\varepsilon_0}\frac{kp^2}{\hbar}L(f)\Delta N. \tag{3.47}$$

As a result, the spontaneous-emission power at the amplifier output in a solid angle dΩ and in the frequency interval (f, f + df) is equal to

$$P = S\hbar f t_s^{-1} N_a L(f)\frac{d\Omega}{4\pi}\,df\int_0^l \exp[2\alpha(f)z]\,dz, \tag{3.48}$$

where S is the cross-sectional area and l is the length of the active medium. Calculating the integral in this formula and introducing the amplifier gain G(f), defined by

$$G(f) = \exp[2\alpha(f)l], \tag{3.49}$$

we obtain

$$P = \hbar f\frac{N_a L(f)}{2\alpha(f)\,t_s}[G(f)-1]\frac{S\,d\Omega}{4\pi}\,df. \tag{3.50}$$

It is known from quantum theory of radiation that the lifetime of an excited state of the atom, due to spontaneous emission, is determined by the relation (see, e.g., [7, 8]),

$$t_s^{-1} = \frac{4}{3}\frac{k^3p^2}{\hbar}. \tag{3.51}$$

Using expression (3.47) for the gain of the medium, we can express t_s in terms of α(f). As a result we have

$$t_s = \frac{3}{32}\frac{\lambda^2 L(f)\,\Delta N}{\alpha(f)\,\varepsilon_0\pi}. \tag{3.52}$$

Substituting this expression in (3.50), we obtain

$$P = \frac{4}{3}\hbar f\frac{N_a}{N_a - N_b}[G(f)-1]\frac{S\,d\Omega}{\lambda^2}\,df. \tag{3.53}$$

The total spontaneous-emission power is obtained by integrating this expression over the solid angle dΩ and over the frequency:

$$P = \frac{4}{3}\hbar\omega\,\Delta f[G(\omega)-1]\frac{N_a}{N_a - N_b}\left\{\frac{S\Omega}{\lambda^2}\right\}. \tag{3.54}$$

We have introduced here the amplifier gain G(ω) at the center of the spectral line, the gain bandwidth Δf, and the total solid angle Ω in which the emission from the amplifier output is received.

We note that the expression in the curly brackets of (3.54) is the number of emission-field modes at the output of the laser amplifier. The spontaneous-emission power in one mode is, therefore,

$$P_1 = \frac{4}{3}\hbar\omega\,\Delta f[G(\omega)-1]\frac{N_a}{N_a - N_b}. \tag{3.55}$$

In most amplifying media with large gain one can neglect the lower-working-level population N_b compared with the population N_α of the upper one, and the gain G(ω) is much larger than unity. Under these conditions, expression (3.55) simplifies to

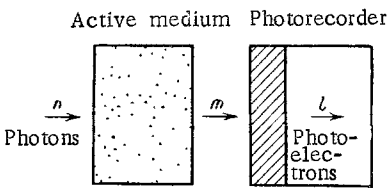


Fig. 3.3

$$P_1 = \frac{4}{3}\hbar\omega\,\Delta f\,G(\omega). \tag{3.56}$$

Thus, the spontaneous-emission power is directly proportional to the power gain of the amplifier. One might draw from this the incorrect conclusion that increasing the gain of a laser amplifier does not increase the sensitivity. A more attentive examination shows, however, that this is not so, and an increase of the gain leads to noticeable increase of the sensitivity. This question is considered in the next subsection.

5. Sensitivity of Receivers with Laser Amplifiers. One of the basic factors that determine the effectiveness of the use of laser amplifiers in receiving devices is the sensitivity. For a radar receiver, the sensitivity problem is solved by plotting the detection characteristics, i.e., by plotting the dependence of the probability of missing an input signal given the probability of a false alarm.

Consider the hypothetical receiver with laser amplifier, shown in Fig. 3.3. Let the investigated-signal field, containing n photons, be incident on the amplifier input. Since the input field in laser radar is always random, the number of incoming photons is also random. We denote by P_n the probability that n photons have entered the amplifier.

In accord with the fundamental physical premises, the amplification in the active medium is also random. Each photon entering the active medium can either be absorbed or create another photon. In quantum-mechanics language it can be said that the created photon belongs to the same linear space as the exciting photon, i.e., it has the same momentum as the exciting photon. The created photon has consequently the same frequency, and "moves" in the same direction as the exciting photon. We denote the probability of photon absorption per unit time by b, and the probability of photon creation per unit time by a. In addition, account must be taken of the spontaneous emission of a photon, with a probability denoted by c. Using the introduced parameters a, b, and c, we can calculate the probability of the appearance, within the recording time, of m photons at the laser amplifier output for an input of n photons. We designate this probability by P_{nm}.

Finally, each of the m photons that land on the photodetector have a certain probability of knocking a photoelectron out of the photocathode. The probability that one photon will lead to the appearance of one photoelectron is equal to the quantum efficiency q of the photocathode. The total probability that the photocathode will emit l photoelectrons if m photons landed on it is determined by the known formula [9]

$$P_{ml} = \frac{m!}{l!\,(m-l)!}\,q^l\,(1-q)^{m-l}. \tag{3.57}$$

A distribution of this type is called binomial.

Equation (3.57) was written under the assumption that the photodetector has no intrinsic noise sources. That this assumption is justified is subject to no doubt when the amplifier gain is high enough for the detector intrinsic noise to play no role. We shall thus consider hereafter only one cause of noise, namely the spontaneous emission of the laser amplifier.

As for the probability distribution P_n for the photons entering the amplifier, we assume that it is of the Poisson type with a mean number μ of entering photons:

$$P_n = \exp(-\mu)\,\frac{\mu^n}{n!}. \tag{3.58}$$

An explicit expression for the probability P_{nm}, which determines the photon multiplication in the amplifier, can be obtained by using the parameters a, b, and c introduced above for the photon creation, annihilation, and spontaneous emission, respectively. In this procedure, the parameter that determines the gain is the time interval t during which the photon is located in the active medium. In this case, the gain G is obviously equal to

$$G = \exp[(a-b)t]. \tag{3.59}$$

The problem of finding the explicit form of the distribution P_{nm} was solved in this formulation in [10]. We shall not repeat here the lengthy intermediate calculations that led to the final result. (Details can be found in [10].) We write the final result:

$$P_{nm} = \frac{(a-b)^L a^m b^n (G-1)^{m+n}}{(aG-b)^{m+n+L}} \sum_{j=0}^{n} \frac{n!\,\Gamma(m+L)}{j!\,(n-j)!\,(m-j)!\,\Gamma(j+L)} \left[\frac{(a-b)^2 G}{ab(G-1)^2}\right]^j. \tag{3.60}$$

In this expression we have introduced the parameter L = c/a, and for the gamma function of integer argument the following relation holds:

$$\Gamma(N) = (N-1)!. \tag{3.61}$$

Taking into account the statements made above concerning the probabilities of the spontaneous and stimulated emission, we arrive at the conclusion that the parameter L is equal to the ratio of the number of modes in which detection of the spontaneous emission takes place to the number of signal-radiation modes.

The final probability P_l for the appearance of l photoelectrons at the photodetector output is equal to [11]

$$P_l = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} P_n P_{nm} P_{ml}. \tag{3.62}$$

In the absence of a signal at the entry to the amplifier, i.e., at μ = 0, we obtain from (3.62) with allowance for (3.57), (3.58), and (3.60), the probability of the appearance of l noise electrons at the photodetector output:

$$P_{\text{no}}(l) = \frac{1}{l!\,(L-1)!} \left(\frac{q}{1-q}\right)^l \left(\frac{1-d}{G-d}\right)^L \sum_{m=l}^{\infty} \left[\frac{(G-1)(1-q)}{G-d}\right]^m \frac{(m+L-1)!}{(m-l)!}, \tag{3.63}$$

where we have introduced the notation d = b/a and, in addition, assumed that the parameter L is an integer. The last assumption does not really restrict the analysis that follows for, in practice, one deals always with integer values of the parameter L.

At μ ≠ 0 we obtain the probability of the photodetector emitting l photoelectrons if a signal and noise are simultaneously present:

$$P_{\text{s+no}}(l) = \frac{\exp(-\mu)}{l!} \left(\frac{q}{1-q}\right)^l \left(\frac{1-d}{G-d}\right)^L \sum_{m=l}^{\infty} \frac{m!\,(m+L-1)!}{(m-l)!} \left[\frac{(G-1)(1-q)}{G-d}\right]^m \times$$
$$\times \sum_{n=0}^{m} \left[\frac{\mu(G-1)d}{G-d}\right]^n \sum_{j=0}^{n} \frac{[(1-d)^2 G/(G-1)^2]^j}{j!\,(m-j)!\,(n-j)!\,(j+L-1)!}. \tag{3.64}$$

In single-threshold recording of the signal from the photodetector output, the probability F of a false alarm is

$$F = 1 - \sum_{l=0}^{l_0} P_{\text{no}}(l), \tag{3.65}$$

where l_0 is the preassigned threshold number of photoelectrons. The probability β of missing a signal is equal to

$$\beta = \sum_{l=0}^{l_0} P_{\text{s+no}}(l). \tag{3.66}$$

The detection characteristics of a radar receiver comprise the dependences of the missed-signal probability β on the average value of the input signal μ at a given false-alarm probability F. To construct the detection characteristics it is necessary to solve, given the false-alarm probability, Eq. (3.65) with respect to l_0 for each average value of the number μ

of entering photons. The result is the dependence $l_0 = f(\mu)$. This dependence is substituted in (3.66) and the missed-signal probability β is determined. To be able to perform these calculations it is necessary to simplify the explicit form of the expressions (3.63) and (3.64).

We begin with expression (3.63) for the probability of appearance of noise electrons. We introduce the notation

$$A=\frac{(G-1)(1-q)}{G-d}. \tag{3.67}$$

Equation (3.63) then takes the form

$$P_{\text{s+no}}(l)=\left(\frac{q}{1-q}\right)^l \frac{1}{l!\,(L-1)!}\left(\frac{1-d}{G-d}\right)^L \sum_{m=l}^{\infty} A^m \frac{(m+L-1)!}{(m-l)!}. \tag{3.68}$$

Making the following change of the summation index

$$k=m-l, \tag{3.69}$$

we obtain

$$P_{\text{no}}(l)=\left(\frac{q}{1-q}\right)^l \frac{1}{l!\,(L-1)!}\left(\frac{1-d}{G-d}\right)^L A^l \sum_{k=0}^{\infty} A^k \frac{(k+l+L-1)!}{k!}. \tag{3.70}$$

The infinite sum in the right-hand side of (3.70) can be calculated in explicit form. In fact, let us consider an infinite sum of the form

$$\sum_{k=0}^{\infty} A^k \frac{(n+k)!}{k!}. \tag{3.71}$$

By direct calculation we can verify that the following equality holds:

$$\sum_{k=0}^{\infty} A^k \frac{(n+k)!}{k!}=\frac{\partial^n}{\partial A^n}\sum_{k=0}^{\infty} A^k. \tag{3.72}$$

If A < 1, both series in this equation converge. The series in the right-hand side is a geometric progression and its sum is

$$\sum_{k=0}^{\infty} A^k=\frac{1}{1-A}. \tag{3.73}$$

Differentiating the right-hand side of (3.73) n times, we obtain

$$\sum_{k=0}^{\infty} A^k \frac{(n+k)!}{k!}=\frac{n!}{(1-A)^{n+1}}. \tag{3.74}$$

Using this equality, we can reduce (3.70) ultimately to the form:

$$P_{\text{no}}(l)=\left(\frac{\varkappa}{1+\varkappa}\right)^l \frac{1}{(1+\varkappa)^L}\frac{(L+l-1)!}{l!\,(L-1)!}, \tag{3.75}$$

where we have introduced the additional notation

$$\varkappa=\frac{q(G-1)}{1-d}. \tag{3.76}$$

We can carry out analogously the summation in (3.64). These lengthy but straightforward calculations are relegated to Appendix III. We write directly the final result:

TABLE 2. Threshold Numbers l_0 of Photoelectrons for Different Values of the Parameters κ and L

κ \ L	5	10^1	10^2	10^3	10^4
$2\cdot10^{-3}$	3	3	6	13	45
$2\cdot10^{-2}$	5	6	13	46	273
$2\cdot10^{-1}$	12	15	49	280	2241
$2\cdot10^{0}$	60	80	343	2406	21240
$2\cdot10^{1}$	539	749	3288	23660	211148
$2\cdot10^{2}$	5461	7916	33730	238612	2117291

TABLE 3. Dependence of Missed-Signal Probability β on the Average Number μ of Incoming Photons

L \ κ	$2\cdot10^{-3}$	$2\cdot10^{-2}$	$2\cdot10^{-1}$	$2\cdot10^{0}$	$2\cdot10^{1}$	$2\cdot10^{2}$
10^1	0,992 340 0,295 2385 0,138 3067 0,0587 3749 0,00872 5112	0,995 55 0,204 386 0,0662 496 0,0180 606 0,00898 661	0,998 13 0,316 80 0,0716 107 0,0100 134 0,00326 148	0,999 6 0,786 24 0,143 42 0,0172 54 0,00499 60	0,999 5 0,569 27 0,140 38 0,0147 49 0,00363 54	0,999 5 0,544 29 0,113 41 0,0357 47 0,00905 53
10^2	0,996 545 0,503 2727 0,120 4363 0,0354 5454 0,00895 6544	0,997 112 0,544 563 0,159 789 0,0264 1014 0,00907 1127	0,999 28 0,489 144 0,235 173 0,0849 202 0,00517 260	0,999 14 0,507 71 0,0707 100 0,0153 114 0,00240 128	0,999 12 0,527 64 0,0649 90 0,0123 103 0,00161 115	0,999 13 0,531 68 0,216 82 0,054 96 0,0873 109
10^3	0,997 1108 0,565 5542 0,170 7759 0,02883 9976 0,00999 11085	0,999 261 0,532 1307 0,0907 1830 0,0237 2091 0,00475 2352	0,999 79 0,497 399 0,199 479 0,0498 559 0,00770 639	0,999 40 0,507 202 0,191 243 0,0410 283 0,00493 324	0,999 36 0,513 182 0,178 219 0,0325 255 0,00304 292	0,999 38 0,507 193 0,158 231 0,0235 270 0,00162 308

Note. In each column, the number on the left is β and on the right μ.

$$P_{s+no}(l)=\exp\left[-\frac{\mu Gq}{\kappa(1+\kappa)}\right]\frac{\kappa^l}{(1+\kappa)^{L+l}}(L+l-1)!\sum_{j=0}^{l}\left[\frac{\mu Gq}{\kappa(1+\kappa)}\right]^j\frac{1}{j!(l-j)!(L+j-1)!}. \qquad (3.77)$$

Using this expression, and also Eq. (3.75) for the probability of appearance of noise electrons, we calculate the threshold number of photoelectrons l_0 and the detection characteristics for a given false-alarm probability. To this end it is necessary to specify the actual values of the parameters d and q. The parameter $d = b/a$ is actually the ratio of the populations of the lower and upper levels of the working transition in the active medium. For example, for an active medium in the form of Nd^{3+} glass, a typical value of this parameter is d = 0.1. We assume a photocathode quantum efficiency $q = 2\cdot10^{-5}$, which corresponds approximately to the quantum efficiency of oxygen–cesium photocathodes at the wavelength λ = 1.06 μm. In Table 2 are listed the threshold numbers l_0 of the photoelectrons calculated for the indicated values of the parameters d and q at a specified false-alarm probability $F = 10^{-6}$. Here we used, for the parameters κ and L, the values of greatest interest from the practical point of view. In Figs. 3.4-3.7 and in Table 3 are given, for the same values of the parameters d, q, and κ and for the same false-alarm probability, the computer-calculated detection characteristics, i.e., the dependence of the missed-signal probability β on the average number μ of incoming photons.

Analysis of the results shows that an increase of the parameter κ leads to an improvement of the detection characteristics. Turning to Eq. (3.76) for the explicit form of this parameter, we conclude that it can be increased either by increasing the quantum efficiency q of the photodetector or by increasing the amplifier gain G. Thus, the use of a laser amplifier is in fact equivalent to increasing the quantum efficiency of the photodetector.

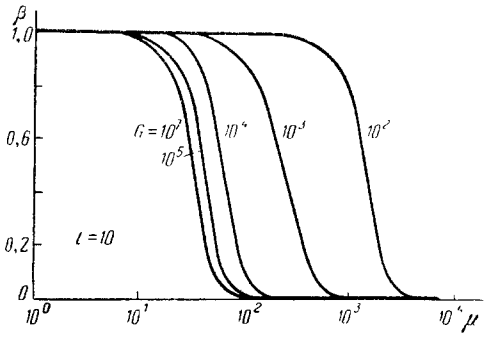


Fig. 3.4

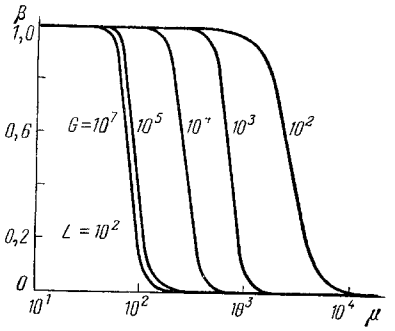


Fig. 3.5

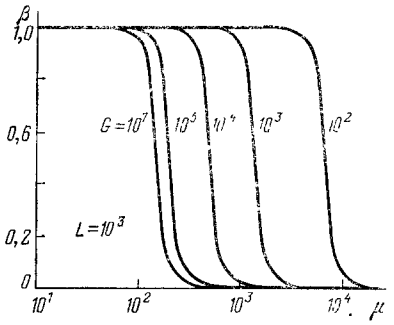


Fig. 3.6

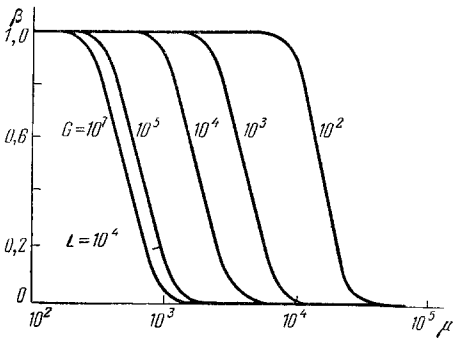


Fig. 3.7

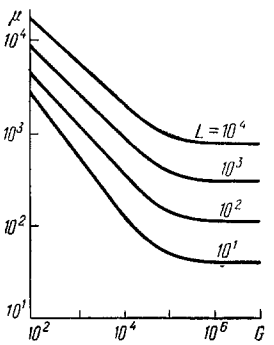


Fig. 3.8

The results listed in Table 3 indicate that to ensure maximum sensitivity it is necessary to satisfy the condition

$$\varkappa \geqslant 1.$$

It follows, therefore, that at low quantum efficiency of the photodetector, and at sufficiently high population inversion in the active medium (i.e., at d << 1), the gain of the laser amplifier should be chosen to satisfy the condition

$$G \geqslant q^{-1}.$$

Figure 3.8 shows the dependence of the average number μ of the incoming photons needed to reach values $F = 10^{-6}$, $\beta = 10^{-1}$ on the amplifier gain G. Just as in the preceding calculations, d = 0.1, $q = 2 \cdot 10^{-5}$. It can be seen that it is meaningless to increase the gain above 10^5. Thus, for practical purposes the following expression is recommended for the optimum gain:

$$G = 2/q.$$

It is interesting that the relation obtained for the gain is independent of L and can thus be used for laser amplifiers of arbitrary design.

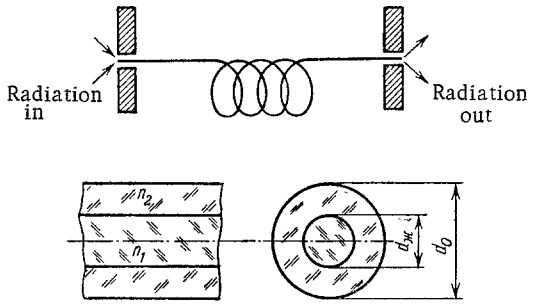


Fig. 3.9

TABLE 4. Principal Characteristics of Neodymium Glasses

Characteristic	Glass brand			
	KGSS-3	KGSS-11	GLS-2	LGS-36
Refractive index	1,5400	1,5485	1,5280	1,5780
Neodymium ion density, 10^{20} cm^{-3}	1,0	6,0	1,9	2,0
Density, g/cm^3	2,92	2,97	2,70	2,78
Duration of excited state, msec	0,60	0,45	0,55	0,45

6. Concrete Types of Laser Amplifiers. We consider now actual amplifying media and actual types of laser amplifiers. We shall consider only media designed for amplification of optical signals at infrared wavelengths, since the use of such amplifiers is not advantageous for visible light by virtue of the sufficient quantum efficiency of the available photodetectors.

One of the active media with the largest gains is the He—Xe gas mixture. The resonant wavelength of this mixture is $\lambda = 3.508$ μm. A traveling-wave He—Xe laser amplifier 0.5 m long has an amplification coefficient 10^3 m^{-1} at the center of the gain band. This corresponds to a gain $\alpha = 0.138$ cm^{-1}. However, despite this high gain, He—Xe amplifiers are not widely used in laser radars for two reasons.

First, they have too narrow a gain band (about 100 MHz). This raises difficulties when amplifying broadband signals and signals that are Doppler-shifted in Frequency.

Second, the existing lasers, which operate at 3.508 μm, are of low power and do not ensure a sufficiently large range.

Highly promising is CO_2 mixed with N_2, He, or H_2 as a buffer gas. The luminescence spectrum of such a mixture consists of several gain bands, each ≈50 MHz wide at a pressure less than 1 torr. These bands are equidistant in the wavelength range 10.51-10.67 μm. With increasing pressure in the mixture, the gain band broadens. The ratio of the increment of the gain bandwidth to the pressure increment is 5 MHz/torr. Thus, to obtain a gain band with a width of about 10^9 Hz ($\Delta\lambda = 0.3$ nm at $\lambda = 10.6$ μm) the pressure mixture must be about 200 torr. At such pressures it is impossible to produce an autonomous discharge in large volumes, and a nonautonomous discharge must be resorted to [12] . The ensuing technical difficulties limit the use of CO_2 amplifiers in receiving devices. The gain of a CO_2 amplifier varies with the regime and ranges from $5\cdot10^{-3}$ cm^{-1} for an autonomous discharge to $5\cdot10^{-2}$ cm^{-1} for a nonautonomous discharge pumped by an electron beam.

Among the solid media, the most important material for laser amplifiers is glass activated with Nd^{3+} ions (neodymium glass). The principal characteristics of the glasses employed are listed in Table 4. The gain bandwidth in neodymium-glass is of the order of 10^{11} Hz ($\Delta\lambda = 0.37$ nm at a wavelength $\lambda = 1.06$ μm), which is sufficient for the reception of radar signals.

TABLE 5. Principal Parameters of Existing Active Fiber-Optics Waveguides

Material		Length, mm	Core diameter, μm	Cutoff parameter	Number of modes	Level angle of one fiber end face, deg	Pump energy, J/mode	Pump energy, Gain, J	Gain	
core part	sheath								maximum	minimum
KGSS-3	K-15	1500	4,1	1,52	1	12	2000 * / 1000	2000 * / 1000	8000	350
GLS-2	KGSS-11	1500	4,7	3,52	4	12	280 / 180	1120 / 720	800	65
LGS-36	KGSS-7	1500	6,4	5,76	72	16	70 / 60	840 / 720	7000	900
LGS-36	K-15	1500	8,3	9,15	26	16	49 / 22	1280 / 600	500	150
GLS-2	K-8	1500	5,5	3,46	4	12	180 / 125	720 / 500	300	50
Garnet	Quartz glass	1500	22	79	1410	20	2 / 1,7	2800 / 2400	1400	300

*The upper and lower numbers are the maximum and average gains, respectively.

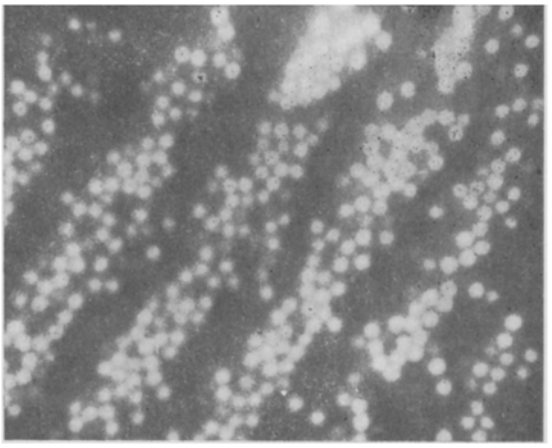


Fig. 3.10

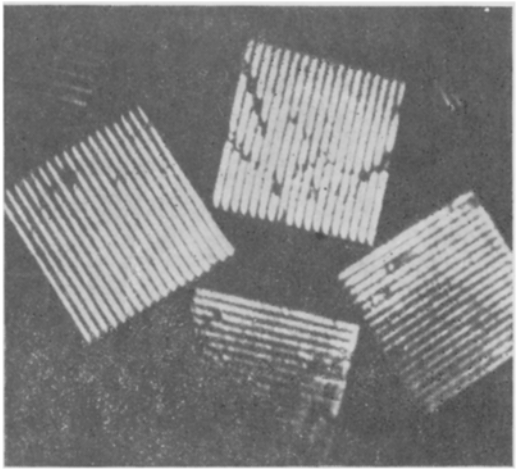


Fig. 3.11

One of the most promising laser amplifiers, from the point of view of use in receivers, is a fiber-optics amplifier. Such a laser amplifier constitutes an optic fiber waveguide (Fig. 3.9) in which the core is glass activated with Nd^{3+}. Optical pumping produces inverted population in the core of this waveguide. The light wave propagates mainly in the core of the waveguide and is therefore effectively amplified there.

Fiber-optics laser amplifiers offer a number of advantages. First, they yield a large stable gain along with a large bandwidth. Second, fiber-optics amplifiers have relatively large input angular apertures at small dimensions. Third, the large ratio of the surface area to the volume improves the cooling when operating at high pulse repetition frequency.

The most advantageous is an operating regime in which only one transverse mode can propagate along the waveguide. This is called the single-mode regime [13]. For the single-mode regime to take place, the following condition must be satisfied (see Fig. 3.9):

$$R = \pi d_c (n_1^2 - n_2^2)^{1/2} \lambda^{-1} \leqslant 2{,}405. \qquad (3.78)$$

The parameter R is called the cutoff parameter. For example, if the core part of the fiber has a diameter d_c = 5.5 μm and is made of KGSS-3 glass with refractive index n_1 = 1.5400, while the sheath is made of K-15 glass and has a refractive index n_2 = 1.5335, then R = 2.304, i.e., condition (3.78) is satisfied.

To obtain the maximum gain it is necessary to eliminate the feedback produced by reflection from the end faces of the fiber. A simple and effective method for eliminating feedback is to polish one of the end surfaces at a definite angle θ. The lower limit of this angle is determined by the condition that the limiting meridional ray not experience total internal reflection from this end face. These conditions take the form:

$$\arccos\frac{n_2}{n_1}<\theta<\arcsin\frac{1}{n_1}-\arccos\frac{n_2}{n_1}.$$

As a rule, the angle θ is 10-20°.

The parameters of the existing active fiber waveguides are listed in Table 5.

Finally, mention must also be made of the possibility of producing cables of active fibers. If the input and output of the cable are matched, it is possible to transmit over such a fiber cable an image and amplify simultaneously its brightness. Experiments show that a fiber-optics laser image amplifier can increase the signal/noise ratio in the recorded image [14].

Figure 3.10 shows a photograph of the output end face of a fiber image amplifier where the image of a target with alternating black and white stripes is amplified. The intensity gain is about 10^3. A certain irregularity of the image is due to the nonuniformity of the optical pumping of the fibers and to cross-coupling of the fibers. The dark dips in places where a white stripe should be are attributable to nonuniform pumping. White spots in dark intervals are the consequence of interfiber coupling.

Much attention is paid of late to amplifiers in which the active medium is based on dye solutions. Such amplifiers have very high gains, on the order of several reciprocal centimeters. This permits a reduction of the working length of the amplifier, which leads, in turn, to low losses and small distortions in image amplification. For example, an amplifier based on a solution of rhodamine 6Zh in ethanol has high gain in the yellow-green region of the spectrum. Figure 3.11 shows selected targets amplified in such an amplifier. The gain is $2\cdot10^3$ when pumped with a neodymium laser with frequency doubling and with a power density 2 MW/cm^2. The maximum resolving power is 120 lines/mm.

3.2. Parametric Converters of Radiation Frequency

1. Principles of Parametric Conversion of Radiation Frequency. A promising method of increasing the efficiency of laser-signal reception is parametric conversion of the frequency in a nonlinear crystal. The gist of this method is that the field of the reflected radar signal interacts in a nonlinear crystal with the pump field, as a result of which the energy of the signal field is transformed into the energy of another field whose frequency is equal to the sum of the frequencies of the original and of the pump fields. This permits the energy of an infrared field to be converted into energy of a visible-light field and be recorded with high-sensitivity receivers such as photomultipliers.

The first parametric conversion in a quartz crystal was implemented by Franken and coworkers [15] in 1961. They generated in their experiments the second harmonic of laser radiation. Parametric conversion can be particularly helpful when it comes to recording the image of an object in the infrared band. By parametric frequency conversion it is possible to transform an image with sufficiently high resolution. The first to transfer an image from the infrared to the visible region of the spectrum was Midwinter [16].

Let us consider the basic principles of field interaction in a nonlinear medium. Maxwell's equations are of the form

$$\begin{aligned}&\operatorname{rot}\boldsymbol{H}=\boldsymbol{J}+\frac{\partial\boldsymbol{D}}{\partial t},\quad \operatorname{rot}\boldsymbol{E}=-\frac{\partial\boldsymbol{B}}{\partial t},\\&\operatorname{div}\boldsymbol{D}=0,\qquad \operatorname{div}\boldsymbol{B}=0,\end{aligned}\tag{3.79}$$

where

$$\boldsymbol{D}=\varepsilon_0\boldsymbol{E}+\boldsymbol{P},\quad \boldsymbol{B}=\mu_0\boldsymbol{H},\quad \boldsymbol{J}=\sigma\boldsymbol{E}.$$

The polarization $\mathbf{P}(\mathbf{r}, t)$ of the medium can be divied into two parts: the polarization $\mathbf{P}^{\mathrm{L}}(\mathbf{r}, t)$ linear in the field and the polarization $\mathbf{P}^{\mathrm{NL}}(\mathbf{r}, t)$ nonlinear (quadratic) in the field. In the most general form, these components can be written in the following form (for each i-th component):

$$P_i^{\mathrm{L}}(\boldsymbol{r}, t) = \varepsilon_0 \int_0^{\infty} dt_1 \int d^2 r_1 \varkappa_{ij}(t_1, \boldsymbol{r}_1) E_j(t - t_1, \boldsymbol{r} - \boldsymbol{r}_1), \tag{3.80}$$

$$P_i^{\mathrm{NL}}(\boldsymbol{r}, t) = \varepsilon_0 \int_0^{\infty} dt_1 dt_2 \int d^2 r_1 d^2 r_2 \chi_{ijk}(t_1, t_2, \boldsymbol{r}_1, \boldsymbol{r}_2)\ E_j(t - t_1, \boldsymbol{r} - \boldsymbol{r}_1) E_k(t - t_1 - t_2, \boldsymbol{r} - \boldsymbol{r}_1 - \boldsymbol{r}_2), \tag{3.81}$$

where $\varkappa_{ij}$ and χ_{ijk} are the tensors of the linear and nonlinear (quadratic) dielectric susceptibility.

Transforming to the frequency representation, we can express the linear component of the polarization in the form

$$P_i^{\mathrm{L}}(\boldsymbol{r}, t) = \varepsilon_0 \int_{-\infty}^{+\infty} d\omega \int \varkappa_{ij}(\omega, \boldsymbol{k}) E_j(\omega, \boldsymbol{k}) \exp[i(\omega t - \boldsymbol{k}\boldsymbol{r})] d^2 k, \tag{3.82}$$

where

$$E_j(\omega, \boldsymbol{k}) = (2\pi)^{-3} \int E_j(\boldsymbol{r}, t) \exp[-i(\omega t - \boldsymbol{k}\boldsymbol{r})] dt\, d^2 r, \tag{3.83}$$

$$\varkappa_{ij}(\omega, \boldsymbol{k}) = (2\pi)^{-3} \int \varkappa_{ij}(t, \boldsymbol{r}) \exp[-i(\omega t - \boldsymbol{k}\boldsymbol{r})] dt\, d^2 r. \tag{3.84}$$

In most cases, the spatial dispersion can be neglected; then

$$P_i^{\mathrm{L}}(\boldsymbol{r}, t) = \varepsilon_0 \int_{-\infty}^{+\infty} \varkappa_{ij}(\omega) E_j(\omega, \boldsymbol{r}) \exp(i\omega t)\, d\omega. \tag{3.85}$$

The relaxation time of the linear dielectric susceptibility is short (on the order of 10^{-15} sec). Considering quasimonochromatic signals, one can therefore neglect in (3.85) the temporal dispersion of the dielectric susceptibility and write

$$P_i^{\mathrm{L}}(\boldsymbol{r}, t) = \varepsilon_0 \varkappa_{ij}(\omega) E_j(\boldsymbol{r}, t) \exp[i(\omega t - \boldsymbol{k}\boldsymbol{r})], \tag{3.86}$$

where $E_j(\mathbf{r}, t)$ is a function of the time and of the coordinates and varies slowly compared with ω^{-1} and k^{-1}.

In practice, however, one encounters frequently cases when the approximation (3.86) cannot be used. This occurs when the field correlation time $\tau \lesssim 10^{-12}$ sec. For a wavelength $\lambda = 1.06$ μm, this correlation time corresponds to a spectrum width $\Delta\lambda = 3$ nm. A comparable spectrum width is possessed by certain solid-state pulsed lasers. In laser radar, however, the radiation used is appreciably more monochromatic, and we shall therefore use hereafter the approximation (3.86), which yields simple analytic results.

For the nonlinear component of the polarization we have, accordingly,

$$P_i^{\mathrm{NL}}(\boldsymbol{r}, t) = \int_{-\infty}^{+\infty} d\omega \int P_i^{\mathrm{NL}}(\omega, \boldsymbol{k}) \exp[i(\omega t - \boldsymbol{k}\boldsymbol{r})] d^2 k, \tag{3.87}$$

where

$$P_i^{\mathrm{NL}}(\omega, \boldsymbol{k}) = \varepsilon_0 \int_{-\infty}^{+\infty} d\omega_1 \int \chi_{ijk}(\omega, \boldsymbol{k}, \omega_1, \boldsymbol{k}_1) E_j(\omega - \omega_1, \boldsymbol{k} - \boldsymbol{k}_1) E_k(\omega_1, \boldsymbol{k}_1) d^2 k_1, \tag{3.88}$$

$$\chi_{ijk}(\omega, \boldsymbol{k}, \omega_1, \boldsymbol{k}_1) = \int \chi_{ijk}(t_1, t_2, \boldsymbol{r}_1, \boldsymbol{r}_2) \exp[-i(\omega t_1 + \omega_1 t_2 - \boldsymbol{k}\boldsymbol{r}_1 - \boldsymbol{k}_1 \boldsymbol{r}_2)] dt_1 dt_2 d^2 r_1 d^2 r_2. \tag{3.89}$$

Relations (3.87) and (3.88) show that the nonlinearity described by the tensor χ_{ijk} leads to interaction of three fields with frequencies ω_1, ω_2, and ω_3, and the relations between the frequencies and the wave vectors of these fields can be written in the form

$$\omega_3 = \omega_1 + \omega_2, \qquad \boldsymbol{k}_3 = \boldsymbol{k}_1 + \boldsymbol{k}_2. \tag{3.90}$$

These interactions were named three-frequency interactions. We shall consider hereafter only such interactions.

The nonlinear part of the polarization is smaller by several orders of magnitude than the linear. Therefore, allowance for the spatial and temporal dispersions of the tensor χ_{ijk} would lead to terms of higher orders of smallness, which are inessential for our analysis. We can therefore neglect the spatial and temporal dispersions of the tensor χ_{ijk} and write the nonlinear component of the polarization in a quasimonochromatic approximation:

$$P_{3i}^{\mathrm{NL}}(\boldsymbol{r}, t)=\varepsilon_0\chi_{ijk}(\omega_3=\omega_1+\omega_2)E_{1j}(\boldsymbol{r}, t)E_{2k}(\boldsymbol{r}, t)\exp[i(\omega_1+\omega_2)t-i(\boldsymbol{k}_1+\boldsymbol{k}_2)\boldsymbol{r}], \tag{3.91}$$

where $\mathbf{E}_1(\mathbf{r}, t)$ and $\mathbf{E}_2(\mathbf{r}, t)$ are slowly varying amplitudes pertaining to fields with frequencies ω_1 and ω_2, respectively; $\mathrm{P}^{\mathrm{NL}}{}_3(\mathbf{r}, t)$ is the nonlinear component of the polarization and its frequency is ω_3; the symbol $\chi_{ijk}(\omega_3 = \omega_1 + \omega_2)$ denotes the tensor of the nonlinear dielectric susceptibility that pertains to an interaction of form (3.90).

For interaction of the type

$$\omega_1=\omega_3-\omega_2, \quad \omega_2=\omega_3-\omega_1 \tag{3.92}$$

the corresponding nonlinear polarization components are

$$P_{1i}^{\mathrm{NL}}(\boldsymbol{r}, t)=\varepsilon_0\chi_{ijk}(\omega_1=\omega_3-\omega_2)E_{3j}(\boldsymbol{r}, t)E_{2k}^*(\boldsymbol{r}, t)\exp[i(\omega_3-\omega_2)t-i(\boldsymbol{k}_3-\boldsymbol{k}_2)\boldsymbol{r}], \tag{3.93}$$

$$P_{2i}^{\mathrm{NL}}(\boldsymbol{r}, t)=\varepsilon_0\chi_{ijk}(\omega_2=\omega_3-\omega_1)E_{3j}(\boldsymbol{r}, t)E_{1k}^*(\boldsymbol{r}, t)\exp[i(\omega_3-\omega_1)t-i(\boldsymbol{k}_3-\boldsymbol{k}_1)\boldsymbol{r}]. \tag{3.94}$$

We now turn to Maxwell's equations (3.79). Taking the curl of both sides of the second equation of system (3.79) and substituting the expression for rot H from the first equation of the same system, we get

$$\nabla^2\boldsymbol{E}=\mu_0\sigma\frac{\partial\boldsymbol{E}}{\partial t}+\mu_0\varepsilon_0\frac{\partial^2\boldsymbol{E}}{\partial t^2}+\mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{L}}}{\partial t^2}+\mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}, \tag{3.95}$$

where we have used the relations rot rot $\mathbf{E}$ = ∇ div $\mathbf{E} - \nabla^2\mathbf{E}$, div $\mathbf{E}$ = 0. Substituting in the obtained equation expression (3.86) for linear polarization, we have

$$\nabla^2\boldsymbol{E}=\mu_0\sigma\frac{\partial\boldsymbol{E}}{\partial t}+\mu_0\varepsilon_0(1+\varkappa)\frac{\partial^2\boldsymbol{E}}{\partial t^2}+\mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}, \tag{3.96}$$

where $\varkappa$ is the tensor of the linear dielectric susceptibility.

We now make the following remark. In practical realization of parametric frequency conversion one usually chooses the polarizations of the signal field and of the pump field in such a way that they correspond to to waves of the ordinary (o) type (see [17] for more details). Yet the transformed wave turns out to be extraordinary (e). Such an interaction ensures satisfaction of the condition of spatial locking [18] and is designated as an oo—e interaction.

Equation (3.96) is valid for each of the interacting waves. In the general case the obtained system of equations turns out to consist of equations of type (3.96), which are related not only through the nonlinear polarization $\mathbf{P}^{\mathrm{NL}}$, but also through the tensor of the linear dielectric susceptibility κ. Waves of type o and e are two types of waves that can propagate in a crystal independently of each other. Therefore, if o (or e) waves are taken to be the interacting ones, then the coupling of the equations via κ is eliminated and the entire system as a whole is simplified. In this case, each o (or e) wave satisfies the similar equation

$$\nabla^2\boldsymbol{E}=\mu_0\sigma\frac{\partial\boldsymbol{E}}{\partial t}+\mu_0\varepsilon_0\varepsilon\frac{\partial^2\boldsymbol{E}}{\partial t^2}+\mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}, \tag{3.97}$$

where $\varepsilon = 1 + \kappa$ is the dielectric constant of the crystal for a wave of type o (or e).

We express the field $\mathbf{E}$ in the form

$$\boldsymbol{E}=\boldsymbol{e}A(\boldsymbol{r}, t)\exp(i\omega t-i\boldsymbol{k}\boldsymbol{r}), \tag{3.98}$$

where $\mathbf{e}$ is the field-polarization vector. Waves with frequencies ω_1, ω_2, and ω_3 will be labeled with the appropriate subscripts 1, 2, and 3. Substituting (3.98) in (3.97) and confining ourselves to the first-order approximation, we obtain

$$\begin{aligned}
\boldsymbol{e}_1 2i\boldsymbol{k}_1 \nabla A_1 &= -\mu_0\sigma_1\omega_1\boldsymbol{e}_1 A_1 + \mu_0\varepsilon_0\boldsymbol{\chi}\boldsymbol{e}_3\boldsymbol{e}_2 A_3 A_2^*\omega_1^2 \exp(-i\boldsymbol{\Delta r}),\\
\boldsymbol{e}_2 2i\boldsymbol{k}_2 \nabla A_2 &= -\mu_0\sigma_2\omega_2\boldsymbol{e}_2 A_2 + \mu_0\varepsilon_0\boldsymbol{\chi}\boldsymbol{e}_3\boldsymbol{e}_1 A_3 A_1^*\omega_2^2 \exp(-i\boldsymbol{\Delta r}),\\
\boldsymbol{e}_3 2i\boldsymbol{k}_3 \nabla A_3 &= -\mu_0\sigma_3\omega_3\boldsymbol{e}_3 A_3 + \mu_0\varepsilon_0\boldsymbol{\chi}\boldsymbol{e}_1\boldsymbol{e}_2 A_1 A_2\omega_3^2 \exp(i\boldsymbol{\Delta r}),
\end{aligned} \tag{3.99}$$

where χ is the tensor of the nonlinear dielectric susceptibility, and for the detuning vector we introduce the notation

$$\boldsymbol{\Delta} = \boldsymbol{k}_3 - \boldsymbol{k}_2 - \boldsymbol{k}_1. \tag{3.100}$$

Neglecting the anisotropic shift of the e wave, and also assuming that all the waves propagate approximately along the z axis, we obtain

$$\boldsymbol{k}\nabla A \approx k\frac{\partial A}{\partial z}. \tag{3.101}$$

With this remark taken into account, system (3.99) is ultimately reduced to the form

$$\begin{aligned}
\frac{\partial A_1}{\partial z} &= -\alpha_1 A_1 - i\beta_1 A_3 A_2^* \exp(-i\boldsymbol{\Delta r}),\\
\frac{\partial A_2}{\partial z} &= -\alpha_2 A_2 - i\beta_2 A_3 A_1^* \exp(-i\boldsymbol{\Delta r}),\\
\frac{\partial A_3}{\partial z} &= -\alpha_3 A_3 - i\beta_3 A_1 A_2 \exp(i\boldsymbol{\Delta r}),
\end{aligned} \tag{3.102}$$

where (j = 1, 2, 3)

$$\begin{aligned}
&\alpha_j = \frac{1}{2}\sigma_j\left(\frac{\mu_0}{\varepsilon_0\varepsilon_j}\right)^{1/2}, \qquad \beta_1 = \frac{1}{2}\omega_1\boldsymbol{e}_1\boldsymbol{\chi}\boldsymbol{e}_2\boldsymbol{e}_3\left(\frac{\mu_0\varepsilon_0}{\varepsilon_1}\right)^{1/2},\\
&\beta_2 = \frac{1}{2}\omega_2\boldsymbol{e}_2\boldsymbol{\chi}\boldsymbol{e}_3\boldsymbol{e}_1\left(\frac{\mu_0\varepsilon_0}{\varepsilon_2}\right)^{1/2}, \quad \beta_3 = \frac{1}{2}\omega_3\boldsymbol{e}_3\boldsymbol{\chi}\boldsymbol{e}_1\boldsymbol{e}_2\left(\frac{\mu_0\varepsilon_0}{\varepsilon_3}\right)^{1/2}.
\end{aligned} \tag{3.103}$$

We continue the analysis under the assumption that the pump field A_2 changes insignificantly during the interaction. This assumption is justified in the case when the input field of the signal A_1 is much weaker than the pump field A_2. In this case, the field generated at the summary frequency A_3 is also considerably weaker than the pump field. Furthermore, to simplify the subsequent calculation, we set the parameters α_1, α_2, and α_3, which describe the absorption of the radiation inside the crystal, equal to zero. Under the assumptions made we get

$$\begin{aligned}
\frac{\partial A_1}{\partial z} &= -i\beta_1 A_3 A_2^* \exp(-i\boldsymbol{\Delta r}),\\
\frac{\partial A_3}{\partial z} &= -i\beta_3 A_1 A_2 \exp(i\boldsymbol{\Delta r}).
\end{aligned} \tag{3.104}$$

Finally, we make the last assumption:

$$\boldsymbol{\Delta r} = \Delta z, \tag{3.105}$$

i.e., we assume that the detuning vector $\boldsymbol{\Delta}$ is directed along the z axis. Then, differentiating the first equation of system (3.104) with respect to z and substituting the value of $\partial A_3/\partial z$ from the second equation, we obtain a differential equation for A_3:

$$\frac{\partial^2 A_3}{\partial z^2} - i\Delta\frac{\partial A_3}{\partial z} + \gamma A_3 = 0, \quad \gamma = \beta_1\beta_3|A_2|^2. \tag{3.106}$$

Under the initial conditions

$$A_1(z)_{z=0} = A_1(0), \quad A_3(z)_{z=0} = 0 \tag{3.107}$$

the solution of (3.106) takes the form

$$A_3(z) = -i\beta_3 A_1(0) A_2 \exp\left(i\frac{1}{2}\Delta z\right)\left(\frac{\Delta^2}{4} + \gamma\right)^{-1/2}\sin\left(\sqrt{\frac{\Delta^2}{4} + \gamma}\, z\right). \tag{3.108}$$

We shall be interested further in the power $P_3(z) = |A_3(z)|^2$ of the converted radiation. From (3.108) we obtain

$$P_3(z)=\frac{\beta_3}{\beta_1}|A_1(0)|^2\gamma\left(\frac{\Delta^2}{4}+\gamma\right)^{-1}\sin^2\left(\sqrt{\frac{\Delta^2}{4}+\gamma}\,z\right). \tag{3.109}$$

As a rule, the conversion coefficient, defined by the dimensionless parameter $\eta = \gamma z^2$, is much smaller than unity. Taking this into account, we can rewrite (3.109) in a more lucid form:

$$P_3(z)=\frac{\beta_3}{\beta_1}|A_1(0)|^2\gamma z^2\left[\frac{\sin(\Delta z/2)}{\Delta z/2}\right]^2. \tag{3.110}$$

The function in the square brackets in the right-hand side of this expression is well known (Fig. 3.12). It reaches a maximum value, unity, at

$$\Delta=0. \tag{3.111}$$

This condition is called the spatial-locking condition.

If the spatial-locking condition is satisfied, the power of the converted radiation at the combined frequency is

$$P_3(z)=\frac{\beta_3}{\beta_1}P_1(0)\gamma z^2, \tag{3.112}$$

where $P_1(0)$ is the power of the signal radiation at the frequency ω_1. In this case, the conversion coefficient

$$\eta=\frac{P_3(z)}{P_1(0)}=\frac{\beta_3}{\beta_1}\gamma z^2 \tag{3.113}$$

is proportional to the square of the nonlinear-interaction region. It follows from (3.110), however, that with increasing z the conditions on the permissible detuning Δ become more stringent.

Spatial synchronization is an important factor in practical realization of the method of parametric frequency conversion. We shall not discuss this question, however, since it pertains mainly to crystal optics [19].

Concluding the analysis of the general principles of parametric frequency conversion, notice must be taken of the following. Equation (3.113) is an expression for the power conversion coefficient. In some cases it is convenient to introduce the photon-number conversion coefficient:

$$\eta'=\frac{N_3}{N_1}, \tag{3.114}$$

where N_3 is the number of photons produced per unit time at the frequency ω_3, and N_1 is the number of photons of frequency ω_1 entering the converter per unit time. In accord with (3.113) we obtain

$$\eta'=\frac{e_3\chi e_1 e_2}{e_1\chi e_2 e_3}\left(\frac{\varepsilon_1}{\varepsilon_3}\right)^{1/2}\gamma z^2. \tag{3.115}$$

This expression does not depend explicitly on the frequencies ω_1 and ω_3. In conversion to higher frequency, the photon-number conversion coefficient is always smaller than the power conversion coefficient.

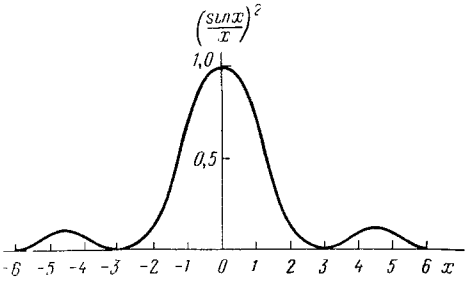


Fig. 3.12

TABLE 6. Experimental Values of the Power Conversion Coefficients for Certain Nonlinear Crystals

Crystal	Conversion coeff.	Pump power density, MW/cm²	Reduced conversion coeff., 10^{-9} cm²/W
RDP	0,089	27,5	3,2
RDA	0,066	16,4	4,0
DKDP	0,082	20	4,1
$LiIO_3$	0,079	7,0	12

Table 6 lists the experimental power conversion coefficients obtained for several nonlinear crystals in wave conversion from $\lambda_1 = 1.15$ μm to $\lambda_3 = 0.43$ μm at a pump wavelength $\lambda_2 = 0.69$ μm [20]. Also given are the values of the reduced power conversion coefficient and of the pump-power density.

2. Image Conversion in Nonlinear Crystals. The parametric frequency conversion method makes possible conversion of not only single-mode radiation, but also conversion of images from the infrared into the visible region of the spectrum. Let us examine the theory of such a conversion on the basis of the results of [21].

Figure 3.13 shows a schematic model of parametric image conversion. The transformed image at the frequency ω_1 is shaped by some optical system in the z_1 plane. A plane wave of frequency ω_2 propagates along the z axis in a nonlinear crystal. Interaction of the signal wave at frequency ω_1 with the pump wave of frequency ω_2 generates a converted wave of frequency ω_3. It will be shown below that this wave shapes a converted image in some plane z_2.

We assume, for the sake of argument, that the nonlinear anisotropic crystal of length z_0 and width d is uniaxial, while the pump wave and the converted-signal waves are ordinary (o) and the converted wave is extraordinary (e). The arrangement of the axes of the coordinate system is shown in Fig. 3.13. We assume that the principal optical axis of the crystal lies in the xz plane (Fig. 3.14).

We agree to assign to the field of frequency ω_j the subscript j. Then the spatial spectrum $E_1(\mathbf{k}_1, z_1)$ of the field of the signal image $E_1(\mathbf{r}, z_1)$ in the z_1 plane can be represented in the form

$$E_1(k_{1x}, k_{1y}, z_1) = \frac{1}{2\pi}\int E_1(x, y, z_1)\exp(ik_{1x}x + ik_{1y}y)\,dx\,dy, \tag{3.116}$$

where the integration is over the entire z_1 plane.

We introduce formally the transfer function $K(k_x, k_y)$ of the nonlinear crystal, defined as follows:

$$E_3(k_{3x}, k_{3y}, z_2) = K(k_{1x}, k_{1y})E_1(k_{1x}, k_{1y}, z_1), \tag{3.117}$$

where $E_3(k_{3x}, k_{3y}, z_2)$ is the spatial spectrum of the converted field at the frequency ω_3 in the z_2 plane.

The wave vectors $\mathbf{k}_1$ and $\mathbf{k}_3$ and the pump wave vector $\mathbf{k}_2$ should satisfy the spatial synchronism condition

$$\Delta = \boldsymbol{k}_3 - \boldsymbol{k}_2 - \boldsymbol{k}_1 = 0. \tag{3.118}$$

Exact satisfaction of this condition is possible only for definite values of the components k_{1x}, k_{1y}, k_{3x}, and k_{3y}. Therefore, at exact synchronism only one plane wave can be converted, i.e., only one resolution element. The conversion of other plane waves making up the image takes place at a different value of the detuning Δ and, consequently, the efficiency of conversion of these components is lower.

Thus, the nonlinear crystal acts as a filter for the spatial frequencies. The larger the passband of this filter, i.e., the less critical the condition (3.118) to changes of the components k_{1x} and k_{1y}, the more resolution elements can be converted by the crystal. Low

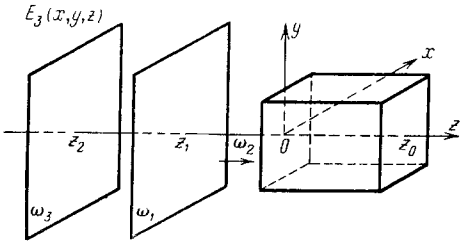


Fig. 3.13

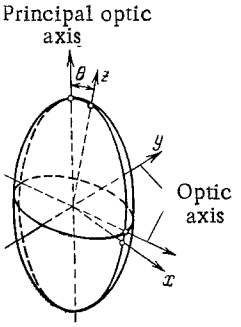


Fig. 3.14

criticality of the synchronism condition is reached when the signal-radiation wave vector makes a small angle with the direction of the pump wave vector upon satisfaction of the exact synchronism condition. Under these conditions we can write in first-order approximation (j = 1, 2, 3)

$$\Delta = k_{3z} - k_{2z} - k_{1z}, \quad k_{jz} = f_j(k_{jx}, k_{jy}), \tag{3.119}$$

$$k_{3x} = k_{1x} + k_{2x}, \quad k_{3y} = k_{1y} + k_{2y}. \tag{3.120}$$

The function f_j in (3.119) is obtained by expanding the dependences of k_{jz} on k_{jx} and k_{jy} in terms of small parameters, namely the ratios k_{jx}/k_j, k_{jy}/k_y. Allowance for the detuning with respect to the x and y components of the wave vectors would lead to terms of next order of smallness in the parameters k_{jx}/k_j and k_{jy}/k_j. Exact synchronism with respect to the x and y components of the wave vectors was therefore assumed in Eqs. (3.120).

To find the spatial distribution of the field $E_3(x, y, z_2)$ in the z_2 plane we must take the inverse Fourier transform of (3.117):

$$E_3(x, y, z_2) = \frac{1}{2\pi} \exp(-ik_{2x}x - ik_{2y}y) \int E_1(k_{1x}, k_{1y}, z_1) K(k_{1x}, k_{1y}) \exp(-ik_{1x}x - ik_{1y}y)\, dk_{1x} dk_{1y}. \tag{3.121}$$

In the derivation of this expression we have used the equations in (3.120). Next, using the convolution theorem and (3.116), we obtain

$$E_3(x, y, z_2) = \exp(-ik_{2x}x - ik_{2y}y) \int E_1(x_1, y_1, z_1)\, \Gamma(x - x_1, y - y_1, z_1, z_2)\, dx_1 dy_1, \tag{3.122}$$

where

$$\Gamma(x, y, z_1, z_2) = \frac{1}{(2\pi)^2} \int K(k_{1x}, k_{1y}) \exp(-ik_{1x}x - ik_{1y}y)\, dk_{1x}\, dk_{1y}. \tag{3.123}$$

The function $\Gamma(x, y, z_1, z_2)$ is thus the scattering function that determines the resolving power of the parametric conversion of the image. The function $\Gamma(x, y, z_1, z_2)$ provides a complete description of parametric frequency conversion in the image conversion regime.

To find the explicit form of the scattering function (3.123) we must solve the truncated equations (3.104) written out for each plane wave component of the image. The truncated equations (3.104) were written for the amplitudes of the plane waves $A_1(k_{1x}, k_{1y}, z)$, $A_3(k_{3x}, k_{3y}, z)$, defined by the relations

$$\begin{aligned} E_1(k_{1x}, k_{1y}, z) &= A_1(k_{1x}, k_{1y}, z) \exp(-ik_{1x}x - ik_{1y}y - ik_{1z}z), \\ E_3(k_{3x}, k_{3y}, z) &= A_3(k_{3x}, k_{3y}, z) \exp(-ik_{3x}x - ik_{3y}y - ik_{3z}z). \end{aligned} \tag{3.124}$$

It is necessary here to satisfy each of the equalities

$$k_{1x}^2 + k_{1y}^2 + k_{1z}^2 = k_1^2, \quad k_{3x}^2 + k_{3y}^2 + k_{3z}^2 = k_3^2, \tag{3.125}$$

where k_1 and k_3 are the moduli of the wave vectors inside the crystal. In accord with the previously obtained solution (3.108), we have

$$A_3(k_{3x}, k_{3y}, z_0) = -i\beta_3 A_1(k_{1x}, k_{1y}, 0) A_2(k_{2x}, k_{2y}) \exp\left(i\frac{1}{2}\Delta z_0\right)\left(\frac{\Delta^2}{4} + \gamma\right)^{-1/2} \sin\left(\sqrt{\frac{\Delta^2}{4} + \gamma}\, z_0\right). \tag{3.126}$$

Let a plane wave of energy E_2 propagate along the z axis. Then $k_{2x} = k_{2y} = 0$ and $A_2(k_{2x}, k_{2y}) = A_2$. It is now necessary to express the amplitudes $A_3(k_{3x}, k_{3y}, z_0)$ and $A_1 \cdot (k_{1x}, k_{1y}, 0)$ in terms of the corresponding values of the field E_3 and E_1 and recalculate the obtained expressions in the planes z_2 and z_1, respectively. These operations are performed with the aid of relations (3.124) and (3.125). It must only be borne in mind here that when the fields are recalculated to the planes z_2 and z_1 it is necessary to substitute the values of the wave vectors outside the crystal, k_{30} and k_{10}, respectively, while the values of the components k_{jx} and k_{jy} do not change on going through the crystal–air interface. As a result, we obtain

$$E_3(k_{3x}, k_{3y}, z_2) = -i\beta_3 A_2\left(\frac{\Delta^2}{4}+\gamma\right)^{-1/2}\sin\left(\sqrt{\frac{\Delta^2}{4}+\gamma}\,z_0\right)\times$$
$$\times\exp\left[i\left(\frac{1}{2}\Delta - k_{3z}\right)z_0\right]\exp\left[-i\frac{1}{2}(k_{1x}^2+k_{1y}^2)\left(\frac{z_1}{k_{10}}-\frac{z_2-z_0}{k_{30}}\right)\right]E_1(k_{1x}, k_{1y}, z_1). \tag{3.127}$$

In the derivation of this expression we used the smallness of the parameters k_{jx}/k_j and k_{jy}/k_j ($j = 1, 3$). An inessential phase factor independent of k_{1x} and k_{1y} was omitted. The phase factor $\exp(-ik_{3z}z_0)$ describes the propagation of the converted wave inside the crystal.

In accordance with the definition of the transfer function $K(k_x, k_y)$, we obtain from (3.117) and (3.127)

$$K(k_x, k_y) = -i\beta_3 A_2\left(\frac{\Delta^2}{4}+\gamma\right)^{-1/2}\sin\left(\sqrt{\frac{\Delta^2}{4}+\gamma}\,z_0\right)\exp\left[i\left(\frac{1}{2}\Delta - k_{3z}\right)z_0 - i\frac{1}{2}(k_x^2+k_y^2)\left(\frac{z_1}{k_{10}}-\frac{z_2-z_0}{k_{30}}\right)\right]. \tag{3.128}$$

The detuning Δ that enters in this expression depends on k_x and k_y. To find this dependence, k_z must be expressed in terms of k_x and k_y. As shown in Appendix IV, all the interacting waves satisfy the inequality

$$k_z = k - \operatorname{tg}\alpha\, k_x - \frac{1+\mu}{2k}k_x^2 - \frac{1+\nu}{2k}k_y^2, \tag{3.129}$$

where α is the anisotropy angle of the crystal for a given direction, k is the modulus of the wave vector in the crystal in the direction of the z axis, and the parameters μ and ν are defined by the relations

$$1+\mu = \frac{1+\varepsilon\cos^2\theta}{1+\varepsilon\sin^2\theta} - 4\operatorname{tg}^2\alpha, \quad 1+\nu = \frac{1+\varepsilon}{1+\varepsilon\sin^2\theta},$$
$$1+\varepsilon = \frac{n_o^2}{n_e^2}, \qquad \operatorname{tg}\alpha = \frac{\varepsilon\cos\theta\sin\theta}{1+\varepsilon\sin^2\theta}. \tag{3.130}$$

In these formulas n_o and n_e are the refractive indices of the ordinary and extraordinary waves, respectively, while θ is the angle between the z axis and the crystal optical axis.

Substituting (3.129) in (3.119), we obtain

$$\Delta = \rho k_x^2 + 2\rho\varkappa k_x + \delta k_y^2 - (k_1 + k_2 - k_3),$$
$$\rho = \frac{1+\mu_1}{2k_1} - \frac{1+\mu_3}{2k_3}, \quad \delta = \frac{1+\nu_1}{2k_1} - \frac{1+\nu_3}{2k_3}, \tag{3.131}$$
$$\varkappa = \frac{1}{2\rho}(\operatorname{tg}\alpha_1 - \operatorname{tg}\alpha_3).$$

In the case of an interaction of the oo–e type, the wave with frequency ω_1 is ordinary and we must put $\varepsilon_1 = 0$. Consequently, $\mu_1 = \nu_1 = \tan\alpha_1 = 0$. We shall assume hereafter that the synchronism condition is satisfied at $k_x = k_y = 0$. Then $k_1 + k_2 - k_3 = 0$, and the expression for the detuning Δ assumes a shorter form

$$\Delta = \rho k_x^2 + 2\rho\varkappa k_x + \delta k_y^2. \tag{3.132}$$

We substitute this expression in (3.128). If the crystal anisotropy is small (i.e., at $\mu \ll 1$, $\nu \ll 1$), we obtain

$$K(k_x, k_y) = -i\beta_3 A_2\left(\frac{\Delta^2}{4}+\gamma\right)^{-1/2}\sin\left(\sqrt{\frac{\Delta^2}{4}+\gamma}\,z_0\right)\times$$
$$\times\exp\left[i\left(\frac{z_0}{2k_1}+\frac{z_0}{2k_3}-\frac{z_1}{k_{10}}-\frac{z_0}{k_{30}}+\frac{z_2}{k_{30}}\right)\frac{k_x^2+k_y^2}{2} + i\frac{1}{2}z_0(\operatorname{tg}\alpha_1 - \operatorname{tg}\alpha_3)k_x\right]. \tag{3.133}$$

It is known from the theory of optical systems that the best focusing of the obtained image corresponds to a position of the recording plane (the z_2 plane in this case), such that the term containing k^2_x and k^2_y in the argument of the exponential is zero. From this we get the position of the image plane z_2:

$$z_2 = z_1 \frac{k_{30}}{k_{10}} + z_0 - z_0\left(\frac{k_{30}}{2k_1} + \frac{k_{30}}{2k_3}\right). \tag{3.134}$$

In this plane the transfer function $K(k_x, k_y)$ is equal to

$$K(k_x, k_y) = -i\beta_3 A_2 \left(\frac{\Delta^2}{4} + \gamma\right)^{-1/2} \sin\left(\sqrt{\frac{\Delta^2}{4} + \gamma}\, z_0\right) \exp\left[i\frac{1}{2} z_0 (\operatorname{tg}\alpha_1 - \operatorname{tg}\alpha_3) k_x\right]. \tag{3.135}$$

To find the scattering function of the analyzed system it is necessary to take the Fourier transform of (3.135). To obtain an analytic result in this case we assume smallness of the parameter γ that determines the conversion coefficient. Taking the Fourier transform, we obtain the following equation for the scattering function of the system:

$$\Gamma(x, y) = -i\beta_3 A_2 z_0 \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty}\int \left\{\frac{z_0}{4k_s}[(k_x + \varkappa)^2 + k_y^2]\right\}^{-1} \sin\left\{\frac{z_0}{4k_s}[(k_x + \varkappa)^2 + k_y^2]\right\} \exp[-i(x - x_0) k_x - iyk_y]\, dk_x\, dk_y, \tag{3.136}$$

where

$$k_s = \frac{k_3 k_1}{k_3 - k_1}, \quad x_0 = \frac{1}{2} z_0 (\operatorname{tg}\alpha_1 - \operatorname{tg}\alpha_3). \tag{3.137}$$

In the derivation of (3.136) we took into account the condition that the anisotropy of the crystal be small, and left out the terms $z_0 k_S (\tan\alpha_1 - \tan\alpha_3)^2 \ll 1$.

Without dwelling on the intermediate calculations, which are relegated to Appendix V, we write directly the final result of the integration of (3.136):

$$\Gamma(x, y) = -\frac{i}{\pi} k_s \beta_3 A_2 \exp[i(x - x_0)\varkappa]\left\{\frac{\pi}{2} - \operatorname{Si}\left[\frac{k_s (x - x_0)^2 + k_s y^2}{z_0}\right]\right\}, \tag{3.138}$$

$$\operatorname{Si}(x^2) = \int_0^x \frac{\sin t}{t}\, dt.$$

A plot of $\pi/2 - \operatorname{Si}(x^2)$ is shown in Fig. 3.15.

According to the Rayleigh criterion, two pointlike pbjects are regarded as resolvable if the intensity at the central minimum of their combined image does not exceed 0.7 of the intensity at the maximum of the picture. Starting from this we can obtain the following estimate for the resolving power of a parametric image converter:

$$\Delta x = \Delta y \approx 2.1\,(z_0/k_s)^{1/2}. \tag{3.139}$$

From (3.138) it follows also that the converted image is shifted by an amount x_0 along the x coordinate. This "drift" of the image is due to the anisotropy of the crystal.

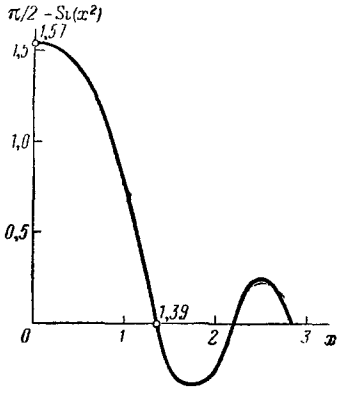


Fig. 3.15

From the practical point of view, the most favorable is a conversion regime in which the converted image is focused directly inside the nonlinear crystal. In the notation of Fig. 3.13 this means that the z_1 plane is located inside the crystal. In this case the inhomogeneities of the refractive index of the crystal and the divergence of the pump radiation are less effective. The number of resolution elements in such a scheme is obviously $(d/\Delta x)^2$, where d is the transverse dimension of the crystal. The plane z_2 of the converted image is located inside the crystal. For example, conversion of an image of wavelength $\lambda_1 = 1.06$ μm at a pump wavelength $\lambda_2 = 0.69$ μm results in an image at a wavelength $\lambda_3 = 0.42$ μm. At a crystal length $z_0 = 1$ cm, we obtain

$$\Delta x = 24 \ \mu\text{m}.$$

At a crystal transverse dimension d = 5 mm, the theoretical number of the converter resolution elements is

$$N = (d/\Delta x)^2 = 4.2 \cdot 10^4.$$

This number is indicative of the resolving power of the converter. No account is taken here of the restrictions imposed by the shaping optical system.

The first experiments on image conversion in a nonlinear crystal were performed by Midwinter [16]. He realized a scheme with an infinitely remote source. In [22] was proposed a more convenient scheme with the source at a shorter distance. Among the other experimental studies of parametric image conversion, notice should be taken of a study where maximum resolving power was obtained in a system with a closely located source, at a wavelength 1.06 μm [23], and an investigation of the "graininess" of the image with a diffuse radiation signal at a wavelength 10.6 μm [24].

3. Sensitivity of Receivers with Parametric Frequency Conversion. The sensitivity of a receiver with signal-frequency conversion is in the general case a complicated characteristic that depends on the noises in the frequency converter and the photodetector, on the transfer functions of these elements, as well as on the statistical properties of the received field. A sufficiently complete analysis of these questions is presented in [25, 26]. We present a simplified analysis that yields approximate results.

The most complete characteristic that permits calculation of the receiver sensitivity is the distribution of the number of photoelectrons emerging from the photodetector and recorded during the observation interval. We assume that the signal photons entering the converter and the noise photons generated by the converter have Poisson distributions. Then the signal photons at the output of the converter have a Bernoulli distribution $P_1(n)$ averaged over a Poisson distribution:

$$P_1(n) = \sum_{m=n}^{\infty} \frac{m!}{(m-n)!\,n!} \eta_{\text{con}}^n (1-\eta_{\text{con}})^{m-n} \exp(-n_s) \frac{n_s^m}{m!}, \tag{3.140}$$

where η_{con} is the photon-number conversion coefficient of the converter and n_s is the average number of signal photons entering the converter. The parameter η_{con} is physically equivalent to the probability of conversion of one photon entering the converter into one output-radiation photon. Simple algebraic transformations yield

$$P_1(n) = \exp(-n_s) \frac{(\eta_{\text{con}} n_s)^n}{n!} \sum_{l=0}^{\infty} \frac{[(1-\eta_{\text{con}}) n_s]^l}{l!} = \exp(-\eta_{\text{con}} n_s) \frac{(\eta_{\text{con}} n_s)^n}{n!}. \tag{3.141}$$

With this expression taken into account, the overall distribution of the number of the photons in the converted signal and noise generated by the converter is a combination of Poisson distributions with mean values $\eta_{\text{con}} n_s$ and n_{no}:

$$P_2(n) = \sum_{l=0}^{n} \exp(-\eta_{\text{con}} n_s) \frac{(\eta_{\text{con}} n_s)^l}{l!} \exp(-n_{\text{no}}) \frac{n^n \text{no}^{-l}}{(n-l)!} = \exp[-(\eta_{\text{con}} n_s + n_{\text{no}})] \frac{(\eta_{\text{con}} n_s + n_{\text{no}})^n}{n!}. \tag{3.142}$$

The summary distribution of the photons at the converter output is also of the Poisson type, with the mean number of photons at the output

$$\langle n\rangle = \eta_{con}\, n_s + n_{no}. \tag{3.143}$$

In the approximation considered, the converter can be treated as a multiple converter of the number of signal photons to which a certain number of noise photons is added. This approximation is valid when the signal power is much less than the pump power.

Each photon leaving the converter lands on the sensitive surface of the photodetector, where it generates, with probability η_q equal to the quantum efficiency of the photocathode, one photoelectron. The distribution of the number of photoelectrons generated by the converter photons is determined in analogy with (3.140):

$$P_3(n) = \sum_{m=n}^{\infty} \frac{m!}{(m-n)!\,n!}\,\eta_q^n (1-\eta_q)^{m-n} P_2(m). \tag{3.144}$$

Carrying out the same transformations as in the derivation of (3.141), we obtain

$$P_3(n) = \exp\left[-(\eta_{con}\, n_s + n_{no})\,\eta_q\right] \frac{[\eta_q(\eta_{con}\, n_s + n_{no})]^n}{n!} \tag{3.145}$$

We recognize now that, in the absence of photons, the photodetector generates a certain sequence of noise electrons, which are due to thermal and field-emission processes in the photocathode. These are called darkness photoelectrons. The darkness photoelectrons have a Poisson distribution with a mean value n_d:

$$P_d(n) = \exp(-n_d)\frac{n_d^n}{n!}. \tag{3.146}$$

The total distribution of the number of photoelectrons, with account taken of the signal and noise photons and of the darkness photoelectrons of the detector, is obtained by combining the two Poisson distributions (3.145) and (3.146). We have already seen [Eq. (3.142)] that the summary distribution is also of the Poisson type with mean value

$$\gamma = \eta_q(\eta_{no} n_s + n_{no}) + n_d. \tag{3.147}$$

From (3.147) follow two important conclusions. First, the conversion coefficient of the converter + photodetector system is determined by the product of the converter coefficient by the photodetector quantum efficiency. Second, the average number of noise photons, referred to the converter input, is

$$\langle n_{no}\rangle = (\eta_q\, n_{no} + n_d)/\eta_q\,\eta_{con}. \tag{3.148}$$

We consider in the same approximation the characteristics of signal detection. In the photoelectron-counting regime, the missed-signal probability is

$$\beta = \sum_{n=0}^{n_0} \exp(-\gamma)\frac{\gamma^n}{n!}, \tag{3.149}$$

where γ is determined by Eq. (3.147) and n_0 is the threshold number of photoelectrons. The false-alarm probability is

$$F = 1 - \sum_{n=0}^{n_0} \exp(-\alpha)\frac{\alpha^n}{n!}, \tag{3.150}$$

where

$$\alpha = \eta_q\, n_{no} + n_d$$

is the average number of noise electrons. We put $\eta_q = 0.1$, $\eta_{con} = 0.06$, $\alpha = 2$. At these values of the parameters the detection characteristics, as functions of n_s, take the form shown in Fig. 3.16. The average number of noise electrons, referred to the input, is equal,

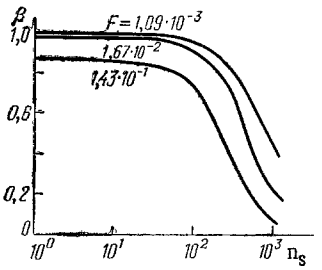


Fig. 3.16

according to (3.148), to 333. In this case, as follows from Fig. 3.16, to ensure the characteristics F = 0.143 and β = 0.1, it is necessary to have at the input an average of 10^3 photons, i.e., the signal/noise ratio must be not less than 3.

The foregoing results are sufficient only for a qualitative estimate of the influence of the converter noise on the receiver sensitivity. In a receiver with parametric frequency conversion there are usually present:

a) spontaneous noise due to multiphoton processes;

b) thermal noise due to fluctuations of the electromagnetic fields in the nonlinear medium;

c) an external background;

d) fluctuations of the pump power;

e) technical noise due to imperfection of the receiver;

f) the photodetector noise.

From among the spontaneous multiphoton processes that contribute to the noise in the summary frequency, notice should be taken of the following:

1) Spontaneous decay of pump photons of frequency ω_2 into a photon at the signal frequency ω_1 and a photon of some supplementary frequency ω, followed by conversion of ω_1 into the output-radiation frequency ω_3:

$$\omega_2 = \omega_1 + \omega, \quad \omega_1 + \omega_2 = \omega_3.$$

2. Generation of the second harmonic of the pump, followed by decay of the harmonic photons into photons of the sum frequency and of a certain difference frequency:

$$\omega_2 + \omega_2 = 2\omega_2, \quad 2\omega_2 = \omega_3 + \omega.$$

3. Four-photon decay of two pump photons into photons of sum and difference frequencies.

$$\omega_2 + \omega_2 = \omega_3 + \omega.$$

The first two processes yield the summary frequency as a result of successive three-photon processes due to the quadratic dielectric susceptibility of the nonlinear crystal. The third is a four-photon process and is due to the cubic dielectric susceptibility. At low pump divergence, the processes of the first and second type predominate, since the cubic dielectric susceptibility is smaller by two or three orders than the quadratic susceptibility.

The most intense among the technical noises are those due to insufficient selection of the pump signal. The pump-suppression requirements, characterized by the suppression coefficient α at the output of the nonlinear crystal, are exceedingly stringent:

$$\alpha = P_1/P_2 \approx 10^{-16}.$$

Three methods are used to suppress the pump:

a) spectral filtering;

b) spatial filtering;

c) combined method.

In the case of conversion from the near infrared, when the frequencies of the pump and of the converted radiation are greatly separated (λ_2 = 0.69-1.06 μm, λ_3 = 0.42-0.53 μm), spatial filtering is, as a rule, sufficient, and sometimes even without the use of interference filters. When converting from the far infrared band, the pump frequency and that of the converted signal are very close (λ_2 = 0.69-1.06 μm, λ_3 = 0.65-0.96 μm) and spectral filtering alone is not enough.

Spatial filtering is used only in two-dimensional interaction, when the directions of the different beams at the output of the linear crystals are different, and there is no need for complicated spatial-filtering schemes. It usually suffices to use a corrected lens that focuses the pump signal and the output radiation on different points, and a diaphragm that separates only the output radiation. More complicated space-filtering schemes can also be used, in which account is taken of the fine structure of the converted signal.

In principle, in the absence of additional effects such as scattering of the pump in the crystal and in the optical elements, and reflection of the pump from various surfaces, any one method would be sufficient. The presence of these effects, however, calls for the combined methods, with use of both spatial and spectral filtering with the aid of narrow-band interference filter. In this case, the most desirable interactions would be those in which the output direction differences are largest.

CHAPTER 4

HETERODYNE METHOD OF PROCESSING OPTICAL FIELDS

Among the known types of photoreceivers intended for the reception of coherent laser signals, the highest sensitivity in the infrared region is possessed by heterodyne receivers. The gist of the heterodyne method is that the received laser signal is mixed (interferes) on the sensitive area of the photodetector with the coherent (reference) field of the heterodyne. The photocurrent component due to the interference of the signal and reference fields is in this case a signal of intermediate frequency equal to the difference between the signal-field and reference-field frequencies. The amplitude of this signal is proportional to the product of the amplitudes of the signal and reference fields. Thus, an effect similar to amplification of the received field takes place. This amplification permits the photocurrent due to a weak signal to be increased to a level at which the influence of the intrinsic noise of the photodetector becomes negligibly small.

Compared with other photodetection methods, the heterodyne method is the most rational when it comes to recording laser signals in the middle infrared region of the spectrum, for which there are at present no photodetectors with internal amplification. In addition, the heterodyne method permits measurement of the frequency of the recorded field and a detailed spectral analysis of the radar signal.

4.1. Spatial Matching of the Fields in Heterodyne Receivers

1. Optical Mixing on the Photodetector Surface. We consider a signal field $E_s(\mathbf{r}, t)$ and a reference heterodyne field $E_r(\mathbf{r}, t)$ incident on the sensitive area of the photodetector. When describing these fields, we shall hereafter separate the field amplitude $A(\mathbf{r}, t)$ that varies slowly in time

$$\begin{aligned} E_s(\boldsymbol{r}, t) &= A_s(\boldsymbol{r}, t)\exp(i\omega_s t), \\ E_r(\boldsymbol{r}, t) &= A_r(\boldsymbol{r}, t)\exp(i\omega_r t). \end{aligned} \tag{4.1}$$

In these expressions ω_s and ω_r are the frequencies of the signal and reference light fields, respectively. The photocurrent at the photodetector output is proportional to the total field power incident on the sensitive area:

$$i_{ph} = \frac{\eta e}{h\nu}\int_S I(\boldsymbol{r}, t)\, d^2r, \tag{4.2}$$

where e is the electron charge, η is the quantum efficiency of the photodetector, h is Planck's constant, ν is the average field frequency, and $I(\mathbf{r}, t)$ is the field intensity. The integration is over the photodetector area. Since the optical frequencies $\nu_s = \omega_s/2\pi$ and $\nu_r = \omega_r/2\pi$ are high (of the order of 10^{13} Hz), and the difference $\nu_s - \nu_r$ does not exceed as a rule several tens of MHz, there is no need to differentiate between ν_s and ν_r in the calculation of the photon energy.

The field intensity $I(\mathbf{r}, t)$ is equal to

$$I(\boldsymbol{r}, t) = |E_s(\boldsymbol{r}, t) + E_r(\boldsymbol{r}, t)|^2. \tag{4.3}$$

Substituting Eqs. (4.1) in this expression, we get

$$\begin{aligned} I(\boldsymbol{r}, t) = I_s(\boldsymbol{r}, t) + I_r(\boldsymbol{r}, t) + \\ + A_s(\boldsymbol{r}, t)A_r^*(\boldsymbol{r}, t)\exp[i(\omega_s - \omega_r)t] + A_s^*(\boldsymbol{r}, t)A_r(\boldsymbol{r}, t)\exp[-i(\omega_s - \omega_r)t], \end{aligned} \tag{4.4}$$

where $I_s(\mathbf{r}, t)$ and $I_r(\mathbf{r}, t)$ are the intensities of the signal and reference fields. Using (4.4), we obtain the explicit form of the expression for the photocurrent:

$$\begin{aligned} i_{ph} = \frac{\eta e}{h\nu}\int_S [I_s(\boldsymbol{r}, t) + I_r(\boldsymbol{r}, t)]\, d^2r + \frac{\eta e}{h\nu}\exp[i(\omega_s - \omega_r)t]\int_S A_s(\boldsymbol{r}, t)A_r^*(\boldsymbol{r}, t)\, d^2r + \\ + \frac{\eta e}{h\nu}\exp[-i(\omega_s - \omega_r)t]\int_S A_s^*(\boldsymbol{r}, t)A_r(\boldsymbol{r}, t)\, d^2r. \end{aligned} \tag{4.5}$$

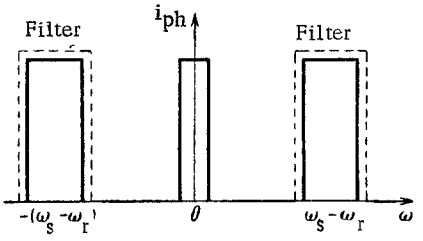


Fig. 4.1

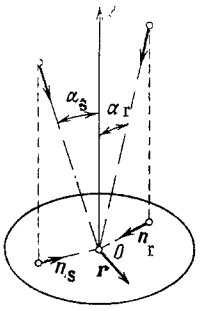


Fig. 4.2

The first term of this expression describes the slow fluctuations of the photocurrent, while the two others describe the alternating photocurrent components with frequency $\Omega = \omega_s - \omega_r$ The spectral composition of the photocurrent is shown schematically in Fig. 4.1.

If the intermediate frequency Ω is high enough to keep the spectral components from overlapping, the last two terms in (4.5) can be separated by radiotechnical means using a narrowband filter, as shown in Fig. 4.1. This separation is practically always possible, since the spectral width of the first term of (4.5) is usually several tens of kilohertz, the spectral width of the two other terms does not exceed several megahertz, and the intermediate frequency $\Omega/2\pi$ is chosen equal to tens of megahertz. We shall be interested hereafter only in the last two terms of (4.5). We therefore leave out the first term without further remarks. In this section we consider only the spatial dependence of the field amplitude. We assume therefore that the field is monochromatic and disregard the time dependence of the field amplitude. Taking the foregoing remarks into account, we rewrite (4.5) in the form

$$i_{ph}=\frac{\eta e}{h\nu}\exp\left[i\left(\omega_s-\omega_r\right)t\right]\int_S A_s(\boldsymbol{r})\,A_r^*(\boldsymbol{r})\,d^2r+\frac{\eta e}{h\nu}\exp\left[-i\left(\omega_s-\omega_r\right)t\right]\int_S A_s^*(\boldsymbol{r})\,A_r(\boldsymbol{r})\,d^2r. \tag{4.6}$$

2. Directivity Pattern of Heterodyne Receiver. We consider the dependence of the intermediate-frequency signal amplitude on the spatial distribution of the reference and signal fields. It follows from (4.6) that the complex amplitude of the photocurrent is

$$a=2\frac{\eta e}{h\nu}\int_S A_s(\boldsymbol{r})\,A_r^*(\boldsymbol{r})\,d^2r. \tag{4.7}$$

The factor $2\eta e/h\nu$, in this formula will be designated by B. This parameter will not play a significant role in the analysis that follows.

We consider first a particular case when a signal plane wave and a reference plane wave of the heterodyne are incident on a planar surface of the photodetector at different angles (Fig. 4.2). In the photodetector plane, the spatial distributions of the signal and heterodyne fields are given by

$$\begin{aligned}A_s(\boldsymbol{r})&=a_s\exp\left(ik_s\alpha_s\boldsymbol{n}_s\boldsymbol{r}\right),\\ A_r(\boldsymbol{r})&=a_r\exp\left(ik_r\alpha_r\boldsymbol{n}_r\boldsymbol{r}\right),\end{aligned} \tag{4.8}$$

where k_s and k_r are the wave numbers of the signal and reference fields. Expressions (4.8) are valid for small inclination angles α_s and α_r. The amplitudes a_s and a_r of the signal and heterodyne plane waves are real.

Substituing expressions (4.8) in (4.7) we obtain

$$a=Ba_sa_r\int_S\exp\left[i\boldsymbol{r}\left(k_s\alpha_s\boldsymbol{n}_s-k_r\alpha_r\boldsymbol{n}_r\right)\right]d^2r. \tag{4.9}$$

The integration is over the entire photodetector surface. As a result of the frequency difference between the reference and signal radiation, the wave number k_s differs from the wave number k_r

$$k_s=k_r+\Delta k,\quad \Delta k=2\pi\Delta\nu/c, \tag{4.10}$$

where $\Delta\nu$ is the difference between the frequencies of the signal and reference fields. The quantity Δk can be neglected if the following inequality holds:

$$\alpha \Delta k R \ll \pi/2, \tag{4.11}$$

where R is the characteristic dimension of the photodetector, and α is the larger of the two inclination angles α_s and α_r. The inequality (4.11) follows directly from (4.9). In practice, the typical parameter values are

$$\alpha < 10^{-1} \text{ rad}, \quad \Delta\nu < 10^8 \text{ Hz}, \quad R < 10^{-3} \text{ m},$$

from which we get

$$\alpha \Delta k R < 2 \cdot 10^{-4} \ll \pi/2.$$

Thus, inequality (4.11) is satisfied with large margin, so that we can put in (4.9)

$$k_s = k_r = k. \tag{4.12}$$

We then obtain

$$a = B a_s a_r \int_S \exp[ikr(\alpha_s n_s - \alpha_r n_r)]\, d^2r. \tag{4.13}$$

The integration is easily carried out in the particular case of a circular aperture. Denoting by φ the angle between the vectors $\mathbf{r}$ and $\alpha_s \mathbf{n}_s - \alpha_r \mathbf{n}_r$ we write expression (4.13) in the form

$$a = B a_s a_s \int_0^{2\pi} \int_0^R \exp(ikr\,|\,\alpha_s n_s - \alpha_r n_r\,|\cos\varphi)\, r\, dr\, d\varphi, \tag{4.14}$$

where R is the photodetector radius. We obtain ultimately

$$a = B a_s a_r S \cdot 2 \frac{J_1(kR\,|\,\alpha_s n_s - \alpha_r n_r|)}{kR\,|\,\alpha_s n_s - \alpha_r n_r\,|}, \tag{4.15}$$

where $S = \pi R^2$ is the photodetector area and J_1 is a Bessel function of first order. If the propagation vectors of the reference and signal waves lie in the same plane, the unit vectors $\mathbf{n}_s$ and $\mathbf{n}_r$ coincide and we get

$$a = B a_s a_r S \cdot 2 \frac{J_1[kR(\alpha_s - \alpha_r)]}{kR(\alpha_s - \alpha_r)}. \tag{4.16}$$

The amplitude of the photocurrent at the intermediate frequency depends thus on the mismatch angle $\alpha_s - \alpha_r$ between the propagation directions of the signal and reference waves. The function $2J_1(x)/x$ is well known. In order for the amplitude of the photocurrent to be not less than half the maximum value, the following condition must be satisfied

$$kR\,|\alpha_s - \alpha_r| < 2.2, \tag{4.17}$$

or in another form

$$|\alpha_s - \alpha_r| < 0.7\lambda/D, \tag{4.18}$$

where D = 2R is the photodetector diameter.

Equation (4.18) determines in fact the width of the directivity pattern of the heterodyne receiver, while Eq. (4.16) determines the form of the directivity pattern. The maximum of the directivity pattern corresponds to a signal-wave arrival direction such the angles α_s and α_r are equal. In this case, as follows from (4.16), the amplitude of the photocurrent at the intermediate frequency is independent of the absolute values of the angles α_s and α_r.

We consider now the matching system frequently used in practice and illustrated in Fig. 4.3, where the photodetector is in the focal plane of the lens. This scheme is convenient because it makes it possible to focus on a small photodetector the energy gathered on a large surface of the focusing lens. Let us find the directivity pattern of the heterodyne receiver in this case. In the arguments that follow, we shall assume that the photodetector dimensions are large enough compared with the focused light spot.

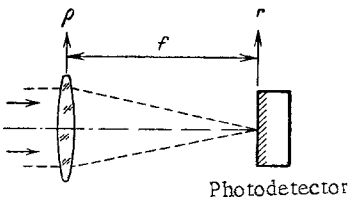


Fig. 4.3

It is known that in the paraxial approximation the field in the focal plane of a lens is the Fourier transform of the field at the entrance aperture of the lens [1]:

$$A(\boldsymbol{r}) = c \exp\left(i \frac{k}{2f} r^2\right) \int A(\boldsymbol{\rho}) \exp\left(-i \frac{k}{f} \boldsymbol{r}\boldsymbol{\rho}\right) d^2\rho. \tag{4.19}$$

Here c is a constant coefficient that will be determined later, f is the focal length of the lens, and the integration is over the plane ρ.

We express the complex amplitude of the photocurrent at the intermediate frequency [Eq. (4.7)] in terms of the field at the entrance aperture of the focusing lens. To this end we substitute (4.19) in (4.7):

$$a = B|c|^2 \int A_s(\boldsymbol{\rho}_1) A_r^*(\boldsymbol{\rho}_2) \exp\left[-i \frac{k}{f} \boldsymbol{r}(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)\right] d^2r\, d^2\rho_1 d^2\rho_2. \tag{4.20}$$

The integration with respect to d^2r yields a delta-function, as a result of which we arrive at the expression

$$a = B|c|^2 \left(2\pi \frac{f}{k}\right)^2 \int A_s(\boldsymbol{\rho}_1) A_r^*(\boldsymbol{\rho}_2) \delta(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)\, d^2\rho_1 d^2\rho_2. \tag{4.21}$$

Integrating further with respect to $d^2\rho_1$, we obtain

$$a = B|c|^2 \lambda^2 f^2 \int A_s(\boldsymbol{\rho}) A_r^*(\boldsymbol{\rho})\, d^2\rho. \tag{4.22}$$

We determine now the coefficient c. To this end we write down an expression for the total power W_1 of an arbitrary field $A(\mathbf{r})$ incident on the photodetector through the focusing lens:

$$W_1 = \int |A(\boldsymbol{r})|^2 d^2r. \tag{4.23}$$

The integration is carried out here over the entire **r** plane. The total power of the field passing through the aperture of the lens is equal to

$$W_2 = \int |A(\boldsymbol{\rho})|^2 d^2\rho. \tag{4.24}$$

In analogy with the derivation of (4.22), we obtain

$$\int |A(\boldsymbol{r})|^2 d^2r = |c|^2 \lambda^2 f^2 \int |A(\boldsymbol{\rho})|^2 d^2\rho. \tag{4.25}$$

By virtue of the energy conservation law, we have

$$W_1 = W_2,$$

whence we get, apart from an inessential phase factor,

$$c = 1/\lambda f. \tag{4.26}$$

Substituting this expression in (4.22) we get

$$a = B \int A_s(\boldsymbol{\rho}) A_r^*(\boldsymbol{\rho})\, d^2\rho. \tag{4.27}$$

The integration in this formula is over the surface of the focusing lens. Comparison of (4.27) with (4.7) shows that from the mathematical viewpoint the complex amplitude of the photocurrent at the intermediate frequency can be calculated by integrating a product of

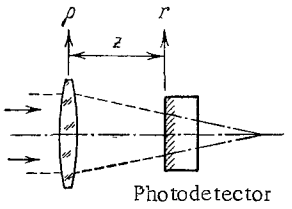


Fig. 4.4

complex-conjugate fields over the entrance aperture of the focusing lens. In certain cases this turns out to be more convenient than calculation of the integral directly over the photodetector surface.

We assume now that the photodetector was moved away from the focal plane into some intermediate plane (Fig. 4.4). We shall show that after this displacement the complex amplitude of the photocurrent does not change and is determined by Eq. (4.27). This result is valid, of course, only when the dimensions of the photodetector are so large that it spans over the entire light beam.

To prove the foregoing statement we consider the transformation that relates the field in the plane ρ of the focusing length and in an arbitrary plane $\mathbf{r}$ behind it. It is known from the theory of optical systems that the most general form of this transformation [1] is

$$A(\boldsymbol{r}) = c\exp(i\alpha r^2)\int A(\boldsymbol{\rho})\exp(i\beta\rho^2 + i\gamma\boldsymbol{\rho}\boldsymbol{r})\,d^2\rho, \tag{4.28}$$

where c is a constant coefficient; α, β, γ are parameters that depend on the radiation wavelength λ and on the distance z and f. The integration in this formula is carried out in the plane ρ.

We consider now an integral over the plane $\mathbf{r}$

$$I = \int A_s(\boldsymbol{r})\,A_r^*(\boldsymbol{r})\,d^2r. \tag{4.29}$$

Using the transformation (4.28) we can recalculate this integral to the ρ plane:

$$I = |c|^2\int A_s(\boldsymbol{\rho}_1)\,A_r^*(\boldsymbol{\rho}_2)\exp[i\beta(\rho_1^2 - \rho_2^2) + i\gamma(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)\boldsymbol{r}]\,d^2\rho_1 d^2\rho_2 d^2r. \tag{4.30}$$

When the complex amplitudes of the fields are multiplied, the phase factor $\exp(i\alpha\mathbf{r}^2)$ becomes equal to unity. Integration with respect to the coordinate $\mathbf{r}$ leads to the appearance of the delta-function $\delta(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)$ under the integral sign in (4.30):

$$I = c_1\int A_s(\boldsymbol{\rho}_1)\,A_r^*(\boldsymbol{\rho}_2)\exp[i\beta(\rho_1^2 - \rho_2^2)]\,\delta(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)\,d^2\rho_1 d^2\rho_2, \tag{4.31}$$

where c_1 is a new constant. Integrating with respect to the coordinate ρ_2 we obtain ultimately

$$I = c_1\int A_s(\boldsymbol{\rho})\,A_r^*(\boldsymbol{\rho})\,d^2\rho. \tag{4.32}$$

From the same energy considerations as before, the constant c_1 is equal to unity.

We thus have

$$\int A_s(\boldsymbol{r})\,A_r^*(\boldsymbol{r})\,d^2r = \int A_s(\boldsymbol{\rho})\,A_r^*(\boldsymbol{\rho})\,d^2\rho. \tag{4.33}$$

Taking (4.27) into account, we arrive at the conclusion that the complex amplitude of the photocurrent at the intermediate frequency does not depend on the plane in which the photodetector is located. This conclusion is valid, of course, only when the photodetector area is large enough to span completely the entire light beam.

3. Spatial Matching at Incomplete Spatial Coherence of the Field. We consider now the peculiarities of the spatial matching of the field for incomplete (partial) spatial coherence of received radiation. We shall not consider here the time dependence of the receiver output signal, and will average the various quantities over an ensemble of random realizations of the spatial distribution of the field on the input aperture. As shown in the preceding subsection, an analysis of the spatial matching of the fields on the photodetector surface is equivalent to analysis of the spatial matching on the input aperture.

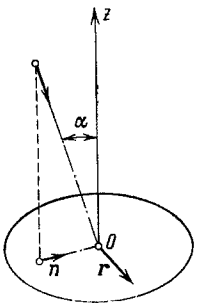


Fig. 4.5

We note first the following. Since the phase of the reflected field in each realization has the same probability of being equal to any value in the interval $(0,2\pi)$, the complex amplitude of the photocurrent (4.27) averaged over the realization is zero. Therefore in the analysis of the influence of the degree of spatial coherence of the field on the output signal of the heterodyne receiver we shall consider the mean squared modulus of the complex amplitude of the photocurrent. This approach is convenient also because the mean squared modulus of the complex photocurrent amplitude is uniquely connected with the power W of the photocurrent at the intermediate frequency:

$$\frac{1}{2}\langle|a|^2\rangle = W.$$

Therefore, investigating the mean squared modulus of the complex photocurrent amplitude, we shall actually investigate the average power of the photocurrent at the intermediate frequency. In accord with (4.27) we obtain

$$\langle|a|^2\rangle = B^2\left\langle\int A_s(\boldsymbol{r}_1)\,A_r^*(\boldsymbol{r}_1)\,A_s^*(\boldsymbol{r}_2)\,A_r(\boldsymbol{r}_2)\,d^2r_1 d^2r_2\right\rangle, \tag{4.34}$$

with the integration carried out over the entrance aperture.

In practice the reference-radiation field can usually be regarded as regular. The averaging extends then only over the signal field $A_S(\mathbf{r})$. As a result we obtain

$$\langle|a|^2\rangle = B^2\int\langle A_s(\boldsymbol{r}_1)\,A_s^*(\boldsymbol{r}_2)\rangle\,A_r^*(\boldsymbol{r}_1)\,A_r(\boldsymbol{r}_2)\,d^2r_1 d^2r_2. \tag{4.35}$$

The mean value that enters in the integrand of this expression is the spatial correlation function $R(\mathbf{r}_1, \mathbf{r}_2)$ of the signal field

$$\langle A_s(\boldsymbol{r}_1)\,A_s^*(\boldsymbol{r}_2)\rangle = R(\boldsymbol{r}_1,\ \boldsymbol{r}_2). \tag{4.36}$$

With allowance for this, (4.35) takes the form

$$\langle|a|^2\rangle = B^2\int R(\boldsymbol{r}_1,\ \boldsymbol{r}_2)\,A_r^*(\boldsymbol{r}_1)\,A_r(\boldsymbol{r}_2)\,d^2r_1 d^2r_2. \tag{4.37}$$

Let us analyze several particular cases on the basis of this formula. We consider first the case of deterioration of the spatial coherence owing to the turbulence of the atmosphere. Assume that in absence of turbulence the signal field is a plane wave incident on the input aperture at an angle α to the normal. In the presence of turbulence this field can then be represented as follows (Fig. 4.5):

$$A_s(\boldsymbol{r}) = a_s\exp(ik\alpha\boldsymbol{r}\boldsymbol{n})\exp[i\varphi(\boldsymbol{r})], \tag{4.38}$$

where $\varphi(r)$ is a random function that describes in first-order approximation the turbulence of the atmosphere (see Chap. 1). As for the reference field of the heterodyne, we assume that it comprises a plane wave that is normally incident on the surface of the input aperture. In this case we can write

$$A_r(\boldsymbol{r}) = a_r. \tag{4.39}$$

Substituting (4.38) in (4.35) we obtain

$$\langle|a|^2\rangle = B^2|a_r|^2|a_s|^2\int\exp[ik\alpha(\boldsymbol{r}_1-\boldsymbol{r}_2)\boldsymbol{n}]\times F(|\boldsymbol{r}_1-\boldsymbol{r}_2|)\,d^2r_1 d^2r_2, \tag{4.40}$$

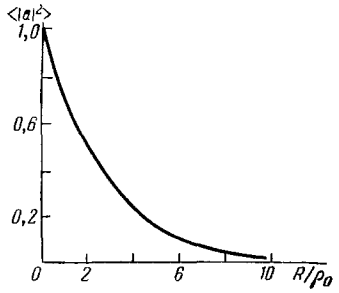


Fig. 4.6

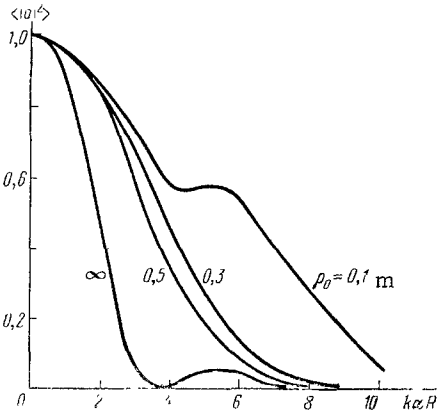


Fig. 4.7

where F is the known correlation function given by Eq. (1.180):

$$F(\rho) = \exp\left[-(|\rho|/\rho_0)^{5/3}\right]. \tag{4.41}$$

In the absence of turbulence $\rho_0 = \infty$ and $F(\rho) = 1$. The mean-squared modulus of the photocurrent amplitude is then

$$\langle |a|^2 \rangle = B^2 |a_s|^2 |a_r|^2 S^2 \left[2\frac{J_1(k\alpha R)}{k\alpha R}\right]^2, \tag{4.42}$$

where R is the aperture radius. The case considered corresponds to a spatially coherent field. Therefore, the directivity of the heterodyne receiver, which is determined by the dependence on the angle α, corresponds to Eq. (4.16), as it should.

In the presence of turbulence, $F(\rho) \neq 1$. In this case it is impossible to integrate in (4.40) in analytic form. We consider, therefore, certain results of numerical integration of (4.40) under the condition that the function $F(\rho)$ is given by (4.41).

Calculations show that a decrease of the field coherence radius, which is characterized by the parameter ρ_0, leads to a decrease of the output-signal amplitude (at $\alpha = 0$). This dependence is shown in Fig. 4.6. The turbulence of the atmosphere does not exert a noticeable influence on the heterodyne receiver in the case when

$$\rho_0 \gg R. \tag{4.43}$$

The condition (4.43) makes it undesirable to choose receiving-aperture dimensions larger than ρ_0, since from that instant on the increase of the gathered light power is offset by a deterioration of the spatial matching. On the other hand, a decrease of ρ_0 causes the directivity of the heterodyne receiver to decrease.

The directivity pattern of a heterodyne receiver is determined by the dependence of the amplitude of the output signal on the plane-wave arrival angle α. It follows from (4.40) that the dependence of the amplitude of the output signal on the angle α is somewhat similar to the Fourier transform of the correlation function $F(\rho)$. Therefore the narrower the function $F(\rho)$ the wider the directivity pattern of the heterodyne receiver. This is illustrated by the calculated data shown in Fig. 4.7 for an aperture 1 m in diameter. In this case when the field coherence radius is much smaller than the aperture size, the angle of the field of view of the heterodyne receiver is given by

$$\Delta\alpha \approx \lambda/\rho_0. \tag{4.44}$$

It must be emphasized that this formula determines the average width of the directivity pattern. The instantaneous directivity pattern (i.e., in one realization) can have a width either smaller or larger than given by Eq. (4.44).

When laser radiation is reflected from an extended object, a deterioration of the spatial coherence of the received field takes place also in the absence of atmospheric turbulence. The correlation function of the field at the aperture in the case of a diffusely reflecting object is given by Eq. (1.89), and in the case of a specularly reflecting object by (1.125). We consider by way of example diffuse reflection of an object. The correlation function of the field at the aperture is then

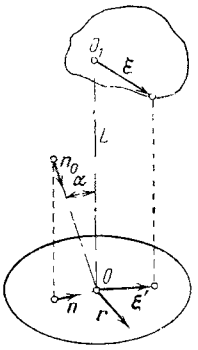


Fig. 4.8

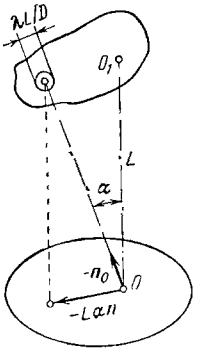


Fig. 4.9

$$\langle A(\boldsymbol{r}_1)A^*(\boldsymbol{r}_2)\rangle = \frac{1}{\lambda^2L^2}\exp\left[i\frac{k}{2L}(\boldsymbol{r}_1^2-\boldsymbol{r}_2^2)\right]\int I(\xi)\exp\left[-i\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\xi\right]d^2\xi, \tag{4.45}$$

where I(ξ) is the distribution of the intensity on the surface of the object, and the integration is over the surface of the object.

We shall find it convenient to transform from the vector ξ to the vector ξ' which is the projection of the vector ξ on the plane of the receiving aperture (Fig. 4.8). In this case one can change from integration over the surface of the object to integration over the plane of the receiving aperture. As shown in Chap. 1, for a diffusely reflecting object such a transition is carried out without taking into account the Jacobian of the coordinate transformation:

$$\langle A(\boldsymbol{r}_1)A^*(\boldsymbol{r}_2)\rangle = \frac{1}{\lambda^2L^2}\exp\left[i\frac{k}{2L}(\boldsymbol{r}_1^2-\boldsymbol{r}_2^2)\right]\int I(\xi')\exp\left[-i\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\xi'\right]d^2\xi', \tag{4.46}$$

and the integral is calculated over the plane of the receiving aperture. We substitute this expression in (4.35):

$$\langle|a|^2\rangle = B^2\frac{1}{\lambda^2L^2}\int\int\exp\left[i\frac{k}{2L}(\boldsymbol{r}_1^2-\boldsymbol{r}_2^2)\right]A_0^*(\boldsymbol{r}_1)A_0(\boldsymbol{r}_2)\int I(\xi')\exp\left[-i\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\xi'\right]d^2\xi'd^2r_1d^2r_2. \tag{4.47}$$

The factor $\exp\left[i\frac{k}{2L}(\boldsymbol{r}_1^2-\boldsymbol{r}_2^2)\right]$ in this expression describes the sphericity of the wave arriving from the object.

To obtain maximum matching the reference wave must also be spherical, with a wavefront radius equal to that of the arriving signal wave:

$$A_r(\boldsymbol{r}) = a_r\exp\left(i\frac{k}{2L}\boldsymbol{r}^2+ik\boldsymbol{n}_r\boldsymbol{r}\right), \tag{4.48}$$

where $\mathbf{n_r}$ is a unit vector in the direction of arrival of the reference wave. When recording radiation reflected from remote objects, the sphericity of the arriving wave can as a rule be neglected. For example, at a receiving aperture diameter D = 10 cm (typical value for a heterodyne receiver), at a distance L = 10 km, and a wavelength λ = 10.6 μm the maximum value of the argument in the exponential of (4.48) is 0.094. We shall, nevertheless, take into account the sphericity of the arriving wave, for this does not introduce additional complications in the calculations that follow.

We substitute (4.48) in (4.47):

$$\langle|a|^2\rangle = B^2|a_r|^2\frac{1}{\lambda^2L^2}\int\exp\left[-ik\alpha(\boldsymbol{r}_1-\boldsymbol{r}_2)\boldsymbol{n}\right]I(\xi')\exp\left[-i\frac{k}{L}(\boldsymbol{r}_1-\boldsymbol{r}_2)\xi'\right]d^2\xi'd^2r_1d^2r_2. \tag{4.49}$$

The meanings of **n** and α are clear from Fig. 4.8. Grouping the integrals with respect to the argument **r** we obtain

$$\langle|a|^2\rangle = B^2|a_r|^2\frac{1}{\lambda^2L^2}\int I(\xi')\left\{\left|\int\exp\left[-ik\left(\alpha\boldsymbol{n}+\frac{\xi'}{L}\right)\boldsymbol{r}\right]d^2r\right|^2\right\}d^2\xi'. \tag{4.50}$$

The inner integral is evaluated over the surface of the receiving aperture. Assume that the aperture is large enough so that the inner integral in (4.50) can be approximated by a delta function:

$$\int\exp\left[-i2\pi\left(\frac{\alpha\boldsymbol{n}}{\lambda}+\frac{\xi'}{\lambda L}\right)\boldsymbol{r}\right]d^2r = \delta\left(\frac{\alpha\boldsymbol{n}}{\lambda}+\frac{\xi'}{\lambda L}\right). \tag{4.51}$$

For an approximation to be valid, it is necessary to have

$$RD/\lambda L \gg 1 \tag{4.52}$$

or

$$r\lambda L/D \ll R, \tag{4.53}$$

where R is the characteristic dimension of the change of the function $I(\xi')$ and D is the diameter of the receiving aperture. Equation (4.51) is approximate, and is better the more accurate inequality (4.53). Actually, the left-hand side of (4.51), which is a function of the argument ξ', has always a finite width of the order of $\lambda L/D$.

The physical meaning of condition (4.53) is that the resolving power of the receiving aperture is sufficient to distinguish a single element of the target image. Taking into account the assumption made, expression (4.50) takes the form

$$\langle |a|^2 \rangle = B^2 |a_r|^2 \frac{S}{\lambda^2 L^2} \int I(\xi') \delta\left(\frac{\alpha n}{\lambda} + \frac{\xi'}{\lambda L}\right) d^2\xi' = B^2 |a_r|^2 S \int I(\xi') \delta(L\alpha n + \xi') d^2\xi', \tag{4.54}$$

where S is the area of the receiving aperture. We ultimately get

$$\langle |a|^2 \rangle = B^2 |a_r|^2 S I(-L\alpha n). \tag{4.55}$$

The result admits of a simple physical interpretation. A heterodyne receiver with sufficiently large receiving aperture gathers energy, in the absence of atmospheric turbulence, from only a small region of the target surface, and the direction to this region coincides with the direction of arrival of the reference wave (Fig. 4.9). The size of this region is of the order of $\lambda L/D$. Thus, the angle of the field of view of the heterodyne receiver is equal to the diffraction angle λ/D independently of the length of the target, and consequently independent of the coherence radius of the field at the receiving aperture.

This conclusion can be drawn from general consideration. In fact, the field at the aperture is a superposition of plane waves (we neglect the sphericity) from different points of the target surface. The heterodyne receiver is linear in the received field, and the signals from the individual plane waves coming from the target are added up in it independently. The directivity pattern of the heterodyne receiver therefore remains unchanged in this case.

If the spatial coherence of the field becomes worse because of atmospheric turbulence, the situation is different. In this case the coherent properties of the field deteriorate not as a result of addition of random waves, but as a result of multiplication of the arriving plane wave by a random phase factor. The directivity pattern of the heterodyne receiver turns out therefore to depend on the coherence radius of the field.

4.2. Signal/Noise Ratio in the Heterodyne Receiver

The question of the sensitivity of a heterodyne receiver is quite complicated. To answer it we must know many technical small details concerning the stability of the heterodyne, the processing of the signal from the output of the photodetector, the connection diagram of the photodetector itself, its type, etc. All these questions, however, are obviously technical and are outside the scope of the present book. We confine ourselves therefore in this section to consideration of the influence of the fluctuations of the received field on the signal/noise ratio in the heterodyne receiver and derive an expression for its maximum sensitivity.

1. Effect of Atmospheric Turbulence on the Signal/Noise Ratio. Let us examine how radar-field fluctuations due to atmospheric turbulence influence the signal/noise ratio in a heterodyne receiver. By signal/noise ratio q at the output of a heterodyne receiver we mean the reciprocal of the relative fluctuations of the power W of the signal at intermediate frequency:

$$q = \frac{\langle W \rangle^2}{\langle W^2 \rangle - \langle W \rangle^2}. \tag{4.56}$$

The intrinsic noise of the photodetector, the photocurrent shot noise, and other noise types not connected directly with the fluctuations of the received field will not be taken into account here.

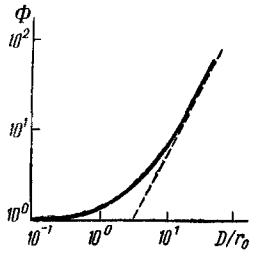


Fig. 4.10

The question of the influence of the atmosphere on the fluctuations of the photocurrent in a heterodyne receiver were considered in great detail in a number of papers. As principal among them we can single out [2-4]. A theoretical analysis shows that the following relation is valid for q [3]:

$$q = \{\exp[4\alpha C_\zeta(0)]\,\Phi(D/r_0) - 1\}^{-1}, \tag{4.57}$$

where α, $C_\zeta(0)$, r_0 are parameters; D is the diameter of the receiving aperture. The parameter r_0 in (4.57) characterizes the spatial coherence of the field passing through a layer of turbulent atmosphere and determines the phase-fluctuation structure function, equal to

$$6.88\,(|\boldsymbol{r}|/r_0)^{5/3}. \tag{4.58}$$

The parameter α in the exponential of (4.57) can be assumed equal to unity for most cases of practical importance. The asymptotic expression for this parameter is the following:

$$\alpha = \begin{cases} 1, & D \ll \sqrt{\lambda L}, \\ 0, & D \gg \sqrt{\lambda L}, \end{cases} \tag{4.59}$$

where L is the length of the turbulent route.

The parameter $C_\zeta(0)$ introduced in [3] describes the fluctuations of the complex amplitude of the field passing through a layer of turbulent atmosphere.

The function $\Phi(D/r_0)$ was calculated theoretically. A plot of this function is shown in Fig. 4.10. The characteristic region of variation of the function $\Phi(D/r_0)$ corresponds to an argument $D/r_0 \sim 1$. The region of argument values $D/r_0 < 1$ corresponds to a situation when the field is spatially coherent within the limits of the receiving aperture. The region $D/r_0 > 1$ corresponds to the case when many regions of spatial coherence of the field are spanned by the receiving aperture. In this region the noise due to the turbulence of the atmosphere increases rapidly with increasing ratio D/r_0. This growth is due to the increase of the number independent noise components, whose role is assumed by individual elements of the spatial coherence on the receiving aperture. Thus, an increase of the diameter of the receiving aperture above a certain definite value leads to a strong increase of the noise. The start of such a increase of the noise corresponds to a receiving-aperture diameter equal in order of magnitude to the size of the region of the spatial coherence of the field.

2. Influence of Reflection from the Target on the Signal/Noise Ratio. Let us assess now the value of the signal/noise ratio at the output of a heterodyne receiver in the case when the input is a random field reflected from the target. We confine ourselves right away to a point target, i.e., a target of such a size and located at such a distance that the receiving aperture does not resolve this target. It follows from the statements made in Sec. 4.1 that such a target can be situated completely in the field of view of the heterodyne receiver. On the other hand, as follows from the van Zittert–Zernike theorem, a field reflected from such a target is spatially coherent within the limits of the receiving aperture. We can therefore continue the analysis without considering the dependence on the spatial coordinate in the received field. Finally, we assume that the heterodyne radiation field does not fluctuate and is regular. Under these assumptions we obtain the signal/noise ratio q. The complex amplitude of the photocurrent at the intermediate frequency is

$$a = BA_s A_r^*, \tag{4.60}$$

where A_s is the random complex amplitude of the signal field, A_r is the complex amplitude of the heterodyne field and is a regular quantity, and B is a coefficient constant for the given

photoreceiver. The power of the output signal is equal to

$$W=\frac{1}{2}B^2|A_s|^2|A_r|^2. \tag{4.61}$$

Averaging (4.61) over the ensemble of realization we have

$$\langle W\rangle=\frac{1}{2}B^2|A_r|^2\langle|A_s|^2\rangle. \tag{4.62}$$

The quantity $\langle W^2\rangle$ in (4.56) is expressed in the form

$$\langle W^2\rangle=\frac{1}{4}B^4|A_r|^4\langle|A_s|^4\rangle. \tag{4.63}$$

Substituting (4.62) and (4.63) in (4.56) we get

$$q=\frac{\langle|A_s|^2\rangle^2}{\langle|A_s|^4\rangle-\langle|A_s|^2\rangle^2}. \tag{4.64}$$

We assume next that the received field A_s is Gaussian. This assumption calls for some supplementary clarification. In the case of a diffusely reflecting target the Gaussian character of the reflected field is obvious, since, by virtue of the central limit theorem, the presence of a large number of independent emitters leads to normalization of the distribution law of the reflected field. However, by far not all types of targets can be regarded as diffusely reflecting, all the more at the wavelength 10.6 μm, at which the heterodyne detection is used most. Frequently the reflection is from only several specular points on the target surface. In this case the statistical distribution of the reflected field differs generally speaking from Gaussian. This question was discussed in Subsec. 1 of Sec. 1.2, where it was shown that the assumption that the field is Gaussian is not excessively restrictive.

Taking the foregoing into account, we can express the quantity $\langle|A_s|^4\rangle$ in terms of $\langle|A_s|^2\rangle$. To this end we must use relation (2.17), which is valid for a Gaussian random quantity:

$$\langle A_sA_s^*A_sA_s^*\rangle=2\langle|A_s|^2\rangle^2. \tag{4.65}$$

We then obtain ultimately from (4.64)

$$q=1. \tag{4.66}$$

In practice, of course, an intermediate-frequency signal undergoes supplementary processing, as a reult of which the signal/noise ratio increases by many hundreds of times. Equation (4.66) shows in fact how large the fluctuations are in the received field in its entire frequency band.

3. Sensitivity Limit of Heterodyne Receiver. We consider now the noise produced in photoheterodyning as well as the maximum attainable sensitivity. The mean-squared shot-noise current in the intermediate-frequency passband filter at the output of the receiver is equal to

$$\langle i^2_{sh}\rangle=2e\beta\left[\frac{\eta e}{h\nu}(P_h+P_s+P_b)+I\right]. \tag{4.67}$$

where P_s and P_h are the powers of the signal and of the heterodyne at the input to the photodetector, e is the electron charge, P_b is the background-radiation power in the band of the input optical filter, I is the darkness current, and β is the transmission bandwidth of the intermediate-frequency filter.

The mean-squared thermal noise current in the band of intermediate-frequency filter at the output of the photodetector is equal to

$$\langle i^2_T\rangle=4\frac{\beta}{R}k(T_{sh}+T_{if}), \tag{4.68}$$

where T_{sh} is the temperature of the detector, T_{if} is the equivalent noise temperature at the input of the intermediate-frequency filter, k is Boltzmann's constant, and R is the load resistence. Since the noises are not correlated, their total power is

$$P=(\langle i^2_{sh}\rangle+\langle i^2_T\rangle)R. \tag{4.69}$$

We assume now that the input and heterodyne signals do not fluctuate and the wavefronts are ideally matched. Then, taking (4.7) and (4.67)-(4.69) into account we obtain the following expression for the signal/noise ratio:

$$q=2\left(\frac{\eta e}{h\nu}\right)^2 P_s P_h\left\{2e\beta\left[\frac{\eta e}{h\nu}(P_s+P_b+P_h)+I\right]+4\frac{\beta}{R}k(T_{sh}+T_{if})\right\}^{-1}. \quad (4.70)$$

With increasing heterodyne power this ratio approaches asymptotically the value

$$q_{max}=\frac{\eta P_s}{\beta h\nu}. \quad (4.71)$$

Thus, the equivalent noise power, i.e., the minimum power of the input signal that ensures q = 1 in a band β = 1 Hz, is equal to

$$\mathcal{N}=\frac{(P_s)_{min}}{\beta}=\frac{h\nu}{\eta}\ \text{W/Hz}.$$

This value characterizes the sensitivity limit of a heterodyne receiver, and its numerical value for 10.6 μm wavelength is

$$\mathcal{N}=\frac{1{,}9\cdot 10^{-20}}{\eta}\ \text{W/Hz}.$$

It must be taken into account that expression (4.70) was obtained for the case when the noise introduced by the heterodyne is purely shot noise. In practice there can arise situations when the heterodyne radiation is modulated by a noise component (fluctuations of the radiation power) at frequencies that lie in the band of the intermediate-frequency filter. Additional noise appears then as a result of direct detection of the heterodyne radiation. Taking into account this noise component, the noise power at the output of the photodetector increases more rapidly with increasing heterodyne power than the usefule-signal power. Therefore, the signal/noise ratio has a maximum at a strictly defined heterodyne radiation power, and this maximum is, naturally, lower than in the limiting case. The spectral density of the heterodyne-power fluctuations decreases as a rule with increasing frequency. Therefore, by choosing a sufficiently high intermediate frequency it is possible to weaken significantly the influence of the heterodyne power fluctuations on the sensitivity of a heterodyne photoreceiver.

4. Concrete Types of Heterodyne Receivers. The photodetectors most widely used for heterodyne receivers are the ternary compounds HgCdTe and PbSnTe, which are solid solutions of two compatible compounds (e.g., CdTe and HgTe). Both materials have the qualities needed for high-speed and high-sensitivity infrafred detectors with sensitivity in the wavelength range 8-13 μm. The photodetectors are cooled with liquid nitrogen. The quantum efficiency of PbSnTe photodetectors is close to unity, while that of HgCdTe photodetectors does, as a rule, not exceed 0.5. However, HgCdTe photodetectors are more widely used at present because of a well-worked-out technology for their manufactore and their good high-frequency characteristics. Comparative characteristics of HgCdTe and PbSnTe photodetectors working at a wavelength 10.6 μm are given in Table 7 [5]. Quadrant HgCdTe photodetectors produced for search and tracking of targets by laser radars have an equiavlent noise power 10^{-19} W/Hz and a passband up to 1.5 GHz [6].

An essential part of a heterodyne receiver is the laser heterodyne. It must satisfy definite stability, bandwidth, and frequency-tuning-rate requirements. The better of the laser heterodynes for 10.6 μm have a tuning range up to 1.5 GHz at a tuning rate 85 MHz/sec and a frequency stability up to 10^9 [7, 8].

TABLE 7. Characteristics of Heterodyne HgCdTe and PbSnTe Photodetectors

Material	Area, mm^2	Intermediate frequency, MHz	Equivalent noise power, 10^{-19} W/Hz
HgCdTe	0,05	0,1	1,0
PbSnTe	0,2	1,0	0,5

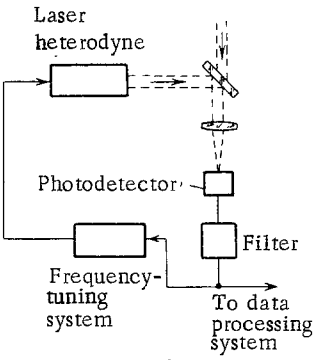


Fig. 4.11

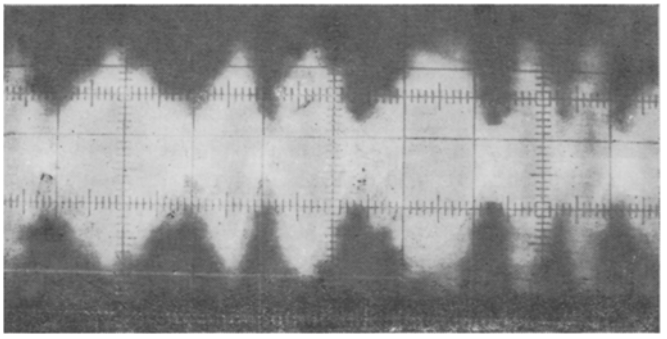


Fig. 4.12

A simplified diagram of a heterodyne receiver is shown in Fig. 4.11. To maintain the intermediate frequency within the passband of the filter, a special frequency tuning circuit is usually introduced. The electric signal at the output of this circuit is proportional to the deviation of the intermediate frequency from the central frequency of the filter. This signal is fed to the control input of the laser heterodyne and shifts its emission frequency by an amount needed to maintain the intermediate frequency within specified limits.

As indicated in Subsec. 2 of the present section, when radiation reflected from a diffusely scattering object is recorded, the relative photocurrent fluctuations at the output of a heterodyne receiver is equal to unity. This fact is well illustrated by the oscillogram, shown in Fig. 4.12, of a signal at intermediate frequency 0.465 MHz, obtained at the output of the heterodyne receiver recording radiation diffusely reflected from an object at a wavelength 10.6 μm (each major division on the oscillogram is equal to 0.1 sec). In [9] are reported detailed experimental investigations of signal fluctuations in a heterodyne receiver. Effects were investigated connected both with the random character of the reflection from the object and with the turbulence of the atmosphere. It was observed that in practically all the cases the signal amplitude had a deep modulation whose character is illustrated in Fig. 4.12.

4.3. Effect of Primary Optical-System Aberrations on the Quality of Heterodyne Reception

1. Derivation of Expression for the Interference Integral. In real heterodyne receivers the optical beams are formed with the aid of a large number of optical components, each having aberrations. The aberrations of optical systems affect adversely the heterodyne matching. It is therefore necessary to know the maximum aberration coefficients that are acceptable in the optical systems of a heterodyne receiver.

We shall consider heterodyne matching in the focal plane. It can be shown that the results remain in force also for other types of matching.

In the case considered (Fig. 4.13) the heterodyne matching is effected on the axis of optical system 1 in a plane that coincides with the focal plane of an ideal aberration-free optical systems (both systems are assumed identical in the absence of aberrations). This can always be done by a suitable displacement of the light beam along the system axis. At the same time, the distance between the axes of two systems may turn out to be an unavoidable defect of the specific mechanical construction.

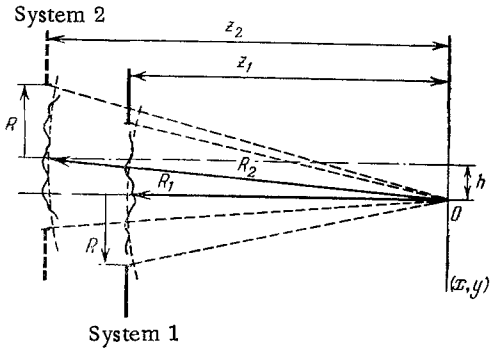


Fig. 4.13

The field in the matching plane is equal to

$$E_m(x,\ y) = \frac{A_m k_m}{2\pi z_m} \exp\left[i\,\frac{k_m}{R_m}(x^2+y^2)\right] \int F_m(u_m,\ v_m) \exp\left[-i\,\frac{k_m}{R_m}(xu_m + yv_m)\right] du_m dv_m. \tag{4.72}$$

Here A_m is the field amplitude; k_m is the wave number, $F_m(u_m, v_m)$ is the complex amplitude of the perturbation at the exit pupil of the optical system; the subscript m = 1 or 2 labels the optical system to which the variable pertains. The remaining notation is clear from Fig. 4.13.

When two concrete fields are combined, the interference (heterodyne) integral takes the form

$$G = \int E_1(x,\ y)\, E_2^*(x,\ y)\, dx\, dy + \int E_1^*(x,\ y)\, E_2(x,\ y)\, dx\, dy. \tag{4.73}$$

Substituting in this expression the integral (4.72), we get

$$G = \frac{A_1 A_2 k_1 k_2}{(2\pi)^2 z_1 z_2} \int \exp\left[i\,(x^2+y^2)\left(\frac{k_1}{R_1} - \frac{k_2}{R_2}\right)\right] F_1(u_1,\ v_1) \times$$
$$\times F_2^*(u_2,\ v_2) \exp\left[-i\,(xu_1 + yv_1)\,\frac{k_1}{R_1} + i\,(xu_2 + yv_2)\,\frac{k_2}{R_2}\right] du_1 dv_1 du_2 dv_2 dx\, dy + \text{c.c.} \tag{4.74}$$

The difference $k_1 - k_2$ of the wave numbers does not exceed, as a rule, 10^{-7} of their mean value, so that we can put $k_1 = k_2$ in (4.74). In addition, for the best matching it is necessary to satisfy the condition $R_1 = R_2$, $z_1 = z_2$. This condition ensures a maximum signal/noise ratio at the output of the photoreceiver that records the signal. Under these conditions the integral in (4.74) is finally reduced to the form

$$G = \frac{A_1 A_2 k^2}{(2\pi)^2 z^2} \int F_1(u,\ v)\, F_2^*(u,\ v)\, du\, dv + \text{c.c.} \tag{4.75}$$

Expressions (4.75) relate the aberrations of the system 1 and 2 with the quality of the heterodyne matching, which is determined by the interference integral G at constant A_1, A_2, k, z.

We represent the function F(u, v) in the form

$$F(u,\ v) = \exp\left[ik\Delta(u,\ v)\right],$$

where Δ(u, v) is the so-called aberration function, equal to the deviation of the wave front from ideal [7]. Equation (4.75) takes then the form

$$G = \frac{A_1 A_2 k^2}{(2\pi)^2 z^2}\, 2 \int \cos\left[k\Delta_1(u,\ v) - k\Delta_2(u,\ v)\right] du\, dv. \tag{4.76}$$

Two obvious conclusions follow directly from (4.76). First, G reaches a maximum when there are no aberrations:

$$G_{\max} = 2\,\frac{A_1 A_2 k^2}{(2\pi)^2 z^2}\, S = 2BS, \tag{4.77}$$

where $B = A_1 A_2 k^2/(2\pi)^2 z^2$, and S is the area of the exit pupil. Second, G reaches a maximum also when the aberrations of the systems 1 and 2 are equal.

Since the matching is effected on the axis of system 1, the only system-1 third-order aberration that contributes to (4.76) is spherical aberration. The reason is that out of all types of primary aberrations only spherical aberration can exist on the system axis. The matching for system 2 is realized at a distance h from its axis, so that all five third-order aberrations will contribute to (4.76), namely spherical aberration, astigmatism, field curvature, distortion, and coma.

It is customary in diffraction theory of aberrations to expand the aberration function $\Delta(u, v)$ in a series [10]. Following [10], we introduce the polar coordinates

$$u = \rho \cos \varphi, \quad v = \rho \sin \varphi \tag{4.78}$$

and expand the function $\Delta(u, v)$:

$$\Delta_1(u, v) = \Delta_1(\rho) = C_{11}\rho^4,$$
$$\Delta_2(u, v) = \Delta_2(\rho, \varphi, h) = C_{21}\rho^4 + C_{22}h^2\rho^2 \cos^2\varphi + C_{23}h^2\rho^2 + C_{24}h^3\rho \cos\varphi + C_{25}h\rho^3 \cos\varphi. \tag{4.79}$$

We have confined ourselves in expansion (4.79) to the terms that contribute to (4.76). These terms describe, respectively, spherical aberration, astigmatism, field curvature, distortion, and coma, while the corresponding coefficients C_{ij} are called the aberration coefficients.

2. Dependence of Interference Integral on the Aberration Coefficients. We proceed now to a separate analysis of the effect of each aberration on the quality of the heterodyne matching and determine the requirements that must be imposed on the coefficients of the corresponding aberrations.

Let system 1 have spherical aberration with a coefficient C_{11}, and system 2 spherical aberration with a coefficient C_{21}. Then

$$ik(\Delta_1 - \Delta_2) = ik(C_{11} - C_{21})\rho^4. \tag{4.80}$$

Substituting (4.80) in the expression for G and transforming to polar coordinates under the integral sign, we obtain

$$G = 4\pi B \int_0^R \cos\left[k(C_{11} - C_{21})\rho^4\right] \rho \, d\rho. \tag{4.81}$$

Using the representation for the Fresnel cosine integral [11]

$$\mathscr{C}(x) = \frac{2}{\sqrt{2\pi}} \int_0^x \cos t^2 \, dt,$$

we can reduce (4.81) to the form

$$G = 2BS \sqrt{\frac{\pi}{2}} \frac{\mathscr{C}(\sqrt{\lambda})}{\sqrt{\lambda}}, \tag{4.82}$$

where $\lambda = k|C_{11} - C_{21}|R^4$. This relation is plotted in Fig. 4.14.

We agree to regard an aberration as admissible if, at a given value of the aberration coefficient, the interference integral G is decreased to not less than one-half. Equation (4.82) leads then to an expression for the admissible aberration coefficient in the case considered:

$$|C_{11} - C_{21}| < 2{,}62/kR^4. \tag{4.83}$$

Condition (4.83) shows that the important role in heterodyne matching is played in this case not by the values of each aberration separately, but by the difference between the aberrations of the two systems. From (4.83) follows also a condition for the admissible values of the coefficient of spherical aberration in the case when system 1 (or system 2) is ideal and aberration-free:

$$|C_{i1}| < 2{,}62/kR^4. \tag{4.84}$$

We consider now the case when system 1 has spherical aberration and system 2 has astigmatism and curvature of field. We then have

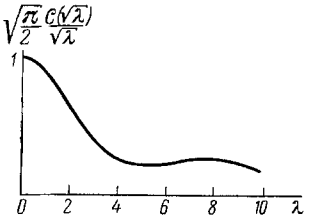


Fig. 4.14

$$ik(\Delta_1 - \Delta_2) = ik(C_{11}\rho^4 - C_{22}h^2\rho^2\cos^2\varphi - C_{23}h^2\rho^2). \tag{4.85}$$

The interference integral turns out in this case to be

$$G = 2\pi B\int_0^{R^2}\cos\left[k\left(C_{11}u^2 - \left(C_{23} + \frac{1}{2}C_{22}\right)h^2u\right)\right]J_0\left(\frac{1}{2}kC_{22}h^2u\right)du, \tag{4.86}$$

where J_0 is a Bessel function of order zero. In the particular case when system 2 has only curvature of field ($C_{22} = 0$), the integral (4.86) is expressed in terms of the Fresnel cosine and sine integrals:

$$G = 2BS\left(\frac{\pi}{2}\right)^{1/2}\frac{1}{\mu}\left[\mathscr{C}\left(\mu^{1/2} - \frac{1}{2}\frac{\xi}{\mu^{1/2}}\right)\cos\left(\frac{\xi^2}{\mu}\right) + \right.$$
$$\left. + \mathscr{S}\left(\mu^{1/2} - \frac{1}{2}\frac{\xi}{\mu^{1/2}}\right)\sin\left(\frac{\xi^2}{\mu}\right) + \mathscr{C}\left(\frac{\xi}{2\mu^{1/2}}\right)\cos\left(\frac{\xi^2}{\mu}\right) + \mathscr{S}\left(\frac{\xi}{2\mu^{1/2}}\right)\sin\left(\frac{\xi^2}{\mu}\right)\right], \tag{4.87}$$

where we have introduced the dimensionless parameters $\mu = kC_{11}R^4$, $\xi = kC_{23}h^2R^2$. The region of admissible aberration coefficients for this case was determined with a computer and is shown in Fig. 4.15. In the limiting case $C_{11} = 0$, when there is no spherical aberration in system 1, an analytic result can be obtained by using the asymptotic expression for the Fresnel integral [1]:

$$\mathscr{C}(x) + i\mathscr{S}(x) = \frac{1}{\sqrt{2}}\exp\left(i\frac{\pi}{4}\right) + (ix\sqrt{2\pi})^{-1}\exp(ix^2). \tag{4.88}$$

Substituting this expansion in (4.87) we get

$$G = 2BS\sin(kC_{23}h^2R^2)/kC_{23}h^2R^2. \tag{4.89}$$

This dependence is well known. Decreasing the value of G by one-half corresponds to a value 1.89 for the argument. The expression for the admissible values of the aberration coefficient is in this case

$$|C_{23}| < 1.89/kh^2R^2. \tag{4.90}$$

In the other particular case, when the system 2 has only astigmatism ($C_{23} = 0$), Eq. (4.86) takes the form

$$G = 2BS\int_0^1\cos\left(\mu x^2 - \frac{1}{2}\eta x\right)J_0\left(\frac{1}{2}\eta x\right)dx, \tag{4.91}$$

where $\mu = kC_{11}R^4$, $\eta = kC_{22}h^2R^2$. Computer calculation by Eq. (4.91) yields for the admissible values of the aberration coefficients of the closed region shown in Fig. 4.16. If in this case $C_{11} = 0$, the following inequality should hold

$$|C_{22}| < 3.6/kh^2R^2. \tag{4.92}$$

Let now system 1 have spherical aberration and system 2 distortion. Then

$$ik(\Delta_1 - \Delta_2) = ik(C_{11}\rho^4 - C_{24}h^3\rho\cos\varphi). \tag{4.93}$$

The expression for the interference integral will be

$$G = 2BS\cdot 2\int_0^1\cos(\mu x^4)J_0(\nu x)x\,dx, \tag{4.94}$$

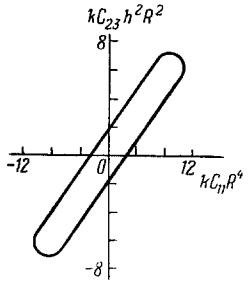


Fig. 4.15

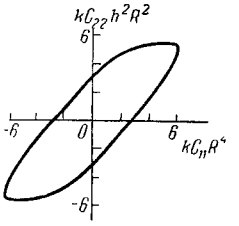


Fig. 4.16

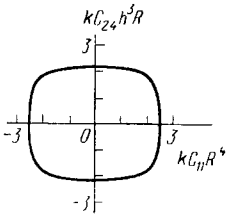


Fig. 4.17

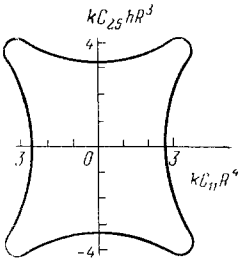


Fig. 4.18

where $\mu = kC_{11}R^4$, $\nu = kC_{24}h^3R$. The computer-calculated region of admissible values of the aberration coefficients is shown in Fig. 4.17. In the particular case $C_{11} = 0$ we have from (4.94)

$$G = 2BS \cdot 2J_1(kC_{24}h^3R)/kC_{24}h^3R. \tag{4.95}$$

This relation is well known, and it follows directly from (4.95) that the distortion of the second system is admissible under the condition

$$|C_{24}| < 2.21/kh^3R. \tag{4.96}$$

We consider, finally, the last case, when system 2 has coma and system 1 has spherical aberration. Then

$$ik(\Delta_1 - \Delta_2) = ik(C_{11}\rho^4 - C_{25}h\rho^3\cos\varphi). \tag{4.97}$$

By introducing dimensionless parameters we can reduce the interference integral to the form

$$G = 2BS \cdot 2\int_0^1 \cos(\mu x^4) J_0(\theta x^3)\, x\, dx, \tag{4.98}$$

where $\mu = kC_{11}R^4$, $\theta = kC_{25}hR^3$. Numerical computer calculation yields the plot of Fig. 4.18. In the particular case $C_{11} = 0$ a simpler condition should be satisfied, namely

$$|C_{25}| < 3.5/khR^3. \tag{4.99}$$

It should be noted in conclusion that the results are valid only if the second system has one of the indicated aberrations. If this system has simultaneously several aberrations, the relations derived no longer hold. It is possible, however, to single out the predominant aberration and to carry out for it approximate calculations. We note finally that it is theoretically possible to consider also the general case by constructing a tolerance hypersurface in the six-dimensional space of the various types of aberration. The cases considered above are then the intersections of this hypersurface with the corresponding planes.

4.4. Measurement of Object Velocity

1. Null Method of Velocity Measurement. The heterodyne detection method permits highly accurate measurement of the Doppler shift of the frequency of the radiation reflected from

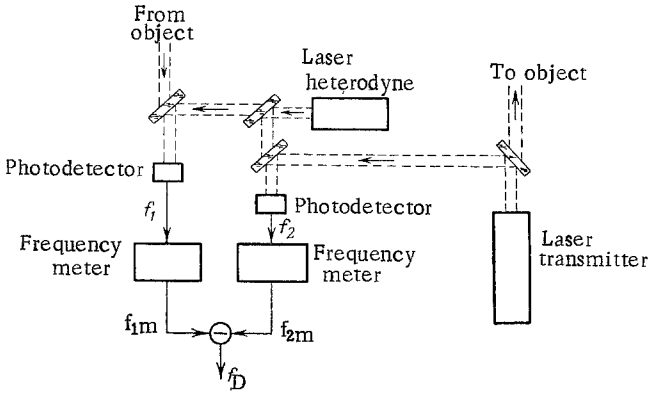


Fig. 4.19

the object, and thus measure the radial component of the velocity of a moving object [12]. The potentially high radial-velocity measurement accuracy inherent in the laser Doppler method is limited by the frequency fluctuations of the laser transmitter. To decrease the effect of these fluctuations, a null method is used. The gist of the method is the following (Fig. 4.19). The heterodyne radiation is mixed on two photodetectors with the transmitter radiation and with the reflected radiation. This results in beats at the outputs of the photodetectors, with frequencies

$$f_1 = f_h - (f_{tr} + f_D), \quad f_2 = f_h - f_{tr}, \tag{4.100}$$

where f_h is the heterodyne radiation frequency, f_{tr} is the transmitter radiation frequency, and f_D is the Doppler frequency shift given by

$$f_D = 2\frac{V}{c} f_{tr}. \tag{4.101}$$

In this expression, V is the radial component of the object velocity relative to that of light.

All the quantities in (4.100) are, generally speaking, time-dependent. It is necessary then to take into account the time delay due to the transmitter-signal propagation to the object and back:

$$f_1(t) = f_h(t) - f_{tr}\left(t - 2\frac{L}{c}\right) - 2\frac{V}{c} f_{tr}\left(t - 2\frac{L}{c}\right), \tag{4.102}$$
$$f_2(t) = f_h(t) - f_{tr}(t),$$

where L is the distance to the object.

The intermediate frequencies $f_1(t)$ and $f_2(t)$ are measured with digital frequency meters. The measured values of the corresponding frequencies at the outputs of the frequency meters are proportional to the integrals of these frequencies during the observation interval T:

$$f_{1m} = \frac{1}{T}\int_0^T f_1(t)\,dt, \quad f_{2m} = \frac{1}{T}\int_0^T f_2(t)\,dt. \tag{4.103}$$

The integral law in this expression is a consequence of the fact that the frequency meters operate by counting the sinusoid half-cycles during a fixed observation interval.

The measured values f_{1m} and f_{2m} are subtracted one from the other and yield an estimate of the Doppler frequency shift:

$$f_D = f_{2m} - f_{1m}. \tag{4.104}$$

We substitute into this expression Eqs. (4.102) and (4.103):

$$f_D = \frac{1}{T}\int_0^T \left[f_{tr}\left(t - 2\frac{L}{c}\right) - f_{tr}(t) + 2\frac{V}{c} f_{tr}\left(t - 2\frac{L}{c}\right)\right] dt. \tag{4.105}$$

We set the distance of the object equal to zero. The Doppler-shift estimate corresponding to this case is then

$$f_{D_0} = 2\frac{V}{c}\frac{1}{T}\int_0^T f_{tr}(t)\,dt. \tag{4.106}$$

The quantity f_{D_0} is obviously equal to the Doppler frequency shift averaged over the observation time interval.

We introduce the deviation of the transmitter frequency, defined as

$$\Delta f(t) = f_{tr}(t) - f_{tr_0}, \tag{4.107}$$

where f_{tr_0} is the transmitter frequency averaged over the ensemble of realizations. We assume that the relative fluctuations of the transmitter frequencies are small:

$$|\Delta f(t)/f_{tr_0}| \ll 1. \tag{4.108}$$

In laser Doppler radars this condition is satisfied as a rule with a large margin. For example, for a transmitter at a wavelength λ = 10.6 μm the value of the ratio (4.108) is of the order 10^{-9}-10^{-10}. We can therefore write with no less an accuracy

$$f_{D0} \approx 2\frac{V}{c} f_{tr0}, \tag{4.109}$$

which coincides with the initial expression (4.101). Actually, the relative error in the measurement of the Doppler shift will be even less than that given by (4.108), inasmuch as in accord with (4.106) the frequency fluctuations become averaged out over the observation interval. This question will be considered in greater detail in the next subsection.

At L ≠ 0 the measurement error increases because expression (4.105) acquires a nonzero difference

$$f_{tr}\left(t - 2\frac{L}{c}\right) - f_{tr}(t). \tag{4.110}$$

Since V/c << 1 at all times, the difference (4.110) can become comparable with the third term in (4.105) and lower substantially the measurement accuracy. Therefore, the time delay due to propagation of the laser signal to the object and back is a significant parameter in the evaluation of the accuracy of the null method.

2. Influence of Fluctuations of the Sounding-Radiation Frequency on the Velocity-Measurement Accuracy. In the measurement an estimate of the Doppler frequency shift in accord with Eq. (4.104) is generated. We obtain now the mean value of this estimate and its variance. The mean value of the estimate is obtained from (4.105) by averaging the integrand. It follows directly from the definition (4.107) that

$$\langle \Delta f(t) \rangle = 0, \quad \langle f_{tr}(t) \rangle = f_{tr_0}, \tag{4.111}$$

where the angle brackets denote averaging over the ensemble of the realizations. From this we get

$$\langle f_D \rangle = \frac{1}{T}\int_0^T \left[\left\langle f_{tr}\left(t - 2\frac{L}{c}\right)\right\rangle - \langle f_{tr}(t)\rangle + 2\frac{V}{c}\left\langle f_{tr}\left(t - 2\frac{L}{c}\right)\right\rangle\right] dt = 2\frac{V}{c} f_{tr_0}. \tag{4.112}$$

The mean value of the employed estimate of the Doppler frequency shift is thus equal to the true value determined by Eq. (4.101). In fact, by averaging (4.101) and taking (4.111) into account, we obtain an expression that coincides with (4.112).

We consider now the measurement error, which is equal to the difference between the estimated Doppler shift and its value averaged over the observation interval:

$$\delta f = f_D - 2\frac{V}{c}\frac{1}{T}\int_0^T f_{tr}\left(t - 2\frac{L}{c}\right) dt. \tag{4.113}$$

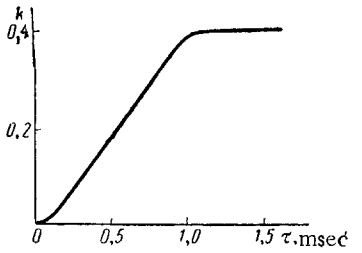


Fig. 4.20

The second term in the right-hand side of this expression differs from (4.106) only in the presence of a time delay. Substituting (4.105) in (4.113) we get

$$\delta f = \frac{1}{T}\int_0^T \left[f_{\text{tr}}\left(t - 2\frac{L}{c}\right) - f_{\text{tr}}(t)\right] dt. \tag{4.114}$$

Since <δf> = 0, it is expedient to investigate the quantity <δf²>. Direct multiplication and averaging yields

$$\langle \delta f^2 \rangle = \frac{1}{T^2}\int_0^T \int_0^T [\langle f_{\text{tr}}(t_1 - \tau) f_{\text{tr}}(t_2 - \tau)\rangle + \langle f_{\text{tr}}(t_1) f_{\text{tr}}(t_2)\rangle - \langle f_{\text{tr}}(t_1 - \tau) f_{\text{tr}}(t_2)\rangle - \langle f_{\text{tr}}(t_1) f_{\text{tr}}(t_2 - \tau)\rangle]\, dt_1 dt_2, \tag{4.115}$$

where the delay time is designated for brevity as

$$\tau = 2L/c.$$

We define now the correlation function of the transmitter emission frequency as

$$R(t_1, t_2) = \langle f_{\text{tr}}(t_1) f_{\text{tr}}(t_2)\rangle. \tag{4.116}$$

Since the process considered is stationary, we have the obvious equality

$$\langle f_{\text{tr}}(t_1 + \tau) f_{\text{tr}}(t_2 + \tau)\rangle = R(t_1 - t_2). \tag{4.117}$$

Taking this into account, Eq. (4.115) assumes the simpler form

$$\langle \delta f^2 \rangle = \frac{1}{T^2}\int_0^T \int_0^T [2R(t_1 - t_2) - R(\tau + t_1 - t_2) - R(\tau + t_2 - t_1)]\, dt_1 dt_2. \tag{4.118}$$

For the calculations that follow we must specify the explicit form of the transmitter-frequency correlation function. Using (4.107) and (4.111) we obtain

$$R(\tau) = \langle f_{\text{tr}}(t + \tau) f_{\text{tr}}(t)\rangle = f_{\text{h}0}^2 + \langle \Delta f(t + \tau)\, \Delta f(t)\rangle. \tag{4.119}$$

For the second term in this formula, it is natural to introduce the notation

$$\langle \Delta f^2 \rangle \rho(\tau) = \langle \Delta f(t + \tau)\, \Delta f(t)\rangle, \tag{4.120}$$

where ρ(τ) is the normalized correlation function of the frequency deviation and takes on the values

$$\rho(\tau) = 1 \quad \text{at} \quad \tau = 0, \tag{4.121}$$

$$\rho(\tau) \approx 0 \quad \text{at} \quad \tau \to \infty. \tag{4.122}$$

With this taken into account, we can rewrite (4.118) in the more convenient form

$$k = \frac{\langle \delta f^2 \rangle}{\langle \Delta f^2 \rangle} = \frac{1}{T^2}\int_0^T \int_0^T [2\rho(t_1 - t_2) - \rho(\tau + t_1 - t_2) - \rho(\tau + t_2 - t_1)]\, dt_1 dt_2, \tag{4.123}$$

where the parameter k so introduced is the dimensionless relative frequency-measurement error.

We assume for ρ(τ) the Gaussian approximation:

$$\rho(\tau) = \exp(-0.6931\tau^2/\tau_f^2), \tag{4.124}$$

where τ_f is the correlation time of the frequency fluctuations. Such a normalization ensures a fall-off of the correlation function by a factor of two relative to the maximum value at $\tau = \tau_f$. Analytic calculations by Eq. (4.123) under the condition (4.124) do not yield a lucid final result, and we resort therefore to numerical methods. Figure 4.20 shows the dependence of the relative error k of the frequency measurement on the delay time $\tau = 2L/c$ for a frequency correlation time $\tau_f = 0.1$ msec and a recording time T = 1 msec. As expected, the measurement error increases with increasing τ. The growth of the error continues all the way to values of τ equal to T and not to τ_f as might seem. The point is that in a null method the frequency fluctuations are subtracted over the extent of the entire measurement interval T. Therefore, if a correlation exists between the frequencies f_1 and f_2 within this interval, the measurement error will be lower than its maximum value. This maximum value is

$$k_{\max} = \frac{2}{T^2}\int_0^T\int_0^T \rho(t_1 - t_2)\,dt_1 dt_2 = 4.2\frac{\tau_f}{T} \tag{4.125}$$

at $\tau_f << T$.

3. Connection between Frequency Fluctuations and the Field Coherence Time. The results obtained in subsections 1 and 2 of the present section pertain to the case when the correlation time of the transmitter-frequency fluctuations are assumed known. Then, knowing τ_f, we can approximate the frequency-fluctuation correlation function by some known (say, Gaussian) function and find the relative frequency-measurement error k. Actually, however, when lasers are certified one measures not τ_f but the coherence time of the laser emission frequency or the width of the laser-emission spectrum. In order that the foregoing analysis be of practical interest, it is necessary to find the connection between the correlation time τ_f of the laser-emission frequency fluctuations and the width of the laser-emission spectrum. The laser emission field can be represented in the form [Eq. (1.12)]

$$E(t) = E_0\cos[\omega_0 t + \varphi(t) + \varphi_0]. \tag{4.126}$$

The instantaneous cyclic frequency of the field is

$$\omega(t) = \omega_0 + \frac{d\varphi}{dt} = 2\pi f(t), \tag{4.127}$$

where f(t) is the frequency introduced above. The deviation of the cyclic frequency is

$$\Omega(t) = \frac{d\varphi}{dt} = 2\pi\Delta f(t). \tag{4.128}$$

It is a stationary random process with zero mean value. The phase shift φ_0 in (4.126) is a random quantitiy that is constant in a given realization and is uniformly distributed over the interval $(0, 2\pi)$.

We consider the random phase advance during the time interval between t and t + τ:

$$\chi(t) = \varphi(t+\tau) - \varphi(t) = \int_t^{t+\tau}\Omega(t)\,dt. \tag{4.129}$$

The mean square of this quantity is

$$F(\tau) = \langle[\varphi(t+\tau) - \varphi(t)]^2\rangle = \langle\Omega^2\rangle\int_t^{t+\tau}\int_t^{t+\tau}\rho(t_1 - t_2)\,dt_1 dt_2, \tag{4.130}$$

where $\langle\Omega^2\rangle = 4\pi^2\langle\Delta f^2\rangle$ is the variance of the cyclic-frequency fluctuations.

Assume that $\Omega(t)$, meaning also $\chi(t)$, is a Gaussian random process. This assumption is well-founded. In fact, the transmitter frequency is altered by drifts of the effective cavity length and by mirror vibrations, due to action of a large number of various independent perturbations on the transmitter. Therefore, the processes $\Omega(t)$ and $\chi(t)$ can be regarded as normal, the probability density for $\chi(t)$ being

$$P(\chi) = [2\pi F(\tau)]^{-1/2}\exp\left[-\frac{\chi^2}{2F(\tau)}\right]. \tag{4.131}$$

We seek now the correlation function of the field. We have

$$C(\tau)=\langle E(t)\,E(t+\tau)\rangle=\frac{1}{2}E_0^2\langle\cos[\omega_0\tau+\varphi(t+\tau)-\varphi(t)]\rangle+\frac{1}{2}E_0^2\langle\cos[\omega_0(2t+\tau)+\varphi(t+\tau)+\varphi(t)+2\varphi_0]\rangle. \quad (4.132)$$

We note that by virtue of the uniform distribution φ_0 the second term in the right-hand side of this expression is zero. Consequently,

$$C(\tau)=\frac{1}{2}E_0^2\cos(\omega_0\tau)\langle\cos\chi\rangle-\frac{1}{2}E_0^2\sin(\omega_0\tau)\langle\sin\chi\rangle. \quad (4.133)$$

Averaging over the sine function by the symmetric distribution (4.131) yields zero. Hence

$$C(\tau)=\frac{1}{2}E_0^2\cos(\omega_0\tau)\langle\cos\chi\rangle. \quad (4.134)$$

Substituting (4.131) in this expression, we obtain [13]

$$C(\tau)=\frac{1}{2}E_0^2\cos(\omega_0\tau)\exp\left[-\frac{F(\tau)}{2}\right]. \quad (4.135)$$

This expression, with account taken of (4.130), reduces to the form

$$C(\tau)=\frac{1}{2}E_0^2\cos(\omega_0\tau)\exp\left[-\frac{1}{2}\langle\Omega^2\rangle\int_t^{t+\tau}\int_t^{t+\tau}\rho(t_1-t_2)\,dt_1dt_2\right]. \quad (4.136)$$

We now make the change of variables

$$t_1-t_2=z,\quad t_1=u. \quad (4.137)$$

The Jacobian of this transformation is equal to unity. Therefore, the double integral in (4.136) can be represented in terms of the new variable in the form

$$\int_t^{t+\tau}\left[\int_{u-t-\tau}^{u-t}\rho(z)\,dz\right]du. \quad (4.138)$$

We introduce a new function

$$\Phi(y)=\int_0^y\left[\int_0^x\rho(z)\,dz\right]dx \quad (4.139)$$

and examine its basic properties. It must be noted first of all that the derivative of this function at zero is zero. In fact,

$$\left(\frac{d\Phi}{dy}\right)_{y=0}=\left(\int_0^y\rho(z)\,dz\right)_{y=0}=0. \quad (4.140)$$

Next, the function $\Phi(y)$ is even:

$$\Phi(y)=\Phi(-y). \quad (4.141)$$

Indeed,

$$\Phi(-y)=\int_0^{-y}\left[\int_0^x\rho(z)\,dz\right]dx=-\int_0^y\left[\int_0^{-x}\rho(z)\,dz\right]dx=-\int_0^y\left[-\int_0^x\rho(-z)\,dz\right]dx=\int_0^y\left[\int_0^x\rho(z)\,dz\right]dx=\Phi(y). \quad (4.142)$$

With the aid of the finction $\Phi(-y)$ introduced in this manner we transform the double integral (4.138):

$$\int_t^{t+\tau}\left[\int_{u-t-\tau}^{u-t}\rho(z)\,dz\right]du=2\Phi(\tau). \quad (4.143)$$

As a result, the field correlation function (4.136) turns out to equal

$$C(\tau)=\frac{1}{2}E_0^2\cos(\omega_0\tau)\exp[-\langle\Omega^2\rangle\Phi(\tau)]. \quad (4.144)$$

From this expression follow two important conclusions. First, a field with a Gaussian distribution of the frequency fluctuations cannot have a Lorentz spectrum. In fact, a Lorentz spectrum corresponds to a coherence function with a slowly varying part of the form exp $(-|\tau|)$. It follows from (4.144) that in this case

$$\Phi(\tau) = \text{const} \cdot |\tau|.$$

But this is impossible, since in accord with (4.140) $\Phi(\tau)$ must have a zero derivative at zero.

Second, in the approximation considered the field cannot have a Gaussian spectrum. A Gaussian spectrum corresponds to a slowly varying part of the coherence function in the form exp $(-\tau^2)$. Then

$$\Phi(\tau) = \text{const} \cdot \tau^2$$

As follows from (4.139), we should have here

$$\rho(z) = \text{const},$$

which is impossible, since the correlation function of a random process with a zero mean value should tend to zero with increasing argument.

We obtain now the explicit form of the spectrum $G(\omega)$ of the field. By definition we have

$$G(\omega) = \frac{2}{\pi}\int_0^\infty C(\tau)\cos(\omega\tau)\,d\tau. \tag{4.145}$$

Substituting (4.144) in this expression and neglecting the rapidly varying terms in the integrand, we obtain the following expression for the spectrum over the positive frequencies

$$G(\omega) = \frac{1}{2\pi}E_0^2\int_0^\infty \exp[-\langle\Omega^2\rangle\Phi(\tau)]\cos[(\omega-\omega_0)\tau]\,d\tau. \tag{4.146}$$

We transform for convenience to a centered spectral density

$$G(\Delta\omega) = G(\omega_0+\Delta\omega), \quad \Delta\omega = \omega-\omega_0. \tag{4.147}$$

Then

$$G(\Delta\omega) = \frac{1}{2\pi}E_0^2\int_0^\infty \exp[-\langle\Omega^2\rangle\Phi(\tau)]\cos(\Delta\omega\,\tau)\,d\tau. \tag{4.148}$$

Using this expression, we find the connection between the spectrum width $\delta\omega$ and the correlation time τ_f of the frequency fluctuations. To this end, we approximate the correlation function $\rho(\tau)$ by some known function with specified parameter τ_f, use (4.148) to obtain the form of the spectrum $G(\Delta\omega)$ and obtain from the latter the spectrum width $\delta\omega$. By definition, we have

$$\delta\omega = G^{-1}(0)\int_0^\infty G(\Delta\omega)\,d(\Delta\omega), \quad \tau_f = \int_0^\infty \rho(\tau)\,d\tau. \tag{4.149}$$

For $\rho(\tau)$ we introduce the approximating function

$$\rho(\tau) = \exp(-|\tau|/\tau_f). \tag{4.150}$$

Then, after simple calculation, we obtain

$$\Phi(\tau) = \tau\tau_f - \tau_f^2(1-\exp(-\tau/\tau_f)) \tag{4.151}$$

at $\tau \geq 0$. For the region $\tau < 0$ it is necessary to use Eq. (4.141).

We obtain now an expression for $\delta\omega$. In accord with (4.149) we have

$$\delta\omega = \frac{1}{2\pi}E_0^2G^{-1}(0)\int_0^\infty \exp[-\langle\Omega^2\rangle\Phi(\tau)]\int_0^\infty \cos(\Delta\omega\,\tau)\,d(\Delta\omega)\,d\tau = \frac{1}{4}E_0^2G^{-1}(0). \tag{4.152}$$

It is now necessary to find G(0). To this end we substitute (4.151) in (4.148):

$$G(0)=\frac{1}{2\pi}E_0^2\int_0^\infty \exp\left\{-\langle\Omega^2\rangle\left[\tau\tau_f-\tau_f^2\left(1-\exp\left(-\frac{\tau}{\tau_f}\right)\right)\right]\right\}d\tau. \tag{4.153}$$

After rather obvious transformations, we obtain

$$G(0)=\frac{1}{2\pi}E_0^2\tau_f\varepsilon(p), \tag{4.154}$$

where we have introduced the notation

$$p=\langle\Omega^2\rangle\tau_f^2, \tag{4.155}$$

$$\varepsilon(p)=e^p\int_0^1 x^{p-1}e^{-px}dx. \tag{4.156}$$

From (4.152) and (4.154) we obtain the final expression for the width of the field spectrum

$$\delta\omega=\frac{\pi}{2\tau_f\varepsilon(p)} \tag{4.157}$$

and for the field coherence time

$$T_{\text{coh}}=4\tau_f\varepsilon(p). \tag{4.158}$$

At low values of the parameter p it is convenient to use an asymptotic representation of the function ε(p):

$$\varepsilon(p)=\frac{1}{p}+\Gamma(p)\sum_{k=1}^{\infty}\frac{p^k}{\Gamma(k+p+1)}. \tag{4.159}$$

The derivation of this expression is relegated to Appendix VI. At p << 1 we can put

$$\varepsilon(p)\approx\frac{1}{p}.$$

Equation (4.157) is then reduced to the more lucid form

$$\delta\omega=\frac{\pi}{2}\langle\Omega^2\rangle\tau_f. \tag{4.160}$$

It is of interest to note that this estimate of the field spectrum width coincides fully with the estimate obtained in [13] in a somewhat different manner.

Thus, Eqs. (4.157) and (4.158) allow us to relate the results obtained in Subsec. 2 of the present section to the experimentally measured laser-emission spectral width.

CHAPTER 5

METHODS OF SPECKLE INTERFEROMETRY AND ADAPTIVE OPTICS

In the preceding chapters we considered methods of interpreting optical fields whose realization calls for the use of lasers and for which high coherence of the laser radiation is obligatory. The present chapter is devoted to speckle interferometry and to adaptive optics, which first came into use in astronomy as a method of improving the resolving power of large optical telescopes under conditions of atmospheric turbulence and dealing with radiation of natural origin. These methods can be successfully used in laser radars in those cases when no rapid measurements in real times are needed.

5.1. Speckle Interferometry Method

1. General Description of the Method. The speckle-interferometry method was proposed by Labeyrie in 1970 [1] and its high effectiveness was demonstrated experimentally later [2, 3]. We consider first the gist of this method and then proceed to analyze it in detail.

Light from a star or from another object, after passing through a layer of turbulent atmosphere, is focused by the receiving telescope in the focal plane and produces a distorted image in the form of a smeared spot (Fig. 51). The angular dimension of this spot is approximately λ/ρ_0, where λ is the average wavelength of the arriving radiation, and ρ_0 is the coherence radius, introduced in Chap. 1, of the wave front [4, 5]. If the width of the emission spectrum is small enough (see Subsec. 2) and the exposure time necessary to record the image does not exceed the "freezing" time of the atmosphere, individual "grains," frequently called "speckles," will be distinctly recorded in the image. This speckled structure is attributed to interference of the waves coming from different sections of the receiving aperture.

In this method, when observing an object having angular dimensions not larger than one second of angle, there is recorded in the focal plane of the telescope a series of short-exposure photographs. The time of each exposure should be shorter than the atmosphere "freezing" time, and the time interval between two neighboring exposures, on the contrary, should be much longer than this freezing time. The speckled structure of a series of photographs that follow one another varies, a fact attributed to the change of the state of the atmosphere ahead of the telescope aperture (Fig. 5.2).

The spatial distribution of the intensity $I_n(\mathbf{r})$ at each n-th photograph can be expressed in terms of the undistorted intensity distribution $I(\mathbf{r})$ obtained by an ideal optical system with infinite resolving power, and through the scattering function $H(\mathbf{x})$ of the atmosphere + telescope system (see Subsec. 3). The latter, naturally, depends on the number of the photograph. As a result, we obtain

$$I_n(\boldsymbol{r}) = \int I(\boldsymbol{\rho})\, H(\boldsymbol{\rho} - \boldsymbol{r})\, d^2\rho. \tag{5.1}$$

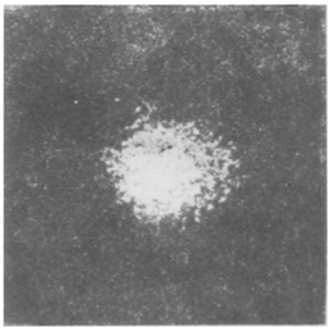


Fig. 5.1

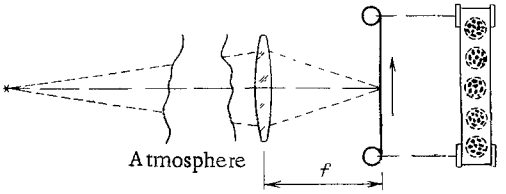


Fig. 5.2

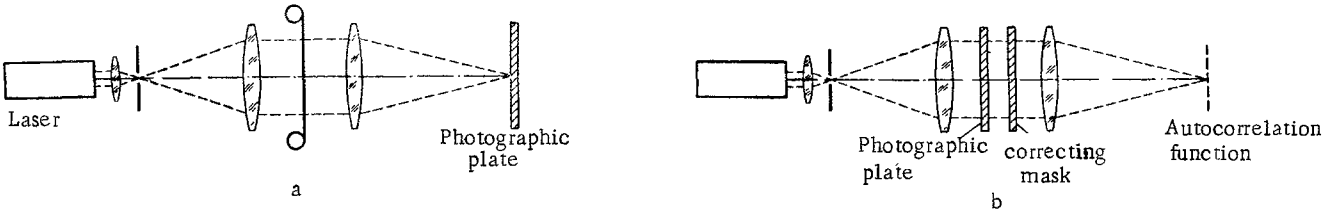


Fig. 5.3

The succeeding processing of the recorded set of images consists of taking the Fourier transform of each image $I_n(\mathbf{r})$. This is done by optical means, as shown in Fig. 5.3a. As a result, the square of the modulus of the spatial spectrum of the image is recorded in the focal plane of the last lens, namely,

$$|V_n(\xi)|^2 = |V(\xi)|^2 |\tau_n(\xi)|^2, \tag{5.2}$$

where $V(\xi)$, $V_n(\xi)$, and $\tau_n(\xi)$ are the spatial spectra of the distorted image and of the atmosphere + telescope system. The function $\tau_n(\xi)$ is called the frequency–contrast characteristic of the system.

By advancing in succession the roll film, all the spatial image spectra obtained in a run are summed on one photographic plate. As a result, the plate is thus a record of the mean-squared modulus of the spatial spectrum of the image

$$\langle |V_n(\xi)|^2 \rangle = |V(\xi)|^2 \langle |\tau_n(\xi)|^2 \rangle. \tag{5.3}$$

The mean-squared modulus of the spatial spectrum of the scattering function of the atmosphere + telescope system $\langle |\tau_n(\xi)|^2 \rangle$ can be obtained in a similar fashion when observing a star that is close to the object, has a comparable brightness, and has an angular dimension much smaller than that of the object. The results of these measurements are used to prepare a correcting mask whose transmission coefficient is equal to

$$\langle |\tau_n(\xi)|^2 \rangle^{-1}. \tag{5.4}$$

The correcting mask is superimposed on the photographic plate with the spatial distribution (5.3), and an inverse Fourier transform is produced (Fig. 5.3b). The result is the square of the modulus of the autocorrelation function of the undistorted image of the object

$$\left| \int I(x) I(x+r) \, d^2x \right|^2. \tag{5.5}$$

This distribution can be used to measure the angular dimensions of the objects and to estimate their shapes. (Similar questions were considered in Chap. 2.)

The resolving power of the speckle-interferometry method is determined by the maximum spatial-frequency modulus $|\xi|_{\max}$, for which the right-hand side of (5.3) is still different from zero. It will be shown below (in Subsec. 3) that for an ideal optical system in the absence of noise we have

$$|\xi|_{\max} = \frac{2\pi D}{\lambda f}, \tag{5.6}$$

where f is the equivalent focal length of the telescope, D is its diameter, and λ is the radiation wavelength. The angular resolving power is therefore λ/D, i.e., equal to the diffractive value. Under real conditions, the recording process is always accompanied by noise which, in fact, determines the minimum distinguishable (i.e., the "zeroth") level of the right-hand side

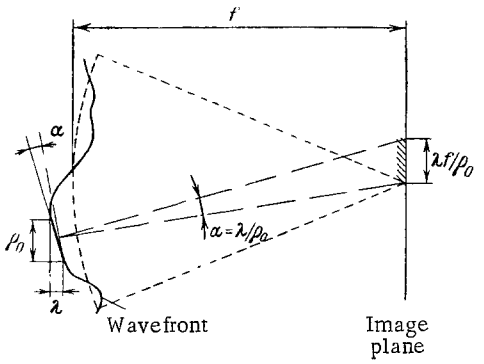


Fig. 5.4

of (5.3). In this case, the resolving power will naturally be less than the diffractive value.

For a more detailed analysis of the speckle-interferometry method it is necessary first to become acquainted with the phenomena that determine the graininess of the image in the focal plane of the telescope.

2. Interference of Fields in the Focal Plane of a Telescope. We consider the field produced at the output of the telescope by a remote pointlike monochromatic source (Fig. 5.4). In the absence of atmospheric turbulence, the wave front of this field is spherical and converges in the image plane. The recorded pattern is in this case the known Airy pattern with a principal maximum of width $\lambda f/D$, where λ is the wavelength, f the focal length of the telescope, and D the diameter of the main mirror. In the presence of atmospheric turbulence the wave front becomes distorted. These distortions can be characterized by a field-coherence radius ρ_0. The physical meaning of this parameter, which was introduced in Chap. 1, is that ρ_0 is the distance in the receiving aperture plane over which the fluctuations of the path difference amount on the average to one wavelength.

Each "almost" flat section of the distorted wave front will propagate in a direction making an angle α with the direction in which this section would propagate in the absence of turbulence (Fig. 5.4). The average absolute value of this angle is, in accord with the definition, λ/ρ_0. As a result, the average dimension of the illuminated region in the image plane is $\lambda f/\rho_0$. Therefore, ρ_0 can also be interpreted as the diameter of an aberration-free lens whose diffractive resolving power is equal to that of a large telescope operating under the conditions of an atmosphere with a given turbulence [6].

The intensity distribution inside the illuminated spot is determined during the time that the atmosphere is "frozen" by the interference of the waves coming from different points of the receiving aperture of the telescope (Fig. 5.5). We consider two sections of the aperture, symmetrical about the axis, and of size ρ_0. The waves from these sections add up in the focal plane, and the path difference at an arbitrary point x is, in first-order approximation,

$$s \approx dx/f + s_0, \tag{5.7}$$

where s_0 is the path difference at the point x = 0 and is determined by the wave-front distortions due to the turbulence of the atmosphere. The distance between the interference fringes produced by the two considered sections is

$$\Delta x = f\lambda/d. \tag{5.8}$$

The distribution of the intensity in the spot is determined by the random locations of the set of interference patterns produced by different sections of the aperture. The result is the random grainy picture shown in Fig. 5.1. The minimum "grain" size is determined by the minimum possible distance between the interference maxima, and is equal, according to (5.8), to $f\lambda/D$, where D is the telescope diameter.

Assume now that the object is not a point, but is extended with an angular dimension γ (Fig. 6). In this case, the interference patterns produced by each of the emitting points on the object surface add up. If γ is less than the angular dimension of the isoplanatism zone, usually about one second of angle, the atmosphere distorts equally the waves coming

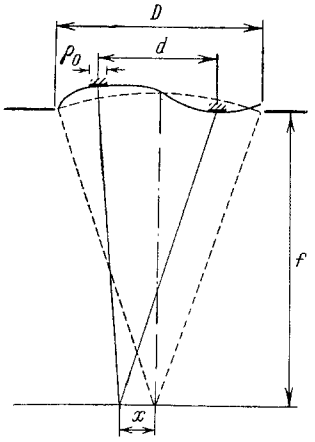


Fig. 5.5

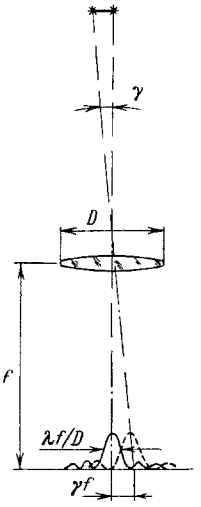


Fig. 5.6

from all the points of the object. The corresponding interference patterns in the focal plane of the telescopes are then identical but shifted relative to one another. The maximum shift is, obviously, γf.

Assume that the radiation on the surface of the object is spatially incoherent, i.e., the phases of all the radiating points are mutually uncorrelated. This is the case in astronomy, where the observed objects are stars. This assumption is sometimes valid also in laser radar. For example, if the spectrum width of the second harmonic of a λ = 1.06-μm laser is 1.0 nm, the coherence length is only 0.25 mm. In this case, the radiation reflected from the target can be regarded as spatially incoherent.

As a result of the incoherent addition of all the interference patterns from the individual points on the objects, the grains of size smaller than γf will vanish. The dimension of the entire illuminated region will not change (provided, of course, that $\gamma \ll \lambda/\rho_0$). Turning again to Eq. (5.8), we arrive at the conclusion that the minimum size γf of the remaining "grains" corresponds to interference of sections on the receiving aperture that are spaced λ/γ apart. At the same time, as follows from the Van-Zittert–Zernike theorem (see Chap. 1), the size of the coherence region of the field produced by an extended spatially incoherent source on the surface of the receiving aperture is also λ/γ. It follows, therefore, that in the considered case the minimum "grain" size in the recorded image corresponds to diffraction not by the entire receiving aperture, but by the element of spatial coherence of the received field.

We examine now the monochromaticity condition under which speckle interferometry can operate. This calls, first of all, for good visibility of the grainy structure of the image, both when the object image is recorded and when the image of the reference point source is recorded. When the radiation wavelength is changed by Δλ the interference pattern is shifted. The shift of "grains" of different size will be different, since they correspond to interference of receiving-aperture sections separated by different distances d. By varying the left- and right-hand sides of (5.8) we obtain the dependence of the displacement δ on d:

$$\delta(d) = f\frac{\Delta\lambda}{d}. \tag{5.9}$$

When recording the image of an incoherent object with angular dimension γ, the minimum "grain" size, as we have seen, is γf. When the wavelength is changed by Δλ, the shift of the interference pattern should be much smaller than the "grain" size, whence

$$\delta(\lambda/\gamma) = f\frac{\Delta\lambda}{\lambda/\gamma} \ll \gamma f$$

or

$$\frac{\Delta\lambda}{\lambda} \ll 1. \tag{5.10}$$

When recording the image of a reference source, the minimum "grain" size is λf/D, whence

$$\delta(D) = f\frac{\Delta\lambda}{D} \ll \frac{\lambda f}{D}$$

or

$$\frac{\Delta\lambda}{\lambda} \ll 1. \tag{5.11}$$

Thus, in both cases, an identical condition, (5.10) or (5.11), which is necessary but not sufficient, must be satisfied. Another condition is that within the limits of the aperture the maximum path-difference fluctuations introduced by the atmosphere must not exceed the coherence length of the received radiation. From this we readily find [4] that

$$\Delta\lambda < \frac{\lambda^2}{aD}. \tag{5.12}$$

At α = 1 second of angle, D = 5 m, and λ = 500 nm we obtain $\Delta\lambda$ < 10 nm. Condition (5.12) is, as a rule, decisive when the radiation bandwidth is chosen.

For the analysis that follows we need some information from the theory of linear optical systems, which will be considered in the next subsection.

3. Scattering Function and Frequency-Contrast Characteristic of a Linear Optical System. Let a remote object be located in the **r** plane and let its image be recorded in the focal plane **x** of an ideal aberration-free lens placed in the plane ρ. If the object is located in the Fraunhofer zone relative to the lens aperture (see Chap. 1), then

$$E(\rho) = c\int E(r)\exp\left(i\frac{k}{L}r\rho\right)d^2r, \tag{5.13}$$

where $E(\rho)$ is the field in the lens plane, $E(\mathbf{r})$ is the field in the object plane, $k = 2\pi/\lambda$ is the wave number, and c is a proportionality coefficient. The field in the focal plane of the lens is [7]

$$E(x) = c_1\int E(\rho)\exp\left(-i\frac{k}{f}\rho x\right)d^2\rho, \tag{5.14}$$

where f is the focal length of the lens and c_1 is a coefficient of no further importance. The integration in this equation is in the plane of the lens. Substituting (5.13) in (5.14) we obtain

$$E(x) = c_2\int E(r)\,h_0\left(\frac{f}{L}r - x\right)d^2r, \tag{5.15}$$

where

$$h_0(x) = \int \exp\left(-i\frac{k}{f}x\rho\right)d^2\rho, \tag{5.16}$$

and $c_2 = cc_1$ is a new constant coefficient. For a lens of radius R the integral (5.16) is easily calculated and leads to the known Airy function

$$h_0(x) = \pi R^2 2J_1\left(\frac{R}{\lambda f}|x|\right)\Big/\frac{R}{\lambda f}|x|. \tag{5.17}$$

Changing the scale in (5.15) along the **r** coordinate:

$$u = \frac{f}{L}r, \tag{5.18}$$

we can write this expression in a more convenient form:

$$E(x) = c_3\int E(u)\,h_0(u-x)\,d^2u, \tag{5.19}$$

where $c_3 = c_2(f/L)^{-2}$. We shall omit hereafter this inessential constant coefficient. The function $h_0(\mathbf{x})$ is called the scattering function of an ideal coherent optical system. It comprises the field distribution produced in the focal plane of a lense by a remote point source.

If the field $E(\mathbf{u})$ is spatially coherent, the following intensity distribution will be recorded on a photographic film (or plate)

$$|E(x)|^2 = \left|\int E(u)\,h_0(u-x)\,d^2u\right|^2. \tag{5.20}$$

If, however. the field E(u) is spatially incoherent, the film record will comprise the intensity averaged over the recording time

$$I(x)=\int\langle E(u_1)E^*(u_2)\rangle h_0(u_1-x)\,h_0^*(u_2-x)\,d^2u_1d^2u_2. \tag{5.21}$$

Since by definition E(u) is spatially incoherent, i.e.,

$$\langle E(u_1)E^*(u_2)\rangle=I(u_1)\delta(u_1-u_2), \tag{5.22}$$

we have

$$I(x)=\int I(u)H(u-x)\,d^2u, \tag{5.23}$$

where I(u) is the average intensity distribution over the object, and H is the so-called scattering function of an incoherent optical system:

$$H(x)=|h_0(x)|^2. \tag{5.24}$$

As already indicated, in astronomy the field in the object plane is always spatially incoherent, but in laser radar this holds true only for a sufficiently broad sounding-radiation spectrum. We shall consider hereafter only this case.

We introduce the image spatial spectrum given by

$$V(\xi)=\int I(x)\exp(i\xi x)\,d^2x, \tag{5.25}$$

where ξ is the spatial frequency. Taking the Fourier transforms of the left- and right-hand sides of (5.23), we obtain

$$V(\xi)=V'(\xi)\tau(\xi), \tag{5.26}$$

where V'(ξ) is the spatial spectrum of the image undistorted by the optical system and τ(ξ) is the so-called frequency–contrast characteristic of the system:

$$\tau(\xi)=\int H(x)\exp(i\xi x)\,d^2x. \tag{5.27}$$

It is known that image details with different dimensions are unequally converted by the optical system. As a rule, the contrast of the large details in the produced image are reproduced without change, whereas the contrast of the small details in the transmitted image decreases. Details with characteristic dimension Δx correspond to a modulus of the spatial frequency

$$|\xi|=\frac{2\pi}{\Delta x}, \tag{5.28}$$

and the contrast of these details in the initial image is determined by the modulus of the spatial spectrum V'(ξ) of the image. **The frequency-contrast characteristic τ(ξ) determines then the contrast of the corresponding details in the transformed image. This characteristic can be either positive or negative. Negative values correspond to the so-called reversal of the contrast** [8].

We obtain now the frequency–contrast characteristic of an ideal optical system. We substitute (5.24) in (5.27):

$$\tau_0(\xi)=\int h_0(x)\,h_0^*(x)\exp(i\xi x)\,d^2x. \tag{5.29}$$

$h_0(x)$ satisfies the relation (5.16) in which the integration is over the aperture of the lens. It is possible to introduce formally in (5.16) the entrance-pupil function of an ideal optical system, using the formula

$$\Pi_0(\rho)=\begin{cases}1, & |\rho|\leqslant R,\\ 0, & |\rho|>R,\end{cases} \tag{5.30}$$

and extending the integration limits to infinity. We then obtain

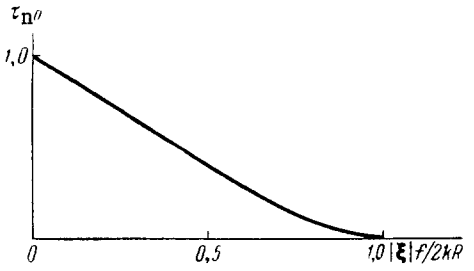


Fig. 5.7

$$h_0(x)=\int \Pi_0(\rho)\exp\left(-i\frac{k}{f}x\rho\right)d^2\rho, \tag{5.31}$$

where the integration is over the entire ρ plane. Substituting this expression in (5.29) and integrating with respect to **x** and **ρ**, we get

$$\tau_0(\xi)=\int \Pi_0(\rho)\,\Pi_0^*\left(\rho+\frac{f}{k}\xi\right)d^2\rho. \tag{5.32}$$

It is customary to use the normalized frequency–contrast characteristic

$$\tau_n(\xi)=\frac{\tau(\xi)}{\tau(0)}=\frac{\int \Pi(\rho)\,\Pi^*\left(\rho+\frac{f}{k}\xi\right)d^2\rho}{\int|\Pi(\rho)|^2\,d^2\rho}. \tag{5.33}$$

For an ideal optical system with an entrance-pupil radius R we have [8]

$$\tau_{n0}(\xi)=1-\frac{2}{\pi}\arcsin\left(\frac{|\xi|f}{2kR}\right)-\frac{2}{\pi}\frac{|\xi|f}{2kR}\left[1-\left(\frac{|\xi|f}{2kR}\right)^2\right]^{1/2}. \tag{5.34}$$

A plot of this function is shown in Fig. 5.7. It can be seen that the frequency–contrast characteristic is equal to zero for all spatial frequencies higher than 2kR/f. This value, in accord with (5.28), determines the minimum size of the image details that can be resolved by an ideal optical system

$$\Delta x_{\min}=\frac{\lambda f}{D},$$

which corresponds, as expected, to the diffractive resolving power.

Assume now that a turbulent-atmosphere layer is located ahead of the entrance pupil of an ideal optical system. If the size of the investigated object is such that it is located entirely in the isoplanatism region, the influence of the atmosphere reduces to placing in front of the telescope an amplitude–phase transparency with a complex transmission coefficient exp ψ(ρ) (see Subsec. 2 of Sec. 2.1). The entrance-pupil function of the atmosphere + telescope system takes now a somewhat different form

$$\Pi(\rho)=\begin{cases}\exp\psi(\rho), & |\rho|\leqslant R,\\ 0, & |\rho|>R.\end{cases} \tag{5.35}$$

We confine ourselves for simplicity in this subsection to consideration of only phase fluctuation of the atmosphere, when $\psi(\rho)=i\varphi(\rho)$, and investigate the behavior of the frequency–contrast characteristic in the region of low and high spatial frequencies. We point out first of all that at $|\xi|f/k<\rho_0$ the frequency–contrast characteristic is close to its ideal value in the absence of atmosphere. In fact,

$$\begin{aligned}\tau(\xi)&=\int \Pi_0(\rho)\,\Pi_0^*\left(\rho+\frac{f}{k}\xi\right)\exp\left[i\varphi(\rho)-i\varphi\left(\rho+\frac{f}{k}\xi\right)\right]d^2\rho\approx\\ &\approx\int \Pi_0(\rho)\,\Pi_0^*\left(\rho+\frac{f}{k}\xi\right)\exp\left[i\varphi(\rho)-i\varphi(\rho)\right]d^2\rho=\int \Pi_0(\rho)\,\Pi_0^*\left(\rho+\frac{f}{k}\xi\right)d^2\rho,\end{aligned} \tag{5.36}$$

which coincides with expression (5.32) for the frequency–contrast characteristic of an ideal optical system.

At $\rho_0<|\xi|f/k<2R$, the frequency–contrast characteristic differs strongly from ideal and can take on both positive and negative values. We obtain the mean value of τ(ξ) in this region. We have

$$\langle\tau(\xi)\rangle=\int\Pi_0(\rho)\,\Pi_0^*\left(\rho+\frac{f}{k}\xi\right)\left\langle\exp\left[i\varphi(\rho)-\right.\right.$$
$$\left.\left.-i\varphi\left(\rho+\frac{f}{k}\xi\right)\right]\right\rangle d^2\rho\approx\int\Pi_0(\rho)\,\Pi_0^*\left(\rho+\frac{f}{k}\xi\right)\langle\exp[i\varphi(\rho)]\rangle\times\left\langle\exp\left[-i\varphi\left(\rho+\frac{f}{k}\xi\right)\right]\right\rangle d^2\rho=0 \tag{5.37}$$

by virtue of the uniform distribution of φ over the interval $(0, 2\pi)$. This circumstance is illustrated in Fig. 5.8a, which shows two different realizations of normalized frequency–contrast characteristics, τ_{n1}, τ_{n2}, and their mean value $\langle\tau_n\rangle$. For comparison, the dashed line shows τ_{no} for an ideal optical system.

We return now to the averaging process described in Subsec. 1 and see what happens if the quantity averaged is not the square of the modulus of the spatial spectrum of the image (5.2), but the image (5.1) itself. It might seem that it is possible to reproduce in similar fashion the undistorted image of the object with diffractive resolving power. This is not so, however. In fact, the averaging (5.1) is equivalent to averaging of the spatial spectra

$$\langle V_n(\xi)\rangle=V(\xi)\langle\tau_n(\xi)\rangle. \tag{5.38}$$

It follows from (5.57) that $\langle\tau_n(\xi)\rangle = 0$ at $|\xi| > k\rho_0/f$. This means that we can construct a correcting mask $\langle\tau_n(\xi)\rangle^{-1}$ and reconstruct the initial spatial spectrum only in the interval $0 \leq |\xi| \leq k\rho_0/f$. It hence follows that the angular resolving power is, in this case, λ/ρ_0, i.e., it is determined as before by the atmospheric turbulence. The favorable aspect of this reconstruction is only the slight improvement in the contrast of those fine details whose angular dimension is not smaller than λ/ρ_0.

The situation is different when it comes to averaging the square of the modulus of the spatial spectrum. In this case, $\langle|\tau_n(\xi)|^2\rangle$ differs from zero all the way to the maximum spatial frequencies $|\xi| = kD/f$, where D is the diameter of the telescope, so that the diffractive resolving power can be reached. This is illustrated in Fig. 5.8b.

4. Mean Value of the Modulus of the Frequency–Contrast Characteristic. In accordance with (5.35), we have

$$\langle|\tau(\xi)|^2\rangle=\int\Pi_0(\rho_1)\,\Pi_0^*\left(\rho_1+\frac{f}{k}\xi\right)\Pi_0^*(\rho_2)\,\Pi_0\left(\rho_2+\frac{f}{k}\xi\right)\times$$
$$\times\left\langle\exp\left[\psi(\rho_1)+\psi^*(\rho_2)+\psi^*\left(\rho_1+\frac{f}{k}\xi\right)+\psi\left(\rho_2+\frac{f}{k}\xi\right)\right]\right\rangle d^2\rho_1 d^2\rho_2. \tag{5.39}$$

We separate in the complex function $\psi(\rho)$ the real and imaginary parts which are responsible for the amplitude and phase fluctuations of the field:

$$\psi(\rho)=a(\rho)+i\varphi(\rho). \tag{5.40}$$

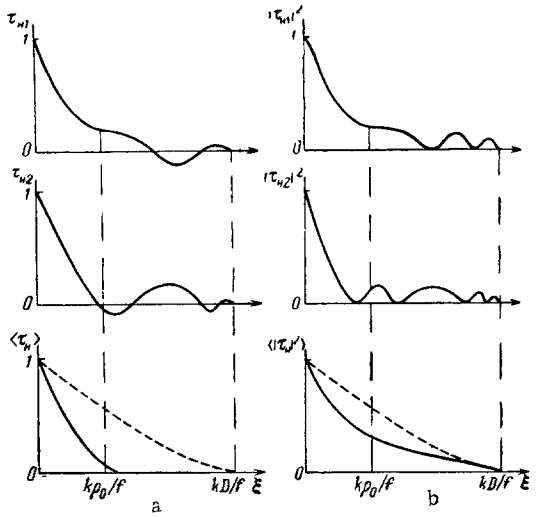


Fig. 5.8

As noted in Chap. 1, it can be assumed that $a(\rho)$ and $\varphi(\rho)$ are mutually independent Gaussian random functions. In this case,

$$\left\langle \exp\left[\psi(\rho_1)+\psi^*(\rho_2)+\psi^*\left(\rho_1+\frac{f}{k}\xi\right)+\psi\left(\rho_2+\frac{f}{k}\xi\right)\right]\right\rangle = \left\langle \exp\left[a(\rho_1)+a(\rho_2)+a\left(\rho_1+\frac{f}{k}\xi\right)+a\left(\rho_2+\frac{f}{k}\xi\right)\right]\right\rangle \times \left\langle \exp\left\{i\left[\varphi(\rho_1)-\varphi(\rho_2)-\varphi\left(\rho_1+\frac{f}{k}\xi\right)+\varphi\left(\rho_2+\frac{f}{k}\xi\right)\right]\right\}\right\rangle. \tag{5.41}$$

It was shown in Chap. 1 that the following relation holds under the condition $\langle\varphi(\rho)\rangle = 0$

$$\langle \exp[i\varphi(\rho)]\rangle = \exp\left[-\frac{1}{2}\langle\varphi^2(\rho)\rangle\right]. \tag{5.42}$$

We can similarly obtain

$$\langle \exp(a(\rho))\rangle = \exp[\langle a(\rho)\rangle]\exp\left[\frac{1}{2}\langle(a(\rho)-\langle a(\rho)\rangle)^2\rangle\right]. \tag{5.43}$$

We introduce for $a(\rho)$ and $\varphi(\rho)$ the correlation functions

$$\begin{aligned} C_a(\rho_1-\rho_2) &= \langle[a(\rho_1)-\langle a\rangle][a(\rho_2)-\langle a\rangle]\rangle, \\ C_\varphi(\rho_1-\rho_2) &= \langle[\varphi(\rho_1)\varphi(\rho_2)]\rangle. \end{aligned} \tag{5.44}$$

Applying (5.43) to the first factor in the right-hand side of (5.41), we obtain, after simple transformations,

$$\begin{aligned} &\left\langle \exp\left[a(\rho_1)+a(\rho_2)+a\left(\rho_1+\frac{f}{k}\xi\right)+a\left(\rho_2+\frac{f}{k}\xi\right)\right]\right\rangle = \\ &= \exp\left[4\langle a\rangle + 2C_a(0) + 2C_a(\rho_1-\rho_2) + 2C_a\left(\frac{f}{k}\xi\right) + C_a\left(\rho_1-\rho_2-\frac{f}{k}\xi\right) + C_a\left(\rho_1-\rho_2+\frac{f}{k}\xi\right)\right]. \end{aligned} \tag{5.45}$$

Similarly, using (5.42), we get

$$\begin{aligned} &\left\langle \exp\left\{i\left[\varphi(\rho_1)-\varphi(\rho_2)-\varphi\left(\rho_1+\frac{f}{k}\xi\right)+\varphi\left(\rho_2+\frac{f}{k}\xi\right)\right]\right\}\right\rangle = \\ &= \exp\left[-2C_\varphi(0) + 2C_\varphi(\rho_1-\rho_2) + 2C_\varphi\left(\frac{f}{k}\xi\right) - C_\varphi\left(\rho_1-\rho_2-\frac{f}{k}\xi\right) - C_\varphi\left(\rho_1-\rho_2+\frac{f}{k}\xi\right)\right]. \end{aligned} \tag{5.46}$$

We now introduce the correlation function for $\psi(\rho)$

$$C_\psi(\rho_1-\rho_2) = \langle[\psi(\rho_1)-\langle\psi\rangle][\psi^*(\rho_2)-\langle\psi^*\rangle]\rangle. \tag{5.47}$$

Expanding the brackets in the right-hand side, we arrive, after simple transformations, at the relation

$$C_\psi(0) - C_\psi(\rho_1-\rho_2) = \langle|\psi|^2\rangle - \langle\psi(\rho_1)\psi(\rho_2)\rangle. \tag{5.48}$$

In Chap. 1, in the derivation of the expression for the correlation function of a field passing through a layer of turbulent atmosphere, we obtained

$$\exp\{-[\langle|\psi|^2\rangle - \langle\psi(\rho_1)\psi^*(\rho_2)\rangle]\} = \exp[-(|\rho_1-\rho_2|/\rho_0)^{5/3}], \tag{5.49}$$

where ρ_0 is defined by (1.179) or (1.182). For the case of cosmic objects

$$\rho_0 = (1.45C_n^2k^2L)^{-3/5}, \tag{5.50}$$

where C^2_n is the structure constant of the fluctuations of the atmosphere's refractive index, $k = 2\pi/\lambda$ is the wave number, and L is the length of the section of the atmosphere. Comparing (5.49) with (5.48), we get

$$C_\psi(0) - C_\psi(\rho_1-\rho_2) = (|\rho_1-\rho_2|/\rho_0)^{5/3}. \tag{5.51}$$

We express C_ψ in terms of C_a and C_φ. To this end we substitute (5.40) in (5.47) and expand the brackets. Since the mean value of $\varphi(\rho)$ is zero by definition, we have

$$C_\psi(\rho_1-\rho_2) = C_a(\rho_1-\rho_2) + C_\varphi(\rho_1-\rho_2). \tag{5.52}$$

We now introduce the structure function of the random quantity $\mathbf{x}(\rho)$:

$$\mathscr{D}_x(\rho_1-\rho_2)=\langle|x(\rho_1)-x(\rho_2)|^2\rangle. \tag{5.53}$$

It is easy to verify that

$$\begin{aligned}\mathscr{D}_\psi(\rho_1-\rho_2)&=2C_\psi(0)-2C_\psi(\rho_1-\rho_2),\\ \mathscr{D}_a(\rho_1-\rho_2)&=2C_a(0)-2C_a(\rho_1-\rho_2),\\ \mathscr{D}_\varphi(\rho_1-\rho_2)&=2C_\varphi(0)-2C_\varphi(\rho_1-\rho_2).\end{aligned} \tag{5.54}$$

Whence, taking (5.52) into account, we get

$$\mathscr{D}_\psi(\rho_1-\rho_2)=\mathscr{D}_a(\rho_1-\rho_2)+\mathscr{D}_\varphi(\rho_1-\rho_2). \tag{5.55}$$

In addition, in accord with (5.51),

$$\mathscr{D}_\psi(\rho_1-\rho_2)=2(|\rho_1-\rho_2|/\rho_0)^{5/3}. \tag{5.56}$$

Before we proceed to a direct calculation of the mean-squared frequency–contrast characteristic of the system, we must make one more remark. We shall show that

$$\langle a\rangle=-C_a(0). \tag{5.57}$$

We start from the fact that relation (5.49) is valid. Then, obviously,

$$\langle\exp[\psi(\rho_1)+\psi^*(\rho_2)]\rangle=\exp\{-[C_\psi(0)-C_\psi(\rho_1-\rho_2)]\}. \tag{5.58}$$

Substituting (5.40) in the left-hand side of this equality and averaging independently of the right-hand side, we obtain

$$\exp[2\langle a\rangle+C_a(0)-C_\varphi(0)+C_a(\rho_1-\rho_2)+C_\varphi(\rho_1-\rho_2)]. \tag{5.59}$$

On the other hand, this expression should equal the right-hand side of (5.58). Hence, with allowance for (5.52), we should get (5.57).

After all this lengthy but necessary reasoning, we proceed to calculate $|\tau<(\xi)|^2>$. Multiplying (5.45) by (5.46) and combining C_α and C_φ, we get

$$\langle|\tau(\xi)|^2\rangle=\int\Pi_0(\rho_1)\Pi_0^*\left(\rho_1+\frac{f}{k}\xi\right)\Pi_0^*(\rho_2)\Pi_0\left(\rho_2+\frac{f}{k}\xi\right)Q(\rho_1-\rho_2,\ \xi)\exp[4C_a(0)]\,d^2\rho_1 d^2\rho_2, \tag{5.60}$$

$$\begin{aligned}Q(\rho_1-\rho_2,\ \xi)=\exp\Big\{-\Big[\mathscr{D}_\psi\Big(\frac{f}{k}\xi\Big)+\mathscr{D}_\psi(\rho_1-\rho_2)+\\ +\frac{1}{2}\mathscr{D}_\psi\Big(\rho_1-\rho_2-\frac{f}{k}\xi\Big)+\frac{1}{2}\mathscr{D}_\psi\Big(\rho_1-\rho_2+\frac{f}{k}\xi\Big)\Big]+\mathscr{D}_\varphi\Big(\rho_1-\rho_2-\frac{f}{k}\xi\Big)+\mathscr{D}_\varphi\Big(\rho_1-\rho_2+\frac{f}{k}\xi\Big)\Big\}.\end{aligned} \tag{5.61}$$

When observing cosmic objects at close to zenith angle, the effective length of the atmospheric section is $L \lesssim 10$ km. In this case, for any point of the atmospheric section the Fresnel number of the receiving aperture is much larger than unity. Thus, at $L = 10$ km, $D = 1$ m, and $\lambda = 0.5$ μm, we have $D^2/\lambda L = 200$. In this case, the following equation, which we present without proof [9], is valid:

$$\mathscr{D}_\psi(\rho_1-\rho_2)=\mathscr{D}_\varphi(\rho_1-\rho_2). \tag{5.62}$$

Taking this into account, we have

$$Q(\rho_1-\rho_2,\ \xi)=\exp\Big[-\mathscr{D}_\psi\Big(\frac{f}{k}\xi\Big)-\mathscr{D}_\psi(\rho_1-\rho_2)+\frac{1}{2}\mathscr{D}_\psi\Big(\rho_1-\rho_2-\frac{f}{k}\xi\Big)+\frac{1}{2}\mathscr{D}_\psi\Big(\rho_1-\rho_2+\frac{f}{k}\xi\Big)\Big], \tag{5.63}$$

with $Q(\rho_1-\rho_2,\ 0)=1$. We shall analyze subsequently the normalized frequency–contrast characteristic, equal to

$$\begin{aligned}\langle|\tau_n(\xi)|^2\rangle=\int\Pi_0(\rho_1)\Pi_0^*\left(\rho_1+\frac{f}{k}\xi\right)\Pi_0^*(\rho_2)\Pi_0\left(\rho_2+\frac{f}{k}\xi\right)\times\\ \times Q(\rho_1-\rho_2,\ \xi)\,d^2\rho_1 d^2\rho_2\Big[\int|\Pi_0(\rho)|^2\,d^2\rho\Big]^{-2},\\ \langle|\tau_n(0)|^2\rangle=1.\end{aligned} \tag{5.64}$$

In accordance with (5.63), (5.54), and (5.51), the explicit expression for $Q(\rho_1 - \rho_2, \xi)$ is

$$Q(\rho_1-\rho_2,\ \xi)=\exp\left[-2\left(\frac{f|\xi|}{k\rho_0}\right)^{5/3}-2\left(\frac{|\rho_1-\rho_2|}{\rho_0}\right)^{5/3}+\left(\frac{\left|\rho_1-\rho_2-\frac{f}{k}\xi\right|}{\rho_0}\right)^{5/3}+\left(\frac{\left|\rho_1-\rho_2+\frac{f}{k}\xi\right|}{\rho_0}\right)^{5/3}\right], \quad (5.65)$$

with ρ_0 defined by Eq. (5.50). Numerical integration using Eqs. (5.64) and (5.65) was carried out in [9], using the parameters of the Mount Palomar Observatory, namely $r_0 \equiv 2.1\rho_0 = 0.13$ m at a wavelength $\lambda = 0.5$ μm, and receiving-telescope diameters 0.15, 1.5, and 5.0 m. To analyze the results it is convenient to introduce two dimensionless parameters:

$$q=\frac{f|\xi|}{kD}, \qquad \alpha=\frac{D}{\rho_0}. \quad (5.66)$$

The values of the parameter α for the three indicated telescope diameters at $\rho_0 = 0.062$ m are, respectively, 2.457, 24.57, and 73.08. Figure 5.9 shows the dependence of $\langle|\tau_n|^2\rangle$ on q at different values of α. For comparison, the same figure shows the q dependence of the mean squared short-exposure frequency–contrast characteristic [9]

$$\langle\tau_n\rangle^2=\exp\left[-2q^{5/3}\alpha^{5/3}(1-q^{1/3})\right]\tau_{n0}^2(q), \quad (5.67)$$

where the normalized frequency–contrast characteristic $\tau_{n0}(q)$ of an ideal optical system is given by (5.34). The function $\langle\tau_n\rangle$ determines the average quality of the images obtained by ordinary methods within an exposure time shorter than the atmosphere freezing time. The principal result of the analysis is that, for high spatial frequencies at $q >> \alpha^{-1}$, the function $\langle|\tau_n|^2\rangle$ exceeds by several orders the value $\langle\tau_n\rangle^2$ which tends rapidly to zero in the indicated region. This means that the speckle-interferometry method provides a much higher resolution than ordinary observation methods.

5. Asymptotic Expression for $\langle|\tau_n|^2\rangle$. The condition $q >> \alpha^{-1}$ corresponds to the previously discussed (in Subsec. 3) condition $|\xi| >> k\rho_0/f$. In this range of ξ one can obtain a relatively simple asymptotic expression for $\langle|\tau_n|^2\rangle$, which will in fact determine the resolution.

We turn to Eq. (5.64). We note first that

$$\left[\int|\Pi_0(\rho)|^2d^2\rho\right]^{-2}=\frac{16}{\pi^2D^4}. \quad (5.68)$$

We introduce next a new coordinate frame **r**, **ρ**, **q** defined as

$$D\boldsymbol{r}=\rho_1-\rho_2, \quad D\boldsymbol{\rho}=\rho_1+\rho_2, \quad D\boldsymbol{q}=\frac{f}{k}\xi \quad (5.69)$$

and a new function

$$S(\boldsymbol{q},\ \boldsymbol{r})=\int\Pi_0\left[\frac{D}{2}(\boldsymbol{r}+\boldsymbol{\rho})\right]\Pi_0^*\left[\frac{D}{2}(\boldsymbol{r}+\boldsymbol{\rho})+D\boldsymbol{q}\right]\Pi_0^*\left[\frac{D}{2}(\boldsymbol{\rho}-\boldsymbol{r})\right]\Pi_0\left[\frac{D}{2}(\boldsymbol{\rho}-\boldsymbol{r})+D\boldsymbol{q}\right]d^2\rho. \quad (5.70)$$

In this notation, Eq. (5.64) reduces to the form

$$\langle|\tau_n(\boldsymbol{q})|^2\rangle=\frac{16}{\pi^2}\int S(\boldsymbol{q},\ \boldsymbol{r})\exp\left[-2(q\alpha)^{5/3}-2(r\alpha)^{5/3}+(|\boldsymbol{r}-\boldsymbol{q}|\alpha)^{5/3}+(|\boldsymbol{r}+\boldsymbol{q}|\alpha)^{5/3}\right]d^2r, \quad (5.71)$$

where account is taken of the fact that the Jacobian of the transformation (5.69) is equal to D^4. We transform in (5.71) to polar coordinates in the integrand, and reckon the angle θ, which defines the direction of the vector **r**, from the arbitrarily chosen direction of the vector **q** (Fig. 5.10). The intermediate calculations are relegated to Appendix VII, and here we give the final result:

$$\langle|\tau_n(q)|^2\rangle=\frac{64}{\pi^2}\int_0^{\pi/2}d\theta\int_0^{s(q,\theta)}S(q,\ r,\ \theta)\times$$

$$\times\exp\left[-2(q\alpha)^{5/3}-2(r\alpha)^{5/3}+(\sqrt{r^2+q^2-2rq\cos\theta}\,\alpha)^{5/3}+(\sqrt{r^2+q^2+2rq\cos\theta}\,\alpha)^{5/3}\right]rdr, \quad (5.72)$$

where

$$\varepsilon(q,\ \theta) = -q\cos\theta + \sqrt{1 - q^2\sin^2\theta}. \tag{5.73}$$

An analysis of the argument of the exponential in (5.72) shows that under the condition $q\alpha >> 1$ the integrand is close to its maximum value at $r << q$; on the contrary, at $r \gtrsim q$, the integrand is small and can be disregarded in the integration. Taking this circumstance into account, we expand the exponential in powers of the small parameter r/q. When the square root is expanded, the terms linear in r/q cancel one another. Therefore, the expansion of the square root must be carried out up to the terms with the second derivative inclusive. We then get

$$\langle|\tau_n(q)|^2\rangle = \frac{64}{\pi^2}\int_0^{\pi/2} d\theta \int_0^{\varepsilon(q,\theta)} S(q,\ r,\ \theta)\exp\left\{-2(r\alpha)^{5/3}\left[1 - \left(\frac{r}{q}\right)^{1/3}\left(\frac{5}{6} - \frac{5}{18}\cos^2\theta\right)\right]\right\} r\,dr. \tag{5.74}$$

Since we are considering the region $q < 1$, it follows from the condition $q\alpha >> 1$ that $\alpha >> 1$. For the integrand in (5.74) to be close to its maximum we must therefore have $r << 1$. We can thus put in the integrand

$$S(q,\ r,\ \theta) \approx S(q,\ 0,\ \theta) = S(q,\ 0,\ 0). \tag{5.75}$$

However, by virtue of the definition (5.70), the function S(q, 0, 0) is equal to the frequency–contrast characteristic of an ideal system: $S(q, 0, 0) = \tau_0(q)$. Introducing the normalized frequency–contrast characteristic of an ideal system $\tau_{n0}(q)$, we have

$$S(q,\ 0,\ 0) = \frac{\pi}{4}\tau_{n0}(q), \quad \tau_{n0}(0) = 1. \tag{5.76}$$

We note that, as expected, $\varepsilon(q, \theta) = 0$ at $q = 1$ and $\langle|\tau_n(1)|^2\rangle = 0$. If, however, we consider the region $\alpha^{-1} << q < 1$, we have $\alpha\varepsilon(q, \theta) >> 1$ and in this approximation we get from (5.74) the final equation:

$$\langle|\tau_n(q)|^2\rangle \approx \frac{1.916}{\alpha^2}\tau_{n0}(q). \tag{5.77}$$

A more exact expression with terms of next order in $(\alpha q)^{-1}$ is given in [9].

From (5.77) follow directly several important conclusions. First, the mean-squared modulus of the frequency–contrast characteristic of the system, which plays the principal role in the speckle-interferometry method, is proportional to the first power of the frequency–contrast characteristic of an ideal optical system and differs thus from zero up to the limiting values of the spatial frequencies. This means that the speckle interferometry system can yield a near-diffractive resolution.

Second, for each value of the spatial frequency ξ there exists an optimal value of the telescope diameter D, independent of ρ_0 at which $\langle|\tau_n(q)|^2\rangle$ is a maximum. In fact, multiplying the numerator and denominator of (5.77) by q^2 and noting that αq does not depend on D, we conclude that the desired optimum is determined by the maximum of the function $q^2\tau_{n0}(q)$. It is reached at $q = 0.58$, while at q equal to 0.9 of its maximum we get $q_1 = 0.46$ and $q_2 =$

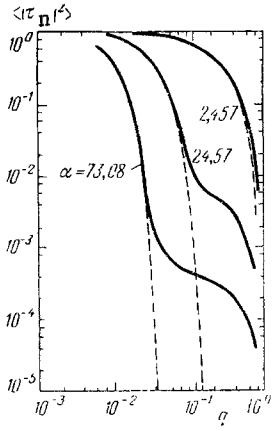


Fig. 5.9

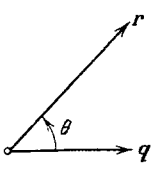


Fig. 5.10

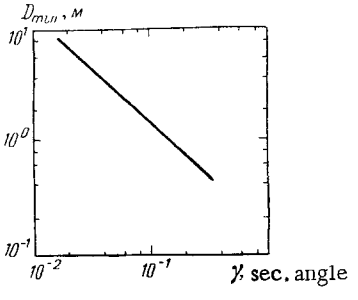


Fig. 5.11

0.70. The latter require, other conditions being equal, a minimum diameter D_{min} of the telescope. If the purpose is to find a minimum telescope diameter D_{min} that would ensure for elements with angle dimension γ a contrast not less than 0.9 of the maximum, then

$$D_{min} = 1.43\lambda\gamma^{-1}. \tag{5.78}$$

This relation is plotted in Fig. 5.11 for λ = 0.5 μm.

6. Experimental Results. By now a rather large amount of experimental material has been accumulated on the use of speckle interferometry for the measurement of starts and of the distances between binary stars with angle dimensions 0.02-0.05 second of angle [2, 3, 10]. To record a set of images one chooses, as a rule, the exposure range 0.125-0.001 sec. Either a television image receiver is chosen with the signal recorded subsequently on video tape, or else a motion-picture camera joined to an image converter. The observation is through a light filter with a wide passband on the order of 30 nm, which ensures for most objectives satisfaction of not only the condition (5.11) but also of (5.12).

Several tens and sometimes even several hundreds (up to 250) [2] of images are recorded. Half of them are images of the observed objects and the other are images of a reference star. The angle distance between the object and the reference star should not exceed 1 sec of angle.

This is followed by preparation of image Fourier-transform negatives, as described in Subsec. 1. The obtained Fourier transforms are averaged by multiple exposure of one photographic plate. This operation is highly critical and photoplates with a special light-sensitive emulsion having a large dynamic range are chosen for them. Analysis of Eq. (5.77) shows that to obtain an approximate resolving power $0.03\lambda/\rho_0$ relative intensities in the range from 1 to $6\cdot10^{-5}$ must be recorded. Photographic materials and standard development have a dynamic range not larger than 10^3. In addition, the inevitable light scattering in the photosensitive emulsion must be eliminated. This is done by using a special opaque screen. To be able to use a screen that is not too large, high-quality optics with low aberration is used. Nonetheless, the central and outer parts of the picture, "spoiled" respectively by the central maximum and by the noise, are not used.

In the first studies [2, 3] the angular dimensions of the objects were determined not from the autocorrelation function (5.5) but from the mean-squared modulus of the spatial spectrum, corrected relative to the reference star. This was done by "fitting" the central part of the profile of the picture to the theoretical profile of a uniform disk. If a picture width at the level 0.5 — w is used for this "fitting," the diameter α of the corresponding uniformly illuminated disk is [3]

$$\alpha = 1.02\lambda_0 f'/wf, \tag{5.79}$$

where λ_0 is the wavelength of the laser used for the Fourier transformation, while f and f' are, respectively, the focal distances of the telescope and of the Fourier-transforming optics.

Binary objects are easily identified by the presence of characteristic bands in the spatial spectrum, and the angle distance between binary objects are determined from the period of the bands.

5.2. Adaptive Optics

1. Principles of Adaptive Wavefront Correction. To compensate for the effect of the atmosphere on the resolving power of large telescopes for observation of cosmic objects, Babcock proposed in 1953 the use of a flexible mirror with feedback [11, 12]. The gist of assumption is the following. It is known (see Sec. 5.1) that for rays entering the receiving aperture within the isoplanatism angle, the turbulence of the atmosphere can be described by the amplitude—phase transparency approximation $\exp[\psi(\rho)]$, where ρ is the coordinate on the receiving aperture. The deterioration of the resolution of telescopes is due principally to phase fluctuations $\exp[i\varphi(\rho)]$.

For an object of arbitrary and unknown shape, it is impossible to separate the amplitude—phase distortions $\psi(\rho)$ due to the atmosphere from the amplitude—phase distribution due to the object. However, if a point source (say a star not resolvable by the telescope) is located in the isoplanatism zone alongside the investigated object, the situation changes. In the absence of the atmosphere the point source produces at the receiving aperture a spherical

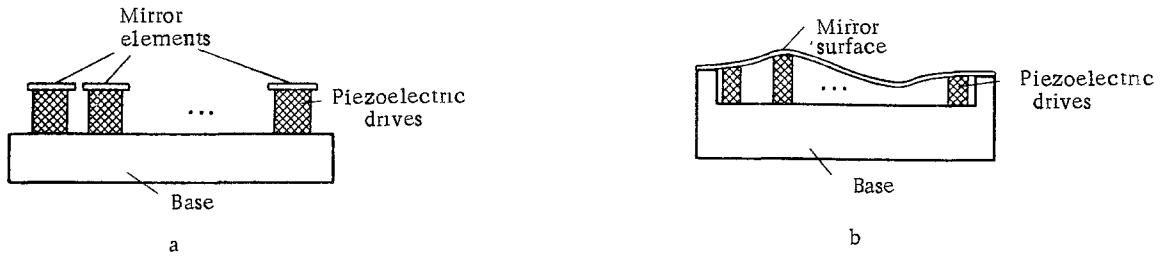


Fig. 5.12

wave front that can be regarded as planar for remote objects (such as stars). Consequently, by separating the radiation from this source and measuring its phase distribution on the receiving aperture, we obtain the distortions due to the atmosphere. Since the point source is located in the isoplanatism zone near the object, the atmospheric phase distortions determined by the factor $\exp[i\varphi(\rho)]$ are identical for the point source as well as for the investigated object. Introducing a phase transparency $\exp[-i\varphi(\rho)]$ in front of the receiving aperture we can thus compensate for the phase fluctuations of the atmosphere and make the resolution of the optical system close to diffractive. The influence of the phase component of the atmospheric turbulence will be compensated for all objects having an adjacent point source in the isoplanatism zone.

The necessary phase correction of the wavefront is effected most conveniently with the aid of a controllable ("flexible") mirror having a variable reflecting-surface relief. Since the state of the atmosphere varies with time, the correction must be effected during the time that the atmosphere is frozen.

The method described is known as the adaptive-optics method. It can be seen that its idea is simple and understandable. Twenty years have elapsed, however, before the start of its development in laboratories. The point is that construction of adaptive optical systems calls for the solution of many major engineering and technical problems. These are outside the scope of the present book, the more so since they are well covered in [13, 14]. We shall therefore analyze only a number of fundamental features that determine the effectiveness of any specific adaptive optic system intended to obtain images of objects.

2. Error of Static Wavefront Correction. The principal element of any adaptive optic system is an adaptive (active) mirror intended to correct the wavefront. Adaptive mirrors were initially constructed in the form of a set of independently controllable (e.g., piezoelectric) "pistons" which displaced individual mirror elements in a vertical direction (Fig. 5.12a) [15]. The initial wavefront correction of such mirrors, however, will be shown below to be subject to large errors. The adaptive mirrors now employed have therefore a flexible reflecting surface (Fig. 5.12b) [13]. The main question that arises in the design of adaptive mirrors is how many individual drives (pushers) are needed to achieve the required wavefront-correction accuracy. Obviously, this depends on the shape of the corrected wave front and on the bending of the adaptive-mirror upon application of the control signal.

Following [16], we assume the adaptive mirror to consist of N discrete drives. When a unit control signal is applied to the j-th drive, the phase will be corrected in the plane of the coordinate ρ. The correction is defined by the so-called response function $R_j(\rho)$ of the j-th drive (Fig. 5.13). If the correcting wave is incident on the adaptive mirror at an angle α, we have

$$R_j(\rho) = 2h_j(\rho)/\lambda \cos\alpha, \tag{5.80}$$

where $h_j(\rho)$ is the displacement produced on the surface of the additive mirror at the point ρ by the j-th drive in response to a unit control signal. The total phase-correction function at the instant of time t is

$$\varphi_0(\rho, t) = \sum_{j=1}^{N} a_j(t) R_j(\rho), \tag{5.81}$$

where $a_j(t)$ is the amplitude of the signal that controls the j-th drive.

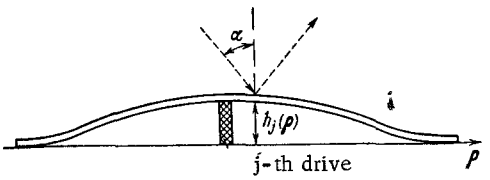


Fig. 5.13

Let the phase distribution of the wave incident on the adaptive mirror be $\varphi(\rho, t)$. Then the distribution of the phase in the reflected wave, i.e., the correction error, is given by

$$\varepsilon(\rho, t) = \varphi(\rho, t) - \varphi_0(\rho, t). \tag{5.82}$$

We consider first the ability of the adaptive mirror to effect static correction of the phase distortions, i.e., how well-prescribed phase distributions can be corrected. We specify phase distributions in the form of tilts, focusings, and third-order aberrations. In the general case, for a mirror in the form of circle with unity radius [7], we have

$$\varphi(\rho, \theta) = C_t\rho\cos\theta + C_f\rho^2 + C_s\rho^4 + C_a\rho^2\cos^2\theta + C_c\rho^3\cos\theta, \tag{5.83}$$

where C_t, C_f, C_s, C_a, and C_c are coefficients that determine, respectively, the tilt, focusing, spherical aberration, astigmatism, and coma; ρ and θ are the polar coordinates in the plane of the mirror. It is convenient to specify each aberration separately in such a way that the mean-squared deviation of the wavefront from a plane within the aperture be equal to the wavelength λ:

$$\left[\frac{1}{\pi}\int_0^{2\pi}\int_0^1 \varphi^2(\rho, \theta)\,\rho\,d\rho\,d\theta\right]^{1/2} = 2\pi. \tag{5.84}$$

In this case, $C_t = 4\pi$, $C_f = 2\sqrt{3}\pi$, $C_s = 2\sqrt{5}\pi$, $C_a = C_c = 4\sqrt{2}\pi$.

The mean-square correction error over the aperture area is defined as

$$\delta^2 = \frac{1}{\pi}\int \varepsilon^2(\rho)\,d^2\rho = \frac{1}{\pi}\int\left[\varphi(\rho) - \sum_{j=1}^{N} a_j R_j(\rho)\right]^2 d^2\rho. \tag{5.85}$$

It is now necessary to minimize δ^2 with respect to all a_j. In a real adaptive system, as we shall see later, this is performed automatically by the electronic control system. At present, however, we must carry out this operation mathematically.

Expanding the brackets in (5.85) we get

$$\delta^2 = \delta_0^2 - 2\sum_{j=1}^{N} a_j b_j + \sum_{j=1}^{N}\sum_{m=1}^{N} a_j a_m d_{jm}, \tag{5.86}$$

where

$$\delta_0^2 = \frac{1}{\pi}\int \varphi^2(\rho)\,d^2\rho, \tag{5.87}$$

$$b_j = \frac{1}{\pi}\int \varphi(\rho)\,R_j(\rho)\,d^2\rho, \tag{5.88}$$

$$d_{jm} = \frac{1}{\pi}\int R_j(\rho)\,R_m(\rho)\,d^2\rho. \tag{5.89}$$

We note that $\delta_0 = 2\pi$ in accordance with the assumed normalization (5.84).

We differentiate δ^2 with respect to a_j and equate the derivative to zero. Then

$$-2b_j + 2\sum_{m=1}^{N} a_m d_{jm} = 0, \tag{5.90}$$

from which we obtain the optimal coefficients a_j:

$$a_j = \sum_{m=1}^{N} b_m d_{jm}^{-1}. \tag{5.91}$$

Here d^{-1}_{jm} is the matrix inverse to d_{jm}. Substitution of (5.91) in (5.86) yields the minimum correction error

$$\delta^2_{\min} = \delta_0^2 - \sum_{j=1}^{N}\sum_{m=1}^{N} b_j d^{-1}_{jm} b_m. \tag{5.92}$$

To continue the calculations we must specify the actual form of the response function $R_j(\rho)$. Assume that all the drives are located inside a circular aperture of unity radius at the nodes of a square net and at its center. Two types of response functions were used in [16] for the calculations:

$$R_j(\rho) = \begin{cases} \left(1 - \frac{|x - x_j|}{r}\right)\left(1 - \frac{|y - y_j|}{r}\right), & |x - x_j| < r, \\ & |y - y_j| < r, \\ 0, & |x - x_j| \geqslant r, \\ & |y - y_j| \geqslant r; \end{cases} \tag{5.93}$$

$$R_j(\rho) = \begin{cases} \left(1 - \frac{(\rho - \rho_j)^2}{r^2}\right)^2, & |\rho - \rho_j| < r, \\ 0, & |\rho - \rho_j| \geqslant r, \end{cases} \tag{5.94}$$

where x and y are Cartesian coordinates in the mirror plane, ρ_j is the coordinate of the j-th drive (its Cartesian coordinates are x_j and y_j), and r is the mesh of the square net of drives.

The response function (5.93) has the form of a pyramid, while (5.94) is bell-shaped. They are shown, respectively, in Figs. 5.14a and 5.14b. A pyramidal response function is possessed by adaptive mirrors whose reflecting surfaces are made up of separate elements. A bell-shaped response function corresponds to adaptive mirrors with deformable reflecting surfaces. Other types of response functions are also possible. such as is shown in Fig. 5.14c, d.

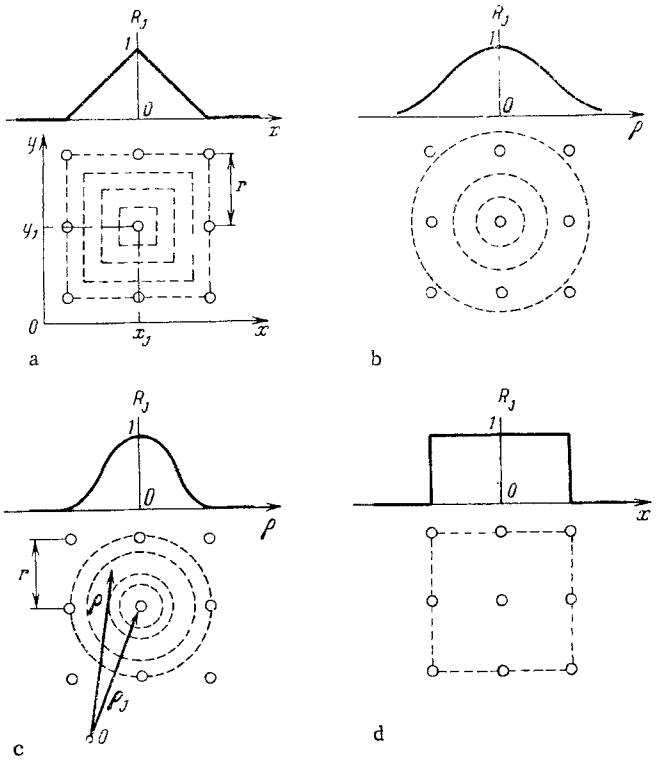


Fig. 5.14

TABLE 8. Wavefront Correction Error for Two Different Response Functions

Type of aberration	Correction error	
	pyramidal response function	bell-shape response function
Tilt	0,0123	0,0338
Focusing	0,0407	0,1180
Astigmatism	0,0254	0,0740
Coma	0,0295	0,0745

The results of the calculations in [16] for the response functions (5.93) and (5.94) are listed in Table 8. The mesh of the net of drives was chosen to be r = 0.105.

3. Correction Error of Atmospheric Distortions of Wavefront. Under real conditions the wavefront of the arriving wave varies randomly with time. We assume that the wavefront is distorted by atmospheric turbulence and obtain the minimum correction error averaged over the ensemble. From (5.92) we get

$$\langle \delta^2_{\min} \rangle = \langle \delta^2_0 \rangle - \sum_{j=1}^{N} \sum_{m=1}^{N} d^{-1}_{jm} \langle b_j b_m \rangle. \tag{5.95}$$

The mean values in this expression are obtained by averaging (5.87) and (5.88):

$$\langle \delta^2_0 \rangle = \frac{1}{\pi} \int \langle \varphi^2(\rho) \rangle \, d^2\rho, \tag{5.96}$$

$$\langle b_j b_m \rangle = \frac{1}{\pi^2} \int R_j(\rho_1) R_m(\rho_2) \langle \varphi(\rho_1) \varphi(\rho_2) \rangle \, d^2\rho_1 d^2\rho_2. \tag{5.97}$$

We note now that the adaptive mirror compensates not for the absolute value of the phase fluctuation at a given point ρ of the aperture, but the deviation of the phase from the value averaged over the aperture

$$\varphi' = \frac{1}{\pi} \int \varphi(\rho) \, d^2\rho. \tag{5.98}$$

The quantity φ' is random and changes from realization to realization; its average over all the realizations is

$$\langle \varphi' \rangle = 0. \tag{5.99}$$

Thus, it is necessary to replace $\varphi(\rho)$ in (5.96) and (5.97) by $\varphi(\rho) - \varphi'$.

We begin with (5.97). In view of the foregoing, we replace the correlation function $\langle \varphi(\rho_1) \varphi(\rho_2) \rangle$ by

$$\langle [\varphi(\rho_1) - \varphi'][\varphi(\rho_2) - \varphi'] \rangle = \langle \varphi(\rho_1) \varphi(\rho_2) \rangle - 2\langle \varphi(\rho) \varphi' \rangle + \langle \varphi'^2 \rangle. \tag{5.100}$$

On the other hand, the phase correlation function $\langle \varphi(\rho_1) \varphi(\rho_2) \rangle$ can be expressed in terms of the phase structure function $\mathscr{D}_\varphi(\rho_1 - \rho_2)$ (see Sec. 5.1):

$$\langle \varphi(\rho_1) \varphi(\rho_2) \rangle = \langle \varphi^2 \rangle - \frac{1}{2} \mathscr{D}_\varphi(\rho_1 - \rho_2), \tag{5.101}$$

where

$$\mathscr{D}_\varphi(\rho_1 - \rho_2) = 2(|\rho_1 - \rho_2|/\rho_0)^{5/3}, \tag{5.102}$$

and ρ_0 is the phase correlation function on the receiving aperture. In analogy with (5.100) we have for $\langle \varphi^2 \rangle$ a new expression:

$$\langle [\varphi(\rho) - \varphi']^2 \rangle = \langle \varphi^2 \rangle - 2\langle \varphi(\rho) \varphi' \rangle + \langle \varphi'^2 \rangle. \tag{5.103}$$

The expression for the structure function $\mathscr{D}_\varphi(\rho_1-\rho_2)$ is unchanged by virtue of its definition.

Substitution of (5.101) in (5.100) yields, when account is taken of (5.103),

$$\langle[\varphi(\rho_1)-\varphi'][\varphi(\rho_2)-\varphi']\rangle=-\frac{1}{2}\mathscr{D}_\varphi(\rho_1-\rho_2)+\langle[\varphi(\rho)-\varphi']^2\rangle \tag{5.104}$$

whence

$$\langle b_j b_m\rangle=\langle[\varphi-\varphi']^2\rangle\frac{1}{\pi^2}\int R_j(\rho)\,d^2\rho\int R_m(\rho)\,d^2\rho-\frac{1}{2\pi^2}\int R_j(\rho_1)R_m(\rho_2)\mathscr{D}_\varphi(\rho_1-\rho_2)\,d^2\rho_1 d^2\rho_2. \tag{5.105}$$

It is convenient to normalize the response function $R_j(\rho)$ to

$$\int R_j(\rho)\,d^2\rho=0. \tag{5.106}$$

This is done trivially, by introducing in (5.93) and (5.94) a supplementary additive coefficient that is of no importance for our analysis and can be left out of all the expressions. Using the indicated normalization, we reduce (5.105) to the form

$$\langle b_j b_m\rangle=-\frac{1}{2\pi^2}\int R_j(\rho_1)R_m(\rho_2)\mathscr{D}_\varphi(\rho_1-\rho_2)\,d^2\rho_1 d^2\rho_2. \tag{5.107}$$

We proceed now to calculate expression (5.96). To this end we integrate (5.103) with respect to $d^2\rho$:

$$\begin{aligned}\frac{1}{\pi}\int\langle[\varphi(\rho)-\varphi']^2\rangle d^2\rho &=\langle\varphi^2\rangle-\frac{2}{\pi^2}\int\int\langle\varphi(\rho)\varphi(\eta)\rangle d^2\rho d^2\eta+\frac{1}{\pi^2}\int\int\langle\varphi(\rho)\varphi(\eta)\rangle d^2\rho d^2\eta=\\ &=\langle\varphi^2\rangle-\frac{1}{\pi^2}\int\int\langle\varphi(\rho)\varphi(\eta)\rangle\,d^2\rho d^2\eta.\end{aligned} \tag{5.108}$$

Using (5.101), we express the integral obtained in terms of the known phase structure function

$$\frac{1}{\pi}\int\langle[\varphi(\rho)-\varphi']^2\rangle d^2\rho=\frac{1}{2\pi^2}\int\mathscr{D}_\varphi(\rho_1-\rho_2)\,d^2\rho_1 d^2\rho_2. \tag{5.109}$$

We turn now to the basic equation (5.95) and reduce it with the aid of (5.107) and (5.109) to the final form

$$\langle\delta^2_{\min}\rangle=\frac{1}{2\pi^2}\int\mathscr{D}_\varphi(\rho_1-\rho_2)\,d^2\rho_1 d^2\rho_2+\frac{1}{2\pi^2}\sum_{j=1}^{N}\sum_{m=1}^{N}d^{-1}_{jm}\int R_j(\rho_1)R_m(\rho_2)\mathscr{D}_\varphi(\rho_1-\rho_2)\,d^2\rho_1 d^2\rho_2, \tag{5.110}$$

where $\mathscr{D}_\varphi(\rho_1-\rho_2)$ is given by **(5.102)**. Before we proceed to the results of the numerical calculation, it is useful to note that $\mathscr{D}_\varphi\to 0$ as $\rho_0\to\infty$; hence $\langle\delta^2_{\min}\rangle\to 0$. This is natural, for the "smoother" the initial wavefront the better it will be approximated by the adaptive mirror.

Numerical calculations using (5.110) were performed in [16] where, besides the already-cited pyramidal and bell-shaped response functions, a Gaussian response function was investigated (Fig. 5.14c):

$$R_j(\rho)=\exp\left[-\frac{(\rho-\rho_j)^2}{r^2}\right], \tag{5.111}$$

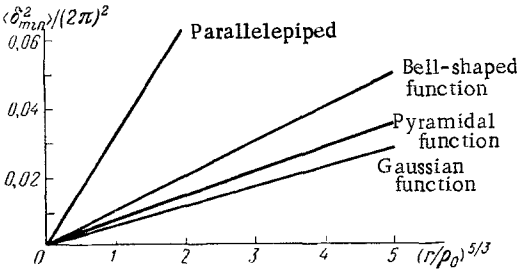


Fig. 5.15

and a response function in the form of a parallelepiped (Fig. 5.14d):

$$R_j(\rho) = \begin{cases} 1, & |x - x_j| < r, \ |y - y_j| < r, \\ 0, & |x - x_j| \geqslant r, \ |y - y_j| \geqslant r. \end{cases} \tag{5.112}$$

It may seem that the last case corresponds to an adaptive mirror of the "piston" type, shown in Fig. 5.12a. This is not so, however, since the effective range of the response function (5.112) extends to include neighboring drives (actuators), so that a cross-coupling is produced.

Figure 5.15 shows plots of the average minimum wavefront correction error $\langle\delta^2_{min}\rangle$ vs the ratio $(r/\rho_0)^{5/3}$. It is interesting to note that, for all the considered response functions, the plot is a straight line and can be approximated by

$$\langle\delta^2_{min}\rangle = \alpha (r/\rho_0)^{5/3}, \tag{5.113}$$

where r is the period of the net at whose sites the drives are located, ρ_0 is the field coherence radius, and the coefficient α depends on the concrete form of the response function of the drive.

4. Adaptation via Sharpness Function. So far we have dealt with the potentially feasible approximation of a wavefront by an adaptive mirror without considering how the adaptation takes place in real devices. At present there exist two basic adaptation methods: via direct measurement of the front and via maximizing the so-called sharpness function. We proceed to examine just the latter method.

Assume the existence of a physical quantity S that can be measured in practice and reaches, for a given object, an absolute maximum when all the phase distortions within the limits of the receiving aperture are zero. If, as before, we denote the signals that control the drives of the adaptive mirrors by a_j (j = 1, 2, ..., N), then S is a function of all the a_j:

$$S = S(a_1, a_2, \ldots, a_N). \tag{5.114}$$

This is called the sharpness function. If any of the arguments a_j is varied, i.e., if one of the control signals is changed somewhat, S changes. From the manner in which S changes (increases or decreases) one can conclude how to vary a_j so as to increase S. By analogous reasoning for all a_j, S can be brought to its maximum value. This state, by virtue of the stipulated property of S, corresponds to compensation for all the phase distortions within the limits of the receiving aperture. The described algorithm of finding the maximum is known as the "gradient rise" algorithm. In the adaptive-optics literature it is frequently called the method of "going up the hill."

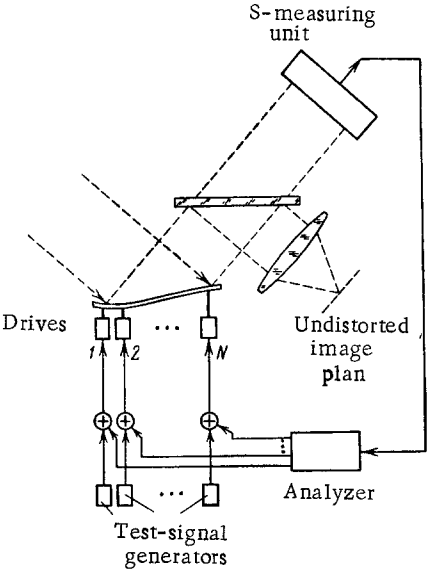


Fig. 5.16

In real devices the described process takes place automatically in accord with the functional diagram shown in Fig. 5.16. Test-signal generators apply small perturbations to the drives. The measured value of the sharpness function S (a_1, a_2, ..., a_N) is analyzed by a special analyzer at whose output are formed N control signals that are summed with the test signal and propel the drives in the required direction. As a result, the field with the corrected wavefront produces an undistorted image of the object.

The construction of the analyzer can vary, depending on the nature of the test signals and on the sequence in which they are generated [13]. A widely used method is to specify harmonic test signals with different frequencies f_j (j = 1, 2, ..., N). The analyzer is then a chain of N parallel filters, each tuned to its own frequency f_j. The amplitudes and phases of the signals are analyzed at the outputs of the filters. Depending on the phase, a control signal is applied to the j-th drive with either "+" or "—" sign, ensuring thereby an increase of the sharpness function and compensation of the phase distortion.

The requirements that the adaptation be fast and that the phase distortions be well compensated call for simultaneously increasing N and f_j. This in turn imposes stringent requirements on the operating speed of the adaptive-mirror drives if they operate independently. To relax the drive operating-speed requirement, another method [17] can be used. As we have seen in Subsec. 1, the distortions of the wavefront can be represented as a sum of several aberrations described by appropriate functions and described by aberration coefficients. We group all N control signals a_j (j = 1, 2, ..., N) in such a way that they can produce a summary response of the adaptive mirror in the form of independent aberration functions, say defocusing, spherical aberration, etc. To describe an arbitrary distortion of the wavefront we need then not N independent control signals, but only few (4 or 5) corresponding to the first terms of the expansion of the aberration function. The total width of the control-signal spectrum, at the same adaptation time as before, can then be greatly decreased, and the drive speed requirements can be suitably relaxed.

Adaptation via the sharpness function is an indirect method of improving the resolution, and requires additional proof that the maximum of the sharpness function corresponds to maximum resolution. Let us consider this question in greater detail.

5. Connection between Sharpness Functions and the Quality Characteristics of an Optical System. One can devise an infinite set of different sharpness functions, but only four, which are the most convenient in practice, have found application. They are listed in Table 9 [17]. The most used is sharpness function S_1. It is the integral of the intensity distribution in the image plane of the optical system, over a small circle having approximately the same dimension as the Airy disk. To realize S_1 it suffices to place in the image plane a photoreceiver, and in front of the latter a diaphragm with small opening. If the opening is smaller than the Airy disk, then

$$S_1 = \int_{\Delta} I(\boldsymbol{r})\, d^2\boldsymbol{r} \approx I(0)\,\Delta, \tag{5.115}$$

where I(0) is the intensity at the center of the picture, and Δ is the area of the opening. We normalize S_1 to unity maximum. Then

$$S_{1n} = I(0)/I_0(0), \tag{5.116}$$

where $I_0(0)$ is the intensity at the center in the absence of phase distortions. In optics it is customary to refer to the function S_{1nor} defined by (5.116) as the Strehl definition or the Strehl coefficient [7, 8] and designated by the letter $\mathscr{D}$. Thus, the normalized sharpness function S_1 is the strehl coefficient $\mathscr{D}$.

We shall show that, in the presence of only phase distortions, the Strehl coefficient, and hence also the sharpness function S_{1nor}, is always less than unity. We note for this purpose that the intensity distribution $I(\mathbf{r})$ is produced by an incoherent point source. It is therefore expressed in terms of the frequency—contrast characteristic $\tau(\boldsymbol{\xi})$ of the system in accord with the equation (see Sec. 5.1)

$$I(\boldsymbol{r}) = \int \tau(\boldsymbol{\xi}) \exp(i\boldsymbol{\xi}\boldsymbol{r})\, d^2\boldsymbol{\xi}, \tag{5.117}$$

where $\boldsymbol{\xi}$ is the spatial frequency. From (5.117) and (5.116) we get

$$\mathscr{D} = \int \tau(\xi)\, d^2\xi \Big/ \int \tau_0(\xi)\, d^2\xi, \tag{5.118}$$

where $\tau_0(\xi)$ is the frequency–contrast characteristic of an ideal optical system without phase distortions. If $\Pi(\rho)$ is the complex function of the exit pupil of an optical with phase distortions, we have, according to (5.32),

$$\tau(\xi) = \int \Pi(\rho)\, \Pi^*\left(\rho + \frac{f}{k}\xi\right) d^2\rho, \tag{5.119}$$

where the integration is over the plane of the entrance aperture, f is the effective focal length of the system, and k is the wave number.

We apply to (5.119) the known Schwarz inequality

$$\left| \int u(x)\, v^*(x)\, dx \right|^2 \leqslant \int u(x)\, u^*(x)\, dx \cdot \int v(x)\, v^*(x)\, dx. \tag{5.120}$$

We then obtain

$$|\tau(\xi)| \leqslant \tau_0(\xi), \tag{5.121}$$

where it is taken into account that $\tau_0(\xi)$ is a nonnegative function. From (5.121) and (5.118) it follows directly that

$$S_{\text{in}} = \mathscr{D} \leqslant 1. \tag{5.122}$$

In the presence of only phase distortions, the Strehl coefficient can be related with the resolution of an optical system. In fact, if d is the diameter of the central maximum of the distribution of the intensity $I(\mathbf{r})$ and determines the resolving power, then

$$\int I(\mathbf{r})\, d^2r \approx I(0)\, \pi d^2/4, \tag{5.123}$$

and, by virtue of the energy conservation law,

$$\int I(\mathbf{r})\, d^2r = \int I_0(\mathbf{r})\, d^2r, \tag{5.124}$$

whence

$$I(0)\, d^2 \approx I_0(0)\, d_0^2, \tag{5.125}$$

where d_0 is the Airy-disk diameter. In accord with the definition (5.116), we have

$$S_{\text{in}} = \mathscr{D} \approx (d_0/d)^2. \tag{5.126}$$

Consequently, if the Strehl coefficient $\mathscr{D}$ is known, the resolving power of the system can be obtained from the formula

$$d \approx d_0 \mathscr{D}^{-1/2}. \tag{5.127}$$

TABLE 9. Comparative Characteristics of Sharpness Functions

Sharpness function	Object	Detector	Sensitivity to wavefront tilt
$S_1 = \int_{\Delta} I(\mathbf{r})\, d^2r$	Pointlike	Single-channel	Yes
$S_2 = \int I^2(\mathbf{r})\, d^2r$	Pointlike extended	Matrix	No
$S_3 = \int I(\mathbf{r})\, M(\mathbf{r})\, d^2r$	Extended of known shape	Single-channel	Yes
$S_4 = \int [I(\mathbf{r}) - I_0(\mathbf{r})]^2 d^2r$	Pointlike, extended of known shape	Matrix	Yes

Thus, an adaptive optical system that maximizes the Strehl coefficient (i.e., the sharpness function S_1) maximizes also the resolving power and brings it closer to the diffractive value.

We consider now the sharpness function S_2. It is remarkable in that it can be applied to a centrosymmetric object of any shape, not necessarily pointlike. By definition,

$$S_2 = \int I^2(\boldsymbol{r})\, d^2r. \tag{5.128}$$

The normalized sharpness function

$$S_{2n} = \int I^2(\boldsymbol{r})\, d^2r \Big/ \int I_0^2(\boldsymbol{r})\, d^2r \tag{5.129}$$

is known in optics as "relative structural content" and is designated by the letter T [8]. It estimates mainly the sharpness of the image with respect to mean-squared values.

We shall show that S_2 reaches its maximum value when there are no phase distortions. To this end we change to the domain of spatial frequencies:

$$\begin{aligned} I(\boldsymbol{r}) &= \int V(\boldsymbol{\xi})\, \tau(\boldsymbol{\xi}) \exp(i\boldsymbol{\xi}\boldsymbol{r})\, d^2\xi, \\ \int I^2(\boldsymbol{r})\, d^2r &= \int V(-\boldsymbol{\xi})\, \tau(-\boldsymbol{\xi})\, V(\boldsymbol{\xi})\, \tau(\boldsymbol{\xi})\, d^2\xi, \end{aligned} \tag{5.130}$$

where $V(\boldsymbol{\xi})$ is the complex spatial spectrum of the image. Since

$$V(-\boldsymbol{\xi}) = V^*(\boldsymbol{\xi}), \quad \tau(-\boldsymbol{\xi}) = \tau^*(\boldsymbol{\xi}), \tag{5.131}$$

we have

$$S_2 = \int I^2(\boldsymbol{r})\, d^2r = \int |V(\boldsymbol{\xi})|^2 |\tau(\boldsymbol{\xi})|^2\, d^2\xi. \tag{5.132}$$

Independently of the shape of the object and its spatial spectrum, we obtain, by virtue of the condition (5.121),

$$\int I^2(\boldsymbol{r})\, d^2r \leqslant \int I_0^2(\boldsymbol{r})\, d^2r \tag{5.133}$$

or

$$S_{2n} \leqslant 1, \tag{5.134}$$

with the maximum value obviously reached when the phase distortions are compensated for. The function S_2 is important in that it makes it possible to carry out adaptation not only for pointlike objects, but also for extended objects of unknown shape.

If the object shape is known, we can use the sharpness function S_3 [where $M(\boldsymbol{r})$ is a mask that duplicates the shape of the object] or S_4 (see Table 9). We note that the normalized function S_4 is called "faithfulness" [8]. An adaptive optical system using the function S_4 is described in [18].

6. Adaptation via Signals of a Wavefront Sensor. Instead of assessing the distortions of a wavefront indirectly from the sharpness function, one can directly measure the relative phase distribution in a set of points of the aperture plane and shape control signals for the compensation of the distortions. The general scheme of the adaptive system will then differ somewhat from the preceding one (Fig. 5.17). The main difference is that the device that measures the sharpness function is replaced by a wavefront sensor [13, 14, 19]. This dispenses with the need for generating test signals.

The wavefront sensor is a complicated optical–mechanical setup, no less complicated than the adaptive mirror itself. Modern sensors work in white light and measure the local tilt angle of the wavefront, i.e., an angle independent of the wavelength. The wavefront tilt angle is in first-order approximation with the derivative of the optical path difference (expressed in wavelength units) with respect to the coordinate in the aperture plane. Integration of the tilt angle yields the optical path difference, calculated relative to the origin, at a given aperture point. This eliminates the uncertainty of the phase measurement.

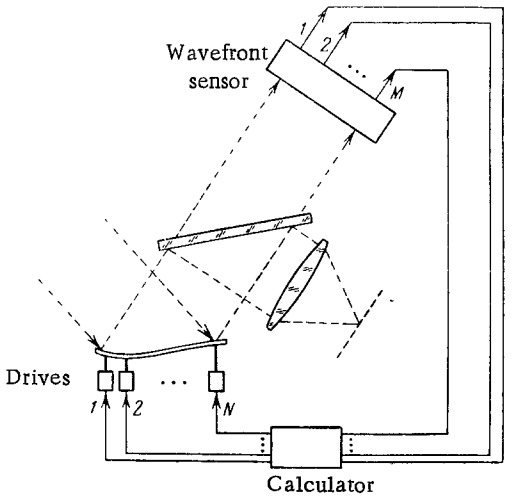


Fig. 5.17

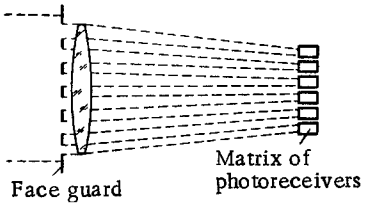


Fig. 5.18

A classical example of a wavefront sensor that measures the tilt angles of a wavefront at specified points of the aperture is the Hartmann sensor (Fig. 5.18) in its various modifications [14]. A matrix of photoreceivers measures the relative shifts of the images of each section of the receiving aperture. These data are used to calculate the average tilt angle of the wavefront within the limits of a given section. The result is a system of partial differential equations for the distribution of the phase $\varphi(x, y)$:

$$\frac{\partial \varphi}{\partial x}(x_i, y_j) = u_{ij}, \quad \frac{\partial \varphi}{\partial y}(x_i, y_j) = v_{ij}. \tag{5.135}$$

There are many different algorithms for numerically integrating this system of equations. These, however, are outside the scope of this book, and we shall not dwell on them. We note only that the corresponding calculations can be performed both in digital and in analog form (see Appendix VIII).

7. Dynamics of Adaptive System. As noted repeatedly, an adaptive optic system must compensate for the wavefront distortions within a time shorter than the freezing time of the atmosphere. The time constant of an adaptive system is a most important parameter alongside the accuracy of the distortion compensation. At the beginning of the present section we considered the static accuracy of the compensation, by an adaptive mirror, of the phase distortions of a received field. We now consider the dynamic properties of an adaptive optic system.

We are interested primarily in the time dependence of the resolution of an adaptive optic system. In accord with (5.127), the resolution is determined by the Strehl coefficient or, equivalently, by the sharpness function S_1. It is therefore natural to consider adaptation via the sharpness function S_1. An additional argument in favor of this choice is that systems with adaptation via the sharpness function S_1 are most widely used at present.

We consider a hypothetical diagram of an adaptive optical system (Fig. 5.19). We assume, for simplicity, that the adaptive mirror is an aggregate of individual independent pistons, as shown in Fig. 5.12a. The distribution of the phase of the reflected waves over all N pistons will be described by the vector

$$\boldsymbol{P} = (p_1, p_2, \ldots, p_N), \tag{5.136}$$

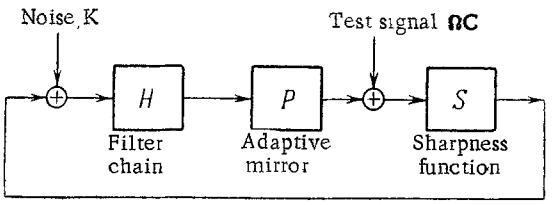


Fig. 5.19

where p_j is the average phase within the confines of the j-th piston. All the p_j will be arbitrarily called "slowly" varying functions of time, in contrast to the rapidly varying test signals $c_j \sin \omega_j t$, the aggregate of which is characterized by the vector

$$\boldsymbol{\Omega C} = (c_1 \sin \omega_1 t,\ c_2 \sin \omega_2 t,\ \ldots,\ c_N \sin \omega_N t), \tag{5.137}$$

where the modulation vector $\boldsymbol{\Omega}$ and the matrix $\mathbf{C}$ of the modulation coefficients are of the form

$$\boldsymbol{\Omega} = (\sin \omega_1 t,\ \sin \omega_2 t,\ \ldots,\ \sin \omega_N t), \tag{5.138}$$

$$\boldsymbol{C} = \begin{pmatrix} c_1 & & 0 \\ & c_2 & \\ & \ddots & \\ 0 & & c_N \end{pmatrix}. \tag{5.139}$$

The sharpness function S depends on $\mathbf{P} + \boldsymbol{\Omega C}$ and on the time:

$$S = S(\boldsymbol{P} + \boldsymbol{\Omega C},\ t). \tag{5.140}$$

Assuming smallness of all the modulation coefficients c_j, we expand the sharpness function in terms of the small parameter $\boldsymbol{\Omega C}$. Confining ourselves to the first term of the expansion, we get

$$S = S(\boldsymbol{P},\ t) + \boldsymbol{\Omega C}\left(\frac{\partial S}{\partial \boldsymbol{P}}\right)^T, \tag{5.141}$$

where the superscript T denotes the transpose of a vector

$$\frac{\partial S}{\partial \boldsymbol{P}} = \left(\frac{\partial S}{\partial p_1},\ \frac{\partial S}{\partial p_2},\ \ldots,\ \frac{\partial S}{\partial p_N}\right). \tag{5.142}$$

The second term in (5.141) is the rapidly varying component of the sharpness function, unlike the first term which varies slowly. The signal S(t) is fed next to the input of a chain of filters, each tuned to its own frequency ω_j (j = 1, 2, ..., N). The first term in (5.141) is not passed by the filters, and the information-containing parameter is then the amplitude $(\partial S/\partial \mathbf{P})\mathbf{C}$ (with account taken of the sign) of the second term. The signals at the filter outputs are synchronously detected to preserve the information on the sign, and are transformed into control signals for the drives of the adaptive mirror. Account must also be taken of the unavoidable noise present in the channel. The slowly varying noise component will be characterized by the vector

$$\boldsymbol{K} = (\xi_1,\ \xi_2,\ \ldots,\ \xi_N) \tag{5.143}$$

with all the ξ_j (j = 1, 2, ..., N) mutually independent random variables.

The phase p_j at each piston is proportional to the control signal. Thus,

$$\boldsymbol{P}(t) = \boldsymbol{P}_0 + \int_{-\infty}^{t} \boldsymbol{Y}(t')\,\boldsymbol{H}(t - t')\,dt', \tag{5.144}$$

where we have introduced a symbol for the vector of the amplitudes of the signals at the filter inputs

$$\boldsymbol{Y}(t) = \frac{\partial S}{\partial \boldsymbol{P}}\boldsymbol{C} + \boldsymbol{K} \tag{5.145}$$

and the matrix of the pulsed responses of the filters

$$\boldsymbol{H}(t)=\begin{pmatrix} h_1(t) & & 0 \\ & h_2(t) & \\ & & \ddots \\ 0 & & h_N(t) \end{pmatrix}. \tag{5.146}$$

The vector $\mathbf{P}_0$ characterizes the initial (unadapted) distribution of the phase of the wave reflected from the adaptive mirror.

Differentiating (5.144) with respect to time under the assumption that all the $h_j(t)$ are exponentially damped functions, we obtain a system for $\mathbf{P}(t)$:

$$\frac{d\boldsymbol{P}(t)}{dt}=\boldsymbol{Y}(t)\,\boldsymbol{H}(0)-\boldsymbol{P}(t)\,\boldsymbol{G}, \tag{5.147}$$

where $\mathbf{G}$ is the matrix of the reciprocals of the filter time constants T_j:

$$\boldsymbol{G}=\begin{pmatrix} T_1^{-1} & & 0 \\ & T_2^{-1} & \\ & & \ddots \\ 0 & & T_N^{-1} \end{pmatrix}. \tag{5.148}$$

We are interested in the $S(t)$ dependence. If we assume beforehand that the adaptation varies more rapidly than the atmosphere, then

$$S(t)=S(\boldsymbol{P}(t)). \tag{5.149}$$

Consequently, to solve our problem we must find the $S(\mathbf{P})$ dependence, solve the system (5.147), and substitute the obtained solution in (5.149).

Let us find $S(\mathbf{P})$. We note for this purpose that, by virtue of (5.116), the sharpness function is proportional to the field intensity $I(0)$ at the center of the image plane. If we confine ourselves to adaptation for a point source, the field at the entrance aperture can be regarded during the entire adaptation process as spatially coherent. Then, without averaging, we can write

$$I(0)=|E|^2, \tag{5.150}$$

where E is the field at the center of the image plane. We assume next that $N \gg 1$. Then the diffraction pattern produced in the image plane by the field reflected from each piston is much wider than the aperture over which the integration is carried out in (5.115). Taking these remarks into account, we obtain an expression for the normalized sharpness function

$$S_{\text{н}}(\boldsymbol{P})=\mathscr{D}=N^{-2}\sum_{m=1}^{N}\sum_{n=1}^{N}\exp[i(p_n-p_m)], \tag{5.151}$$

whence

$$\frac{\partial S}{\partial \boldsymbol{P}}=-2N^{-2}\left[\sum_{m=1}^{N}\sin(p_1-p_m),\ \sum_{m=1}^{N}\sin(p_2-p_m),\ \ldots,\ \sum_{m=1}^{N}\sin(p_N-p_m)\right]. \tag{5.152}$$

Next, we assume for $h_j(t)$ the normalization $h_j(0) = T^{-1}{}_j$ corresponding to a unity transmission coefficient of the filter at the resonance frequency ω_j. System (5.147) is then reduced to the form

$$\frac{d\boldsymbol{P}(t)}{dt}=\frac{\partial S}{\partial \boldsymbol{P}}\boldsymbol{C}\boldsymbol{G}+\boldsymbol{K}\boldsymbol{G}-\boldsymbol{P}(t)\,\boldsymbol{G}. \tag{5.153}$$

This is a closed system of equations, since the explicit form of $\partial S/\partial \mathbf{P}$ is known.

To obtain a solution in explicit form, we assume that all the differences $|p_n - p_m|$ are small. Then

$$\mathscr{D}=1-N^{-2}\sum_{\substack{m,\,n=1\\ m<n}}^{N}(p_n-p_m)^2, \tag{5.154}$$

$$\frac{dp_j}{dt}=-c_jT_j^{-1}2N^{-2}\sum_{m=1}^{N}(p_j-p_m)-T_j^{-1}p_j+T_j^{-1}\xi_j. \tag{5.155}$$

If a calculation error of 20% is admissible, the condition that $|p_n - p_m|$ be small is not very restrictive, and takes the form

$$|p_n-p_m|\leqslant 1.0, \tag{5.156}$$

corresponding to a maximum wavefront distortion 0.16λ. At any rate, the error will decrease in the course of time, since the differences $p_n - p_m$ will tend to zero (in the absence of noise).

As a rule, we can neglect the channel differences and put $c_j = c$, $T_j = T$ (j = 1, 2, ..., N). Then

$$\mathscr{D}=1-N^{-2}\sum_{\substack{n,\,m=1\\ m<n}}^{N}\varphi_{nm}^2, \tag{5.157}$$

$$\frac{d\varphi_{nm}}{dt}=-\frac{2c+N}{NT}\varphi_{nm}+T^{-1}\xi_{nm}, \tag{5.158}$$

where $\varphi_{nm}=p_n-p_m$, $\xi_{nm}=\xi_n-\xi_m$. We now integrate (5.158), substitute the result in (5.157), and average. As a result, we get

$$\langle\mathscr{D}\rangle=1-\frac{1-N^{-1}}{2}\left\{\langle\varphi_0^2\rangle\exp\left(-2\frac{2c+N}{NT}t\right)+w\frac{N}{(2c+N)T}\left[1-\exp\left(-2\frac{2c+N}{NT}t\right)\right]\right\}, \tag{5.159}$$

where $\langle\varphi_0^2\rangle$ is the average initial phase dispersion and $w = \langle\xi^2_j\rangle$ is the spectral density of the phase noise in one channel. Usually, $c \ll 1$, $N \gg 1$. In this case, (5.159) takes the simpler form

$$\langle\mathscr{D}\rangle=1-\frac{1}{2}\left\{\langle\varphi_0^2\rangle\exp\left(-2\frac{t}{T}\right)+\frac{w}{T}\left[1-\exp\left(-2\frac{t}{T}\right)\right]\right\}. \tag{5.160}$$

Since the spectral width of one channel is T^{-1}, the parameter w/T plays the role of the mean-squared noise amplitude in one channel. The role of the adaptation time constant is played by the quantity T/2, half the time constant of the control circuit. A plot of the function $\langle\mathscr{D}(t)\rangle$ is shown in Fig. 5.20, while Fig. 5.21 shows the change of the average resolvable image element <d(t)>, calculated from (5.127). It is typical that in the absence of noise the adaptation leads in final analysis to the diffractive resolution. The noise affects the average resolution adversely. At low noise levels, when w/T << 1, the deterioration of the resolution can be estimated from the formula

$$\delta d\approx d_0\frac{w}{4T}. \tag{5.161}$$

8. Experimental Results. Convincing results that demonstrate the capabilities of adaptive optics were obtained in [20]. The experiments were performed with a telescope of 30-cm diameter both with laser illumination on a route of length 250 m (λ = 0.63 μm) and with an incoherent white-light noise. The adaptation was via the sharpness function S_1 along only one coordinate, using an adaptive mirror of the "piston" type, consisting of six individual elements. The time constant of the system was 60 μsec.

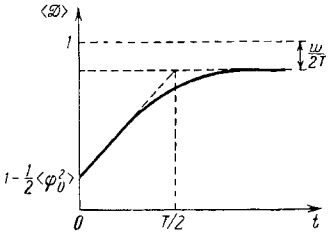


Fig. 5.20

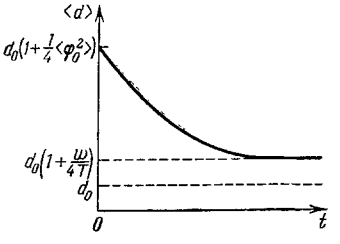


Fig. 5.21

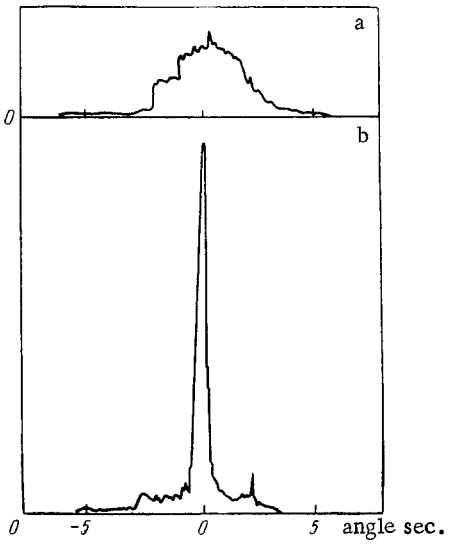


Fig. 5.22

The adaptation process was carried out by successive correction of each of the six drives. One adjustment cycle lasted 8 msec.

Figure 5.22 shows plots [20] of the one-dimensional image of a laser point source of light, shaped through a 250-cm layer of turbulent atmosphere. The ordinates and abscissas are, respectively, the relative illumination and the angular coordinates of the image. Figure 5.22a shows the illumination distribution when the control is turned off and there is no adaptation. The areas under the curves are equal, and the width of the central peak of the adapted image is one-tenth the width of the unadapted image.

Finally, notice must be taken of the successful test of an adaptive optic system using a wavefront sensor [21]. A monolithic piezoelectric mirror was used in these experiments with 21 drives. The distortions were compensated along two axes.

APPENDICES

I. Table of the Functions H(x) and H(x)/x

x	$H(x)$	$H(x)/x$	x	$H(x)$	$H(x)/x$
0,001	$9{,}999 \cdot 10^{-4}$	0,999	1,0	0,919	0,919
0,01	$9{,}999\ 10^{-3}$	0,999	1,5	1,241	0,827
0,10	$9{,}992 \cdot 10^{-2}$	0,999	2,0	1,426	0,713
0,15	0,149	0,998	2,5	1,468	0,587
0,20	0,199	0,997	3,0	1,388	0,463
0,25	0,249	0,995	3,5	1,223	0,349
0,30	0,298	0,993	4,0	1,025	0,256
0,35	0,346	0,989	4,5	0,842	0,187
0,40	0,395	0,987	5,0	0,715	0,143
0,45	0,443	0,983	5,5	0,669	0,122
0,50	0,489	0,979	6,0	0,706	0,118
0,55	0,536	0,975	6,5	0,812	0,125
0,60	0,582	0,970	7,0	0,955	0,136
0,65	0,628	0,965	7,5	1,099	0,147
0,70	0,672	0,959	8,0	1,211	0,151
0,75	0,716	0,954	8,5	1,265	0,149
0,80	0,758	0,948	9,0	1,252	0,139
0,85	0,800	0,941	9,5	1,179	0,124
0,90	0,841	0,935	10,0	1,067	0,107
0,95	0,880	0,927			

II. Proof of Satisfaction of the Reciprocity Principle

Let $G(\mathbf{r}, \mathbf{r}_1)$ be a Green's function. This means that a point source placed at the point $\mathbf{r}_1$ produces at the point $\mathbf{r}$ a field $G(\mathbf{r}, \mathbf{r}_1)$. The function $G(\mathbf{r}, \mathbf{r}_1)$ satisfies the following equation:

$$\Delta G(\boldsymbol{r}, \boldsymbol{r}_1) = -k^2 n^2(\boldsymbol{r})\, G(\boldsymbol{r}, \boldsymbol{r}_1) - 4\pi\delta(\boldsymbol{r} - \boldsymbol{r}_1). \tag{II.1}$$

Similarly, the function $G(\mathbf{r}, \mathbf{r}_2)$ is the field produced at the point $\mathbf{r}$ by a point source located at the point $\mathbf{r}_2$. It is required to prove that

$$G(\boldsymbol{r}_2, \boldsymbol{r}_1) = G(\boldsymbol{r}_1, \boldsymbol{r}_2).$$

The function $G(\mathbf{r}, \mathbf{r}_2)$ satisfies an equation similar to (II.1):

$$\Delta G(\boldsymbol{r}, \boldsymbol{r}_2) = -k^2 n^2(\boldsymbol{r})\, G(\boldsymbol{r}, \boldsymbol{r}_2) - 4\pi\delta(\boldsymbol{r} - \boldsymbol{r}_2). \tag{II.2}$$

We multiply (II.1) by $G(\mathbf{r}, \mathbf{r}_2)$, (II.2) by $G(\mathbf{r}, \mathbf{r}_1)$, and subtract one equation from the other

$$G(\boldsymbol{r}, \boldsymbol{r}_2)\,\Delta G(\boldsymbol{r}, \boldsymbol{r}_1) - G(\boldsymbol{r}, \boldsymbol{r}_1)\,\Delta G(\boldsymbol{r}, \boldsymbol{r}_2) = -4\pi[\delta(\boldsymbol{r} - \boldsymbol{r}_1)\, G(\boldsymbol{r}, \boldsymbol{r}_2) - \delta(\boldsymbol{r} - \boldsymbol{r}_2)\, G(\boldsymbol{r}, \boldsymbol{r}_1)].$$

We integrate the obtained equality over a closed volume V. Then, using Green's theorem, we obtain

$$-\int_S [G(\boldsymbol{r}, \boldsymbol{r}_1)\,\nabla G(\boldsymbol{r}, \boldsymbol{r}_2) - G(\boldsymbol{r}, \boldsymbol{r}_2)\,\nabla G(\boldsymbol{r}, \boldsymbol{r}_1)]\, dS = -4\pi[G(\boldsymbol{r}_1, \boldsymbol{r}_2) - G(\boldsymbol{r}_2, \boldsymbol{r}_1)]. \tag{II.3}$$

The integration is over the closed surface of volume V with outward normal. We choose the volume V to be a sphere of radius R. As $R \to \infty$, the integral in the left-hand side of (II.3) then tends to zero (by virtue of the Sommerfeld radiation condition). From this we get

$$G(\boldsymbol{r}_1, \boldsymbol{r}_2) = G(\boldsymbol{r}_2, \boldsymbol{r}_1),$$

Q.E.D.

III. Summation of the Series in the Equation for $P_{s+no}(l)$

The initial expression for P_{s+no} is of the form

$$P_{s+no}(l) =$$

$$= \frac{\exp(-\mu)}{l!}\left(\frac{q}{1-q}\right)^l\left(\frac{1-d}{G-d}\right)^L \sum_{m=l}^{\infty} \frac{m!(m+L-1)!}{(m-l)!}\left[\frac{(G-1)(1-q)}{G-d}\right]^m \times \sum_{n=0}^{m}\left[\frac{\mu(G-1)d}{G-d}\right]^n \sum_{j=0}^{n} \frac{[(1-d)^2 G(G-1)^{-2}]^j}{j!(m-j)!(n-j)!(j+L-1)!}.$$

We rewrite this expression in a different form

$$P_{s+no}(l) = \exp(-\mu)\frac{1}{l!}\left(\frac{q}{1-q}\right)^l\left(\frac{1-d}{G-d}\right)^L \sum_n \sum_m \sum_j m!(m+L-1)!\left[\frac{(1-q)(G-1)}{G-d}\right]^m\left[\frac{\mu(G-1)d}{G-d}\right]^n \times$$
$$\times\left[\frac{(1-d)^2 G}{(G-1)^2}\right]^j |(m-l)!(m-j)!j!(j+L-1)!(n-j)!|^{-1}. \qquad \text{(III.1)}$$

In this expression the summation is cut off whenever the factorial becomes less than zero, i.e., at $m < l$, j; $n < j$. We introduce the notation

$$A = (1-q)(G-1)(G-d)^{-1}, \quad B = \mu(G-1)d(G-d)^{-1},$$
$$C = (1-d)^2 G(G-1)^{-2}.$$

We sum first over the index n:

$$\sum_{n=j}^{\infty} \frac{B^n}{(n-j)!} = B^j \exp(B).$$

The triple sum S in (III.1) is then equal to

$$S = \exp B \sum_m \sum_j A^m \frac{m!(m+L-1)!(BC)^j}{(m-l)!(m-j)!j!(j+L-1)!}.$$

We sume now over the index m. We have

$$\sum_m \frac{(m+L-1)!\,m!}{(m-j)!(m-l)!} A^m = A^l \frac{\partial^l}{\partial A^l} \sum_m \frac{(m+L-1)!}{(m-j)!} A^m.$$

We introduce in the right-hand side of this expression a new summation index $k = m - j$. We then obtain

$$\sum_m \frac{(m+L-1)!}{(m-j)!} A^m = A^j \sum_k \frac{(k+L+j-1)!}{k!} A^k = A^j \frac{(L+j-1)!}{(1-A)^{L+j}}.$$

The last equation is obtained in the analogy with the procedure used in the main text. The sum S is then equal to

$$S = \exp B A^l \sum_j \frac{(BC)^j}{j!(j+L-1)!} \frac{\partial^l}{\partial A^l}\left[\frac{A^j (L+j-1)!}{(1-A)^{L+j}}\right].$$

The factorial $(j + L - 1)!$ cancels out, and after simple transformations we get

$$S = \exp(B) A^l \frac{\partial^l}{\partial A^l}\left\{\frac{\exp[BCA(1-A)^{-1}]}{(1-A)^L}\right\} = A^l \exp[B(1-C)] \frac{\partial^l}{\partial A^l}\left\{\frac{\exp[BC(1-A)^{-1}]}{(1-A)^L}\right\}. \qquad \text{(III.2)}$$

We calculate the l-th derivative in (III.2). To this end we express the exponential in the form of a series

$$\exp\left(\frac{BC}{1-A}\right) = \sum_{m=0}^{\infty}\left(\frac{BC}{1-A}\right)^m \frac{1}{m!}.$$

Substituting this expression in (III.2), we get

$$\frac{\partial^l}{\partial A}\left\{\frac{\exp\left[BC\,(1-A)^{-1}\right]}{(1-A)^L}\right\}=\frac{1}{(1-A)^{L+l}}\sum_{m=0}^{\infty}\left(\frac{BC}{1-A}\right)^m\frac{(L+m+l-1)\,!}{m!\,(L+m-1)\,!}.$$

The infinite sum is inconvenient for computer calculations; we reduce it therefore to a sum of a finite number of terms. We note for this purpose that

$$\frac{1}{(1-A)^{L+l}}\sum_{m=0}^{\infty}\left(\frac{BC}{1-A}\right)^m\frac{(L+m+l-1)\,!}{m\,!\,(L+m-1)\,!}=\frac{1}{(1-A)^{L+l}}\frac{\partial^l}{\partial A^l}\sum_m\frac{D^{L+m+l-1}}{D^{L-1}m\,!}=\frac{1}{(1-A)^{L+l}\,D^{L-1}}\frac{\partial^l}{\partial A^l}\left(D^{L+l-1}e^D\right),$$

where

$$D=BC\,(1-A)^{-1}.$$

Differentiating and making a trivial replacement of the summation index in the resultant sum, we arrive at the formula

$$\frac{\partial^l}{\partial A^l}\left\{\frac{\exp\left[BC\,(1-A)^{-1}\right]}{(1-A)^L}\right\}=$$

$$=\frac{1}{(1-A)^{L+l}}\exp\left[BC\,(1-A)^{-1}\right]l\,!\,(L+l-1)\,!\sum_{j=0}^{l}\left(\frac{BC}{1-A}\right)^j\frac{1}{j\,!\,(l-j)\,!\,(L+j-1)\,!}.$$

Substituting this formula in expressions (III.2) and (III.1), we obtain the final expression for $P_{s+no}(l)$:

$$P_{s+no}(l)=\exp\left[-\frac{\mu Gq}{\varkappa(1+\varkappa)}\right]\frac{\varkappa^l}{(1+\varkappa)^{L+l}}(L+l-1)\,!\sum_{j=0}^{l}\left[\frac{\mu Gq}{\varkappa\,(1+\varkappa)}\right]^j\frac{1}{j\,!\,(l-j)\,!\,(L+j-1)\,!}.$$

IV. Derivation of the Dependence of k_z on k_x and k_y in Eq. (3.129)

Consider a uniaxial anisotropic crystal. Figure A.1 shows the optical indicatrix of such a crystal. The axes x_1, y_1, and z_1 are the principal axes of the optical indicatrix, and the wave vector of the extraordinary wave is directed along the z axis, which makes an angle θ with the z_1 axis. Small changes of the direction of propagation of the wave (angle $d\theta$) lead to small changes of the refractive index for this wave. The refractive index is determined by the absolute value of the angle θ between the axes z and z_1. In the notation of Fig. A.1 we have

$$d\theta=k_x/k,$$

where k_x is the projection of the vector **k** on the x axis.

The displacements $d\theta$ and $d\varphi$ cause the angle of inclination of the wave vector to the crystal axis z to change from θ to θ_0 (Fig. II.2 — side view as seen from the z_1 axis). The refractive index of the extraordinary wave is expressed in terms of the angle θ_0 in the following manner:

$$n\,(\theta_0)=n_o\,(1+\varepsilon\sin^2\theta_0)^{-1/2},$$

where

$$\varepsilon=n_o^2/n_e^2-1,$$

n_o and n_e are the refractive indices of the ordinary and extraordinary rays.

It is easy to verify that

$$\cos\theta\cos\varphi=\cos\theta_0,$$

whence

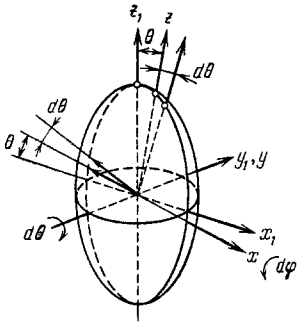


Fig. A.1

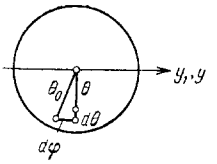


Fig. A.2

$$n(\theta,\ \varphi) = n_0 (1 + \varepsilon \sin^2\theta \cos^2\varphi + \varepsilon \sin^2\varphi)^{-1/2}.$$

We have next

$$k_z = (k_0 n^2 (\theta + d\theta,\ \varphi + d\varphi) - k_x^2 - k_y^2)^{1/2},$$

where k_0 is the modulus of the wave vector in vacuum.

Using the smallness of the angles θ and φ, we get

$$k_z = k + n^{-1}\frac{\partial n}{\partial\theta}k_x + n^{-1}\frac{\partial n}{\partial\varphi}k_y + \left[n^{-1}\frac{\partial^2 n}{\partial\theta^2} + \left(n^{-1}\frac{\partial n}{\partial\theta}\right)^2 - 1\right]\frac{k_x^2}{2k} + \left[n^{-1}\frac{\partial^2 n}{\partial\varphi^2} + \left(n^{-1}\frac{\partial n}{\partial\theta}\right)^2 - 1\right]\frac{k_y^2}{2k}, \qquad \text{(IV.1)}$$

where k is the modulus of the wave vector in the z direction. The corresponding derivatives were written for the initial value of θ and $\varphi = 0$. It is easy to verify that $\partial n/\partial\varphi = 0$ at $\varphi = 0$. The other derivatives are equal to

$$\begin{aligned}
\left.\frac{\partial n}{\partial\theta}\right|_{\varphi=0} &= -n_0\varepsilon \sin\theta\cos\theta\,(1+\varepsilon\sin^2\theta)^{-3/2},\\
\left.n^{-1}\frac{\partial n}{\partial\theta}\right|_{\varphi=0} &= -\varepsilon\sin\theta\cos\theta\,(1+\varepsilon\sin^2\theta)^{-1} = -\operatorname{tg}\alpha,\\
\left.\frac{\partial^2 n}{\partial\theta^2}\right|_{\varphi=0} &= -n_0\varepsilon\,|1 - 2(1+\varepsilon)\sin^2\theta + \varepsilon\sin^4\theta|\,(1+\varepsilon\sin^2\theta)^{-5/2},\\
\left.\frac{\partial^2 n}{\partial\varphi^2}\right|_{\varphi=0} &= -n_0\varepsilon\,(1-\sin^2\theta)\,(1+\varepsilon\sin^2\theta)^{-3/2}.
\end{aligned} \qquad \text{(IV.2)}$$

Substituting formulas (IV.2) in (IV.1), we get (3.129).

V. Calculation of the Scattering Function $\Gamma(x, y)$

The general expression for $\Gamma(x, y)$ is (3.136):

$$\Gamma(x,\ y) = -i\beta_3 A_2 z_0 \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int\left\{\frac{z_0}{4k_s}[(k_x+\varkappa)^2 + k_y^2]\right\}^{-1} \sin\left\{\frac{z_0}{4k_s}[(k_x+\varkappa)^2 + k_y^2]\right\}\exp[-i(x-x_0)k_x - iyk_y]\,dk_x dk_y.$$

Making the change of variables

$$x - x_0 = x_1, \quad k_x + \varkappa = k_{x1},$$

we have

$$\Gamma(x_1,\ y) = -i\beta_3 A_2 z_0 \exp(ix_1\varkappa)\,\frac{1}{(2\pi)^2}\iint_{-\infty}^{+\infty}\frac{\sin[\alpha(k_{x1}^2 + k_y^2)]}{\alpha(k_{x1}^2 + k_y^2)}\exp[-i(x_1 k_{x1} + yk_y)]\,dk_{x1}dk_y,$$

where $\alpha = z_0/4k_S$. We introduce the two vectors

$$\boldsymbol{r} = (x_1,\ y), \quad \boldsymbol{k} = (k_{x1},\ k_y).$$

We then obtain

$$\Gamma(r) = -i\beta_3 A_2 z_0 \exp(ix_1\varkappa)\,\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{\infty}\frac{\sin(ak^2)}{ak^2}\exp(-irk\cos\psi)\,k\,dk\,d\psi,$$

where ψ is the angle between the vectors **r** and **k**. Integrating with respect to $d\psi$, we transform the expression into

$$\Gamma(r) = -i\beta_3 A_2 z_0 \exp(ix_1\varkappa) \int_0^\infty \frac{\sin(ak^2)}{ak^2} J_0(rk)\, k\, dk, \qquad (V.1)$$

where $J_0(rk)$ is a Bessel function of zero order.

We now calculate the integral in (V.1):

$$I(r) = \int_0^\infty \frac{\sin(ak^2)}{ak^2} J_0(rk)\, k\, dk.$$

We differentiate it with respect to the parameter r:

$$\frac{d}{dr} I(r) = -\int_0^\infty \frac{\sin(ak^2)}{ak^2} k^2 J_1(rk)\, dk = -\frac{1}{a}\int_0^\infty \sin(ak^2) J_1(rk)\, dk,$$

where $J_1(rk)$ is a Bessel function of first order. The last integral is equal to

$$\int_0^\infty \sin(ak^2) J_1(rk)\, dk = \frac{1}{r}\sin\left(\frac{r^2}{4a}\right).$$

Hence

$$I(r) = A - \frac{1}{a}\int_0^r \frac{1}{t}\sin\left(\frac{t^2}{4a}\right) dt = A - \frac{1}{2a}\int_0^{r^2/4a} \frac{\sin u}{u}\, du.$$

The integration constant A is obtained from the condition

$$I(0) = \int_0^\infty \frac{\sin(ak^2)}{ak^2} k\, dk = \frac{\pi}{4a}.$$

As a result we get the following expression for I(r):

$$I(r) = \frac{1}{2a}\left[\frac{\pi}{2} - \mathrm{Si}\left(\frac{r^2}{4a}\right)\right], \quad \mathrm{Si}(x) = \int_0^x \frac{\sin u}{u}\, du.$$

Substituting this expression in (V.1), we obtain ultimately

$$\Gamma(r) = -i\beta_3 A_2 k_s \exp(kx_1\varkappa)\frac{1}{\pi}\left[\frac{\pi}{2} - \mathrm{Si}\left(\frac{r^2}{4a}\right)\right).$$

VI. Asymptotic Representation of the Function $\varepsilon(p)$

The initial expression is of the form

$$\varepsilon(p) = \int_0^1 (1-x)^{p-1} \exp(px)\, dx.$$

We seek an asymptotic representation for $\varepsilon(p)$ in the form

$$\varepsilon(p) = \varepsilon_0(p) + \delta(p),$$

where $\varepsilon_0(p)$ is the limit of $\varepsilon(p)$ as $p \to 0$, and $\delta(p)$ is a certain function that must be determined. At $p \ll 1$, we have

$$\varepsilon(p) \approx \int_0^1 (1-x)^{p-1}\,dx + p\int_0^1 (1-x)^{p-1}\,x\,dx.$$

As p → 0 both integrals in the right-hand side of this expression have poles of equal order in their integrands. Consequently, the value of the second integral without the factor p does not exceed that of the first. In the limit as p → 0, we therefore obtain

$$\varepsilon_0(p) = \int_0^1 (1-x)^{p-1}\,dx = \frac{1}{p}.$$

Let us obtain δ(p). We have

$$\delta(p) = \varepsilon(p) - \varepsilon_0(p) = \int_0^1 x^{p-1}\left[\exp\left((1-x)\,p\right) - 1\right]dx = \sum_{k=1}^{\infty}\frac{p^k}{k!}\int_0^1 x^{p-1}(1-x)^k\,dx.$$

The integral under the summation sign is known:

$$\int_0^1 x^{p-1}(1-x)^k\,dx = \frac{\Gamma(k+1)\,\Gamma(p)}{\Gamma(k+p+1)}.$$

From this we get

$$\delta(p) = \sum_{k=1}^{\infty} p^k \frac{\Gamma(p)}{\Gamma(k+p+1)}.$$

Ultimately we have the following expression for ε(p):

$$\varepsilon(p) = \frac{1}{p} + \Gamma(p)\sum_{k=1}^{\infty}\frac{p^k}{\Gamma(k+p+1)}.$$

VII. Derivation of Eq. (5.72)

In the integration with respect to the coordinate $\boldsymbol{\rho}_1$ (or $\boldsymbol{\rho}_2$) the integration region is a circle of diameter D. In the integration with respect to the dimensionless coordinate $\mathbf{r}$ (or $\boldsymbol{\rho}$) the integration region is likewise a circle, but its radius is now 2D/D = 2. Transformation to polar coordinates yields, in accord with Fig. 5.10,

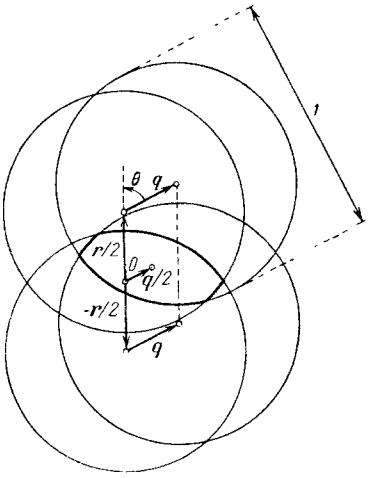


Fig. A.3

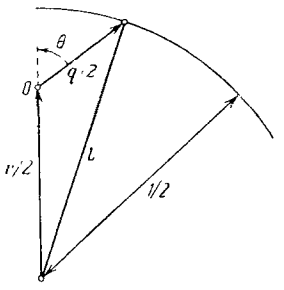


Fig. A.4

$$|\boldsymbol{r} \pm \boldsymbol{q}| = \sqrt{r^2 + q^2 \pm 2rq\cos\theta}, \quad \int d^2r = \int_0^{2\pi}\int_0^1 r\,dr\,d\theta.$$

The function $S(\mathbf{q}, \mathbf{r})$ in the $\mathbf{r}$ plane, however, differs from zero not in the entire circle of diameter 2, but only in a definite region whose boundary is determined by the condition that all four functions Π_0 in (5.70) overlap (see Fig. A.3). The symmetry center of this region is determined by the vector $\mathbf{q}/2$ in the $\mathbf{r}$ plane, and the equation of the boundary (in the first quadrant) is determined from the condition (Fig. A.4)

$$l \equiv \left|\frac{1}{2}\boldsymbol{r}(q, \theta) + \frac{1}{2}\boldsymbol{q}\right| = \frac{1}{2}$$

and is given by

$$r(q, \theta) \equiv \varepsilon(q, \theta) = -q\cos\theta + \sqrt{1 - q^2\sin^2\theta},$$

where, to avoid confusion, we introduced in the text a new function $\varepsilon(q, \theta)$. In the new coordinates the function $S(q, r, \theta)$ has identical integrals with respect to θ in each quadrant. Therefore,

$$\int_0^{2\pi}\int_0^1 (\cdot)\,r\,dr\,d\theta = 4\int_0^{\pi/2} d\theta \int_0^{\varepsilon(q,\theta)} (\cdot)\,r\,dr,$$

and it is this which leads to Eq. (5.72).

VIII. Reconstruction of Wavefront by Least Squares

Consider a square aperture consisting of n × n equal square subapertures at each of which the tilt of the wavefront is measured. We shall identify each subaperture by a pair of indices (i, j), where i is the number of the row and j the number of the column (i, j = 1, 2, ..., n). In the notation (5.135) the average local tilt of the wavefront within the limits of the subaperture is determined by the partial derivatives $u_{i,j}$ and $v_{i,j}$, which we shall refer arbitrarily to the center of the subaperture. The simplest algorithm for calculating the phase is to find in succession the phase increment on moving from one aperture to the other horizontally (subscript i) and vertically (subscript j). Then, however, we obtain two values of the phase in each internal node of the net. To eliminate this ambiguity is is expedient to use the least-squares method.

We seek the distribution of the phase $\varphi_{i,j}$ at the corners of the subapertures. The values of the phase at the lower left corner of an aperture with indices (i, j) will be designated $\varphi_{i,j}$, that in the upper left $\varphi_{i,j+1}$, in the upper right $\varphi_{i+1,j+1}$, and in the lower right $\varphi_{i+1,j}$. Clearly, the number of nodes at which the phase is to be determined is $(n + 1)^2$, and the number of measurements is $2n^2$, i.e., at $n \geq 3$ the number of equations exceeds the number of unknowns.

Assume that all the values of the phase at the corner of the net have been measured. As an estimate of the local tilts of the wavefront at the center of each subaperture we can assume

$$\begin{aligned} u'_{i,j} &= \frac{1}{2}(\varphi_{i+1,j+1} + \varphi_{i+1,j}) - \frac{1}{2}(\varphi_{i,j+1} + \varphi_{i,j}), \\ v'_{i,j} &= \frac{1}{2}(\varphi_{i+1,j+1} + \varphi_{i,j+1}) - \frac{1}{2}(\varphi_{i+1,j} + \varphi_{i,j}). \end{aligned} \tag{VIII.1}$$

We assume that the phase $\varphi_{i,j}$ (i, j = 1, 2, ..., n + 1) were measured well if the deviations of the estimates $u'_{i,j}$ and $v'_{i,j}$ (i, j = 1, 2, ..., n) from the really measured values $u_{i,j}$ and $v_{i,j}$ (i, j = 1, 2, ..., n) are minimal. As a measure of the deviation, we shall use the sum of the squares of the differences

$$\Delta = \sum_{i,j=1}^{n} [(u_{i,j} - u'_{i,j})^2 + (v_{i,j} - v'_{i,j})^2]. \tag{VIII.2}$$

This quantity reaches a minimum at values $\varphi_{i,j}$ that are solutions of the system of equations

$$\frac{\partial \Delta}{\partial \varphi_{p,q}} = 0 \qquad (p, q = 1, 2, \ldots, n+1). \tag{VIII.3}$$

The number of the equations in this system is $(n + 1)^2$, the same as the number of unknowns.

Substitution of (VIII.2) in (VIII.3) yields

$$2\sum_{i,j=1}^{n}\left[(u_{i,j} - u'_{i,j})\frac{\partial u'_{i,j}}{\partial \varphi_{p,q}} + (v_{i,j} - v'_{i,j})\frac{\partial v'_{i,j}}{\partial \varphi_{p,q}}\right] = 0$$
$$(p, q = 1, 2, \ldots, n+1). \tag{VIII.4}$$

Using (VIII.1) we can find the partial derivatives contained in (VIII.4):

$$2\frac{\partial u'_{i,j}}{\partial \varphi_{p,q}} = \delta_{p,i+1}\delta_{q,j+1} + \delta_{p,i+1}\delta_{q,j} - \delta_{p,i}\delta_{q,j+1} - \delta_{p,i}\delta_{q,j},$$
$$2\frac{\partial v'_{i,j}}{\partial \varphi_{p,q}} = \delta_{p,i+1}\delta_{q,j+1} + \delta_{p,i}\delta_{q,j+1} - \delta_{p,i+1}\delta_{q,j} - \delta_{p,i}\delta_{q,j}. \tag{VIII.5}$$

Here δ is the Kronecker delta defined as

$$\delta_{p,q} = \begin{cases} 1, & p = q, \\ 0, & p \neq q. \end{cases}$$

We now substitute (VIII.5) in (VIII.4) and sum over (i, j) at arbitrary (p, q). As a result we arrive at a system of $(n + 1)^2$ equations:

$$(u_{p-1,q-1} - u'_{p-1,q-1}) + (u_{p-1,q} - u'_{p-1,q}) - (u_{p,q-1} - u'_{p,q-1}) -$$
$$- (u_{p,q} - u'_{p,q}) + (v_{p-1,q-1} - v'_{p-1,q-1}) + (v_{p,q-1} - v'_{p,q-1}) - (v_{p-1,q} - v'_{p-1,q}) - (v_{p,q} - v'_{p,q}) = 0$$
$$(p, q = 1, 2, \ldots, n+1).$$

In this system of equations the primed terms are expressed in terms of the known phase values with the aid of Eqs. (VIII.1). For the inner nodes of the net of phase values, i.e., at p, q $\neq$ 1 and n + 1, we have

$$4\varphi_{p,q} - \varphi_{p-1,q-1} - \varphi_{p-1,q+1} - \varphi_{p+1,q+1} - \varphi_{p+1,q-1} =$$
$$= u_{p-1,q-1} + u_{p-1,q} - u_{p,q-1} - u_{p,q} + v_{p-1,q-1} + v_{p,q-1} - v_{p-1,q} - v_{p,q}. \tag{VIII.6}$$

At the edges of the net, at p, q = 1, n + 1, certain u and v terms vanish from (VIII.6), since they correspond to subapertures missing from the set of n^2 elements making up the total aperture. At the same time, for certain values of φ it is necessary to use extrapolated values. There are eight variants of Eqs. (VIII.6) when p, q = 1, n + 1. The system of equations for $\varphi_{p,q}$ can be written in general form by introducing the discretely varying functions

$$f_{p,q} = \begin{cases} 1, & 1 \leqslant p, q < n+1, \\ 0 & \text{in other cases,} \end{cases}$$
$$f'_{p,q} = \begin{cases} 1, & 1 \leqslant p, q \leqslant n+1, \\ 0 & \text{in other cases,} \end{cases} \tag{VIII.7}$$
$$g_{p,q} = \begin{cases} 4, & 1 < p, q < n+1, \\ 1, & p = q = 1,\ p = q = n+1, \\ 2 & \text{in other cases.} \end{cases}$$

We then obtain the following system of $(n + 1)^2$ equations for the $(n + 1)^2$ unknown values of the phases $\varphi_{p,q}$:

$$g_{p,q}\varphi_{p,q} - f_{p-1,q-1}\varphi_{p-1,q-1} - f_{p-1,q+1}\varphi_{p-1,q+1} - f_{p+1,q-1}\varphi_{p+1,q-1} -$$
$$- f_{p+1,q+1}\varphi_{p+1,q+1} = f'_{p-1,q-1}(u_{p-1,q-1} + v_{p-1,q-1}) +$$
$$+ f'_{p-1,q}(u_{p-1,q} - v_{p-1,q}) + f'_{p,q-1}(-u_{p,q-1} + v_{p,q-1}) +$$
$$+ f'_{p,q}(-u_{p,q} + v_{p,q})$$
$$(p, q = 1, 2, \ldots, n+1). \tag{VIII.8}$$

The coefficients in this system are the $2n^2$ measured values of the average tilts of the wavefront at each subaperture. It must be noted that a recurrence procedure cannot be used to solve (VIII.8), so that the usual method of solving equation systems must be used.

We write the system (VIII.8) in a somewhat different form

$$\sum_{p,q=1}^{n+1} (a_{p,q}^{i,j}\varphi_{p,q} - b_{i,j}) = 0, \tag{VIII.9}$$

or

$$\begin{aligned}
a_{1,1}^{1,1}\varphi_{1,1} + a_{1,2}^{1,1}\varphi_{1,2} + \ldots + a_{n+1,n+1}^{1,1}\varphi_{n+1,n+1} &= b_{1,1},\\
a_{1,1}^{1,2}\varphi_{1,1} + a_{1,2}^{1,2}\varphi_{1,2} + \ldots + a_{n+1,n+1}^{1,2}\varphi_{n+1,n+1} &= b_{1,2}.\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{1,1}^{1,n+1}\varphi_{1,1} + a_{1,2}^{1,n+1}\varphi_{1,2} + \ldots + a_{n+1,n+1}^{1,n+1}\varphi_{n+1,n+1} &= b_{1,n+1},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{1,1}^{n+1,n+1}\varphi_{1,1} + a_{1,2}^{n+1,n+1}\varphi_{1,2} + \ldots + a_{n+1,n+1}^{n+1,n+1}\varphi_{n+1,n+1} &= b_{n+1,n+1},
\end{aligned}$$

where

$$a_{p,q}^{i,j} = \begin{cases} -f_{p,q} & \text{at} \quad (p,q) = (i-1,\ j-1),\ (i-1,\ j+1),\ (i+1,\\ & \qquad j-1),\ (i+1,\ j+1),\\ g_{p,q} & \text{at} \quad (p,q) = (i,\ j),\\ 0 & \text{in the remaining cases,} \end{cases}$$

$$b_{i,j} = f'_{i-1,j-1}(u_{i-1,j-1} + v_{i-1,j-1}) + f'_{i-1,j}(u_{i-1,j} - v_{i-1,j}) + f'_{i,j-1}(-u_{i,j-1} + v_{i,j-1}) + f'_{i,j}(-u_{i,j} + v_{i,j}).$$

We note that the diagonal elements $a_{p,q}^{p,q} \neq 0$. Consequently, the solution of system (VIII.9) can be solved by using the Gauss method in its simplest modification. To this end we transform the matrix $a_{p,q}^{i,j}$ into the triangular matrix $c_{p,q}^{i,j}$ in accord with the following algorithm. Let m = 1, k = 1. For i = m + 1, m + 2, ..., n + 1 and j = k + 1, k + 2, ..., n + 1 we calculate the new matrix elements

$$\begin{aligned}
c_{p,q}^{i,j} &= 0 && \text{for } p = m,\ q = k,\\
c_{p,q}^{i,j} &= a_{p,q}^{i,j} + r_{i,j}a_{p,q}^{m,k} && \text{for } (p,q) > (m,k),
\end{aligned}$$

where

$$r_{i,j} = -a_{m,k}^{i,j}/a_{m,k}^{m,k},$$

and represent analogously the right-hand sides of the equations:

$$d_{i,j} = b_{i,j} + r_{i,j}b_{m,k}.$$

Increasing m and k in succession by unity, we repeat the operation of replacing the matrix elements until m and k run through all the values from 1 to n + 1. As a result, we obtain the triangular matrix $c_{p,q}^{i,j}$. The sought solution of system (VIII.9) is then obtained from the formulas

$$\varphi_{n+1,n+1} = d_{n+1,n+1}/c_{n+1,n+1}^{n+1,n+1},$$

$$\varphi_{i,j} = \left(d_{ij} - \sum_{p=1}^{n+1-i}\sum_{q=1}^{n+1-j} c_{p+i,q+j}^{i,j}\varphi_{p+i,q+j}\right)\Big/ c_{i,j}^{i,j}.$$

LITERATURE CITED

CHAPTER 1

1. T. H. Maiman, R. H. Hoskins, I. J. D. Haenens, et al., Phys. Rev., 123, 1151 (1961).
2. T. H. Maiman, Phys. Rev. Lett., 4, 564 (1960).
3. T. H. Maiman, Nature, 187, 493 (1960).
4. L. F. Johnson and K. Nassau, Proc. IRE, 49, 1704 (1961).
5. C. K. N. Patel, Phys. Rev., 136, 1187 (1964).
6. C. K. N. Patel, Appl. Phys. Lett., 7, 15 (1965).
7. D. Gabor, J. Inst. Electron. Eng., 93, 429 (1946).
8. J. A. Ville, Cables et Transmissions, No. 1, 16 (1948).
9. S. M. Rytov, Introduction to Statistical Physics [in Russian], Part I, Chap. VI, Nauka, Moscow (1976).
10. E. Timchmarsh, Introduction to the Theory of Fourier Integrals, Oxford Univ. Press (1937).
11. J. Dugundji, IRE Trans.: Inf. Theory, IT-4, 53 (1958).
12. S. L. Sobolev, Equations of Mathematical Physics [in Russian], Gostekhizdat, Moscow—Leningrad (1947).
13. A. Sommerfeld, Optics, Academic Press (1949).
14. A. Papoulis, Systems and Transforms with Applications in Optics, McGraw-Hill (1968).
15. V. A. Danilychev, O. M. Kerimov, and I. B. Kovsh, "Molecular lasers," in: Radio-Engineering, Adv. in Science, I. A. Boloshin (ed.) [in Russian], Vol. 12, VINITI, Moscow (1977).
16. A. A. Kuriksha, Quantum Optics and Optical Ranging [in Russian], Sov. Radio, Moscow (1973).
17. R. Barakat, Opt. Acta, 21, 903 (1974).
18. J. K. Jao and M. Elbaum, Proc. IEEE, 66, 781 (1978).
19. M. Born and E. Wolf, Principles of Optics, Pergamon Press (1970).
20. E. Jahnke and F. Emde, Tables of Higher Functions [Russian translation], Gostekhizdat, Moscow—Leningrad (1948).
21. S. M. Rytov, Yu. A. Kravtsov, and V. I. Tatarskii, Introduction to Statistical Physics [in Russian], Part II, Nauka, Moscow (1978).
22. A. S. Gurvich, A. I. Kon, V. L. Mironov, and S. S. Khmelevtsev, Laser Radiation in a Turbulent Atmosphere [in Russian], Chap. II, Sec. 4, Nauka, Moscow (1976).
23. R. F. Lutomirsky and H. T. Yura, Appl. Opt., 10, 1652 (1971).
24. Yu. A. Kravtsov and Z. I. Feizulin, Izv. Vyssh. Uchebn. Zaved., Radiofiz., 12, 896 (1969).
25. V. I. Smirnov, A Course of Higher Mathematics, Vol. 4, Chap. IV, Pergamon Press (1964).
26. H. T. Yura, Appl. Opt., 11, 1399 (1972).
27. V. I. Tatarskii, Propagation of Waves in a Turbulent Atmosphere [in Russian], Nauka, Moscow (1967).
28. M. E. Gracheva and A. S. Gurvich, Izv. Akad. Nauk SSSR, Fiz. Atm. Okeana, 16, 1107 (1980).

CHAPTER 2

1. P. V. Shcheglov, Problems of Optical Astronomy [in Russian], Chap. 8, Nauka, Moscow (1980).
2. N. D. Ystinov, I. N. Matveev, B. A. Shiryaev, et al., in: Optical Instrumentation and Holography. First All-Union Scientific—Technical Symposium [in Russian] (1976).
3. A. A. Kuriksha, Radiotekh. Elektron., No. 5, 771 (1968).
4. E. O'Neill, Introduction to Statistical Optics [Russian translation], Mir, Moscow (1966).
5. L. M. Soroko, Principles of Holography and Coherent Optics [in Russian], Nauka, Moscow (1971).
6. I. S. Reed, IRE Trans.: Inf. Theory, IT-17, 149 (1971).
7. M. Born and E. Wolf, Principles of Optics, Pergamon Press (1970).
8. I. N. Matveev, B. A. Shiryaev, and O. V. Kireenko, in: Optical Instrumentation and Holography. First All-Union Scientific—Technical Symposium [in Russian] (1976).
9. J. B. Breckenridge, J. Opt. Soc. Am., 65, 755 (1975).
10. H. W. Wessely and J. O. Bolstad, J. Opt. Soc. Am., 60, 678 (1970).

11. J. W. Goodman, W. H. Huntley, D. W. Jackson, and M. Lehmann, Appl. Phys. Lett., 8, 311 (1966).
12. J. W. Goodman, D. W. Jackson, M. Lehmann, and J. Knotts, Appl. Opt., 8, 1581 (1969).
13. B. J. Thompson, "Applications of holography," in: Laser Applications [Russian translation], Mir, Moscow (1974).

CHAPTER 3

1. W. E. Lamb, Phys. Rev., A134, 1429 (1964).
2. L. D. Landau and E. M. Lifshitz, Quantum Mechanics [in Russian], Chap. IV, Nauka, Moscow (1974).
3. F. A. Hopf and M. O. Scully, Phys. Rev., A179, 399 (1969).
4. A. Icsevgi and W. E. Lamb, Phys. Rev., A185, 517 (1969).
5. A. Papoulis, Systems and Transforms with Applications in Optics, McGraw-Hill (1968).
6. V. P. Konarev, I. N. Matveev, and V. V. Protopopov, Kvantovaya Elektron. (Moscow), 4, 882 (1977).
7. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, Relativistic Quantum Theory [in Russian], Part 1, Chap. V, Nauka, Moscow (1968).
8. E. Fermi, Quantum Mechanics [Russian translation], Mir, Moscow (1968).
9. H. Cramer, Mathematical Methods of Statistics, Princeton Univ. Press (1946).
10. K. Shimoda, H. Takachasi, and C. Townes, J. Phys. Soc. Jpn., 12, 687 (1957).
11. V. P. Konarev, I. N. Matveev, and S. M. Pshenichnikov, Kvantovaya Elektron. (Moscow), No. 7, 86 (1972).
12. V. A. Danilychev, O. M. Kerimov, and I. B. Kovsh, "Molecular lasers," in: Radio Engineering, Adv. in Sciences, I. A. Boloshin (ed.) [in Russian], Vol. 12, VINITI, Moscow (1977).
13. D. Marcuse, Integrated Optics, IEEE Press, New York (1973).
14. V. L. Koval'chuk, V. P. Konarev, I. N. Matveev, and S. M. Pshenichnikov, Prib. Tekh. Eksp., No. 3, 186 (1974).
15. P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Lett., 7, 118 (1961).
16. J. E. Midwinter, Appl. Phys. Lett., 12, 68 (1968).
17. M. Born and E. Wolf, Principles of Optics, Pergamon Press (1970), Chap. 14.
18. S. A. Akhmanov and R. V. Khokhlov, Problems of Nonlinear Optics [in Russian], VINITI, Moscow (1964).
19. F. Zernike and J. E. Midwinter, Applied Nonlinear Optics, Wiley (1973).
20. V. A. Kudryashov, I. N. Matveev, S. M. Pshenichnikov, et al., Prib. Tekh. Eksp., No. 3, 173 (1974).
21. Yu. A. Il'inskii and Yu. A. Yanait, Izv. Vyssh. Uchebn. Zaved., Radiofiz., 13, 37 (1970).
22. E. S. Boronin, M. I. Divlekeev, Yu. A. Il'inskii, et al., Pis'ma Zh. Eksp. Teor. Fiz., 10, 172 (1969).
23. E. S. Voronin, M. I. Divlekeev, Yu. A. Il'inskii, and V. S. Solomatin, Zh. Eksp. Teor. Fiz., 58, 51 (1970).
24. E. S. Voronin, Yu. A. Il'inskii, V. S. Solomatin, and V. V. Shuvalov, Kvantovaya Elektron. (Moscow), No. 4, 115 (1973).
25. Yu. A. Il'inskii and V. M. Petnikova, Kvantovaya Elektron. (Moscow), No. 5, 124 (1972).
26. Yu. A. Il'inskii and V. M. Petnikova, Kvantovaya Elektron. (Moscow), 1, 1133 (1974).

CHAPTER 4

1. A. Papoulis, Systems and Transforms with Applications in Optics, McGraw-Hill (1968).
2. D. L. Fried, Proc. IEEE, 55, 57 (1967).
3. D. L. Fried, IEEE J. Quantum Electron., QE-3, 213 (1967).
4. R. F. Lutomirsky and H. T. Yura, J. Opt. Soc. Am., 61, 482 (1971).
5. T. S. Moss, Infrared Phys., 16, No. 1/2, 29 (1976).
6. D. L. Spears, Infrared Phys., 17, 5 (1977).
7. R. L. Abrams, Appl. Phys. Lett., 25, 304 (1974).
8. J. H. McElroy, N. McAvoy, E. H. Jackson, et al., Trans. IEEE, 65, No. 2, 54 (1977).
9. R. A. Brandewie and W. C. Davis, Appl. Opt., 11, 1526 (1972).
10. M. Born and E. Wolf, Principles of Optics, Pergamon (1970).
11. E. Jahnke and F. Emde, Tables of Functions [Russian translation], Gostekhizdat, Moscow–Leningrad (1948).

12. V. P. Loginov and S. D. Fomenko, Zarubezhnaya Elektron., No. 11, 41 (1978).
13. S. M. Rytov, Introduction to Statistical Radiophysics [in Russian], Part I, Chap. VI, Nauka, Moscow (1976).

CHAPTER 5

1. A. Labeyrie, Astron. Astrophys., 6, 85 (1970).
2. D. Y. Gezari, A. Labeyrie, and R. V. Stachnik, Bull. Am. Astron. Soc., 3, 244 (1971).
3. D. Y. Gezari, A. Lebeyrie, and R. V. Stachnik, Astrophys. J., 173, No. 1, Part 2, LI (1972).
4. A. A. Tikovinin and P. V. Shcheglov, Usp. Fiz. Nauk, 129, 645 (1979).
5. P. V. Shcheglov, Problems of Optical Astronomy [in Russian], Nauka, Moscow (1980).
6. D. L. Fried, J. Opt. Soc. Am., 56, 1372 (1966).
7. M. Born and E. Wolf, Principles of Optics, Pergamon (1970).
8. E. O'Neill, Introduction to Statistical Optics [Russian translation], Mir, Moscow (1966).
9. D. Korff, J. Opt. Soc. Am., 63, 971 (1973).
10. H. A. McAlister, Sky Telescope, 53, 346 (1977).
11. H. W. Babcock, Publ. Astron. Soc. Pac., 65, 229 (1953).
12. H. W. Babcock, J. Opt. Soc. Am., 48, 500 (1958).
13. E. I. Vetrinskii (ed.), in: Adaptive Optics [Russian translation], Mir, Moscow (1980).
14. J. W. Hardy, Proc. IEEE, 66, No. 6, 65 (1978).
15. R. A. Muller and A. Buffington, J. Opt. Soc. Am., 64, 1200 (1974).
16. R. Hudgin, J. Opt. Soc. Am., 67, 393 (1977).
17. V. N. Mahajan, J. Govignon, and R. J. Morgan, Proc. SPIE, 228, 63 (1980).
18. L. E. Schmutz, J. K. Bowker, J. Feinleib, et al., Proc. SPIE, 228, 14 (1980).
19. J. W. Hardy, "Role of active optics in large telescopes," in: Optical Telescopes of the Future [Russian translation], P. V. Shcheglov (ed.), Mir, Moscow (1981).
20. A. Buffington, F. S. Crawford, R. A. Muller, et al., J. Opt. Soc. Am., 67, 298 (1977).
21. J. W. Hardy, J. E. Lefebvre, and C. L. Koliopoulos, J. Opt. Soc. Am., 67, 360 (1977).