DESIGN CONSIDERATIONS FOR OPTICAL HETERODYNE RECEIVERS: A REVIEW

John J. Degnan
Instrument Electro-optics Branch
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771

ABSTRACT

By its very nature, an optical heterodyne receiver is both a receiver and an antenna. Certain fundamental antenna properties of heterodyne receivers are described which set theoretical limits on the receiver sensitivity for the detection of coherent point sources, scattered light, and thermal radiation. In order to approach these limiting sensitivities, the geometry of the optical antenna-heterodyne receiver configuration must be carefully tailored to the intended application. The geometric factors which affect system sensitivity include the local oscillator (LO) amplitude distribution, mismatches between the signal and LO phasefronts, central obscurations of the optical antenna, and nonuniform mixer quantum efficiencies. The current state of knowledge in this area, which rests heavily on modern concepts of partial coherence, is reviewed.

Following a discussion of noise processes in the heterodyne receiver and the manner in which sensitivity is increased through time integration of the detected signal, we derive an expression for the mean square signal current obtained by mixing a coherent local oscillator with a partially coherent, quasi-monochromatic source. We then demonstrate the manner in which the IF signal calculation can be transferred to any convenient plane in the optical front end of the receiver. Using these techniques, we obtain a relatively simple equation for the coherently detected signal from an extended incoherent source and apply it to the heterodyne detection of an extended thermal source and to the back-scatter lidar problem where the antenna patterns of both the transmitter beam and heterodyne receiver must be taken into account. Finally, we consider the detection of a coherent source and, in particular, a distant point source such as a star or laser transmitter in a long range heterodyne communications system.

1. INTRODUCTION

Heterodyne or coherent detection can be advantageous in a variety of applications. Heterodyne receivers have at least two features which are qualitatively different from incoherent (or direct detection) receivers (ref. 1). First of all, the receiving bandwidth is determined by the IF bandwidth which, in principle, can be varied at will to give very high spectral resolution. Secondly, information related to the phase of the radiation signal is retained in the IF output and the outputs of two or more receivers can be correlated to make coherence measurements comparable to the aperture synthesis techniques of radio astronomy.

To achieve high spectral radiation with incoherent or direct detection systems, radiation filters or spectrometers must be utilized and the combination of very narrow bandwidth and high sensitivity (low loss) is usually difficult to realize. In general, a heterodyne receiver will be more sensitive than a direct detection receiver with an equivalent noise equivalent power (NEP) for spectral resolutions below a cutoff bandwidth which depends on the NEP and the infrared wavelength (refs. 1, 2). The sub-Doppler spectral resolution of heterodyne receivers can be exploited to study the molecular constituents and kinematics of remote sources yielding specific information such as altitude profiles of absolute abundance of the species, vertical temperature profiles, and wind velocities (ref. 3). In detecting extraterrestrial thermal sources, the information is gathered by passive heterodyne spectrometers whereas, in our own atmosphere or in planetary atmospheres visited by spacecraft, active backscatter lidars can be employed. In contrast to the above applications where the radiation signal is totally incoherent or only partially coherent, the signal from the laser transmitter in a heterodyne communication system (ref. 4) is coherent except as modified by atmospheric effects (ref. 5). This article attempts to present a unified theory of heterodyne receivers which addresses the optical design considerations for all of these applications.

A representative heterodyne receiver is illustrated in Figure 1. Signal radiation is collected by an optical antenna and focused, along with a local oscillator beam, onto a square-law frequency mixer operating at the radiation frequency. The latter beams have center frequencies ν_S and ν_L and powers P_S and P_L. The two frequencies mix to give an output spectrum centered at the intermediate frequency $\nu_{IF} = \nu_S - \nu_L$ where ν_{IF} is much smaller than the infrared frequencies ν_S and ν_L and typically on the order of a GHz or less. The resulting signal current is amplified by an IF amplifier of bandwidth B_{IF} and rectified by a nominally square-law detector to give a current output proportional to the power in the IF. This is usually input to a low-frequency filter or integrating circuit to further enhance the spectral resolution and/or sensitivity and is then recorded.

Although the present article will address most factors influencing the performance of the receiver in Figure 1, it will emphasize the design of the optical front end of the receiver for a variety of applications and, in particular, the manner in which the optical antenna geometry and local oscillator distribution affect system sensitivity. In Section 2 of this paper, we review the noise processes relevant to the IF signal and discuss the system signal-to-noise in the IF in terms of an as yet undefined mean square signal current. Section 3

briefly outlines the sensitivity improvement achieved by time integration techniques. In Section 4, we address the calculation of the mean square signal current in the mixer plane for a general, partially coherent, quasi-monochromatic source and, in Section 5, demonstrate the manner in which the IF signal calculation can be transferred to any convenient plane in the optical front end of the receiver. In Section 6, we apply the general result to the specific problem of coherently detecting an extended incoherent source. The results of that section are then applied to the heterodyne detection of an extended thermal source in Section 7 and to the backscatter lidar problem in Section 8 and some useful design guidelines are generated. In Section 9, we apply the results of Section 4 to the detection of a spatially coherent source such as a laser transmitter in a heterodyne communications system or a distant point source such as a star.

2. THE SIGNAL-TO-NOISE RATIO OF A HETERODYNE RECEIVER

The power signal-to-noise ratio of a heterodyne receiver is a measure of its sensitivity since setting the ratio equal to one permits calculation of the noise equivalent power (NEP). It is given, in most cases of interest, by (ref. 1)

$$\left(\frac{S}{N}\right)_{power} = \frac{\langle i_M^2 \rangle}{\langle i_S^2 \rangle + \langle i_T^2 \rangle + \langle i_B^2 \rangle + \langle i_J^2 \rangle + \langle i_A^2 \rangle} \tag{2.1}$$

We will leave the calculation of the mean square signal current $\langle i_M^2 \rangle$ to later sections and limit our present discussion to the various noise terms in the denominator of Equation (2.1).

The local oscillator induced shot noise, or quantum noise, $\langle i_S^2 \rangle$ is often the dominant noise if $h\nu >> KT_B$ where T_B is the equivalent blackbody temperature of a thermal source lying inside the antenna pattern of the receiver. Shot noise is due to fluctuations in the rate of arrival of LO photons. If the LO power is much greater than the signal power, the mean square shot noise current is given by

$$\langle i_S^2 \rangle = 2\beta e B_{IF} \bar{i}_{DC}$$

$$= \frac{2\beta e^2 B_{IF}}{h\nu} \iint_D d\vec{r}_D \, \eta_Q(\vec{r}_D) \, I_L(\vec{r}_D) \tag{2.2}$$

where $\bar{i}_{DC}$ is the DC current generated by the LO, e is the electronic charge, B_{IF} is the intermediate frequency bandwidth, and $h\nu$ is the photon energy. The integrand contains the detector quantum efficiency η_Q and the LO intensity I_L which are assumed to vary over the plane of the detector defined by the two-dimensional coordinate $\vec{r}_D$. The parameter β equals 1 for photoemissive mixers

while, for photoconductors, it equals 2 due to fluctuations in the generation and recombination of charge carriers as described by Levinstein (ref. 6).

One can rewrite Equation (2.2) in the more familiar form

$$\langle i_S^2 \rangle = \frac{2\bar{\eta}_Q \beta e^2 B_{IF} P_L}{h\nu} \tag{2.3}$$

if we define an average quantum efficiency $\bar{\eta}_Q$ by

$$\bar{\eta}_Q = \frac{\iint_D d\vec{r}_D \; \eta_Q(\vec{r}_D) \; I_L(\vec{r}_D)}{\iint_D d\vec{r}_D \; I_L(\vec{r}_D)} \tag{2.4}$$

and P_L is the local oscillator power incident on the detector.

Radiation from a thermal source contained within the receiver field of view and the receiver bandwidth B_{IF} will be coherently detected and subject to so-called "heterodyne amplification." In some experiments, such as in passive heterodyne spectrometry, this thermal source is the object of study, while in others it corresponds to unwanted background noise. We will show in later sections that it can be described by the equation

$$\langle i_T^2 \rangle = \frac{2\bar{\eta}_Q \eta_T e^2 B_{IF} P_L}{h\nu (e^{h\nu/KT} - 1)} \tag{2.5}$$

where η_T is an overall efficiency which depends in part on the design of the optical front end.

Fluctuations in background radiation, which spectrally is outside the receiver bandwidth but within the infrared response band of the mixer, will also produce noise currents, given by $\langle i_B^2 \rangle$ in Equation (2.1) as will sources of radiation outside the antenna pattern of the receiver but inside the heterodyne receiving bandwidth. McLean and Putley (ref. 7) have derived expressions for this noise component which are complicated functions of wavelength, spectral interval, detector area and temperature, and field of view. The latter noise is not amplified by the heterodyne process, however, and can be rendered negligible by choosing a large enough local oscillator power and by spatially and spectrally filtering the input radiation.

Two other important sources of noise are Johnson or thermal noise associated with the mixer and the IF amplifier. The mixer noise is given by

$$\langle i_J^2 \rangle = \frac{4KT_M B_{IF}}{R_M} \tag{2.6}$$

where T_M and R_M are the mixer's (or mixer load resistor's) temperature and resistance as seen by the IF amplifier. For most cooled mixers, this would be negligible compared with the amplifier noise given by

$$\langle i_A^2 \rangle = \frac{4KT_A B_{IF}}{MR_A} \tag{2.7}$$

where T_A and R_A are the amplifier's noise temperature and input resistance, and M is a factor less than unity which accounts for impedance mismatches between the mixer and amplifier.

Clearly, other sources of noise exist. "Excess noise" is common in receivers which employ diode laser local oscillators and generally arises from multimode effects or other non-ideal behavior in the LO. Noise can also be introduced at the electrical contacts to the mixer element or by temperature fluctuations in the mixer. These sources are unique to specific systems and will not be considered further here.

With sufficient LO power, most of the above noise sources can be made negligible relative to the quantum noise $\langle i_S^2 \rangle$ and/or the background thermal noise contribution $\langle i_T^2 \rangle$. If the mean square signal current is given by an expression of the form

$$\langle i_M^2 \rangle = 2\bar{\eta}_Q \eta_{HET} \left(\frac{e}{h\nu}\right)^2 P_S P_L \tag{2.8}$$

where P_S is the received signal power and η_{HET} is an as yet undefined heterodyne receiver efficiency, then, under strong LO illumination, the signal-to-noise ratio tends to

$$\left(\frac{S}{N}\right)_{power} = \frac{\langle i_M^2 \rangle}{\langle i_S^2 \rangle + \langle i_B^2 \rangle} = \frac{\eta_{HET} P_S}{h\nu B_{IF}\left\{\beta + \eta_T\left[\exp(h\nu/KT) - 1\right]^{-1}\right\}} \tag{2.9}$$

Setting the latter ratio equal to one and solving for P_S/B_{IF} yields the noise equivalent power per unit bandwidth; i.e.,

$$NEP\,(W/Hz) = \frac{h\nu}{\eta_{HET}}\left\{\beta + \eta_T\left[\exp(h\nu/KT) - 1\right]^{-1}\right\} \tag{2.10}$$

where η_{HET} and η_T both depend on the optical front end geometry. In the quantum noise limit ($h\nu >> KT$), Equation (2.10) reduces to

$$NEP(W/Hz) = \frac{\beta h\nu}{\eta_{HET}} \tag{2.11}$$

whereas, in the thermal limit ($h\nu << KT$), it becomes

$$NEP(W/Hz) = \frac{\eta_T}{\eta_{HET}} KT \tag{2.12}$$

If we include mixer and amplifier Johnson noise, we can write for a general photoconductor

$$NEP(W/Hz) = \frac{2h\nu}{\eta_{HET}} + \frac{\eta_T h\nu}{\eta_{HET}\left[\exp(h\nu/KT) - 1\right]} + \frac{K(T_M + T_A)}{G} \tag{2.13}$$

where G is the "conversion gain" defined by Arams et al. (ref. 8).

3. DETECTION AND TIME INTEGRATION

If the power signal-to-noise ratio in the IF is less than unity, the signal can be detected by integrating the detector output over a sufficiently long period of time. The voltage signal-to-noise ratio at the filter output in Figure 1 is linearly related to the power S/N by the equation (ref. 1)

$$\left(\frac{S}{N}\right)_V = \left(\frac{S}{N}\right)_{power} \frac{1}{\sqrt{2}} \left(\frac{B_{IF}}{B_o}\right)^{1/2} \tag{3.1}$$

The latter equation assumes that the IF amplifier has a rectangular bandpass spectrum (double sideband), the rectifying detector is an ideal square-law device, the final output filter has a noise bandwidth B_o much less than B_{IF} and the power S/N is much less than unity. Smith (ref. 9) has considered the more general case where the IF amplifier is not strictly square-law and does not have a rectangular bandpass spectrum. He has also considered power S/N ratios much greater than unity. If the output filter is a single stage RC circuit such that $B_o = \tau_o/4 = RC/4$, Equation (3.1) becomes

$$\left(\frac{S}{N}\right)_V = \left(\frac{S}{N}\right)_{power} \sqrt{2B_{IF}\tau_o} \tag{3.2}$$

4. COHERENT DETECTION OF A GENERAL QUASI-MONOCHROMATIC SOURCE

We turn now to the calculation of the mean square signal current $\langle i_M^2 \rangle$ for a general quasi-monochromatic source. This problem has been considered previously by Rye (ref. 10) and McGuire (ref. 11). With only minor modification, the derivation given here parallels that of McGuire. If we assume that the detected radiation lies within a frequency bandwidth $\Delta\nu_S$ that is narrow with respect to the center frequency ν_S, the real signal field at the mixer plane can be represented by an expression of the form

$$E_S(\vec{r}_D,t) = \sqrt{2}\ \mathrm{Re}\left\{E_S(\vec{r}_D,t)\ e^{i\omega_S t}\right\} \tag{4.1}$$

where $\omega_S = 2\pi\nu_S$ and the complex signal field envelope $E_S(\vec{r}_D,t)$ at the point $\vec{r}_D$ in the detector plane varies slowly in time relative to the exponential $\exp(i\omega_S t)$. The time dependence of the envelope might reflect the modulated output of a transmitter laser in a heterodyne communications system, the amplitude and phase fluctuations inherent in the signal from an incoherent thermal source or backscatter lidar, or even the effects of atmospheric turbulence on the signal. The envelope, through its dependence on the detector coordinate $\vec{r}_D$, also contains spatially dependent amplitude and phasefront information.

If we represent the LO field by a similar expression, the current out of the square-law mixer is given by

$$i(t) = \frac{e}{2h\nu} \iint_D d\vec{r}_D\ \eta_Q(\vec{r}_D) \left[\mathrm{Re}\left\{E_S(\vec{r}_D,t)\ e^{i\omega_S t} + E_L(\vec{r}_D,t)\ e^{i\omega_L t}\right\}\right]^2 \tag{4.2}$$

where ω_L is the LO center frequency and the integral is over the active detector area. Upon performing the quadratic multiplication of fields in Equation (4.2), we obtain both sum and difference frequencies. High-frequency sum terms varying as $\exp(\pm 2i\omega_S t)$, $\exp(\pm 2i\omega_L t)$, $\exp(\pm i(\omega_S+\omega_L)t)$, lie outside the bandwidth of the mixer and hence can be ignored. The difference terms produce two "DC" currents corresponding to the average signal and local oscillator induced currents and an additional mixing term given by

$$i_M(t) = \left(\frac{2e}{h\nu}\right) \iint_D d\vec{r}_D\ \eta_Q(\vec{r}_D)\ \mathrm{Re}\left\{E_S(\vec{r}_D,t)\ E_L^*(\vec{r}_D,t)\ e^{i\omega_{IF} t}\right\} \tag{4.3}$$

where the IF frequency $\omega_{IF} = \omega_S - \omega_L$. Squaring Equation (4.3) yields

$$i_M^2(t) = \left(\frac{2e}{h\nu}\right)^2 \iint_D d\vec{r}_D \iint_D d\vec{r}_D{}' \; \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D{}')$$

$$\times \mathrm{Re}\left\{E_S(\vec{r}_D,t)\; E_L^*(\vec{r}_D,t)\; e^{i\omega_{IF}t}\right\} \mathrm{Re}\left\{E_S(\vec{r}_D{}',t)\; E_L^*(\vec{r}_D{}',t)\; e^{i\omega_{IF}t}\right\}$$

$$= 2\left(\frac{e}{h\nu}\right)^2 \iint_D d\vec{r}_D \iint_D d\vec{r}_D{}' \; \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D{}')$$

$$\times \left[\mathrm{Re}\left\{E_S(\vec{r}_D,t)\; E_S^*(\vec{r}_D{}',t)\; E_L(\vec{r}_D{}',t)\; E_L^*(\vec{r}_D,t)\right\}\right.$$

$$\left. + \mathrm{Re}\left\{E_S(\vec{r}_D,t)\; E_S(\vec{r}_D{}',t)\; E_L^*(\vec{r}_D,t)\; E_L^*(\vec{r}_D{}',t)\; e^{i2\omega_{IF}t}\right\}\right] \qquad (4.4)$$

If we average the above expression over a time interval T short compared to the coherence times of the signal and local oscillator field (T_S and T_L) but long compared to the IF beat period, T_{IF}, we may write

$$i_M^2(t) = \frac{1}{T}\int_{t-T/2}^{t+T/2} dt\; i_M^2(t)$$

$$\approx 2\left(\frac{e}{h\nu}\right)^2 \iint_D d\vec{r}_D \iint d\vec{r}_D{}' \; \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D{}')$$

$$\times E_S(\vec{r}_D,t)\; E_S^*(\vec{r}_D{}',t)\; E_L(\vec{r}_D{}',t)\; E_L^*(\vec{r}_D,t) \qquad (4.5)$$

since the field envelopes can be viewed as effectively constant over this time interval and hence the terms varying as $\exp(\pm 2i\omega_{IF}t)$ in Equation (4.4) average to zero over an IF beat period. In certain applications, such as passive heterodyne spectrometry of a thermal source, the integration time can be arbitrarily long. The limit of Equation (4.5) as T approaches infinity is then

$$\langle i_M^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{t-T/2}^{t+T/2} dt\; i_M^2(t)$$

$$= 2\left(\frac{e}{h\nu}\right)^2 \iint_D d\vec{r}_D \iint d\vec{r}_D'\; \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D')$$

$$\times \langle E_S(\vec{r}_D,t)\; E_S^*(\vec{r}_D',t)\rangle \langle E_L(\vec{r}_D',t)\; E_L^*(\vec{r}_D,t)\rangle \qquad (4.6)$$

where we have invoked the fact that the signal and local oscillator fields are statistically independent and hence the fourth-order correlation function $\langle E_S(\vec{r}_D,t)\; E_S^*(\vec{r}_D',t)\; E_L(\vec{r}_D',t)\; E_L^*(\vec{r}_D,t)\rangle$ can be written as the product of two second-order functions. The second-order correlation functions can be related to quantities appearing in the theory of partial coherence by noting that the "mutual coherence function" (MCF) of a quasi-monochromatic, stationary optical signal field is defined by (ref. 12)

$$\Gamma_S(\vec{r}_1,\vec{r}_2,\tau) = \langle E_S(\vec{r}_1,t+\tau)\; E_S^*(\vec{r}_2,t)\rangle\, e^{i\omega_S\tau} \qquad (4.7)$$

Under the assumption of cross spectral purity (refs. 12, 13), the spatial and time variables are separable leading to

$$\Gamma_S(\vec{r}_1,\vec{r}_2,\tau) = J_S(\vec{r}_1,\vec{r}_2)\; g(\tau)\; e^{i\omega_S\tau} \qquad (4.8)$$

where $g(0) = 1$ and $J_S(\vec{r}_1,\vec{r}_2)$ is the "mutual intensity function" (MIF) of the signal field. From Equations (4.7) and (4.8) we note that $\langle E_S(\vec{r}_D,t)\; E_S^*(\vec{r}_D',t)\rangle = \Gamma_S(\vec{r}_D,\vec{r}_D',0) = J_S(\vec{r}_D,\vec{r}_D')$ and hence Equation (4.6) can be written in its final form

$$\langle i_M^2 \rangle = 2\left(\frac{e}{h\nu}\right)^2 \iint_D d\vec{r}_D \iint_D d\vec{r}_D'\; \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D')\; J_S(\vec{r}_D,\vec{r}_D')\; J_L(\vec{r}_D',\vec{r}_D) \qquad (4.9)$$

where $J_S(\vec{r}_D,\vec{r}_D')$ and $J_L(\vec{r}_D',\vec{r}_D)$ are the mutual intensity functions of the signal and local oscillator fields in the detector plane. Calculation of the mean square mixing current by means of Equation (4.9) is not always a simple task due to the difficulty in computing $J_S(\vec{r}_D,\vec{r}_D')$ for many sources of practical interest. In ensuing sections, we will demonstrate how the calculation can be carried out in optical planes other than the detector plane and the enormous simplifications that often result.

Before closing this section, it is worthwhile to note two useful properties of the mutual intensity function; i.e.,

$$J_S^*(\vec{r}_D,\vec{r}_D') = \langle E_S^*(r_D,t)\ E_S(\vec{r}_D',t)\rangle = J_S(\vec{r}_D',\vec{r}_D) \tag{4.10}$$

and

$$J_S(\vec{r}_D,\vec{r}_D) = I_S(\vec{r}_D) \tag{4.11}$$

where $I_S(\vec{r}_D)$ is the time averaged signal intensity at the point $\vec{r}_D$.

5. PROPAGATION OF THE MUTUAL INTENSITY FUNCTION

Consider the signal electric field propagating from the antenna plane in Figure 2 to the detector plane. Small angle scalar diffraction theory (ref. 12) gives the electric field in the detector plane; i.e.,

$$E_S(\vec{r}_D,t) = \iint d\vec{r}_A\ P_A(\vec{r}_A)\ E_S\left(\vec{r}_A,t - \frac{r_{AD}}{c}\right)\left(-\frac{ie^{ikr_{AD}}}{\lambda r_{AD}}\right) \tag{5.1}$$

where $k = 2\pi/\lambda$, $P_A(\vec{r}_A)$ is the antenna pupil function and the term in brackets corresponds to a Huygen's wavelet emanating from a point $\vec{r}_A$ in the antenna plane and traveling a distance r_{AD} to a point $\vec{r}_D$ in the mixer plane. Then, from the definition of the mutual intensity function (MIF), it is clear that

$$J_S(\vec{r}_D,\vec{r}_D') \equiv \langle E_S(\vec{r}_D,t)\ E_S^*(\vec{r}_D',t)\rangle$$

$$= \iint d\vec{r}_A \iint d\vec{r}_A'\ P_A(\vec{r}_A)\ P_A(\vec{r}_A')$$

$$\times \left[\frac{e^{ik(r_{AD}-r_{A'D'})}}{\lambda^2 r_{AD} r_{A'D'}}\right]\left\langle E_S\left(\vec{r}_A,t - \frac{r_{AD}}{c}\right)\ E_S^*\left(\vec{r}_A',t - \frac{r_{A'D'}}{c}\right)\right\rangle \tag{5.2}$$

For a stationary process, the time origin is of no consequence and therefore

$$\left\langle E_S\left(\vec{r}_A,t - \frac{r_{AD}}{c}\right)\ E_S^*\left(\vec{r}_A',t - \frac{r_{A'D'}}{c}\right)\right\rangle = \left\langle E_S(\vec{r}_A,t)\ E_S^*\left[\vec{r}_A',t - \frac{(r_{A'D'} - r_{AD})}{c}\right]\right\rangle \tag{5.3}$$

Now, if the transverse dimensions of the antenna and detector pupil are small compared to the coherence length of the signal radiation defined by $l = c/\Delta\nu_S$, the variation of the signal electric field over a time interval $t = (r_{A'D'} - r_{AD})/c$ is negligible and Equation (5.3) is effectively the signal MIF in the antenna plane. Equation (5.2) then becomes the propagation law for the MIF as first derived by Zernike (refs. 12, 14); i.e.,

$$J_S(\vec{r}_D,\vec{r}_D') = \iint d\vec{r}_A \iint d\vec{r}_A'\; P_A(\vec{r}_A)\; P_A(\vec{r}_A') \left[\frac{e^{ik(r_{AD}-r_{A'D'})}}{\lambda^2 r_{AD} r_{A'D'}}\right] J_S(\vec{r}_A,\vec{r}_A') \tag{5.4}$$

If we substitute Equation (4.5) in (4.9) for the mean square signal current and reverse the order of integration, we obtain

$$\langle i_M^2 \rangle = 2\left(\frac{e}{h\nu}\right)^2 \iint d\vec{r}_A \iint d\vec{r}_A'\; P_A(\vec{r}_A)\; P_A(\vec{r}_A')\; J_S(\vec{r}_A,\vec{r}_A') \times \left\{ \iint d\vec{r}_D \iint d\vec{r}_D'\; P_D(\vec{r}_D)\; P_D(\vec{r}_D')\; \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D')\; J_L(\vec{r}_D',\vec{r}_D) \left[\frac{e^{ik(r_{AD}-r_{A'D'})}}{\lambda^2 r_{AD} r_{A'D'}}\right] \right\} \tag{5.5}$$

where $P_D(\vec{r}_D)$ is the mixer pupil function. If we now define an effective local oscillator field given by

$$E_E(\vec{r}_D',t) \equiv \eta_Q(\vec{r}_D')\; E_L(\vec{r}_D',t) \tag{5.6}$$

the corresponding effective MIF is then equal to

$$J_E(\vec{r}_D',\vec{r}_D) = \eta_Q(\vec{r}_D)\; \eta_Q(\vec{r}_D')\; J_L(\vec{r}_D',\vec{r}_D) \tag{5.7}$$

Substituting Equation (5.7) into (5.5) and comparing the resulting expression with the MIF propagation law (5.4), we note that the bracketed term in Equation (5.5) is simply the MIF of the effective local oscillator back-propagated to the antenna plane. We may therefore write for the mean square mixing current

$$\langle i_M^2 \rangle = 2\left(\frac{e}{h\nu}\right)^2 \iint d\vec{r}_A \iint d\vec{r}_A'\; P_A(\vec{r}_A)\; P_A(\vec{r}_A')\; J_S(\vec{r}_A,\vec{r}_A')\; J_E(\vec{r}_A',\vec{r}_A) \tag{5.8}$$

The physical significance of Equation (5.8) is that the calculation of mean square IF signal current can be carried out in any convenient optical plane as first pointed out by Rye (ref. 10). This has practical importance since it is usually easier, for example, to compute the backpropagation of a coherent LO electric field through an optical system than to propagate the MIF of an incoherent source in a forward direction through the system to the mixer. This fact will be well illustrated in later sections.

Although we have considered only free space propagation in the present derivation, the approach is equally valid when intervening optical elements such as lenses, mirrors, and apertures are present. The simple Huygens wavelet in Equation (5.1) is then replaced by an appropriate transmission function for the optical system (refs. 10, 12).

6. HETERODYNE DETECTION OF AN EXTENDED INCOHERENT SOURCE

The expressions derived up to this point have assumed a general, partially coherent, quasi-monochromatic source. We consider now an important practical application in which the signal radiation emanates from an extended incoherent source and propagates to the antenna plane as in Figure 3. The propagation of the MIF proceeds in precisely the same fashion as in the previous section except that there is no coherence between the Huygens wavelets emanating from the infinitesimal sources located at $\vec{r}_S$ and $\vec{r}_S'$. Thus the second-order correlation function in the source-antenna plane version of Equation (5.3) becomes

$$\left\langle E_S(\vec{r}_S,t)\; E_S^*\left[\vec{r}_S',t-\left(\frac{r_{S'A'}r_{SA}}{c}\right)\right]\right\rangle \approx \langle E_S(\vec{r}_S,t)\; E_S^*(\vec{r}_S',t)\rangle = I_S(\vec{r}_S)\;\delta(\vec{r}_S-\vec{r}_S') \tag{6.1}$$

where $I_S(\vec{r}_S)$ is the time averaged radiation intensity at the point $\vec{r}_S$ in the source plane and $\delta(\vec{r}_S-\vec{r}_S')$ is the two-dimensional Dirac delta function. It can be shown that substitution of Equation (6.1) into the source-antenna plane version of Equation (5.2) and performing the double integral over $\vec{r}_S'$ yields the propagation law for the MIF of an incoherent source (ref. 13); i.e.,

$$J_S(\vec{r}_A,\vec{r}_A') = \lambda^2 \iint_S d\vec{r}_S\; I_S(\vec{r}_S)\left[\frac{e^{ik(r_{SA}-r_{SA'})}}{\lambda^2 r_{SA}r_{SA'}}\right] \tag{6.2}$$

where the integral is over the finite dimensions of the source. We may now substitute Equation (6.2) into (5.8) and reverse the order of integration to obtain for the mean square IF signal current

$$\langle i_M^2 \rangle = 2\left(\frac{e}{h\nu}\right)^2 \lambda^2 \iint_S d\vec{r}_S \; I_S(\vec{r}_S)$$
$$\times \left\{ \iint d\vec{r}_A \iint d\vec{r}_A' \; P_A(\vec{r}_A) \; P_A(\vec{r}_A') \left[\frac{e^{ik(r_{SA}-r_{SA'})}}{\lambda^2 r_{SA} r_{SA'}} \right] J_E(\vec{r}_A', \vec{r}_A) \right\} \qquad (6.3)$$

Through use of the MIF propagation law given by Equation (5.4), we recognize the bracketed term in Equation (6.3) as the mutual intensity function of the backpropagated effective local oscillator (BPELO) evaluated at the points $\vec{r}_S = \vec{r}_S'$. But, since $J_E(\vec{r}_S, \vec{r}_S) = I_E(\vec{r}_S)$, the time averaged intensity of the BPELO in the source plane, Equation (6.3) reduces to the relatively simple expression

$$\langle i_M^2 \rangle = 2\left(\frac{e}{h\nu}\right)^2 \lambda^2 \iint_S d\vec{r}_S \; I_S(\vec{r}_S) \; I_E(\vec{r}_S) \qquad (6.4)$$

Thus we have the very useful result that the mean square IF signal current is proportional to the overlap integral of the extended incoherent source intensity with the backpropagated effective LO intensity. In the next two sections, we will apply this result to the detection of thermal radiation and to the backscatter lidar problem.

7. THERMAL SOURCE DETECTION

The total power ΔP radiated into a hemisphere, within the IF bandwidth B_{IF}, from a small area ΔA on a blackbody is

$$\Delta P = \frac{2\pi}{\lambda^2} \frac{h\nu B_{IF}}{\left[\exp(h\nu/KT) - 1\right]} \Delta A \qquad (7.1)$$

Only the power emitted in the direction of the receiver contributes to the signal MIF in the antenna plane. Thus, if the receiver is in a direction normal to the plane of the blackbody, we must multiply the above expression by a factor $1/\pi$ corresponding to the power emitted per steradian in the normal direction. We must also multiply by 1/2 to account for the fact that the heterodyne receiver detects only one polarization component. Thus, the intensity to be substituted into Equation (6.4) is given by

$$I_S = \left(\frac{1}{2}\right)\left(\frac{1}{\pi}\right)\frac{\Delta P}{\Delta A} = \frac{1}{\lambda^2} \frac{h\nu B_{IF}}{\left[\exp(h\nu/KT) - 1\right]} \qquad (7.2)$$

and Equation (6.4) becomes

$$\langle i_M^2 \rangle = \frac{2e^2 B_{IF}}{h\nu\left[\exp(h\nu/KT) - 1\right]} \iint_S d\vec{r}_S \; I_E(\vec{r}_S) \tag{7.3}$$

where the integral is simply the total backpropagated effective LO power subtended by the source.

If the dominant noise mechanism is the LO-induced shot noise given by Equation (2.3), the IF signal-to-noise ratio is

$$\left(\frac{S}{N}\right)_{power} = \frac{\langle i_M^2 \rangle}{\langle i_S^2 \rangle} = \frac{\eta_T}{\beta\left[\exp(h\nu/KT) - 1\right]} \tag{7.4}$$

where η_T is the overall heterodyne receiver efficiency for thermal source detection introduced in Equation (2.5) and defined by

$$\eta_T = \frac{1}{\bar{\eta}_Q P_L} \iint_S d\vec{r}_S \; I_E(\vec{r}_S) \tag{7.5}$$

where $\bar{\eta}_Q$ is the average mixer quantum efficiency defined by Equation (2.4) and P_L is the LO power incident on the detector. If the mixer quantum efficiency is uniform, Equation (7.5) reduces to

$$\eta_T = \frac{\eta_Q}{P_L} \iint_S d\vec{r}_S \; I_L(\vec{r}_S) \tag{7.6}$$

where we have used Equations (5.7) and (4.11). The quantity $I_L(\vec{r}_S)$ is the intensity of the actual backpropagated LO rather than the effective LO. The quantity η_T replaces the mixer efficiency in the corresponding equations in the classic paper by Siegman (ref. 15).

If the source is so large that the backpropagated LO is contained entirely within its disk radius, the integral in Equation (7.6) is simply the total LO power in the source plane. Except for an atmospheric transmission factor η_A, the latter is equal to the backpropagated LO power exiting from the antenna. Thus, the overall heterodyne efficiency (7.6) can be broken down into several components; i.e.,

$$\eta_T = \eta_Q \eta_A \eta_o \eta_R \tag{7.7}$$

where η_o takes into account routine optical losses due to reflections and scattering while η_R is a geometric efficiency which takes into account vignetting, central obstructions, LO phasefront curvature, etc. in the optical

antenna. Numerically, η_R is equal to the fraction of the original LO power which exits from the antenna during backpropagation.

As an illustration, consider the Cassegrain telescope in Figure 4. Let us assume that the mixer is illuminated by the fundamental gaussian mode of the local oscillator laser. If the gaussian mode is not truncated too badly by the mixer perimeter or by the secondary mirror, it remains gaussian until it is truncated by the primary mirror of radius a and centrally obscured by the secondary of radius b in the antenna plane. The geometric efficiency η_R is then given by

$$\eta_R = \frac{2}{\pi\omega^2}\left[2\pi \int_b^a dr\ re^{-2(r/\omega)^2}\right] = e^{-\gamma^2\alpha^2} - e^{-\alpha^2} \tag{7.8}$$

where ω is the gaussian spot radius in the antenna plane and we have defined two parameters (ref. 16) $\alpha = a/\omega$ and $\gamma = b/a$. The geometric efficiency has been plotted as a function of α and γ in Figure 5.

The important thing to note in Figure 5 is that, for a given nonzero value of the linear obscuration ratio $\gamma = b/a$, the optimum efficiency is less than what one would expect based on simple blockage of the incoming radiation by the central obscuration. For example, $\gamma = 0.5$ would imply an areal obscuration efficiency of $1 - \gamma^2$ or 75%. The peak efficiency in Figure 5, however, would only be about 47% if one were to choose an optimum gaussian spot radius corresponding to $\alpha = 1.3$. Nonoptimum choices clearly result in significantly worse performance.

Clearly, to maximize the efficiency of coherent detection of a thermal source which fills the receiver field of view, one wishes to choose an optical geometry which allows the effective backpropagated LO to exit from the telescope with near-unity efficiency. Although this is most easily accomplished with off-axis reflective telescope geometries which eliminate the central obscuration problem, one is not limited to such geometries in general. For example, if we use appropriate masks in the LO beam to create a local oscillator distribution in the mixer plane which matches the Airy pattern of the centrally obscured Cassegrain telescope in Figure 4, the backpropagated LO will form an annular disk in the antenna plane which matches the antenna pupil function and provides unity transmission. This result assumes, of course, that the mixer quantum efficiency is reasonably uniform. The transmission loss of the beam splitter in Figure 4 is included in the optical efficiency η_o.

For such large sources, the efficiency is not sensitive to the wavefront curvature of the LO beam except to the extent that it modifies the LO transmission through the antenna pupil. For example, if one considers two systems, projecting the same gaussian spot size in the antenna plane of Figure 4 but having two different radii of curvature for the LO phasefronts, the fractional transmission and hence the receiver efficiency will be the same. The system with the wider backpropagated LO divergence will detect point sources near the optic axis with less sensitivity but this will be compensated for by the

detection of additional point sources which are beyond the field of view of the receiver with the smaller backpropagated LO divergence. On the other hand, if the source is of limited spatial extent, maximum detection efficiency dictates that the backpropagated LO be contained totally within the source pupil function and hence LO phasefront curvature effects will play a more important role. For small thermal sources in the near field of the receiver, as in a laboratory experiment, this can be accomplished by choosing an optical system which effectively focuses the backpropagated LO onto the target source and provides near-unity transmission efficiency for the backpropagated LO.

8. INCOHERENT BACKSCATTER LIDAR

Consider the lidar system in Figure 6. An outgoing pulse of temporal width δ is transmitted through the atmosphere illuminating the aerosol scatterers in its path. The mixer current at time t is due to radiation scattered at a time $t - R/c$ from a volume defined by the length $c\delta/2$ within the receiver field of view as determined by the backpropagated effective LO intensity. Although the aerosol scatterers are randomly spaced and typically many wavelengths apart, the return is not strictly incoherent since the scatterers within the volume of interest are "frozen" in their positions during the passage of a short laser pulse, thereby producing a coherent or "speckle" component in the return. Thus, based on a single return, one cannot perform the long time average necessary to progress from Equation (4.5) to (4.6) in our derivation of the mean square mixing current $\langle i_M^2 \rangle$. However, if we imagine repeating the lidar experiment many times over the same source volume and obtaining an average current waveform out of the mixer, the coherent component would be expected to average to zero over the ensemble of measurements due to the random relative motions of the scatterers. After averaging a sufficiently large number of current waveforms, we would then be left with the incoherent component. Thus, if the physical process being observed is ergodic, i.e., ensemble averages are equal to time averages, the mean square mixing current will be given by $\langle i_M^2 \rangle$ where the notation now applies to either an ensemble average or time average since the two are equivalent.

With the additional argument given above, we can apply Equation (6.4) to the pulsed backscatter lidar problem. The source intensity function I_S which is now a function of range (Z coordinate) as well as the transverse coordinates, is given by

$$I_S(R,\vec{r}_S) = p\ \rho(R,\vec{r}_S)\left(\frac{c\delta}{2}\right)\frac{d\sigma(\pi)}{d\Omega}\ I_T(R,\vec{r}_S) \tag{8.1}$$

where $I_T(R,\vec{r}_S)$ is the intensity of the coherent transmitter beam at the range R and transverse coordinate $\vec{r}_S$, $d\sigma(\pi)/d\Omega$ is the differential scattering cross section in the backward direction, $c\delta/2$ is the length of the scattering volume, $\rho(R,\vec{r}_S)$ is the density distribution of scatterers, and p is a factor of order unity or less which takes into account depolarization effects. The product $\left[d\sigma(\pi)/d\Omega\right] I_T(R,\vec{r}_S)$ is the power scattered in the backward direction per steradian by a single scatterer located at the coordinates

$(R,\vec{r}_S)$ while the product $\rho(R,\vec{r}_S)(c\delta/2)$ is the number of scatterers per unit cross-sectional area in the source volume. Substituting Equation (8.1) into (6.4) gives

$$\langle i_M^2 \rangle = 2p\left(\frac{c\delta}{2}\right)\frac{d\sigma(\pi)}{d\Omega}\left(\frac{e}{h\nu}\right)^2\lambda^2\iint d\vec{r}_S\ \rho(R,\vec{r}_S)\ I_T(R,\vec{r}_S)\ I_E(R,\vec{r}_S) \tag{8.2}$$

which yields the important result that the mean square signal current is proportional to the overlap integral of three quantities - the coherent transmitter intensity, the backpropagated effective LO intensity, and the density distribution of scatterers. It is useful to note that we have not made the assumption that the transmitted and local oscillator beams are coaxial in deriving Equation (8.2). In fact, the equation can be used for bistatic lidar systems provided the transmitter and receiver optical axes are nearly parallel and an appropriate offset between transmitter and LO beams is included before computing the integral. If the transverse separation between transmitter and receiver is small relative to the spot sizes of the transmitter and BPELO at the range R, the bistatic system can be treated as coaxial to a good approximation.

As a simple numerical example, we now consider the case of gaussian transmitter and local oscillator beams described by

$$I_T(R,\vec{r}_S) = \frac{2P_T}{\pi\omega_T^2(R)}\,e^{-2\left(\frac{r_S}{\omega_T(R)}\right)^2} \tag{8.3}$$

and

$$I_L(R,\vec{r}_S) = \frac{2P_L}{\pi\omega_L^2(R)}\,e^{-2\left(\frac{r_S}{\omega_L(R)}\right)^2} \tag{8.4}$$

where P_T and P_L are the transmitter and local oscillator output powers and $\omega_T(R)$ and $\omega_L(R)$ are the corresponding guassian radii at the range R. Substitution of Equations (8.3) and (8.4) into (8.2) yields

$$\langle i_M^2 \rangle = 2\left(\frac{\eta_Q e}{h\nu}\right)^2\left\{\frac{2\pi p\left[\rho(R)\,c\delta/2\right]\frac{d\sigma(\pi)}{d\Omega}\,\lambda^2 P_T P_L}{\pi^2\left[\omega_T^2(R)+\omega_L^2(R)\right]}\right\} \tag{8.5}$$

where we have assumed a uniform scattering density $\rho(R)$ and a uniform mixer efficiency η_Q. Clearly, $\langle i_M^2 \rangle$ increases with decreasing ω_T and ω_L implying that the signal level will be maximized in a laboratory scattering experiment by focusing the transmitter and backpropagated LO into the sample.

If the scattering volume in the lidar system of Figure 6 lies in the far field of the transmitter and LO beam waists, we can use the approximations

$\omega_T(R) \approx \lambda R/\pi\omega_{TO}$ and $\omega_L(R) \approx \lambda R/\pi\omega_{LO}$ where ω_{TO} and ω_{LO} are the respective beam waists and R is the distance between the waists and the scattering volume (ref. 17). Equation (8.5) then becomes

$$\langle i_M^2 \rangle = 2\left(\frac{\eta_Q e}{h\nu}\right)^2 \left\{\frac{2\pi p\left[\rho(R)\, c\delta/2\right] \frac{d\sigma(\pi)}{d\Omega} P_T P_L}{R^2}\right\} \left(\frac{\omega_{TO}^2 \omega_{LO}^2}{\omega_{TO}^2 + \omega_{LO}^2}\right) \tag{8.6}$$

which exhibits the familiar R^{-2} dependence for the lidar equation. Equations (8.3) and (8.4) suggest the definition of an effective area for the gaussian beam waists given by $A_T = \pi\omega_{TO}^2/2$ and $A_L = \pi\omega_{LO}^2/2$. Further defining an average antenna area $A = (A_T + A_L)/2$ and letting $A_L = \varepsilon\bar{A}$ and $A_T = (2 - \varepsilon)\bar{A}$, Equation (8.6) becomes

$$\langle i_M^2 \rangle = 2\left(\frac{\eta_Q e}{h\nu}\right)^2 \left\{\frac{4P_T P_L \bar{A}}{R^2}\, p\left[\frac{\rho(R)\, c\delta}{2}\right] \frac{d\sigma(\pi)}{d\Omega}\right\} \left[\varepsilon(2 - \varepsilon)\right] \tag{8.7}$$

which has a maximum for $\varepsilon = 1$ given by

$$\langle i_M^2 \rangle_{max} = 2\left(\frac{\eta_Q e}{h\nu}\right)^2 \left\{\frac{4P_T P_L \bar{A}}{R^2}\, p\left[\frac{\rho(R)\, c\delta}{2}\right] \frac{d\sigma(\pi)}{d\Omega}\right\} \tag{8.8}$$

Thus, we have demonstrated that, if we constrain the sum of the transmitter and receiver areas to the value $2\bar{A}$, we obtain a maximum signal when $\varepsilon = 1$ or $A_L = A_T$, i.e., when the antenna areas are matched. To include optical and atmospheric transmission losses, Equation (8.3) should be multiplied by $\eta_A\eta_{TO}$ and Equation (8.4) by $\eta_A\eta_{RO}$ where η_A is the atmospheric transmission for the range R and η_{TO} and η_{RO} are the efficiencies of the transmitter and receiver optical systems.

It should be clear that, just as in the case of thermal source detection, any LO power falling on the mixer that cannot be backpropagated through the receiver optics to the source will contribute to the shot noise but not to the signal current and therefore represents a reduction in system signal-to-noise. Thus, vignetting, central obscurations, and phasefront errors can have a major impact on the lidar efficiency by (1) reducing the transmission of the back propagated LO and (2) influencing the antenna pattern of the backpropagated effective LO in Equation (8.2). The antenna patterns of vignetted, centrally obscured, and decollimated gaussian beams have been computed by Klein and Degnan (ref. 16).

9. COHERENT SOURCE DETECTION

For a spatially coherent source such as a laser or distant star, we can write for the mutual intensity function at the mixer

$$J_S(\vec{r}_D,\vec{r}_D') = \left[\varepsilon_S(\vec{r}_D)\ e^{i\phi_S(\vec{r}_D)}\right]\left[\varepsilon_S(\vec{r}_D')\ e^{-i\phi_S(\vec{r}_D')}\right] \tag{9.1}$$

where ε_S and ϕ_S are real functions which describe the signal amplitude distribution and phasefront in the mixer plane. A similar expression can be written for the laser LO. Substituting Equation (9.1) and the LO equivalent into our general expression for $\langle i_M^2\rangle$ given by Equation (4.9), we obtain for a coherent source

$$\langle i_M^2\rangle = 2\left(\frac{e}{h\nu}\right)^2 \left| \iint_D d\vec{r}_D\ \eta_Q(\vec{r}_D)\ \varepsilon_S(\vec{r}_D)\ \varepsilon_L(\vec{r}_D)\ e^{i\left[\phi_S(\vec{r}_D)-\phi_L(\vec{r}_D)\right]} \right|^2 \tag{9.2}$$

In the trivial case where the mixer efficiency and the signal and LO beams are uniform over the mixer of area A_D, Equation (9.2) reduces to the familiar form

$$\langle i_M^2\rangle = 2\left(\frac{\eta_Q e}{h\nu}\right)^2 P_S P_L \tag{9.3}$$

where $P_S = \varepsilon_S{}^2 A_D$. In the most general case, we can use Equation (2.8) to define a coherent heterodyne efficiency given by

$$\eta_{HET} = \frac{1}{\bar{\eta}_Q P_L P_S} \left| \iint_D d\vec{r}_D\ \eta_Q(\vec{r}_D)\ \varepsilon_S(\vec{r}_D)\ \varepsilon_L(\vec{r}_D)\ \exp\left\{i\left[\phi_S(\vec{r}_D) - \phi_L(\vec{r}_D)\right]\right\} \right|^2 \tag{9.4}$$

where P_S is the total signal power in the mixer plane. Equation (9.4) can also be written in the form

$$\eta_{HET} = \frac{\left| \iint_D d\vec{r}_D\ \eta_Q(\vec{r}_D)\ \varepsilon_S(\vec{r}_D)\ \varepsilon_L(\vec{r}_D)\ \exp\left\{i\left[\phi_S(\vec{r}_D) - \phi_L(\vec{r}_D)\right]\right\} \right|^2}{\left[\iint_D d\vec{r}_D\ \eta_Q(\vec{r}_D)\ \varepsilon_L{}^2(\vec{r}_D)\right]\left[\iint d\vec{r}_D\ \varepsilon_S{}^2(\vec{r}_D)\right]} \tag{9.5}$$

where we have used Equation (2.4) and the explicit expression for P_S. Degnan and Klein (ref. 18) have performed computations of η_{HET} for the case where the signal and LO phasefront curvatures are matched, the mixer efficiency η_Q is uniform and $\varepsilon_S(\vec{r}_D)$ is an Airy pattern formed by a centrally obscured, circular antenna illuminated by a plane wave from a distant point source. As the size of the central obscuration is increased, more incoming radiation is blocked by the obscuration and a smaller fraction of the radiation which reaches the mixer

plane is contained in the central lobe of the signal Airy pattern. Degnan and Klein (ref. 18) considered several illumination profiles for the LO including uniform, gaussian, and an Airy pattern matched to the signal Airy pattern. Their results are summarized in Figure 7. Optimum detection efficiency is achieved when the mixer captures the entire signal Airy pattern and a matched LO is used. In this instance, the receiver efficiency is simply $1 - \gamma^2$ (where γ is the obscuration ratio defined previously for the Cassegrain antenna in Figure 4) corresponding to the areal obscuration loss and represented by the "matched" LO curve in Figure 7. The difference between the ideal or "matched" LO curve and the uniform or gaussian curves corresponds to the heterodyne detection efficiency η_{HET}.

If the mixer is illuminated by a uniform LO, the optimum Airy disk radius (to the first null) is found to be $R_A \approx 1.35R_D$ where R_D is the mixer radius. It should be noted that the Airy disk radius varies with the obscuration ratio for an optical antenna having a given f-number (ref. 18). The optimum efficiency η_{HET} is approximately 83% for no obscuration and falls rapidly as the obscuration ratio is increased even if one chooses an optimum signal spot size. An optimized gaussian LO with waist radius $\omega = 0.64R_A$ and a central Airy signal disk which matches the mixer radius R_D yields greater sensitivity compared to the uniform LO since it more closely matches the intensity distribution of the central Airy disk for the signal. The power contained in the outer rings of the Airy pattern is lost, however, and this accounts for the major difference between the "ideal" matched LO and gaussian LO curves in Figure 7. For a more detailed discussion, and for more general plots of non-optimized geometries, the reader is referred to the original paper by Degnan and Klein (ref. 18).

It is a simple matter to compute the effects of misalignment between the signal and LO beams or of a mismatch between phasefront curvatures using the general expression (9.5). For example, if the two wavefronts are misaligned by an angle θ in the y_D direction illustrated in Figure 2, the exponential argument in Equation (9.5) is

$$\phi_S(\vec{r}_D) - \phi_L(\vec{r}_D) = (\vec{k}_S - \vec{k}_L) \cdot \vec{r}_D \approx k \sin\theta \, y_D$$

where $\vec{k}_S$ and $\vec{k}_L$ are the propagation vectors for the signal and LO beams, $|k_S| \approx |k_L| \approx k = 2\pi/\lambda$, and y_D is the y-component of the vector r_D. For cylindrically symmetric fields, Equation (9.5) reduces to a special case previously derived by Cohen (ref. 19); i.e.,

$$\eta_{HET} = \frac{\left[\int_0^{r_o} dr_D \, r_D \, \eta_Q(r_D) \, \varepsilon_S(r_D) \, \varepsilon_L(r_D) \, J_o(kr_D \sin\theta)\right]^2}{\left[\int_0^{r_o} dr_D \, r_D \, \eta_Q(r_D) \, \varepsilon_L^2(r_D)\right]\left[\int_0^{\infty} dr_D \, r_D \, \varepsilon_L^2(r_D)\right]} \qquad (9.6)$$

where r_o is the radius of the mixer, and we have used $y_D = r_D \cos \phi_D$ and the integral expression for the Bessel function $J_o(z)$, i.e.,

$$J_o(z) = \frac{1}{2\pi} \int_0^{2\pi} d\phi \; e^{iz \sin \phi} \tag{9.7}$$

Cohen (ref. 19) has generated plots of η_{HET} for a variety of source-LO illumination function combinations such as uniform-uniform, Airy-uniform, matched Airy-Airy, uniform-gaussian, and Airy-gaussian. He considered the tolerance of the various combinations to misalignment and allowed for a quadratically varying mixer quantum efficiency. The sensitivity to misalignment for the various combinations varied less than 15% relative to the most sensitive uniform-uniform case given by

$$\eta_{HET} = \eta_Q \left[\frac{2J_o(kr_o \sin \theta)}{kr_o \sin \theta} \right]^2 \tag{9.8}$$

Thus, $\eta_{HET} = \eta_Q$ for no misalignment and $\eta_{HET} = \eta_Q/2$ for $\theta = 0.5\lambda/(2r_o)$ corresponding to a half-wavelength phase difference over the mixer diameter $2r_o$. For a wavelength of 10 μm and a mixer diameter of 200 μm, the misalignment angle at which the detection efficiency is reduced by a factor of 2 is $\theta = 1.4^o$.

For a mismatch in phasefront curvatures, the exponential argument in Equation (9.5) is

$$\phi_S(\vec{r}_D) - \phi_L(\vec{r}_D) = \frac{kr_D^2}{2} \left(\frac{1}{C_S} - \frac{1}{C_L} \right)$$

where C_S and C_L are the curvatures of the signal and LO phasefronts at the mixer plane. For cylindrically symmetric beams, Equation (9.5) reduces to

$$\eta_{HET} = \frac{\left| \int_0^{r_o} dr_D \; r_D \; \eta_Q(r_D) \; \varepsilon_S(r_D) \; \varepsilon_L(r_D) \; \exp\left\{ \frac{ikr_D^2}{2} \left(\frac{1}{C_S} - \frac{1}{C_L} \right) \right\} \right|^2}{\left[\int_0^{r_o} dr_D \; r_D \; \varepsilon_L^2(r_D) \; \eta_Q(r_D) \right] \left[\int_0^{\infty} dr_D \; r_D \; \varepsilon_S^2(r_D) \right]} \tag{9.9}$$

For the uniform-uniform case,

$$\eta_{HET} = \eta_Q \left\{ \frac{\sin\left[\frac{kr_o^2}{4}\left(\frac{1}{C_S} - \frac{1}{C_L}\right)\right]}{\frac{kr_o^2}{4}\left(\frac{1}{C_S} - \frac{1}{C_L}\right)} \right\}^2 \tag{9.10}$$

and $\eta_{HET} = \eta_Q$ for $\Delta\left(\frac{1}{C}\right) = \left(\frac{1}{C_S} - \frac{1}{C_L}\right) = 0$ while $\eta_{HET} = 0$ for $\Delta\left(\frac{1}{C}\right) = 2\lambda/r_o^2$ where r_o is the mixer radius. Thus, if the local oscillator beam has a planar phasefront ($C_L = \infty$), the signal beam phasefront curvature must satisfy $C_S >> r_o^2/2\lambda$.

It should be noted in closing that we have arbitrarily chosen to perform the above calculations in the mixer plane. For a particular antenna or LO geometry, it may be more convenient to perform the computation in some other optical plane as noted previously in Section 5.

10. CONCLUDING REMARKS

This article has attempted to present a unified approach to the calculation of signal-to-noise ratios in optical heterodyne receivers for a variety of important applications. No attempt has been made to give an exhaustive review of the existing literature. The references cited are those which, in the author's opinion, either lend themselves particularly well to the development of the general theory of optical heterodyne receivers given here or have presented numerical results having widespread application. There are, for example, various uncited articles which present calculations of signal-to-noise for very specific incoherent source or backscatter lidar geometries. These have usually employed brute force computational methods that give little insight into the general approach for optimizing system sensitivity. While these provide excellent tests of the general theory, the articles were deemed to be too specialized to be included in the present review.

Clearly, no attention has been paid to the effects of the atmosphere on coherent wave propagation. Although the amplitude and phase fluctuations produced by the atmosphere are inherently included in the complex electric field envelopes introduced in Section 4, no attempt has been made here to give a quantitative assessment of their impact. In the approach taken here, the atmosphere can be viewed as simply another optical element through which the coherent backpropagated effective LO must pass to reach the signal source or vice versa. In the thermal source detection and backscatter lidar problem, the atmosphere presumably modifies the backpropagated effective LO intensity distribution thereby influencing the overlap integral in Equation (6.4). A number of papers in this area have appeared since the early work of Fried (ref. 5) including a rather extensive recent report by Capron et al. (ref. 20) applicable to coherent optical radar.

REFERENCES

1. T.G. Blaney, Space Science Reviews,17, 691 (1975).
2. J.H. McElroy, Applied Optics,11,1619(1972).
3. M.J. Mumma, T. Kostiuk, and D. Buhl, Optical Engineering, 17,50(1978).
4. J.H. McElroy, N. McAvoy, E.H. Johnson, J.J. Degnan, F.E. Goodwin, D.M. Henderson, T. A. Nussmeier, L.S. Stokes. B.J. Peyton, and T. Flattau, Proc. IEEE,65,221(1977).
5. D.L. Fried, Proc. IEEE,55,57(1967).
6. H. Levinstein,Applied Optics,4,639(1965).
7. T.P. McLean and E.H. Putley, RRE Journal,52,5(1965).
8. F.R. Arams, E.W. Sard, B.J. Peyton, and F.P. Pace, IEEE JQE, QE-3,11(1967).
9. R.A. Smith, Proc. IEEE,98,43(1951).
10. B.J. Rye,Applied Optics,18,1390(1979).
11. D. McGuire,Optics Letters,5,73(1980).
12. M.Born and E.Wolf, "Principles of Optics", 5th Ed., Chapt. 10 (Pergamon,New York, 1975).
13. L.Mandel and E.Wolf,Rev. Mod. Phys.,37,231(1965).
14. F.Zernike,Physika,5,791(1938).
15. A.E. Siegman,Applied Optics,5,1588(1966).
16. B.J. Klein and J.J. Degnan,Applied Optics,13,2134(1974).
17. A.E. Siegman,"An Introduction to Lasers and Masers", Chapter 8(McGraw-Hill,New York,1971).
18. J.J. Degnan and B.J. Klein,Applied Optics,13,2397(1974); Erratum,13,2762(1974).
19. S.C. Cohen,Applied Optics,14,1953(1975).
20. B.A.Capron,R.C. Harney, and J.H. Shapiro,"Turbulence Effects on the Receiver Operating Characteristics of a Heterodyne Reception Optical Radar", Project Report TsT-33,Lincoln Laboratory. Massachusetts Institute of Technology(1979).

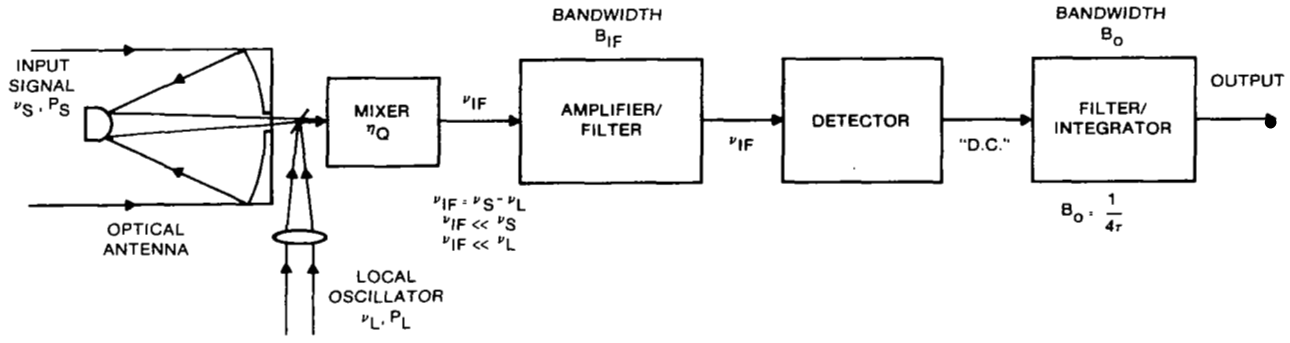

Figure 1.- Block diagram of a representative heterodyne receiver.

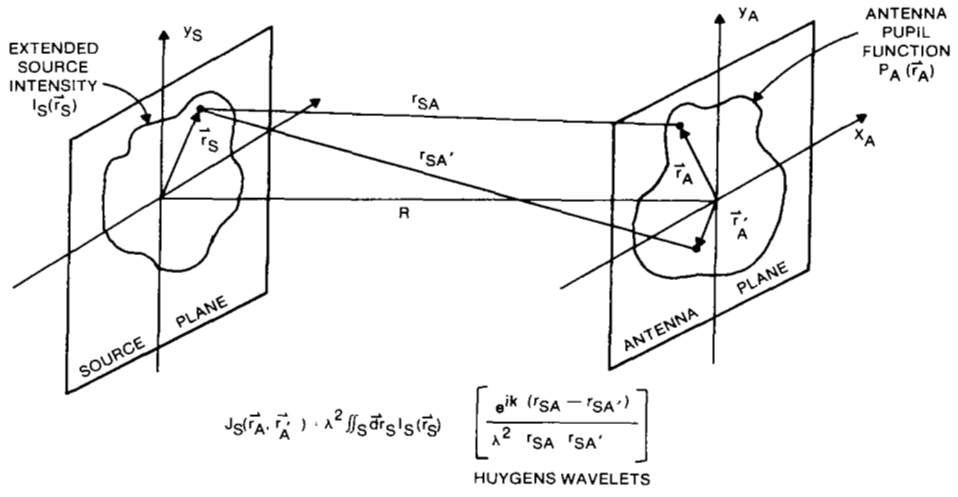

Figure 2.- Huygen's wavelet model for propagation of the mutual intensity function between the antenna and mixer plane.

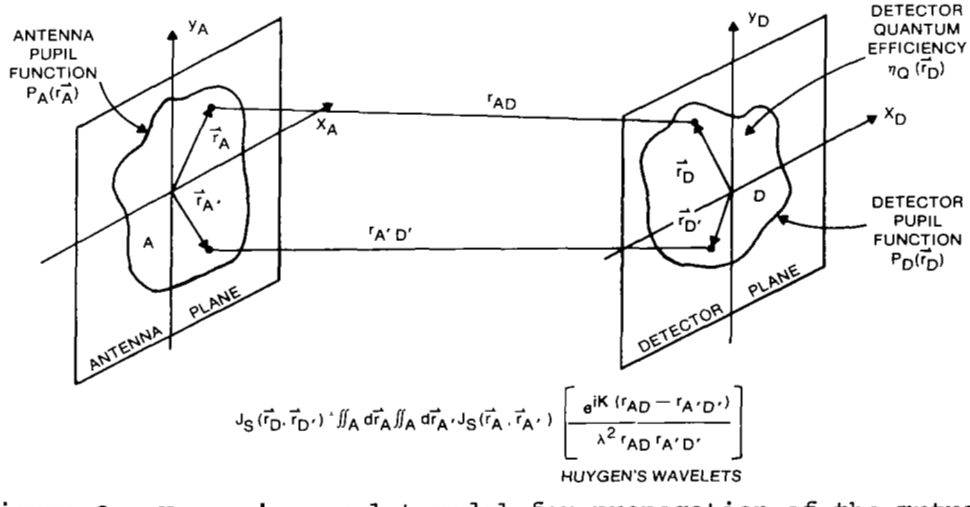

Figure 3.- Huygen's wavelet model for propagation of the mutual intensity from an extended incoherent source.

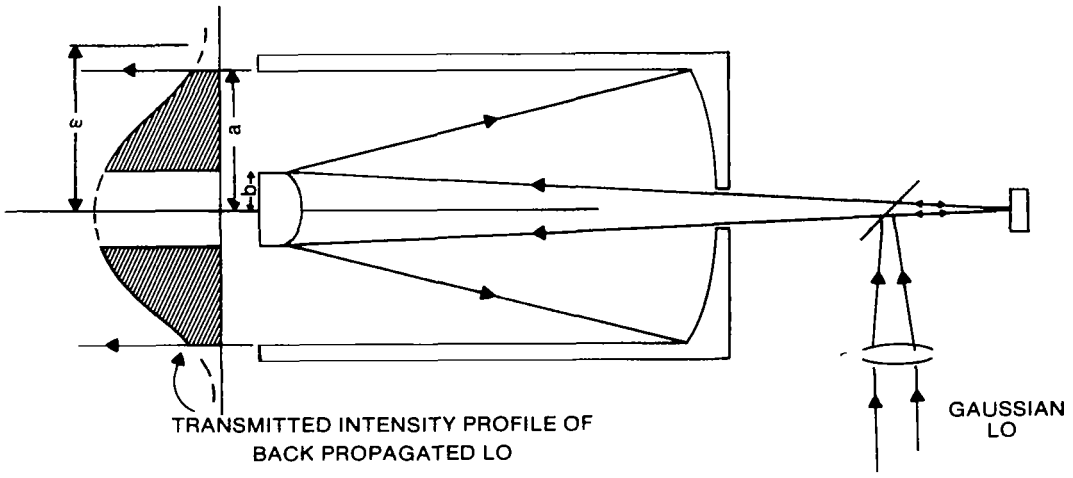

Figure 4.- Backpropagation of a gaussian local oscillator beam through a Cassegrain telescope.

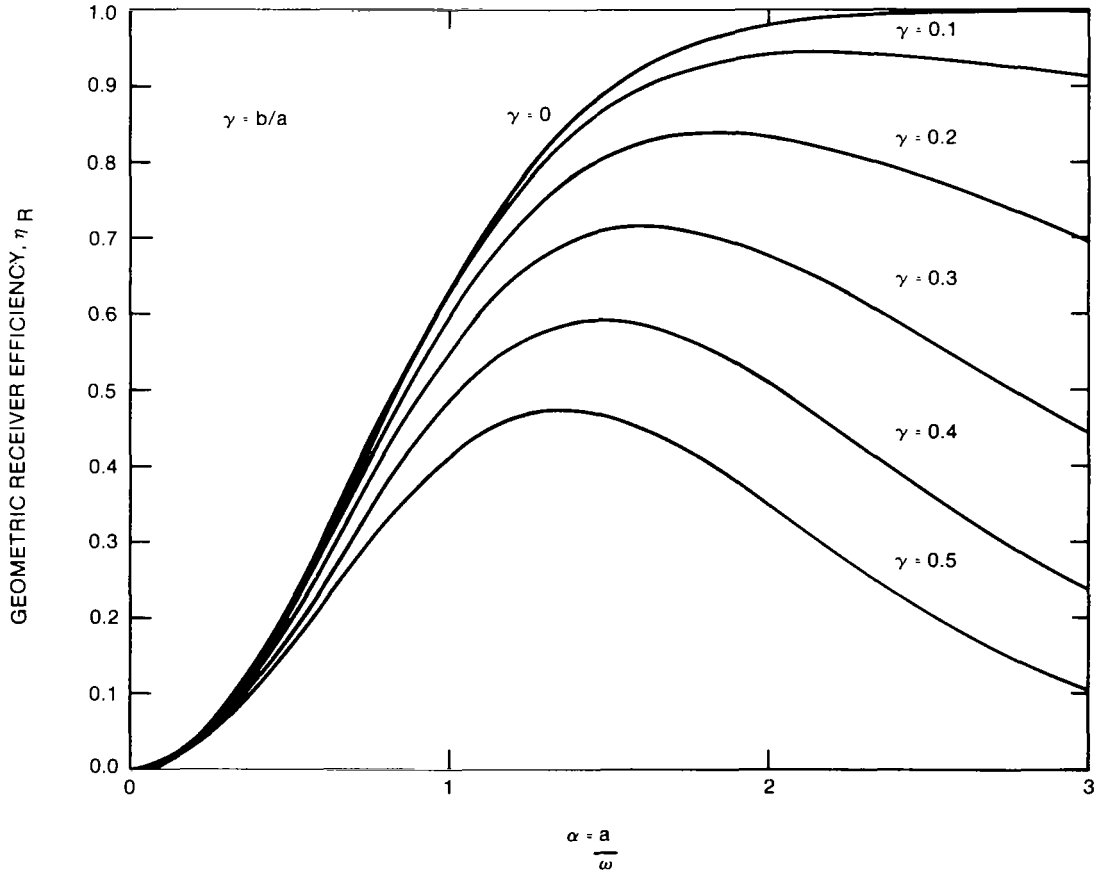

Figure 5.- Geometric receiver efficiency for a large thermal source viewed through a centrally obscured telescope by a mixer illuminated by a gaussian local oscillator beam.

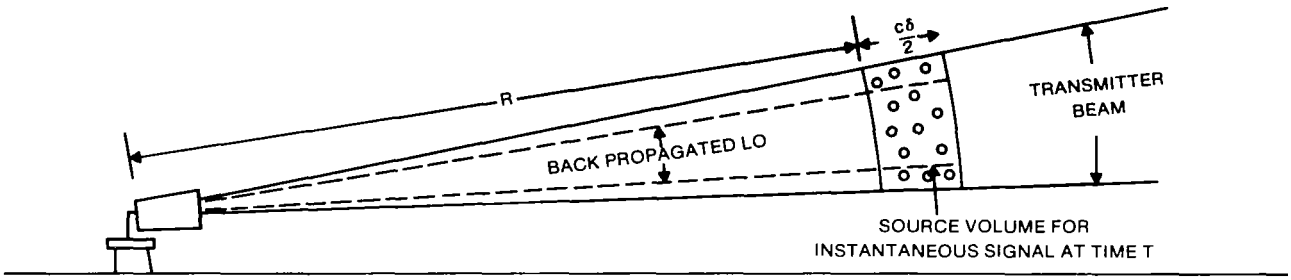

Figure 6.- Functional diagram of a heterodyne incoherent backscatter lidar system.

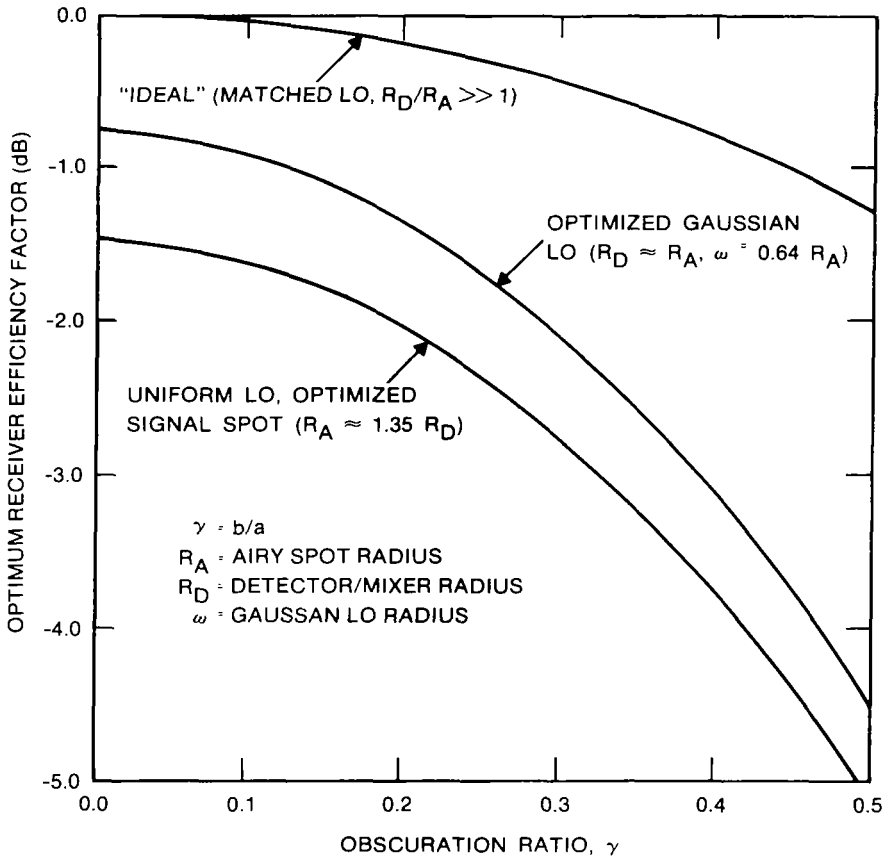

Figure 7.- Maximum receiver efficiency factors in dB for detection of a distant point source by a heterodyne receiver consisting of a general centrally obscured telescope (primary radius a, secondary radius b) as a function of linear obscuration ratio γ = b/a and several optimized LO distributions (uniform, gaussian, and matched Airy).