STUDY OF LASER FREQUENCY STABILITY AND SPECTRAL PURITY

under the direction of

A. E. Siegman

FINAL REPORT

for

NASA Grant NGR 05-020-234

National Aeronautics and Space Administration

Washington, D. C. 20546

for the period

1 March 1968 – 28 February 1969

M. L. No. 1770

May 1969

Microwave Laboratory
W. W. Hansen Laboratories of Physics
Stanford University
Stanford, California

STAFF

NASA Grant NGR 05-020-234

for the period

1 March 1968 – 28 February 1969

PRINCIPAL INVESTIGATOR

A. E. Siegman

RESEARCH ASSISTANTS

R. Arrathoon

K. Manes

S. C. Wang

TABLE OF CONTENTS

ABSTRACT

This program has been concerned with the study and improvement of the short-term spectral purity and the long-term frequency stability of stable laser oscillators.

We have successfully measured the short term frequency phase fluctuations in high quality He-Ne gas lasers, and have measured, for the first time, the ultimate limit on laser spectral width caused by quantum phase noise, as predicted by the Schawlow-Townes formula. Verification of this formula, one of the most basic in laser theory, has only recently come within the range of experimental feasibility.

An additional task of this program was to study and, if indicated, to demonstrate experimentally a new absolute laser frequency stabilization method. The study has been completed and experimental verification is underway, with many problems solved.

Finally we have suggested an important new mechanism to account for the frequency fluctuations in He-Ne 6328 Å lasers, and presented both experimental and theoretical support for this mechanism.

I. COMPLETED FREQUENCY FLUCTUATION MEASUREMENTS

This section of the report summarizes laser frequency fluctuation and quantum phase noise measurements that were completed and published during the period of this contract.

Even in a very high quality, stable, single-frequency laser oscillator, the instantaneous laser phase or frequency will retain some noise fluctuations, due to two different sources. First of all, and usually most important in practical situations, are various external disturbances, by which we mean such things as microphonics, structural vibrations, plasma oscillations and other disturbances in the laser. In addition, as a fundamental or ultimate limiting source of noise, there are quantum phase or frequency fluctuations that arise from the basic process of atomic spontaneous emission in the laser medium. The magnitude of the oscillator spectral broadening due to these ultimate quantum noise sources is given by the so-called Schawlow-Townes formula, indicated at the bottom of Fig. 1.

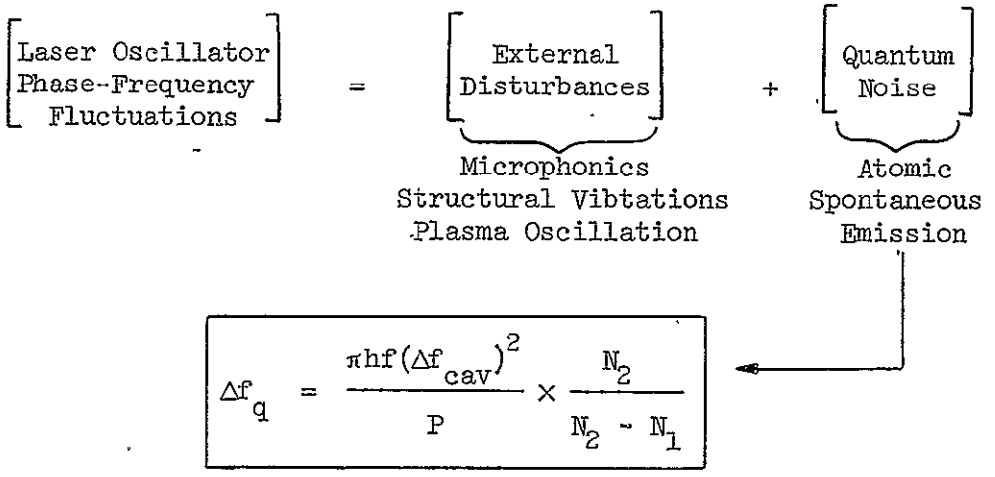


FIG. 1--Frequency fluctuation sources.

One of our primary purposes in this work has been to observe these quantum noise fluctuations in a stable laser oscillator, and thereby to verify this Schawlow-Townes formula. We have achieved this objective; and along the way have also learned some important things about the apparent primary source of external disturbances in the 6328 Å He-Ne laser in particular.

Figure 2 is a schematic showing the essential features of the experimental apparatus for these measurements. As indicated, we heterodyned together two stable single-frequency high-quality He-Ne lasers, which are basically Spectra-Physics Model 119 lasers, shorn of their regular cabinetry and equipped with special power supplies and control circuits. As indicated in the upper sketch, the frequency of laser number 1 is set very close to the edge of the laser oscillation range, in order to obtain stable operation of this laser at very low power level, thereby enhancing the size of the quantum noise disturbances because of the $1/p$ dependence in the Schawlow-Townes formula.

Two sets of experiments were carried out. The first of these used a beat frequency of 30 MHz and commercially available electronics. This work was reported in Phys. Rev. Letters 20, 17 (April 1968), "Observation of Quantum Phase Noise in a Laser Oscillator"; Appendix I of this report. The second set employed a 4.5 MHz beat frequency and was reported in J. Appl. Phys. 40, (February 1969), "Further Measurements of Quantum Phase Noise in a He-Ne Laser"; Appendix II of this report. We will summarize the results of these measurements very briefly here.

The two lasers are heterodyned together on a photomultiplier tube, with a nominal four-and-a-half MHz or 30 MHz difference frequency, and the resulting beat note is fed into a 4.5 MHz or 30 MHz RF discriminator.

EXPERIMENTAL APPARATUS

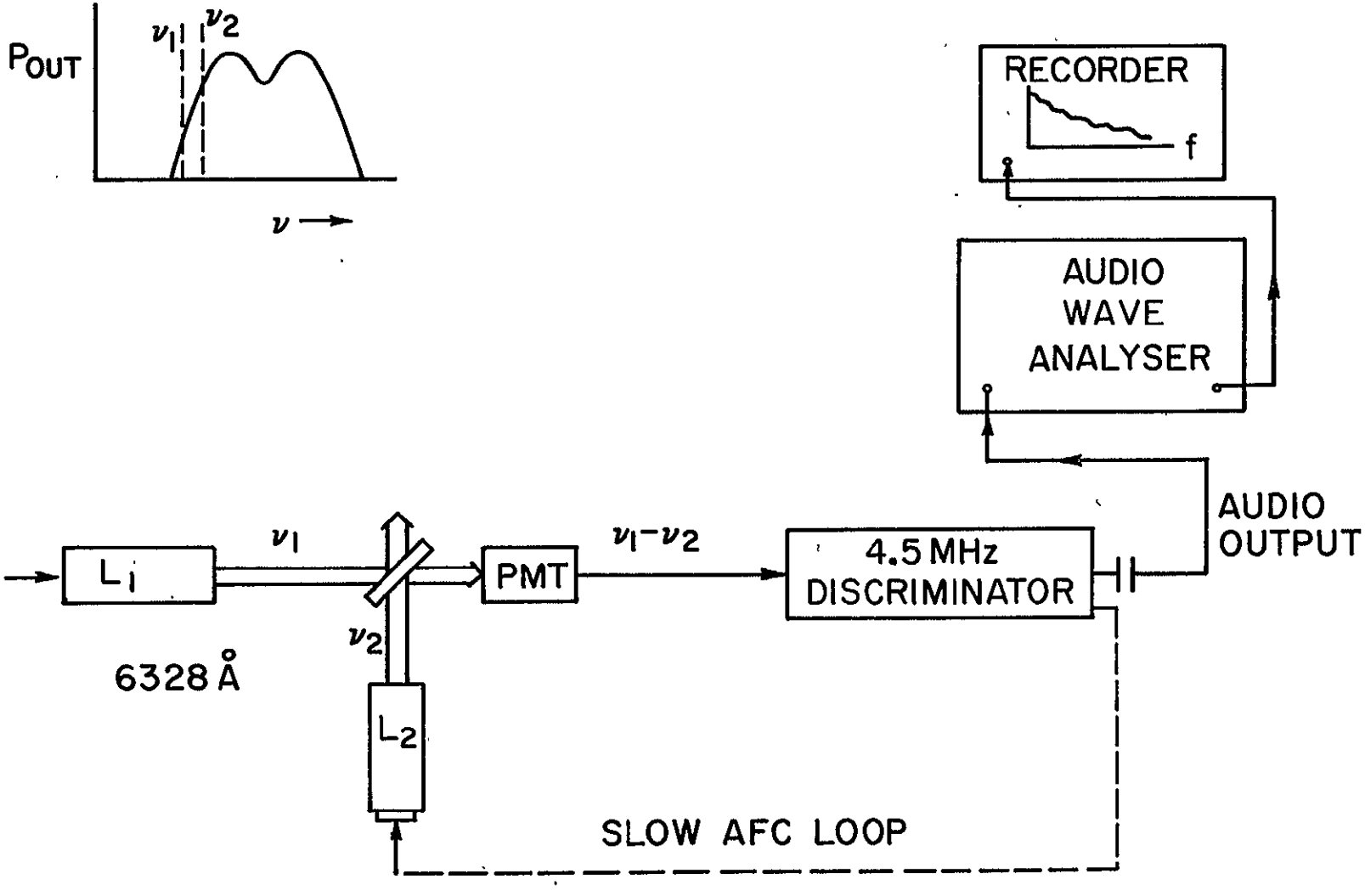


FIG. 2--Schematic showing the essential features of experimental apparatus.

Part of the output of this discriminator is fed back through a slow automatic frequency control loop with a long time constant in order to keep the frequency of laser L_2 spaced away from laser L_1 by 4-1/2 MHz or 30 MHz on the average.

Due to the instantaneous frequency or phase fluctuations in each laser, there is a finite amount of frequency jitter in the beat frequency, with a spectral width and a peak deviation in the 10 KHz range. This frequency jitter is reflected in the audio output of the RF discriminator, which is taken to an audio wave analyzer, whose output vs audio frequency is then recorded on an xy recorder.

Figure 3 illustrates the general character of the discriminator audio output spectrum as recorded on the xy recorder. The effect of external disturbances is essentially to modulate the instantaneous frequency of the lasers, or of the beat note, in a Gaussian noise-like fashion; and the effect of this is to give a $1/f^2$ audio output spectrum in the discriminator output, extending from the cutoff frequency of the AFC loop, which is a small fraction of 1 Hz, upwards. By contrast, the effect of the quantum noise may be described as a random walk of the instantaneous phase of the laser oscillator. This has a different spectral characteristic, and appears as a white noise output in the audio output spectrum of the discriminator. Because of the $1/p$ dependence of the quantum noise term, this quantum noise level increases with decreasing laser power, as indicated on the figure; while the external disturbances represent frequency modulation of the laser by external sources and is independent of the laser's oscillation level.

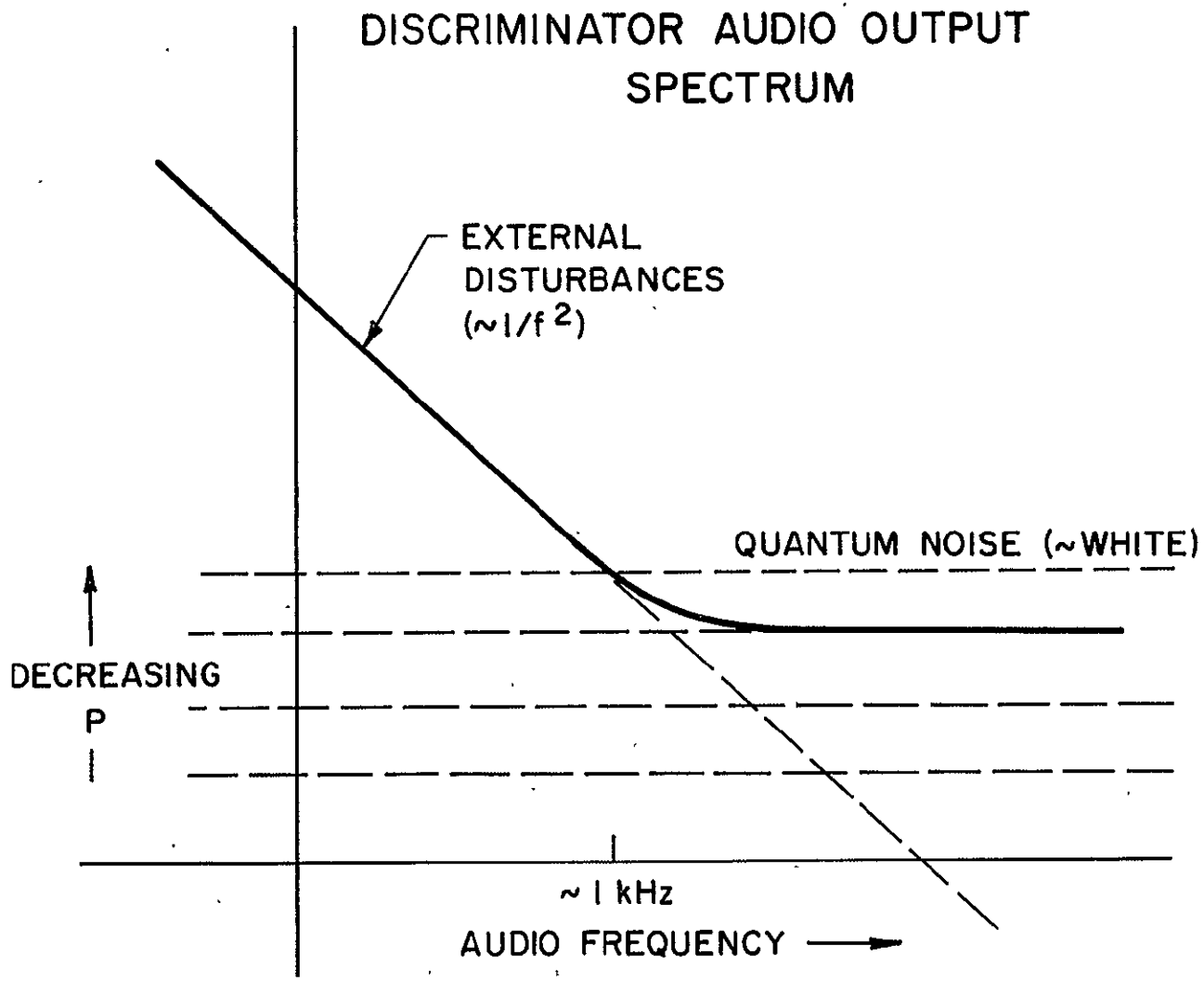

DISCRIMINATOR AUDIO OUTPUT
SPECTRUM
EXTERNAL
DISTURBANCES
(~1/f 2)
QUANTUM NOISE (~WHITE)
DECREASING
P
~ 1 kHz
AUDIO FREQUENCY

FIGURE 3

Figure 4 shows a typical experimental result, indicating how well the experimental measurements agree with the theoretical expectations. Also in agreement with theory, we observe that the external contribution portion of this curve does not change as the oscillation power level of the lasers is changed; whereas the quantum contribution does indeed increase as the inverse of the laser power level. We should emphasize here that in all of our experiments, the frequency fluctuations of the laser and the resulting spectral broadening of the laser output were dominated by the external contribution, which gave a total spectral spread for the lasers, and consequently for the beat note also, of about 10 KHz. The quantum contribution, on the other hand, represented a net broadening contribution of only a few 100 of cycles or less. Nonetheless, this quantum contribution could be separated out and measured precisely because of its different spectral appearance in the outputs of the discriminator.

It should be noted that radio frequency discriminators are in general rather difficult devices, with far from ideal noise properties; and spurious results that appear very much like the expected quantum contribution results can be easily obtained simply as a result of the inherent noise properties of the discriminator. Appropriate steps were taken to correct for these difficulties.

By making a large series of measurements such as the typical example shown in Fig. 4, with the laser L_1 operating at different power levels in each case, we were able to obtain the data on quantum linewidth contribution in Hz vs laser oscillation power in watts shown in Fig. 5. These represent the principle quantum noise experimental results of this report, and show quite clearly the expected $1/p$ dependence predicted by the

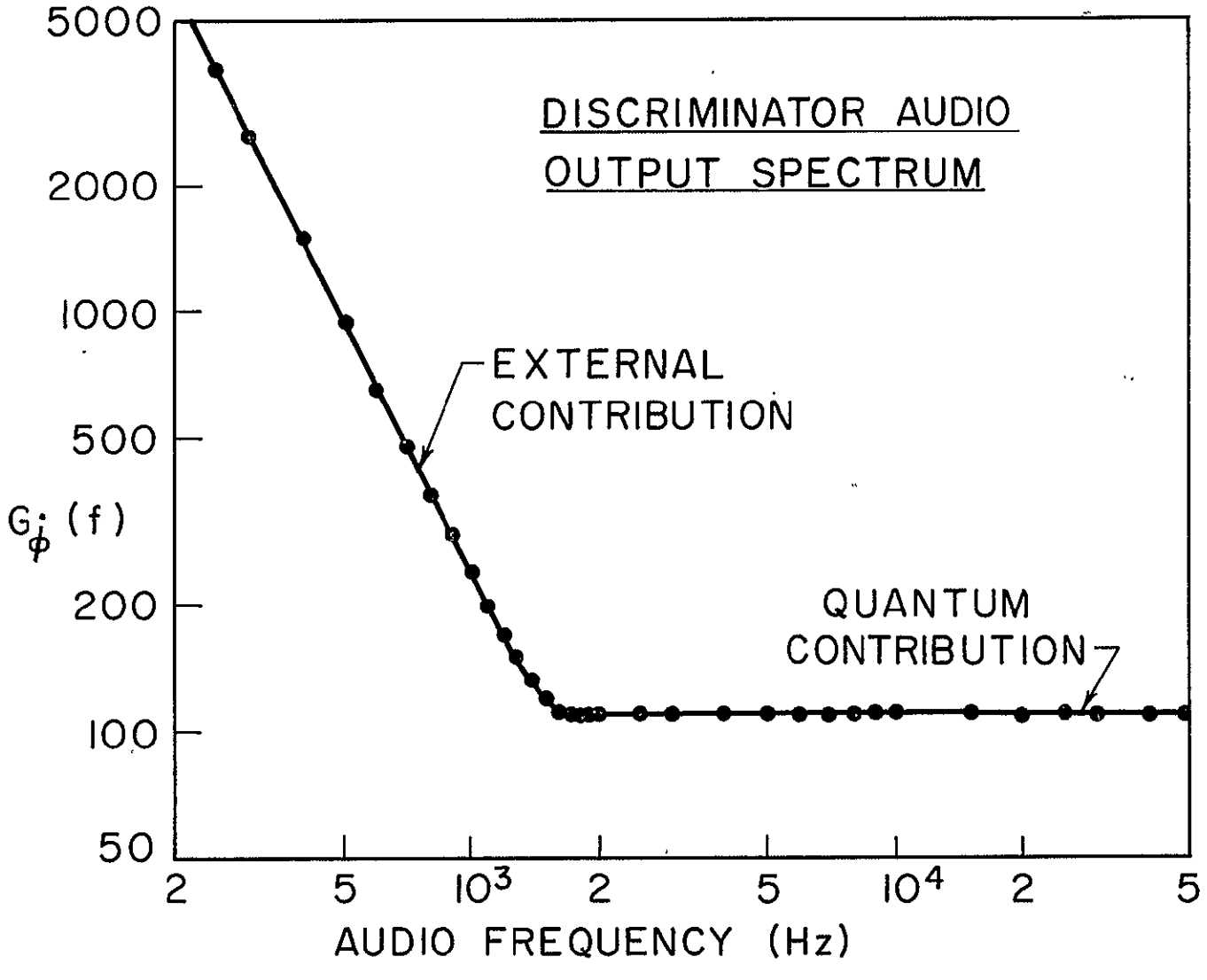

5000
2000
1000
500
200
100
50
$G_\phi(f)$
DISCRIMINATOR AUDIO
OUTPUT SPECTRUM
EXTERNAL
CONTRIBUTION
QUANTUM
CONTRIBUTION
2
5
10^3
2
5
10^4
2
5
AUDIO FREQUENCY (Hz)

FIGURE 4

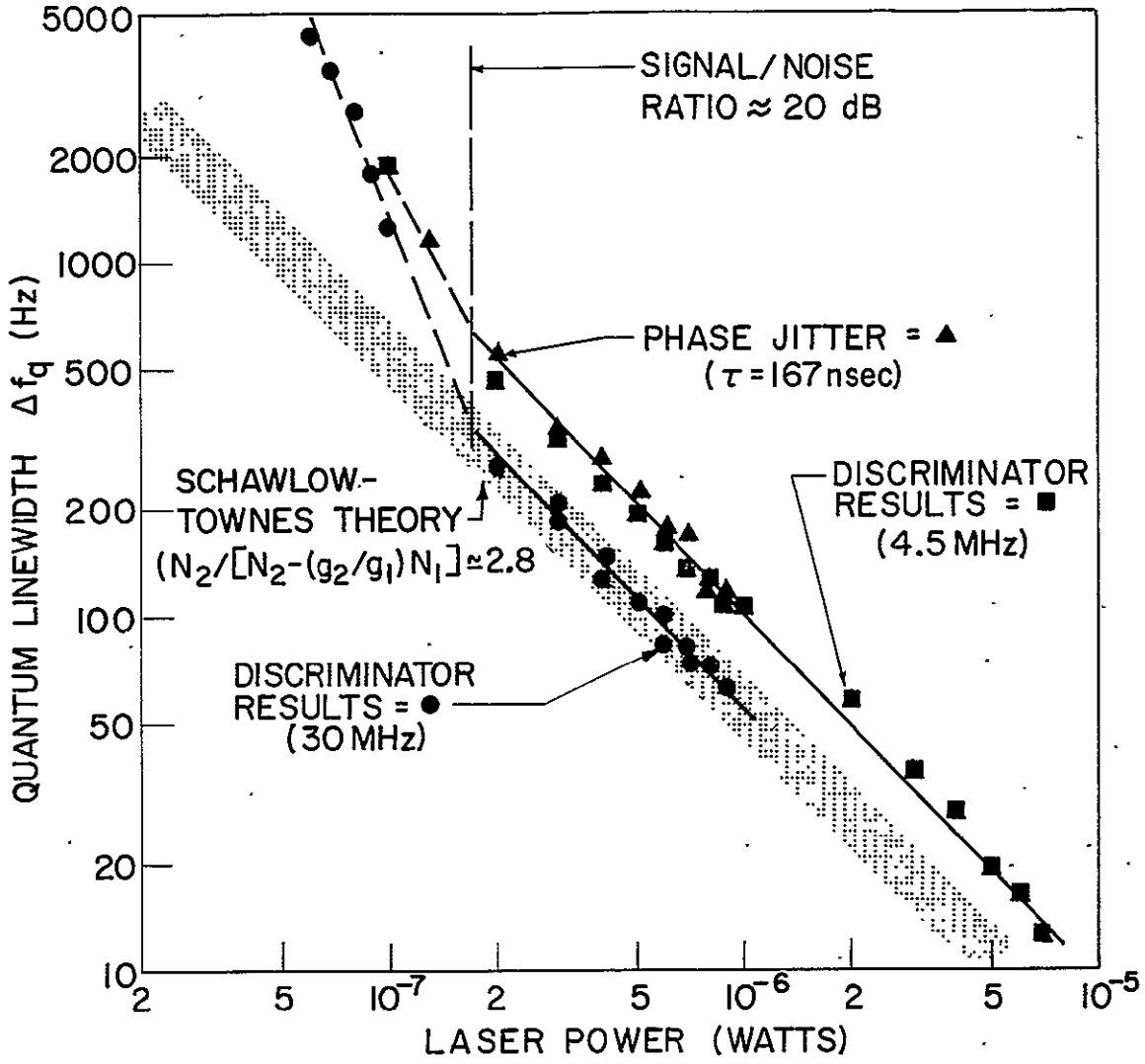


FIG. 5--Data on quantum linewidth contribution in Hz vs laser oscillation power in watts.

Schawlow-Townes formula over a range of nearly two orders of magnitude. The lower points represent the results of the first series of experiments using an RF discriminator centered at 30 MHz. The upper curve represents our more recent results using the 4-1/2 MHz discriminator, together with a few experimental points obtained by measuring the beat note phase jitter using a sampling oscilloscope technique. Various laser tubes and mirrors were changed between the two series of measurements, and the approximate factor of 2 difference may be a good measure of the reproducibility of our results.

The fuzzy line represents the prediction of the Schawlow-Townes theory, taking into account the uncertainties in how well we know the cold cavity bandwidth and other parameters of our laser. The excess noise ratio for our lasers was determined by carrying out a separate series of measurements of the excess amplitude fluctuations of the lasers at low powers, following the method of Fried and House. From such measurements the excess noise factor can be determined directly, and was found to have a value of approximately 3.

Note that the apparent noise contribution begins to increase sharply at about the signal power level at which the signal-to-noise ratio — that is the ratio of the beat signal to the photomultiplier shot noise — decreases below about 20 dB. We believe, however, that this represents excess noise behavior in the discriminator itself as the signal level drops below the necessary threshold level for good fm detection; rather than any real increase in the quantum noise fluctuations.

II. ON-GOING FREQUENCY FLUCTUATION MEASUREMENTS

This section of the report summarizes some additional frequency fluctuation measurements which were started and carried to a fairly advanced level, but not fully completed during the period of this contract.

The Schawlow-Townes formula suggests that Δf_q may be increased to a detectable magnitude by decreasing P and increasing Δf_{cav}. Decreasing P in dc-excited lasers generally leads to an increase in plasma noise which swamps out Δf_q; therefore, rf excitation on rf quieting of the plasma is indicated. Increasing Δf_{cav} is accomplished by lowering the cavity Q with low reflection mirrors. Naturally the laser cannot be made to lase at all unless the threshold requirement is met; i.e., gain must be high enough to overcome mirror transmission. A laser line with one of the highest known gains is $\lambda = 3.508\ \mu$ in He-Xe. He-Xe 3.508 μ lasers produced here recently have shown gains in excess of 100 dB/meter. Further, we have found the He-Xe plasma to be easily controllable over a wide range of Xe pressures.

However, because of the very rapid clean-up of the Xe gas in the discharge, lasers of this type previously have been limited to an operating life measured in hours. We have found that an apparently completely successful solution to this problem is to heavily overfill the laser with Xe, and then condense out the excess Xe in a side arm cooled to near liquid nitrogen temperature. The Xe actually solidifies in the side arm; however, the vapor pressure of Xe at just above liquid nitrogen temperature ($77^{\circ}K$) is 19 to 20 microns of mercury; i.e., optimum for lasing.

Measurements of the ~ 100 MHz wide 3.51 μ transition of the He-Xe laser show no clearly identifiable Lamb dip. The absence of the expected dip may be due to a large homogeneous collision broadening linewidth, of the order of the inhomogeneous doppler width, introduced by the high pressure chosen for maximum lifetime together with the fact that the Xe we have used so far is not isotopicly pure.

One of our objectives in this work is a direct measurement of Δf_q with He-Xe lasers. To get some idea of the experimental desirability of He-Xe over the common 6328 Å He-Ne laser, the Schawlow-Townes prediction is plotted for several lasers in Fig. 6.

Two 30 cm external-mirror He-Xe laser on invar bases have been fabricated and have undergone preliminary experimentation (Fig. 7). Gain per unit length at various partial pressures of Xe, optimum mirror reflectivities, Brewster window materials, detector systems, and alignment difficulties have been some of the problems under study. We have now seen output powers of 2-1/2 milliwatts at 3.508 μ using MgF_2 or CaF_2 Brewster windows and output mirror substrates coated 40% reflective.

Beats have been obtained from these lasers and locked at 200 KHz. The beat was very noisy and showed clear 60 Hz FM. Work is now in progress to quiet the electronics and improve the pressure stabilization on these lasers.

Two dc excited Spectra-Physics 119 lasers, formerly used for quantum noise measurements at 6328 Å were modified for 3.508 He-Xe operation (without return paths to avoid cataphoresis effects in the Xe discharge). Although these had good temperature-compensated invar cavities, they were extremely noisy both in amplitude and phase with power outputs of only

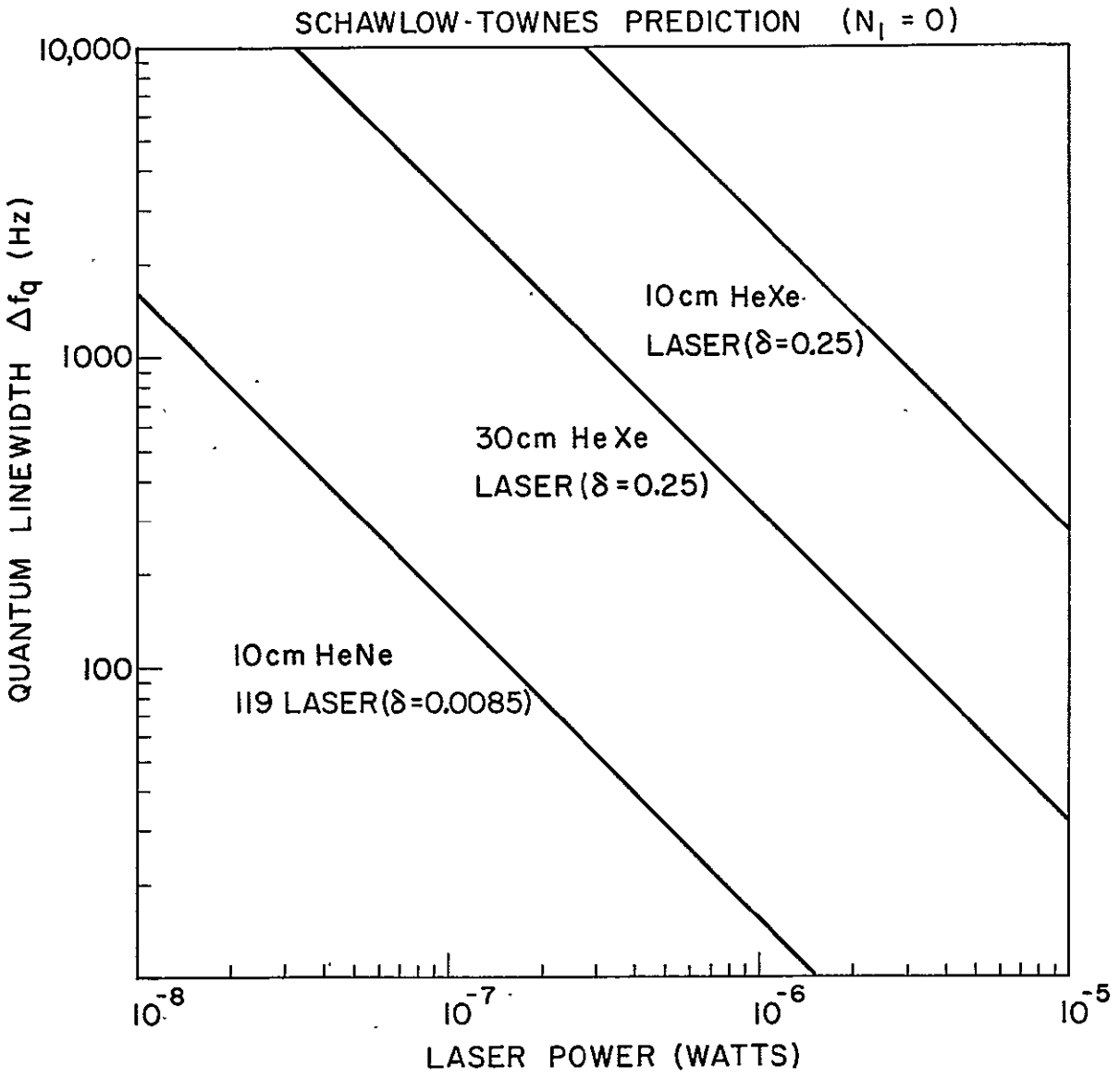


FIG. 6--Plot of S-T formula for several lasers.

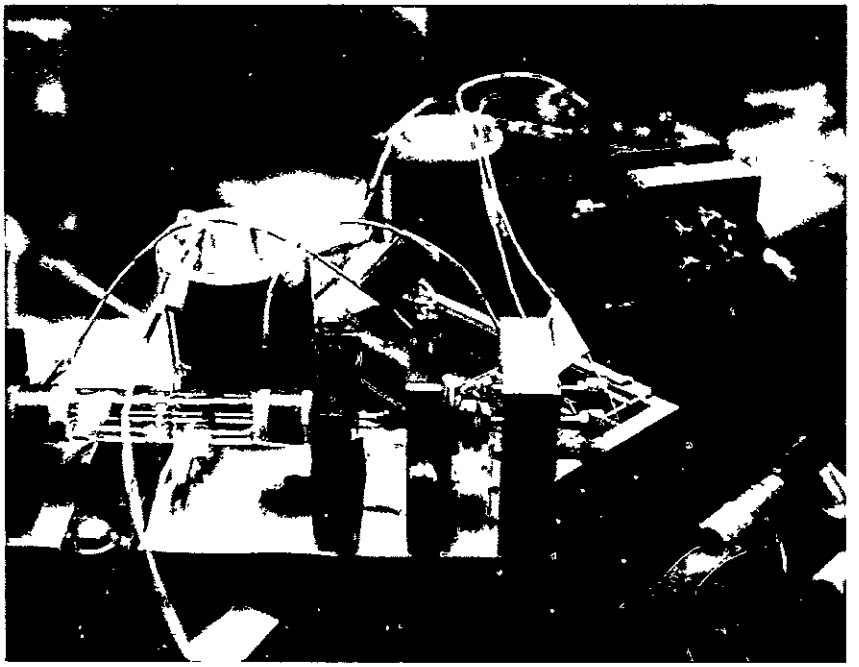


FIG. 7--30 cm external mirror cavity.

20 to 30 microwatts. The bandwidth of the beat was larger than the 1 MHz bandwidth of our detector (about 7 MHz). Conversation with Dr. Gamo of the University of California at Irvine who had a similar laser, corroborated this result. Improvements in plasma tube design and specifically the inclusion of a gas return path might reduce this plasma noise and bring the beat width within the detector bandwidth.

Finally, a quartz ingot internal mirror laser is nearing completion (Fig. 8). This laser could serve as the prototype for internal mirror Cer-Vit lasers providing maximum mechanical and thermal stability. Cer-Vit is a new low-expansion glass ceramic with an expansion coefficient adjustable through zero at $25^{\circ}C$.

Liquid nitrogen cooled InAs photo conductive detectors can achieve 20 dB signal-to-noise ratios at 10^{-9} watt power levels. As can be seen from Fig. 1, such a low-power level should render Δf_q easily measurable. We are confident that measurement of Δf_q over a wide range of P will be achieved, thereby checking the fundamental quantum noise theory.

III. STUDY OF CAUSES OF FREQUENCY FLUCTUATIONS

In Applied Physics Letters, vol. 13, No. 6, pp. 197-199 (September 1968) "Current Pushing of the Oscillation Frequency of a 6328 Å He-Ne Laser" (Appendix III), we reported our results in explaining what seems to be the primary source of the external contributions which still limit the overall laser spectral width of He-Ne 6328 Å lasers. It should be noted that in our work, and also in most other work with 6328 Å lasers that we have seen reported, it appears that as one improves the laser design, and improves the shielding of the laser from acoustic and microphonic disturbances, the laser spectral width decreases to something on

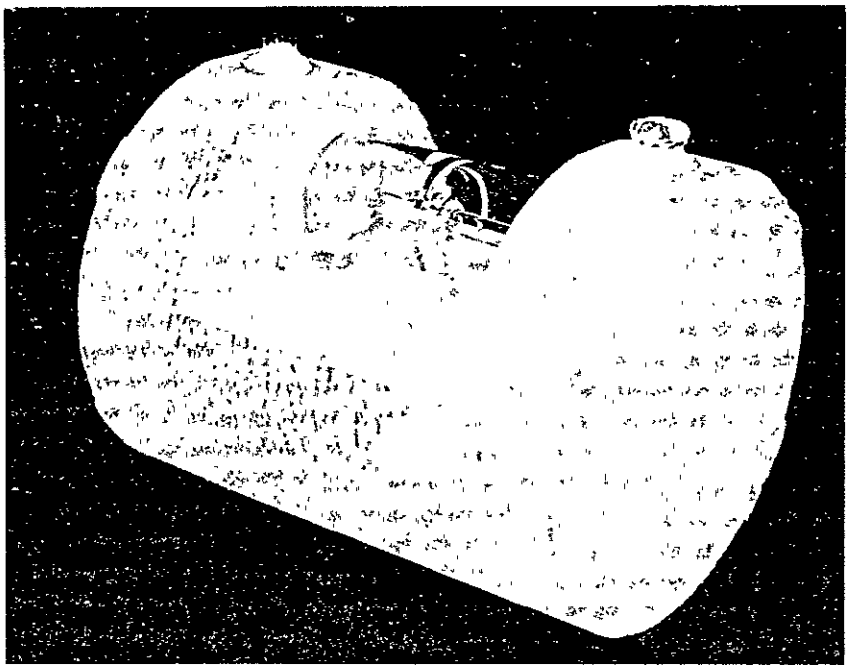


FIG. 8--Ingot laser.

the order of 10 KHz; and then simply does not get any better as further improvements are made. No one seems, for example, to have been able at 6328 Å to achieve the beat note purity of 20 or 30 cycles obtained by Javan, Jaseja, and Townes in their "Wine Cellar Experiments" which were done at 1.15 μ. We now believe that this is associated with an unusually high polarizability of the He-Ne discharge at 6328 Å, caused by the presence of Neon 1-s metastable atoms, and by the resulting modulation of the laser cavity frequency through turbulent disturbances in the laser plasma.

Figure 9(a) shows a simple experiment that was done to investigate the current pushing in our lasers, that is the small shift Δf in instantaneous laser oscillation frequency caused by a small change Δi in the discharge current through the laser tube. Very small shifts caused by small changes in discharge current could be easily measured using the heterodyne technique and the rf discriminator.

Figure 9(b) shows the experimental results for the current pushing figure, that is, the derivative of laser frequency vs laser discharge current, as a function of where the laser's oscillation frequency is tuned within the atomic linewidth. This quantity is fairly large, and absolutely independent of where the laser is tuned within its oscillation range. This indicates that the current pushing effect is not caused by any pushing or pulling effects associated with the lasing atomic transition itself; but must be caused by a background change in the dielectric constant of the plasma discharge, having nothing to do with the laser transition. Similar measurements were made at 1.15 μ also, as indicated on the bottom of the figure, and it was found that the current pushing

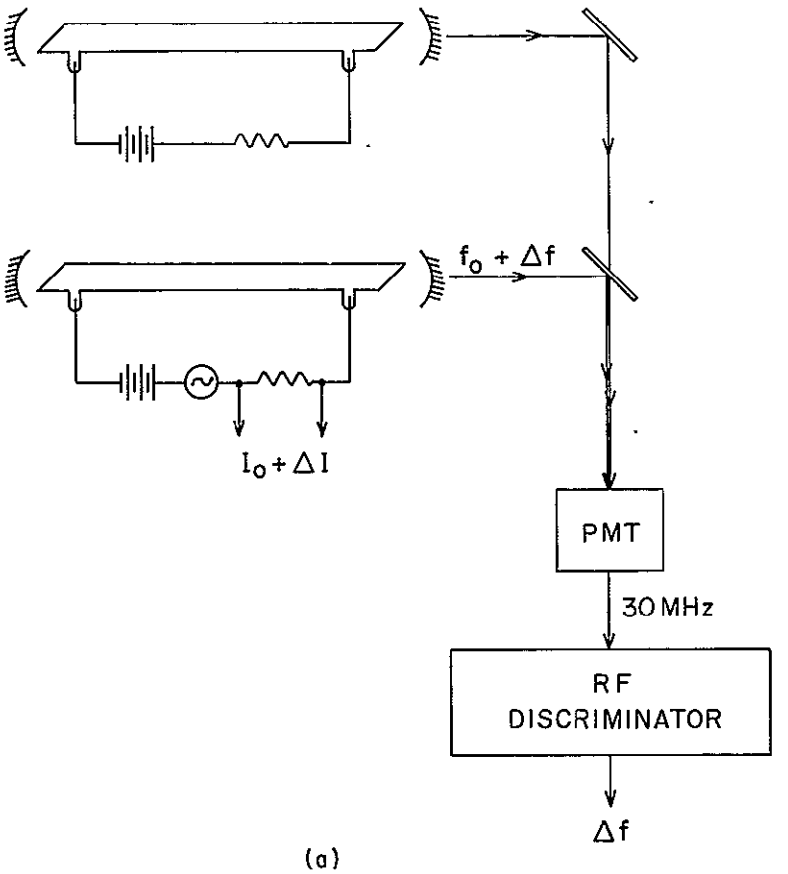
figure in this case was of opposite sign, very much smaller, and not necessarily flat across the atomic line. We will concentrate here on explaining the 6328 Å results.

Various mechanisms for explaining this result were considered. For example, the change in free electron density with increasing current in the laser discharge is many orders of magnitude too small to account for the observed current pushing. What we believe to be the correct explanation for this effect is illustrated in Fig. 10, showing the relevant energy levels of the He and Ne atoms. Notice that there is a collection of lowest excited Ne levels, the 1-s metastable levels, which build up a sizable population in the laser discharge. Furthermore, there are two relatively strong absorptive transitions at 6334 Å and 6402 Å from the most heavily populated of these metastable levels up to certain of the 2p levels which are not directly connected with the laser transition. The numbers in parentheses indicate the oscillator strengths of these two transitions. Increasing the current through the laser increases the populations of the Ne 1-s metastable levels; and the reactive effects of these excited atoms then change the effective dielectric constant of the cavity at the nearby laser transition.

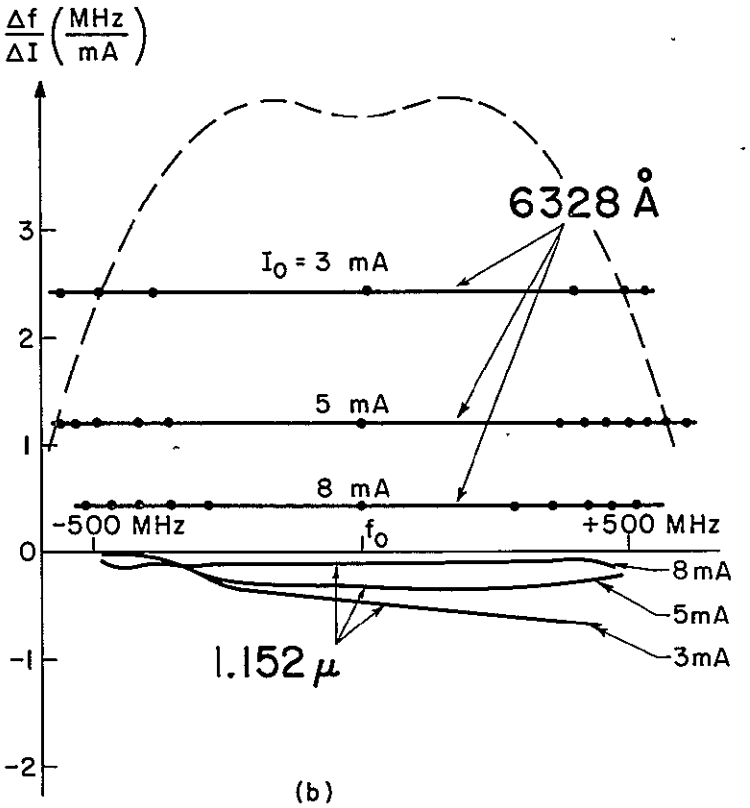
Figure 11 illustrates this point. The laser transition at 6328 Å is far out on the tails of the dispersive curves of these two contributions. Nonethless, evaluation of the expected metastable density, the oscillator strength of these transitions, and the resultant pulling effects make it clear that these two transitions are the culprits. Note that the weaker but much closer 6334 Å transition contributes about two-thirds of the effect while the stronger but more distant 6402 Å transition contributes the remaining one-third. These are absorptive transitions; and

FIG. 9--(a) Heterodyne experiment for measuring small shifts in laser oscillation frequency caused by current modulation or other perturbations.

(b) Shift in laser oscillation frequency with discharge current, df/dI, measured at both 6328 Å and 1.15 μ in the same lasers, as a function of laser frequency tuning within the atomic linewidth, for three different values of dc discharge current I_0


$f_o + \Delta f$
$I_o + \Delta I$
PMT
30 MHz
RF
DISCRIMINATOR
Δf

(a)


$\frac{\Delta f}{\Delta I}\left(\frac{MHz}{mA}\right)$
6328 Å
I_o = 3 mA
5 mA
8 mA
-500 MHz
f_o
+500 MHz
8mA
5mA
3mA
1.152 μ
3
2
1
0
-1
-2

(b)

FIGURE 9

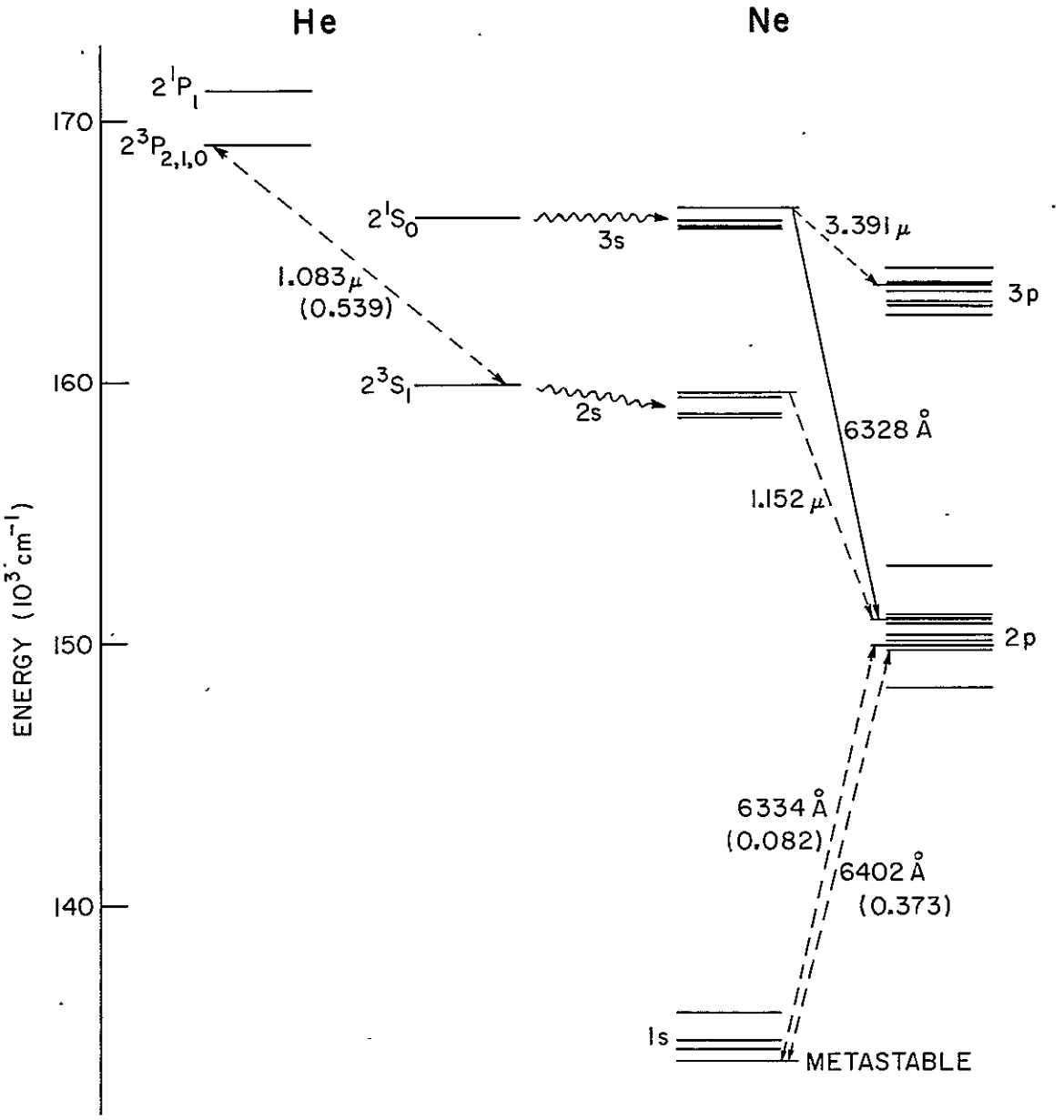


FIG. 10--He-Ne energy diagram.

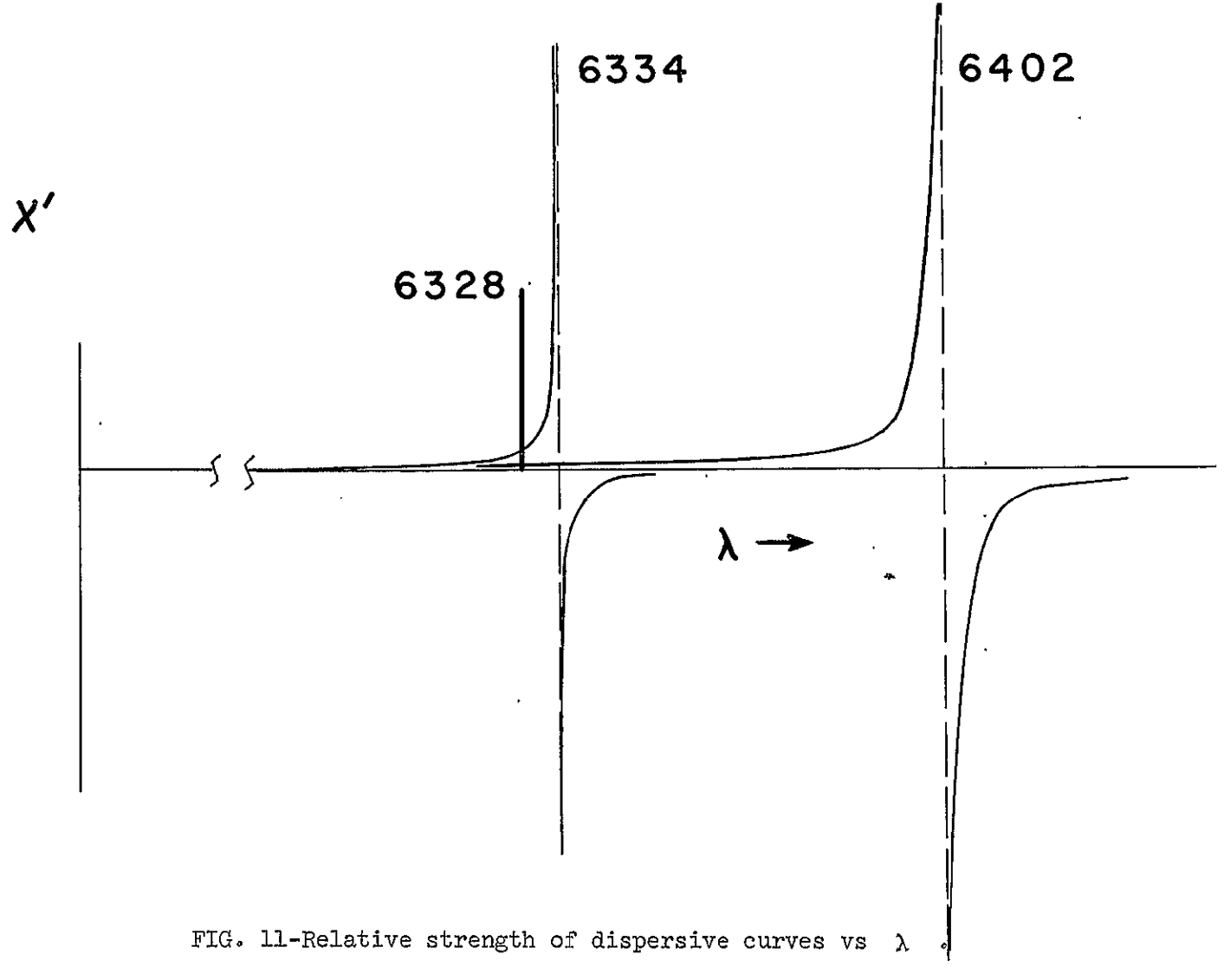


FIG. 11-Relative strength of dispersive curves vs λ

absorptive transition "pushes" a cavity frequency away from it; and hence the shift in the cavity frequency is to shorter wavelength, that is a blue shift with increasing discharge current. An extensive theoretical discussion of this effect is included in Appendix IV.

It is, therefore, our conclusion that the residual spectral broadening in high quality 6328 Å lasers probably results from random modulation of the cavity resonance frequency caused by random noise and turbulance in the laser discharge, acting through this Ne metastable mechanism. Note that in all of our experiments, care was taken to see that no coherent or single-frequency plasma oscillations were present or detectable by any means in the laser apparatus. Even so, there is still a substantial amount of low frequency 1/f type of noise in the plasma discharge. No simple mechanism for eliminating these effects is apparent at present, and this effect must be considered in any future applications of the 6328 Å laser, in particular as an ultra-stable frequency standard or laser communications source.

IV. ABSOLUTE FREQUENCY STABILIZATION OF He-Xe 3.51 μ LASERS

As part of the effort under this grant, we have also considered the long-term absolute frequency stabilization of the He-Xe 3.51 μ laser. We have proposed that a frequency discriminant be obtained by frequency-modulating a portion of a primary laser oscillator's output and passing it through an unsaturated and non-regenerative reference laser amplifier. The FM-to-AM conversion at the modulation frequency in the reference laser amplifier becomes zero only when the primary laser oscillator frequency exactly coincides with the atomic line center of the reference amplifier.

Thus, the FM-to-AM conversion provides a frequency discriminant for a stabilization feedback loop. The basic stabilization scheme is shown in Fig. 12.

This absolute stabilization scheme has the following desirable attributes:

(a) The main laser output is completely free from any applied modulation sidebands or frequency jittering, since only the portion tapped off to the reference amplifier is modulated.

(b) The reference laser amplifier operates in a linear, small-signal condition, thus avoiding any of the complexities which may be associated with saturation, "hole burning" effects, or other second-order effects.

(c) The gas fill, pressure and pumping conditions of the reference laser amplifier can be adjusted for optimum long-term frequency stability and reproducibility; while these parameters in the primary laser oscillator can be adjusted for maximum output power, efficiency, or other considerations independent of frequency stability.

This scheme should be particularly effective with the 3.51 μ He-Xe laser because:

(a) The He-Xe laser at 3.51 μ has high gain, which permits good sensitivity for the present scheme.

(b) Xenon is a heavy molecule, and thus has a narrow doppler linewidth.

(c) The He-Xe plasma is a cool plasma, and this also leads to a narrow dopper line.

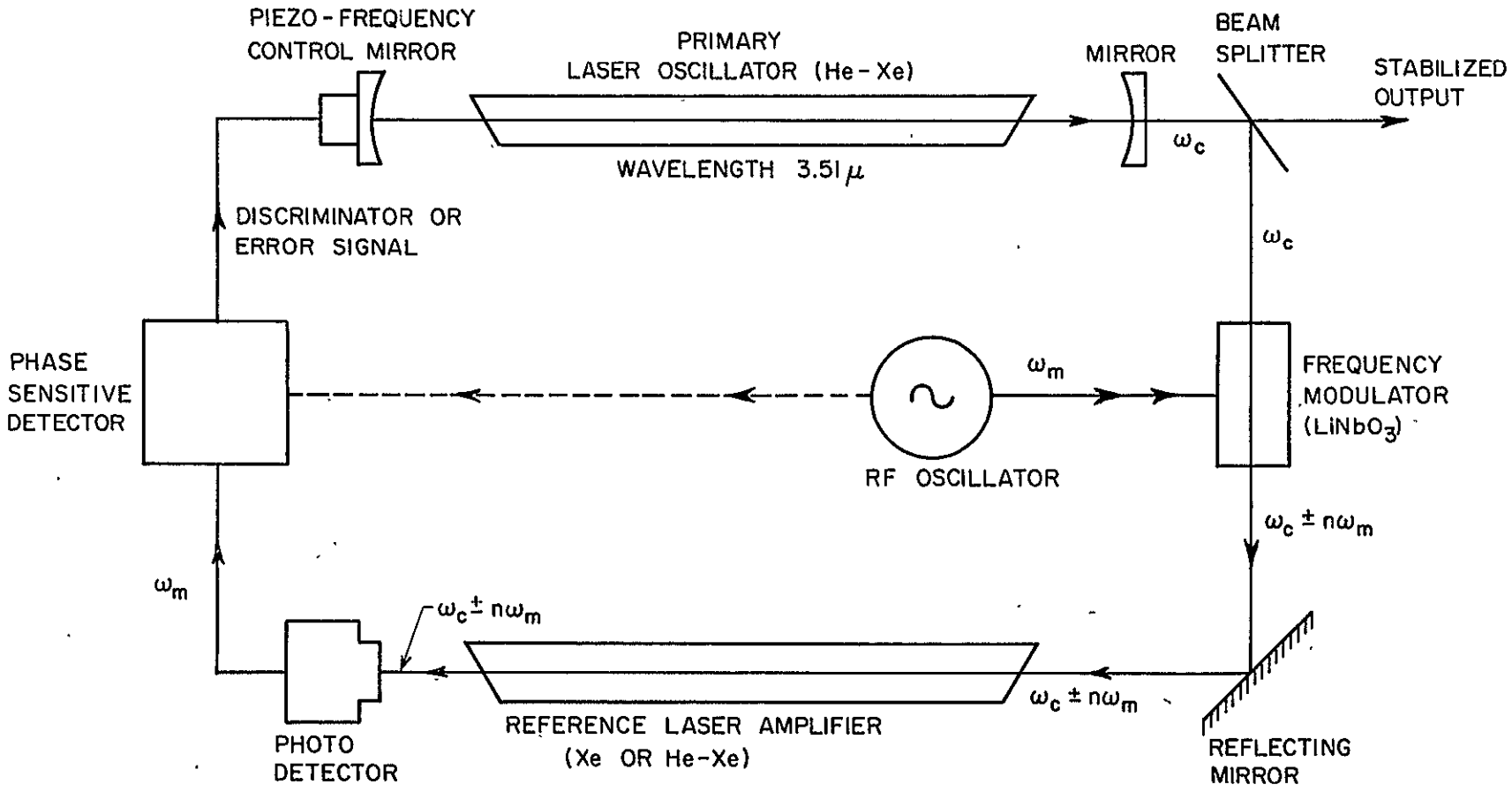

PIEZO - FREQUENCY
CONTROL MIRROR
PRIMARY
LASER OSCILLATOR (He - Xe)
MIRROR
BEAM
SPLITTER
STABILIZED
OUTPUT
WAVELENGTH 3.51 μ
ω_c
DISCRIMINATOR OR
ERROR SIGNAL
ω_c
PHASE
SENSITIVE
DETECTOR
ω_m
FREQUENCY
MODULATOR
($LiNbO_3$)
RF OSCILLATOR
$\omega_c \pm n\omega_m$
ω_m
$\omega_c \pm n\omega_m$
$\omega_c \pm n\omega_m$
PHOTO
DETECTOR
REFERENCE LASER AMPLIFIER
(Xe OR He-Xe)
REFLECTING
MIRROR

FIGURE 12

(d) The He-Xe laser amplifier can be operated at very low pressure, thus having minimal pressure shifts.

(e) The liquid N_2 pressure stabilization method used in our laser tubes should provide long-term stable operation of the lasers, with essentially constant Xe pressure.

Three pure xenon amplifier tubes have already been constructed, one 50 cm long and two 25 cm long. The inner bore diameters are all 5 mm. Choosing this large bore diameter permits easy alignment of the laser beam through the amplifier tubes, while also lowering the gain factor to avoid saturation and regeneration effects. Each tube has a side arm with a reservoir bulb which is cooled down to liquid N_2 temperature in operation, so that the tube is operated at a pressure corresponding to the Xe vapor pressure at that temperature. All of the tubes are designed for dc operation, so that pumping conditions can be controlled. However, the two 25 cm tubes can also be operated under rf excitation. These tubes appear to be satisfactory for our experiments.

To date all components of this system except for the necessary phase sensitive detectors, power amplifiers, and associated feedback equipment have been designed and are under construction or on hand.

Final reference amplifier gain measurements will begin as soon as the laser amplifier and oscillator can be set up in the vault. The frequency modulation system will be tested when the $LiNbO_3$ modulator is constructed. Finally, the whole stabilization scheme will be tested and the short-term and long-term stability measured. From sensitivity calculations, a frequency stabilization as good as one part in $10^{11\sim12}$ is possible with this system. However, the difficulties introduced by various practical considerations, such as interference and imperfect components, can be determined only with the acutal experimental system.

A theoretical discussion of this system was presented in Applied Optics, vol. 5, No. 10, October 1966; and is included for reference in Appendix V.

This completes the final report on Grant NGR 05(020)-234. All of the aims of the original proposal have been met or exceeded and a firm foundation for ongoing and future research has been laid.

APPENDIX

OBSERVATION OF QUANTUM PHASE NOISE IN A LASER OSCILLATOR

by

A. E. Siegman and R. Arrathoon

OBSERVATION OF QUANTUM PHASE NOISE IN A LASER OSCILLATOR*

A. E. Siegman and R. Arrathoon
Department of Electrical Engineering, Stanford University, Stanford, California 94305
(Received 12 February 1968)

By heterodyning together two stable 6328-Å He-Ne lasers, with one laser operating at very low power, quantum phase fluctuations caused by spontaneous emission have been observed. Results, although preliminary, seem in good agreement with the predictions of Schawlow-Townes and others.[1]

Neglecting amplitude fluctuations, the beat note between two laser oscillators may be written $v = V_0 \cos(\omega_0 t + \varphi)$, where ω_0 is the mean frequency and $\varphi(t)$ the randomly varying phase. Since the instantaneous beat frequency is $\omega_0 + \dot{\varphi}(t)$, the power spectral density $G_{\dot{\varphi}}(f)$ of the quantity $\dot{\varphi}(t)$ may be determined with an rf frequency discriminator centered at ω_0, followed by an audio-wave analyzer (Fig. 1). With appropriate instrumentation one can also measure the analytically related mean-square phase jitter $\langle \Delta\varphi^2(\tau)\rangle \equiv \langle [\varphi(t+\tau) - \varphi(t)]^2 \rangle$ as a function of the time interval τ.

In real lasers the random phase variation $\varphi(t)$ includes an "external" contribution $\varphi_e(t)$ due to acoustic noise, structural vibrations, and plasma disturbances, plus a usually much weaker contribution $\varphi_q(t)$ due to quantum noise. In our experiments the external disturbances occur primarily at low audio frequencies, with $G_{\dot{\varphi}e}(f) \sim 1/f^2$ [see Fig. 2(a)] and $\langle \Delta\varphi_e^2(\tau)\rangle \sim \tau^2$. The resulting beat-note power spectral density $G_v(f)$ is Gaussian with a linewidth typically $\Delta f_e \approx 3.5$ kHz in our apparatus, essentially independent of laser power level. By contrast, the quantum contribution should have a white-noise spectrum $G_{\dot{\varphi}q}(f) = 4\pi\Delta f_q$ and a phase jitter $\langle \Delta\varphi_q^2(\tau)\rangle = 2\pi\Delta_q\tau$. The Schawlow-Townes prediction is

$$\Delta f_q = \frac{\pi h f (\Delta f_{\text{cav}})^2}{P} \frac{N_2}{N_2 - (g_2/g_1)N_1},$$

where f is the oscillation frequency, Δf_{cav} the "cold"-cavity bandwidth, P the laser-oscillation power level, and N_2 and N_1 the upper and lower level populations. Except at our lowest power levels, the quantum contribution to the total beat-note spectral density is considerably less than the Gaussian external contributions. The quantum contributions are, however, separable by

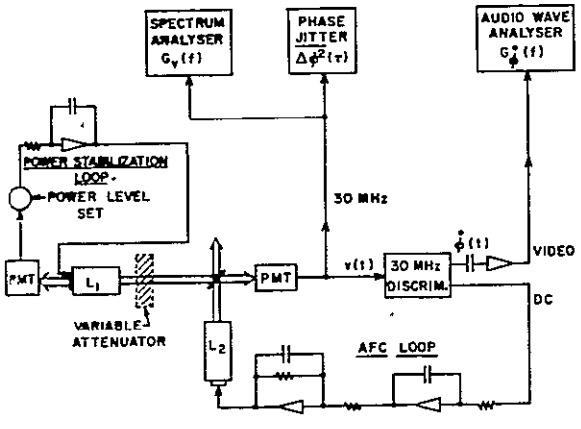


FIG. 1. Block diagram of system used to observe quantum phase fluctuations.

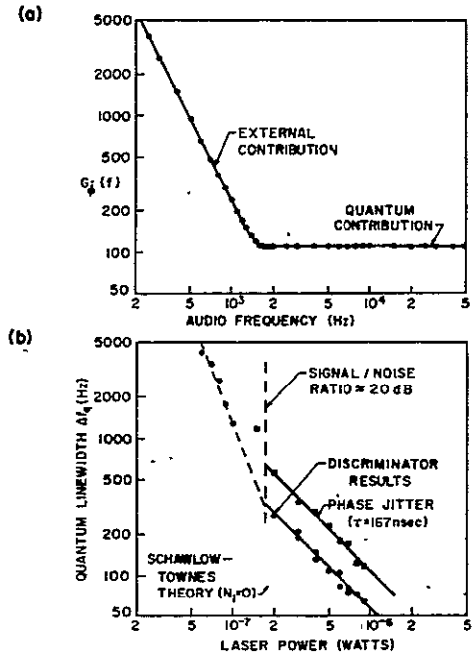


FIG. 2. (a) Typical discriminator noise output spectrum as measured by audio-wave analyzer, showing $1/f^2$ portion due to external disturbances and flat portion due to quantum noise (or, in some cases, to discriminator characteristics). (b) Quantum phase-noise linewidth contribution, as measured by flat discriminator noise level, versus oscillation power level of laser L_1. Also shown are the Schawlow-Townes theoretical result assuming $N_2/(N_2 - N_1) = 1$ and the results of experimental phase-jitter measurements at a single fixed delay $\tau = 167$ nsec.

examining $G_{\dot{\varphi}}(f)$ at high enough frequencies, or $\langle\Delta\varphi^2(\tau)\rangle$ at short enough times.

The experimental apparatus (Fig. 1) is basically the same as reported earlier.[2] Laser L_1 has a relatively high-transmission output mirror (1.7%) to enlarge Δf_{cav} and thus enhance the quantum line broadening. A slow power-stabilization loop provides stable operation at low power levels by piezotuning the cavity close to the edge of its oscillation range.[3] Laser L_2 is locked 30 MHz from L_1 by a slow automatic-frequency-control loop. The signal-to-noise ratio (S/N) of the beat-note depends upon signal power P from L_1 and photocathode quantum efficiency in the usual way. Since these measurements necessarily involved low and decreasing S/N, it was necessary to verify that measured noise linewidth increases at low P did not simply represent equipment characteristics. This was checked following each measurement by operating L_1 at a higher output level, where quantum fluctuations in L_1 should be negligible, and inserting a variable optical attenuator (~30 dB) in the output of L_1 to produce the same post-attenuator power output (and hence S/N conditions) in the succeeding apparatus. Comparison of the two measurements effectively determined the quantum contribution.

The circled points in Fig. 2(b) show the measured white-noise level of $G_{\dot{\varphi}}(f)$, expressed as equivalent linewidth Δf_q, versus oscillation level of L_1. The theoretical curve is the Schawlow-Townes formula[1] taking into account the uncertainty in cavity parameters for L_1 but assuming $N_2/(N_2-g_2N_1/g_1)=1$. The offset between theory and experiment can be accounted for by assuming $N_2/(N_2-g_2N_1/g_1)\approx 3$, not unreasonable for this particular laser system. The square data points represent measurements of $\langle\Delta\varphi^2(\tau)\rangle$ at one fixed value $\tau=167$ nsec, again converted to equivalent quantum linewidth. As also in our earlier work, we are unable to resolve the factor of 2 difference between discriminator and phase-jitter results here, and must continue to attribute it to experimental uncertainties or to some systematic error in one of the measurement techniques. We discount the apparent rapid increase in Δf_q below $P=2\times10^{-7}$ W since in this region the beat-note S/N decreases below 20 dB. At such low power levels, fluctuations in the power stabilization loop could result in an increase in observed noise because of nonlinear power dependence ($1/P$) of the quantum noise.

Our useful measurement range is uncomfortably limited at present by the vanishing quantum contribution at higher values of P and by reduction of the heterodyne S/N at lower values of P. However, we believe that the $1/P$ dependence observed between $P=2\times10^{-7}$ W and $P=9\times10^{-7}$ W most probably represents quantum phase fluctuations in laser L_1. In future experiments more detailed study should be possible by using a high-gain infrared laser transition, reducing the cavity Q, increasing Δf_{cav}, and thus greatly enhancing the quantum noise contribution.

Note added in proof.—Just as this report was completed we received the translation of a Russian Letters journal reporting very similar observations,[4] although at substantially higher power levels P and hence much lower values of Δf_q (~0.1-1.0 Hz). Our only reservations concerning the Russian results have to do with the nonideal discriminator characteristics mentioned earlier; i.e., we find that the observed inherent noise output from our real (nonideal) discriminator can have a flat spectrum that rises as ~$1/P$ due simply to decreasing S/N in the rf bandwidth rather than to any real frequency fluctuations in the beat signal. The results of Ref. 4 imply the availability of a very nearly ideal rf discriminator at $\omega_0=2\pi\times8.4$ MHz.

*Work supported by Joint Services Electronics Program Contract No. Nonr-225(83), and by a National Science Foundation Traineeship for R. Arrathoon.

[1]A. L. Schawlow and C. H. Townes, Phys. Rev. 112, 1940 (1958); for a more recent discussion, cf. M. Lax, in Physics of Quantum Electronics, edited by P. L. Kelley, B. Lax, and P. E. Tannenwald (McGraw-Hill Book Company, Inc., New York, 1966), p. 735.

[2]A. E. Siegman, B. Daino, and K. R. Manes, IEEE J. Quantum Electron. QE-3, 180 (1967).

[3]F. T. Arecchi, A. Berne, A. Sona, and P. Burlamacchi, IEEE J. Quantum Electron. QE-2, 341 (1966).

[4]Yu. N. Zaitsev and D. P. Stepanov, Zh. Eksperim. i Teor. Fiz.—Pis'ma Redakt. 6, 733 (1967) [translation: JETP Letters 6, 209 (1967)].

APPENDIX II

FURTHER MEASUREMENTS OF QUANTUM PHASE NOISE IN A He-Ne LASER

by

R. Arrathoon and A. E. Siegman

FURTHER MEASUREMENTS OF QUANTUM PHASE NOISE IN A He-Ne LASER

R. Arrathoon and A. E. Siegman

Department of Electrical Engineering
Stanford University
Stanford, California

ABSTRACT

Extended measurements of laser frequency fluctuations, using improved instrumentation, have provided further verification of the Schawlow-Townes relation for the spontaneous-emission-limited linewidth of a laser oscillator.

FURTHER MEASUREMENTS OF QUANTUM PHASE NOISE IN A He-Ne LASER*

by

R. Arrathoon and A. E. Siegman

Department of Electrical Engineering
Stanford University
Stanford, California

We recently reported the observation of spontaneous-emission-induced phase fluctuations, or "quantum phase noise", in a He-Ne 6328 Å laser operating at power levels in the low microwatt range.[1] Using slightly modified instrumentation, we have now extended these quantum phase fluctuation measurements over a wider range of laser powers, and have used amplitude fluctuation measurements to determine the excess noise factor appropriate to our lasers. Our new results in general validate our earlier measurements, and give further support to the Schawlow-Townes relation for the spontaneous emission contribution to the linewidth of a laser oscillator.[2]

The Schawlow-Townes formula predicts that the quantum phase noise contribution to the oscillation linewidth of a laser is

$$\Delta f_q = \frac{\pi h f (\Delta f_{cav})^2}{P} \times \left(\frac{N_2}{N_2 - (g_2/g_1) N_1} \right) = \alpha \frac{\pi h f (\Delta f_{cav})^2}{P} \qquad (1)$$

where Δf_q is the full lorentzian linewidth contribution due to this noise source; N_2 and N_1 are the upper and lower laser level populations; P is the laser power level; f is the oscillation frequency; and Δf_{cav} is the "cold" laser cavity bandwidth. The excess noise

factor $\alpha \equiv N_2/[N_2-(g_2/g_1)N_1]$ may be determined by measuring the amplitude fluctuations in the same laser, as shown by Freed and Haus.[3] The audio-frequency amplitude fluctuation spectrum of the laser, measured using a photomultiplier with an appropriate load resistance R_L, may be written as the sum of a shot-noise term and an excess or quantum amplitude noise term

$$S(f) = S_s + S_e(f)$$

$$= R_L^2 \left[2M\Gamma e I_a + 4MeI_a \eta\alpha \left(\frac{\Delta f_{cav}}{\Delta f} \right)^2 \frac{1}{1 + (f/\Delta f)^2} \right] . \qquad (2)$$

In this expression S is the spectral density of the voltage across R_L (in volt^2/Hz); M is the photomultiplier power gain; I_a is the anode dc current; η is the photocathode quantum efficiency; Γ is the shot noise enhancement factor; and Δf ($\approx$ 10 to 30 kHz) is the bandwidth of the amplitude fluctuation spectrum. For lasers operating at low oscillation power levels, the excess amplitude noise term $S_e(f)$ is substantially larger than the shot noise term for frequencies less than Δf.

The excess amplitude noise spectrum $S_e(f)$ was measured for our lasers, using an audio wave analyzer, at frequencies up to approximately 1 MHz. The results were in general agreement with those of Freed and Haus,[3] except for a small additional increase in the amplitude fluctuations at very low frequencies ($\lesssim$ 3 kHz), which we tentatively attribute to plasma disturbances. The mean-square voltage readings $\overline{v_n^2} = S_e B$, where B is the analyzer bandwidth, were multiplied by the appropriate correction factor of 1.13^2 to account for gaussian noise in a sinusoidally

calibrated average-reading voltmeter. The value of the excess noise spectrum for frequencies less than Δf, together with the measured bandwidth Δf of the excess noise fluctuations, in essence measures the excess noise factor α. In fact, if the measured dc voltage V_a across the anode resistance R_L is measured at the same time as the excess noise fluctuations, then the Schawlow-Townes formula may be cast into the extremely simple form

$$\Delta f_q = \frac{\pi}{4} \frac{S_e \Delta f^2}{V_a^2} \quad , \tag{3}$$

in which all quantities are directly measurable. The measured value of this expression was used to provide the theoretical value of Δf_q against which to compare our frequency fluctuation measurements. Using this approach, it is not necessary to make separate determinations of the photocathode quantum efficiency or the photomultiplier gain. However, these quantities were also independently measured to guarantee that all aspects of the experiment were under proper control.

The principal change in the frequency fluctuation instrumentation was the replacement of the previous commercially available 30 MHz FM discriminator (RHG Electronics Inc., Model DT 3006) with a quadrature detector discriminator[4] centered at 4.5 MHz. The 4.5 MHz amplifier, limiter, and discriminator were assembled from Fairchild and RCA integrated circuits. Two stages of preamplification and limiting were constructed from RCA CA-3013 integrated circuits. A subsequent stage consisting of a single Fairchild μA717E integrated circuit supplied further amplification and limiting. This stage, in conjunction with an external quadrature tank,[4] also provided the frequency discrimination.

The narrower bandwidth and somewhat better sensitivity of this system permitted quantum phase fluctuation measurements at higher laser power levels (lower frequency fluctuation levels) than previously possible.

Our total experimental results, together with the Schawlow-Townes prediction as calibrated from the amplitude fluctuation measurements, are presented in Fig. 1. The triangular data points indicate results obtained earlier by a phase jitter technique at a specific delay time $\tau = 167$ nsec . The square data points indicate quantum noise results obtained with the new 4.5 MHz discriminator and i.f. system. Both of these results differ by approximately a factor of two from the earlier measurements made with the 30 MHz discriminator system. The difference may be due to different limiting characteristics in the two discriminators. The Schawlow-Townes predictions, taking into account uncertainties in the evaluation, seem to be in good agreement with our overall results.

The apparent sharp increase in quantum noise below $P = 2 \times 10^{-7}$ watts should probably, as before, be discounted. In this region the signal-to-noise detection ratio has decreased below 20 dB; hence, all of the FM measurement techniques begin to deteriorate rapidly, in accordance with the well-known threshold properties of FM receivers. Also, due to the $1/P$ power dependence of quantum noise, we might expect that small fluctuations in the power stabilization loop at these low levels could result in an augmented noise output.

Progress is currently being made on a high-gain He-Xe 3.51 μ laser system with which we hope to observe quantum phase fluctuations in greater detail and with improved accuracy.

REFERENCES

* This study was supported by the NASA Electronics Research Center, Cambridge, Massachusetts.

1. A. E. Siegman and R. Arrathoon, "Observation of Quantum Phase Noise in a Laser Oscillator," Phys. Rev. Letters 20, 901 (22 April 1968); see also, A. E. Siegman, B. Daino, and K. R. Manes, "Preliminary Measurement of Laser Short-Term Frequency Fluctuations," IEEE J. Quant. Elect. QE-3, 180 (May 1967).

2. A. L. Schawlow and C. H. Townes, Phys. Rev. 112, 1940 (1958); for a more recent discussion, cf. M. Lax, in Physics of Quantum Electronics, edited by P. L. Kelley, B. Lax, and P. E. Tannenwald (McGraw-Hill Book Company, Inc., New York, 1966), p. 735.

3. R. Freed and H. A. Haus, "Measurement of Amplitude Noise in Optical Cavity Masers," Appl. Phys. Letters 6, 85 (March 1965).

4. James N. Giles, ed., Fairchild Semiconductor Linear Integrated Circuits Applications Handbook, Fairchild Semiconductors, Mountain View, California (Library of Congress No. 67-27446).

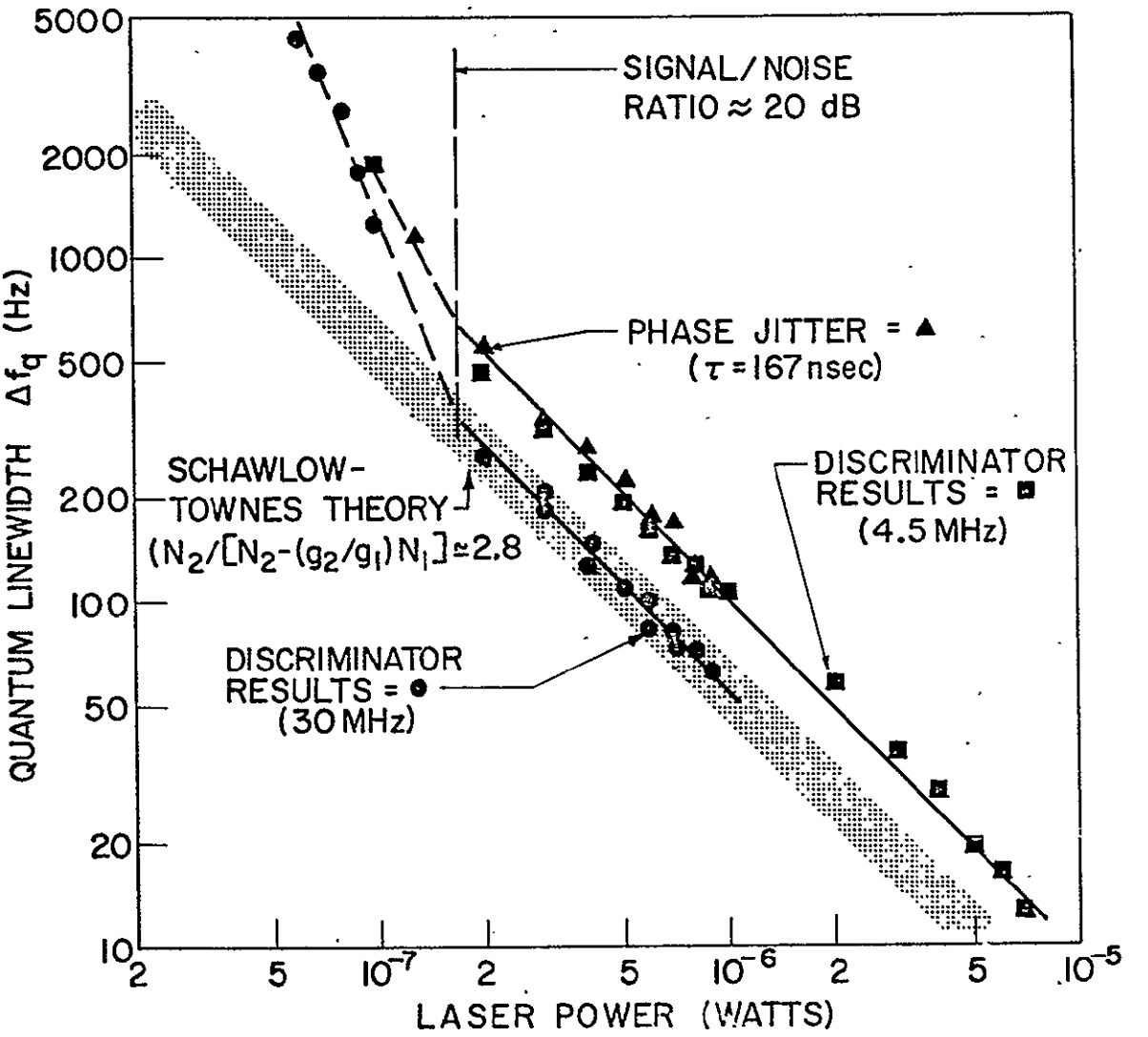


FIG. 1--Quantum phase noise measurements as obtained with the 30 MHz discriminator, the 4.5 MHz discriminator, and the phase jitter instrumentation. The shaded area represents the theoretically expected values based on the evaluation of Eq. (3).

APPENDIX III

CURRENT PUSHING OF THE OSCILLATION FREQUENCY OF A 6328 Å He-Ne LASER

by

R. Arrathoon and A. E. Siegman

CURRENT PUSHING OF THE OSCILLATION FREQUENCY OF A 6328 Å He-Ne LASER*

R. Arrathoon and A. E. Siegman

Stanford University
Stanford, California 94305

ABSTRACT

Increasing the discharge current through a small dc-excited 6328 Å laser changes the index of refraction of the discharge and causes a blue shift of the laser oscillation frequency of ~ 1 MHz/ma . The shift is independent of laser cavity tuning and decreases with increasing dc discharge current. Approximate calculations plus other experiments indicate that this shift represents the dispersive effects of the strong 6334 Å and 6402 Å upward transitions from the neon $1s_5$ metastable level. The residual spectral width and FM noise in stable 6328 Å lasers, especially as compared to 1.15μ lasers, may be due to plasma disturbances acting through this dispersive mechanism, rather than to microphonic disturbances as commonly suggested.

Supported by NASA Electronics Research Center, Cambridge, Massachusetts, under Grant NGR-05-020-234.

CURRENT PUSHING OF THE OSCILLATION FREQUENCY OF A 6328 Å He-Ne LASER

As shown in Fig. 1(a), we heterodyne together the outputs from two small stable dc-excited He-Ne lasers (Spectra Physics Model 119) at a nominal beat frequency of 30 MHz.[1,2] A small sinusoidal audio-frequency modulation ΔI added to the dc discharge current I_0 of one laser produces a corresponding small ac frequency shift Δf in the oscillation frequency of that laser. This shift is measured by passing the beat note into an rf discriminator centered at 30 MHz and observing the video output from the discriminator.

Figure 1(b) plots the resulting frequency-current pushing figure $\Delta f/\Delta I$ for several different values of dc current I_0 as a function of the tuning of the laser center frequency within the atomic linewidth, as measured both at 6328 Å and (by changing mirrors on the same lasers) at 1.15μ. The 6328 Å current pushing in particular is sizeable; decreases with increasing dc current I_0 ; and is constant to within measurement accuracy independent of laser cavity tuning within the atomic linewidth. The last observation indicates that the pushing represents a change in the background index of refraction of the laser discharge, rather than any line shifts or pulling effects associated with the 6328 Å laser atomic transition itself.

We believe that this shift can be attributed to the dispersive (i.e., off-resonant reactive) effects of the 6334 Å and 6402 Å absorptive

transitions from the neon $1s_5$ metastable level,[3] as indicated in Fig. 2. The results of Ladenburg[4] can be combined with more recent He-Ne laser studies to obtain rate equations that predict the neon 1s metastable density versus current in our lasers. These calculated results are in better than order-of-magnitude agreement with the neon 1s densities necessary[5] to account for the observed current pushing. Because the metastable density saturates with increasing current, the slope $\Delta f/\Delta I$ should decrease with increasing I_0 as observed.

As a partial test of this explanation, we focused ultraviolet radiation in the 3000-4000 Å band from a Hg arc lamp through one end mirror into the bore of one of the 6328 Å lasers during operation. This radiation excites atoms from the neon 1s up to the neon 3p levels.[3] As expected we observed a small but measurable downward frequency shift in the laser frequency when the ultraviolet radiation was pulsed on, independent of both tuning and of dc discharge current. We interpret the latter result as meaning that the incident ultraviolet radiation is almost totally absorbed on the 1s-3p transitions independently of the 1s metastable density, so that the frequency shift depends only on the number of incident ultraviolet photons and not on the 1s density or dc discharge current I_0.

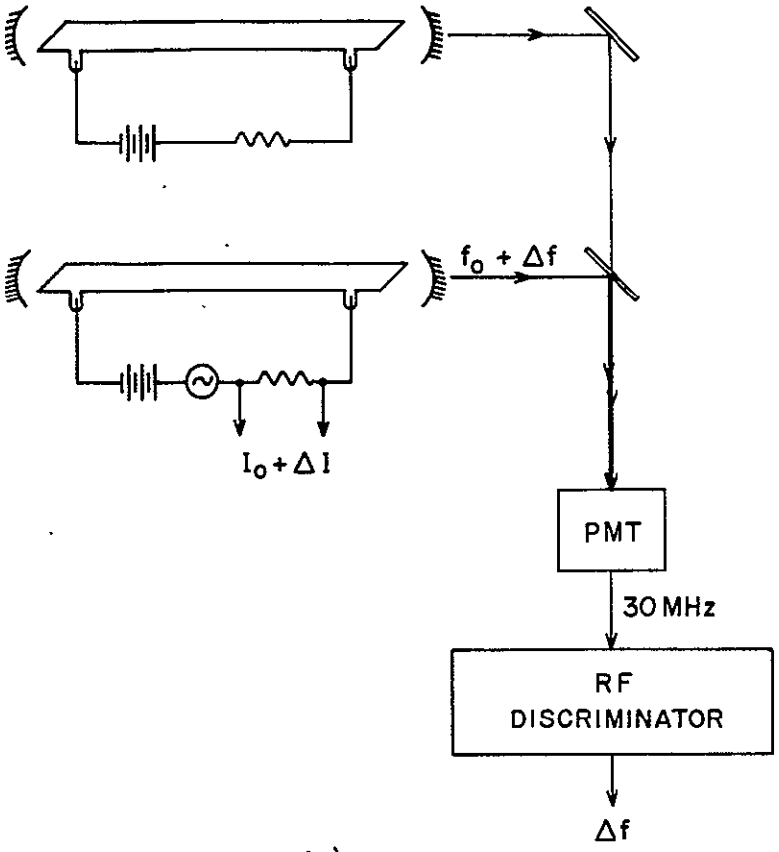
The observed current pushing figures for 1.15μ operation in Fig. 1(b) are seen to be an order of magnitude smaller, and of opposite sign from the 6328 Å results. They also exhibit some variation across the atomic linewidth. We have not pursued these results in the same detail as the 6328 Å results, but earlier studies of the 1.15μ transition[6] suggest that the 1.15μ current pulling may represent a combination of

the strong 1.083μ transition from the metastable 2^3S_1 He level; closely adjacent transitions within the Ne 1.15μ oscillation group; and the more pronounced pulling effects of the 1.15μ transition itself.

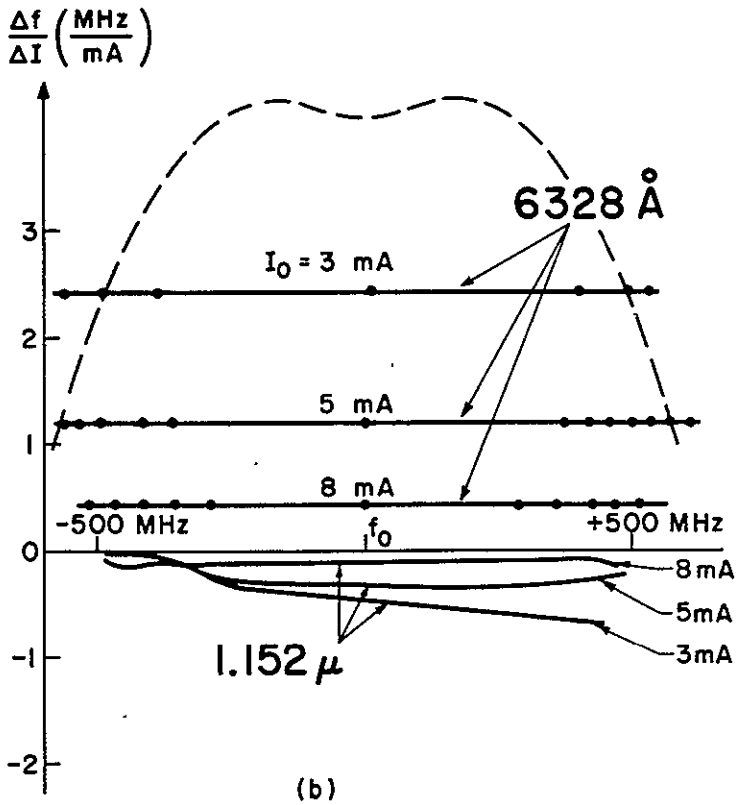
Javan and co-workers, in their "wine cellar" experiments of several years ago,[7,8] observed a beat note between free-running 1.15μ He-Ne lasers with a spectral purity on the order of a few tens of Hz. By contrast, the spectral purities or noise FM bandwidths obtained in beating experiments with 6328 Å lasers generally seem to be from 2 to 3 orders of magnitude larger.[1,2,9] Moreover, in our experience improvements in mechanical design and acoustic isolation do not seem to yield any improvement in this spectral purity beyond a certain point. Under quiet operating conditions, the power spectral density of the instantaneous frequency fluctuation in our lasers ("G_ϕ") exhibits no discrete oscillation spikes, such as might be attributed to plasma oscillations,[10-13] or to mechanical vibration modes. The frequency fluctuation spectrum does show a smooth f^{-2} variation, essentially the same as the f^{-2} noise spectrum we observe in the plasma discharge current itself. (However, the observed discharge current noise spectrum, when multiplied by the measured frequency-current pushing figure, is <u>not</u> sufficient to account for the full observed spectral width of the laser, although both have the same spectral dependence).

We suggest, therefore, that the residual spectral width of present-day high quality dc-excited (and perhaps also rf-excited) 6328 Å lasers may be due to the dispersive effects of random fluctuations in the neon 1s metastable density, caused by plasma disturbances and fluctuations.[10-13] These plasma fluctuations (striations, etc.) are not presently understood

in detail, but in our lasers have an f^{-2} spectrum and are apparently larger inside the discharge than is indicated by external discharge current fluctuations. Substantially greater spectral purity may be obtainable (with care!) in 1.15μ lasers, because the discharge is much less optically polarizable at 1.15μ, and hence the laser frequency is much less perturbed by plasma disturbances.


$f_o + \Delta f$
$I_o + \Delta I$
PMT
30 MHz
RF
DISCRIMINATOR
Δf

(a)


$\frac{\Delta f}{\Delta I}\left(\frac{MHz}{mA}\right)$
6328 Å
$I_o = 3$ mA
5 mA
8 mA
-500 MHz
f_o
+500 MHz
8mA
5mA
3mA
1.152 μ
3
2
1
0
-1
-2

(b)

FIGURE 1

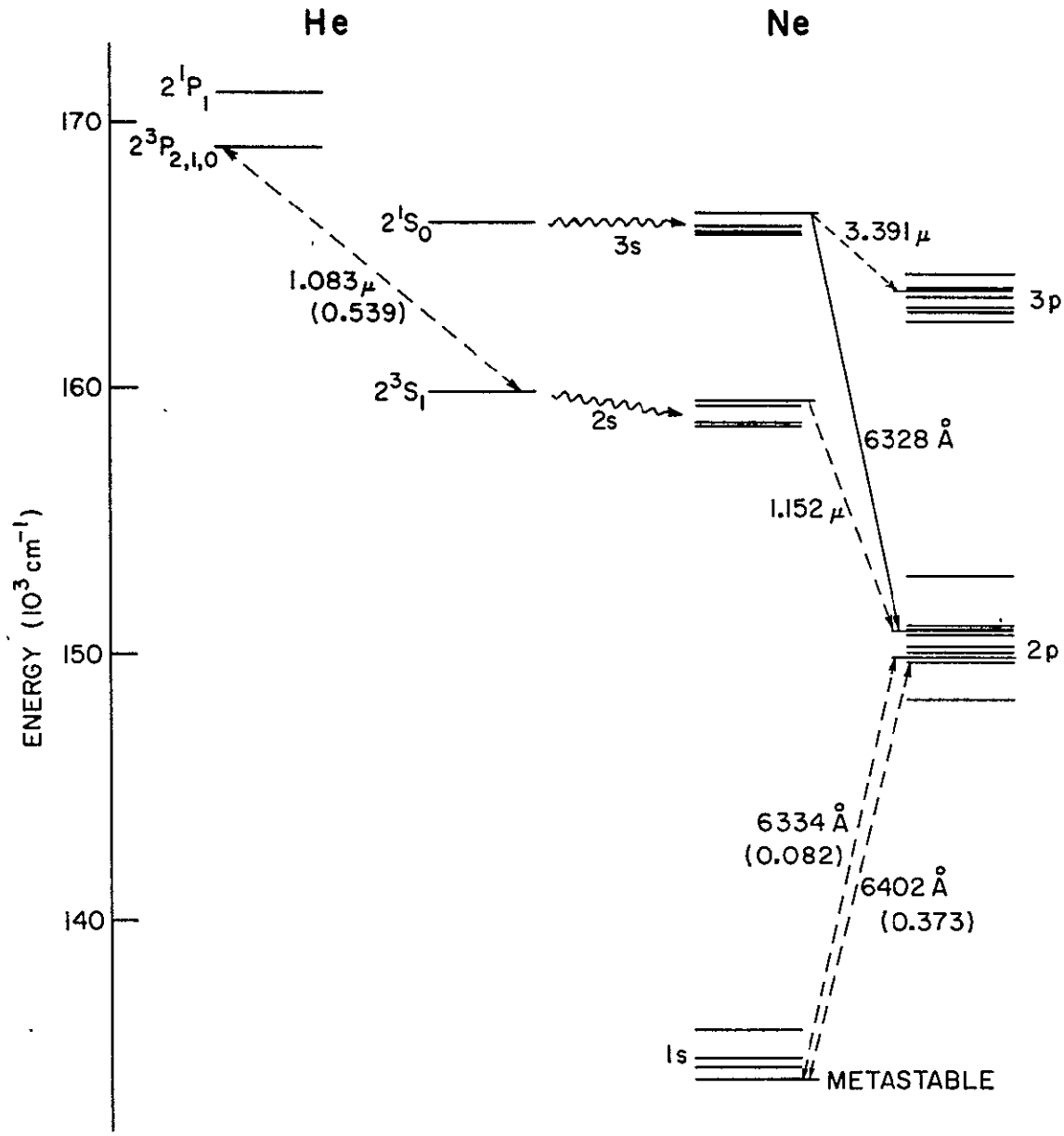


FIG. 2--Helium-neon energy level diagram, showing the transitions believed responsible for the observed frequency shifts.

APPENDIX IV

POSITIVE COLUMN POPULATION CALCULATIONS FOR THE EVALUATION OF DISPERSIVE EFFECTS IN He-Ne LASERS

by

R. Arrathoon

POSITIVE COLUMN POPULATION CALCULATIONS FOR THE EVALUATION OF DISPERSIVE EFFECTS IN He-Ne LASERS

by

R. Arrathoon

Microwave Laboratory
Stanford University
Stanford, California 94305

ABSTRACT

Sizeable shifts in the oscillation frequency of a He-Ne laser with variations in the dc discharge current have been observed. The shifting effects are believed to be due principally to changes in the populations of the helium 2^3S level and the neon $1s_5$ level. A simplified model is proposed for the approximate evaluation of these populations. A slightly modified version of the model, under regenerative conditions, is then used to evaluate the current shifting effects and the results are compared with experimental observations. The approach used is of sufficient generality to be applicable to most He-Ne lasers operating under optimum gain conditions. The results suggest that laser operation at 1.15 μ, rather than 6328 Å, may considerably reduce the effects of current variations on the oscillation frequency. For constant pressure-diameter product conditions, theory predicts that the current-shifting effects will vary approximately inversely with tube diameter.

POSITIVE COLUMN POPULATION CALCULATIONS FOR THE EVALUATION OF DISPERSIVE EFFECTS IN He-Ne LASERS

I. INTRODUCTION

The refraction index of a gas can be appreciably altered by the presence of significant excited state populations. In a gas laser this effect may be substantial for oscillation wavelengths in the vicinity of an atomic or molecular resonance. Changes in the refractive index of the plasma at the laser wavelength may then be effected by fluctuations in the populations of nearby resonance levels. The net result would be to modify the effective interferometer spacing and produce shifts in the laser oscillation frequency.

One method of affecting the level populations is to vary the discharge current. Earlier experiments[1] at 6328 Å, under single mode conditions, showed that low frequency perturbations in the dc discharge current of a small bore He-Ne laser produced blue shifts of the order of 1 MHz/mA. These shifts were essentially constant as a function of oscillation position within the Doppler broadened atomic line. Any pushing or pulling effects associated with the laser transition itself may be expected to be markedly dependent on position within the atomic line. These effects should approach zero at the line center and change sign on either side. Since such a variation was not observed, the shifting was thought to be due to background dispersive effects rather than to any changes in the inhomogeneously broadened atomic line itself; however, an unlikely combination of asymmetric line broadening and shifting could also have produced the observed qualitative characteristics.

As a test of the former hypothesis, population calculations based on experimental data[2,3] and adjusted for the conditions of our particular discharge were performed. These calculations indicated that background dispersive changes at 6328 Å, of the proper magnitude and sign to account for the observed effects, could indeed be produced by changes in the population of the Ne $1s_5$ level acting through the dispersive effect of the 6334 Å and 6402 Å absorptive transitions.

For laser operation at 1.15 μ, excitation induced red shifts were observed that were considerably reduced in magnitude from the 6328 Å effects; moreover, these shifts depended significantly on oscillation position within the Doppler broadened line. The shifting effects did not become zero at the line center, indicating the probable presence of residual background dispersive effects. Very approximate calculations, again based on adjusted experimental data,[4] indicated that the background effects were probably incurred by changes in the population of the He 2^3S level acting through the 1.08 μ absorptive transition.

The proposed mechanisms for the frequency shifting effects appeared to be consistent with the experimental observations[1] and with the very approximate population calculations that were made. Such calculations, however, were not very satisfactory as large adjustments in the available data were necessary in order to fit our particular system. Further, the adjustments were made on the basis of the ambipolar diffusion theory of the positive column, which has only limited validity for most He-Ne lasers. A more general and more accurate set of population calculations applicable to a large variety of He-Ne systems was clearly desirable.

A precise calculation of the populations in a He-Ne system has not as yet been obtained and would certainly involve the simultaneous solution

of a considerable number of rate equations. The problem would be further complicated by the fact that an explicit evaluation of all the terms in each of these equations would probably be very approximate. An alternate approach is to choose a simplified model that has the proper qualitative characteristics and which may be expected to be approximately valid over most of the region of operation. This latter approach has been chosen here in that it may be used to yield an approximate idea of the populations involved, for most He-Ne lasers, with a minimum of calculational effort.

The following section of this paper formulates the pertinent rate equations within the framework of a model that exhibits some of the characteristics that have been observed for He-Ne laser systems. These equations are then used in Section III to determine the level populations under both regenerative and non-regenerative conditions. The current-pushing effects are evaluated in Section IV and the results are compared with experimental data. Section V contains a discussion of some of the implications of this particular formulation.

II. FORMULATION OF THE He-Ne RATE EQUATIONS

The first qualitative description of the populations of some of the He-Ne levels was made by White and Gordon,[4,5] who found their model provided good agreement with the gain characteristics of the 3.39 μ transition. Any qualitative evaluation of these populations requires a much more specific model. In this section the plasma processes that must be considered in the rate equations will be explicitly defined and the simultaneous solution of a limited number of these equations will be used to describe the behavior of the relevant levels. The rate

equations which shall be developed here will be derived under non-regenerative small signal conditions such that stimulated emission terms may be neglected.

In order to justify the simplified model that will shortly be proposed, a brief description of some of the salient characteristics of the level populations, the electron number densities and the electron temperatures in a typical He-Ne discharge will be presented. Measurements by Labuda and Gordon[6] indicated that the electron temperature could be determined solely by the pressure-diameter (pd) product of the tube and that this temperature was current independent over the region of interest. Their measurements may be used to estimate the average electron temperature, for lasers in the optimum gain region of $pd = 3\text{-}4$ torr-mm,[5,7] as approximately $80,000 - 100,000^{\circ}K$. By contrast, a pure neon discharge with typical pressures and diameters will have a considerably lower average electron temperature that might be in the vicinity of $20,000 - 40,000^{\circ}K$.[2,3] As discussed in Section III, the relatively high electron temperature in a He-Ne discharge yields some inherent calculational advantages for the determination of level populations. Labuda and Gordon[6] also found that the electron number density varied linearly with current. The linear dependence of the electron number density on the discharge current, I, provides some justification for the experimentally determined relation proposed by Gordon and White[4,5] that describes the He 2^1S metastable level population. This relation takes the form $AI/(BI + C)$, where A, B and C are constants. At higher currents the current dependent de-excitation term, BI, will exceed the current independent term, C, and saturation

will ensue. The excitation and de-excitation of the He 2^1S metastable level were assumed to be primarily due to collisions with electrons that involve only single step processes proportional to the first power of the electron number density. The explicit electron collisional de-excitation mechanism, however, was not defined.

There are three relevant two-body collision processes involving electrons and excited or unexcited atoms: excitation from the ground state to an excited state, excitation from an excited state to a still higher excited state (stepwise excitation) and ionization from an excited state (stepwise ionization). Each of these processes has a converse counterpart that may be related to the forward process. The current dependent rate of depopulation of the He metastable levels will then be determined by collisional de-excitation to the ground state, stepwise ionization out of the excited state and stepwise excitation out of the excited state. Collisional de-excitation rates, which may be determined from the collisional excitation rates of the next section, are negligible in comparison with current independent loss mechanisms such as diffusion. Stepwise ionization rates, though somewhat larger than collisional de-excitation rates, also appear to be too small to account for the observed saturation characteristics. Stepwise excitation rates out of one excited state into another nearby excited state can, however, be considerably larger than stepwise ionization processes. If there exists a significant loss mechanism out of the nearby state such that most of the atoms lost to the nearby state do not return to the original excited state, stepwise excitation terms must also be included in the rate equations.

An energy level diagram of the helium atom is presented in Fig. 1. The lower four singlet and triplet levels are substantially removed from the next group of higher levels, and direct electron excitation rates from the ground state to these higher levels will be significantly smaller;[8] consequently, we will assume that only these lower four energy levels need be considered. Helium atoms are excited to the 2^1S level by electron collisions and are removed from this level by stepwise excitation. There is also destruction of the 2^1S atoms by diffusion and by resonant transfer collisions with neon ground state atoms that are then excited to the $3s_3$ level. Associated with this latter provess is its converse, production of 2^1S atoms by resonant transfer collisions of Ne $3s_2$ atoms with helium ground state atoms.

For stepwise excitation to be an important process in the net depopulation of the He 2^1S level it is necessary to assume that some significant loss mechanism exists from the 2^1P level so that atoms thus excited do not return to the 2^1S level via 2.06 μ radiative relaxation. One possibility for such a mechanism appears to be direct relaxation of the 2^1P level to the ground state. The 2.06 μ inverse transition rate is in the low microsecond range while Holstein's relations[9] indicate that the trapped inverse rate of the 2^1P to the ground state transition is also in this range. Such calculations, however, are only generally indicative, as Holstein's results are not quite applicable to the laser situation.[10] If there is trapping of the 2.06 μ transition, radiative relaxation to the ground state from the 2^1P level will become even more important. Another loss mechanism for the He 2^1P level appears to be resonant transfer collisions involving higher neon levels, several of which are

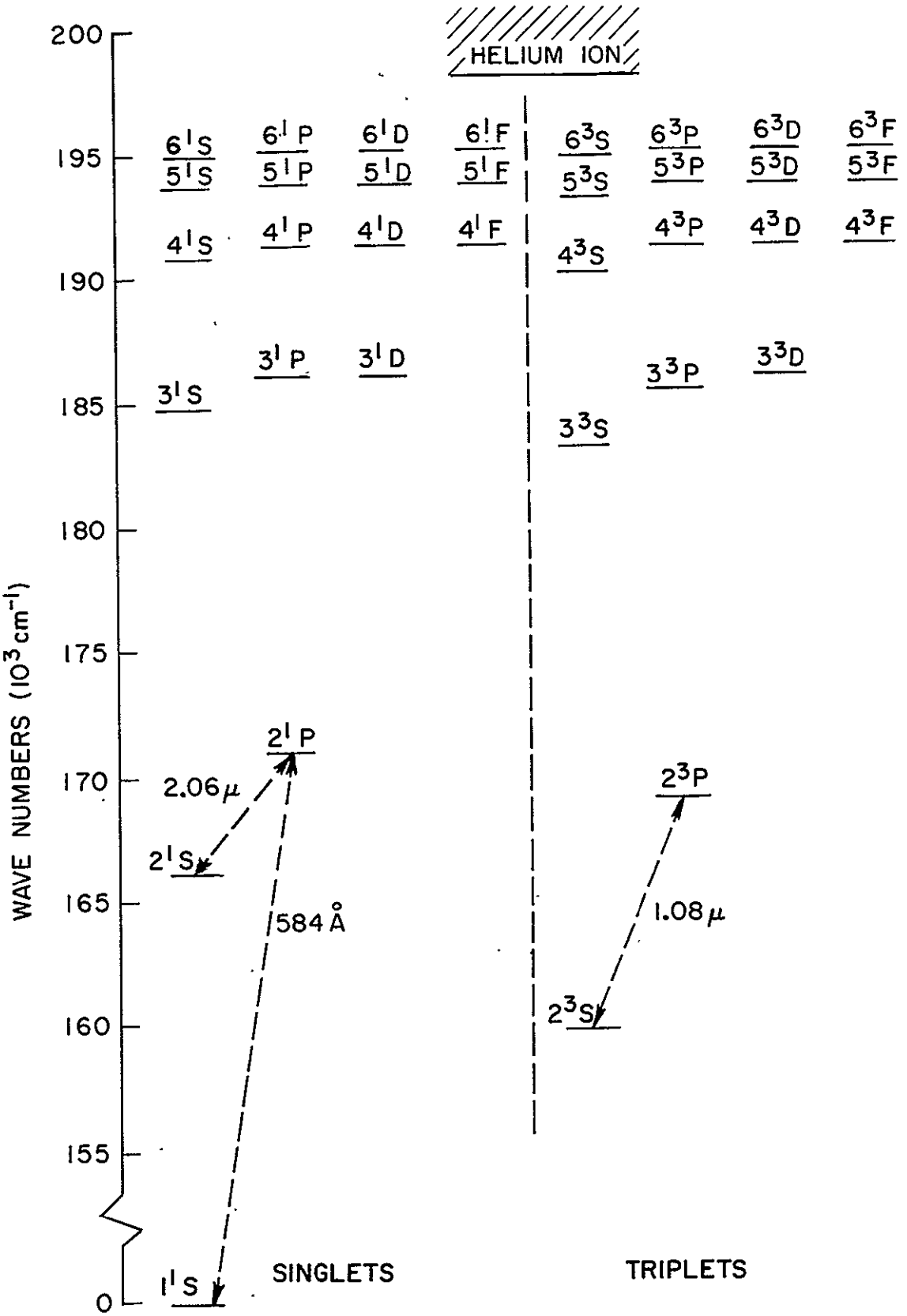


FIG. 1--Energy level diagram of the helium atom.

nearly coincident with the He 2^1P level. Collisions of this kind are believed to be responsible for promoting laser action between the neon 6p-6d transitions.[11]

Spin exchange transitions between optically forbidden levels, such as $2^1S - 2^3P$ stepwise excitation and $2^1S - 2^3S$ stepwise de-excitation, must also be considered. The cross sections for these processes can be quite large and are particularly probable for transitions in which there are no changes in the azimuthal quantum number.[12] In fact enhancement of the 2^3S population by collisional de-excitation from the 2^1S level has been observed in current pulsing experiments[13] and is believed to be one of the reasons for the increased power output of the 1.15 μ laser transition under pulsed operation. Experiments involving He-Ne lasers under steady state conditions, however, indicate that the populations of both the 2^1S and 2^3S levels entail current dependent excitation and de-excitation terms that are only linearly dependent on current.[4,14] Any interaction between the 2S levels would necessarily involve higher powers of current. We must then assume that the upper level is somehow effectively decoupled from the lower level under steady state conditions. This will be the case if the optically forbidden interaction between the triplet and singlet S levels is not as important as the optically allowed interaction between the S states and their respective P states. Alternately, a sufficient condition for the effective de-coupling of the 2^1S level would be that conditions of operation are such that the 2^1S and 2^3S populations are related by a Boltzmann distribution at the electron temperature (for a Maxwellian electron distribution) so that excitation and de-excitation processes between the two levels would cancel. In any case, we will base our model on the experimental observations that

the 2^1S level is essentially de-coupled from the 2^3S level to obtain the following approximate rate equation

$$\frac{dN(2^1S)}{dt} = 0$$

$$= \alpha_1 N_e N_o - \beta_1 N_e N(2^1S) - Z_{1r} N(2^1S) + Z'_{1r} N(3s_2) - D_1 \nabla^2 N(2^1S) \tag{1}$$

where α_1 is the direct electron excitation rate constant $(cm^3 \cdot sec^{-1})$, N_e is the electron number density, N_o is the ground state density, β_1 is the stepwise excitation rate constant, Z_{1r} is the resonant transfer rate, Z'_{1r} is the reverse transfer rate and D_1 is the diffusion coefficient. The various processes that contribute to this rate equation are illustrated in the rate diagram of Fig. 2.

A rate equation may also be written for the $3s_2$ level, in which the primary excitation mechanism is resonance transfer and the primary de-excitation mechanisms are reverse transfer and spontaneous emission. Interaction with the other $3s$ levels may be ignored[15] so that the rate equation becomes

$$\frac{dN(3s_2)}{dt} = 0$$

$$= Z_{1r} N(2^1S) - Z'_{1r} N(3s_2) - A_1 N(3s_2) \quad , \tag{2}$$

where A_1 is the inverse lifetime of the $3s_2$ level. In the vicinity of the tube center Eqs. (1) and (2) may be solved to yield

$$N(2^1S) = \frac{\alpha_1 N_e^o N_o}{\beta_1 N_e^o + Z_{1r}\left\{\frac{A_1}{Z'_{1r} + A_1}\right\} + D_1\left\{\frac{2.405}{R}\right\}^2} \tag{3}$$

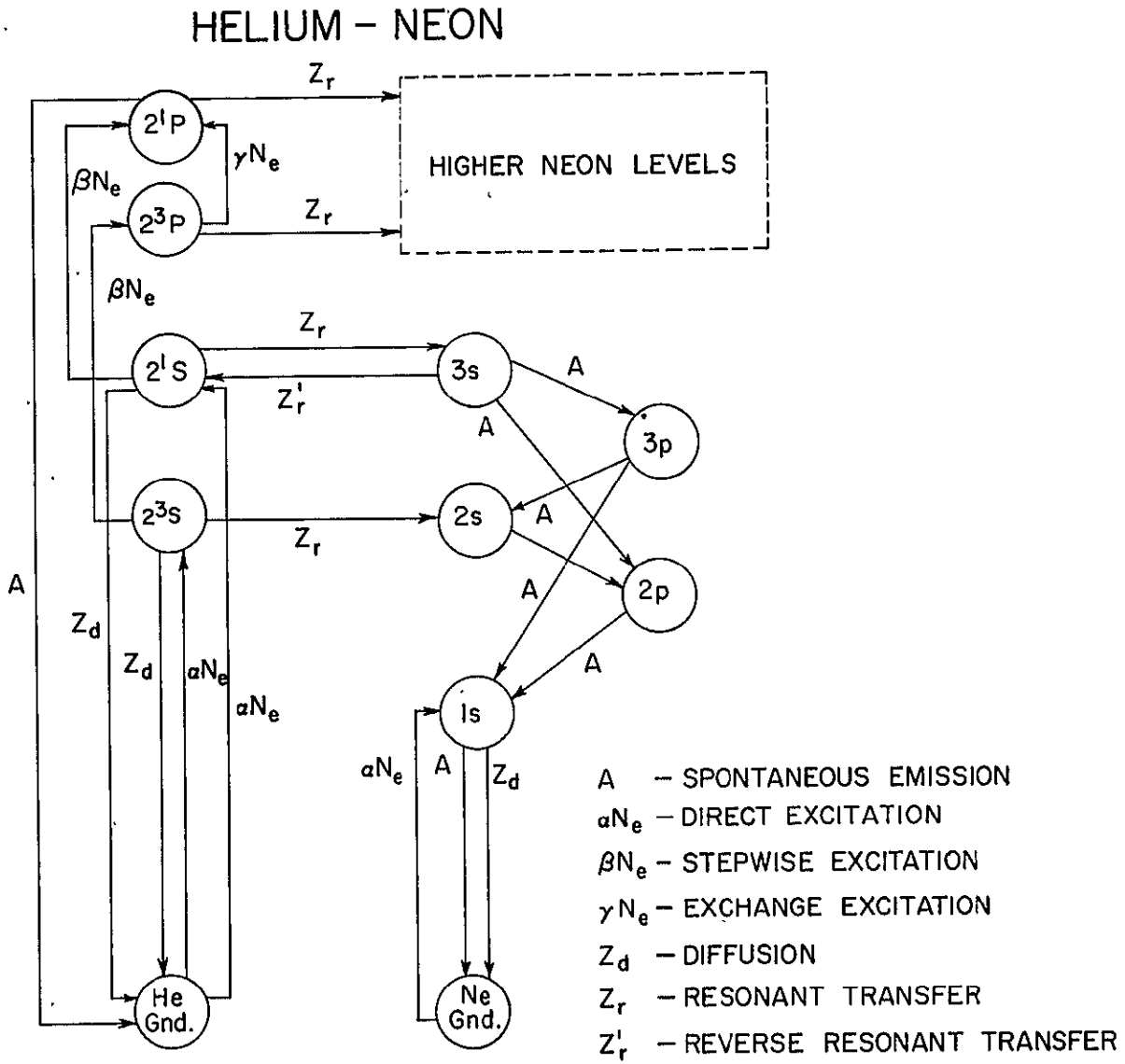


FIG. 2--Rate diagram of the He-Ne system illustrating the principal processes in the formulation of the rate equations described in Section II.

where a 0'th order Bessel function for the electron number density has been assumed, N_e^0 represents the on axis electron number density and R is the tube radius. Since most of the lasing action is confined to the central portion of the discharge tube, these results will be particularly useful when extended to regenerative conditions. The sum of the first two terms in Eq. (2) represents the net production rate at which $3s_2$ atoms are formed due to the presence of 2^1S helium. This rate, E_1, becomes

$$E_1 = Z_{1r}\left(\frac{A_1}{Z'_{1r} + A_1}\right) N(2^1S) \quad . \qquad (4)$$

The 2^3S level may be treated in a similar way. In this case stepwise excitation will occur between the 2^3S and 2^3P levels. It is not clear precisely how the 2^3P level will be depopulated, as there are no allowed transitions between any of the triplets and the singlet ground state; moreover, radiative relaxation at 1.08 μ from the 2^3P state back to the 2^3S state will be stronger than in the case of the 2.06 μ $2^1P - 2^1S$ transition. There will, however, be some atomic collisional de-excitation of the 2^3P level, as this level is separated from the 2^1P level by only about 7 kT. In addition spin exchange electron collisional de-excitation between these closely spaced P states might be expected to be an even more significant interaction mechanism. As in the case of the 2^1P level, there will also be resonant transfer collisions involving the 2^3P level and higher neon levels. The remaining excitation and de-excitation processes are similar to those for the 2^1S level so that the rate equation for the 2^3S level will have the same form as that of Eq. (1).

The resonant transfer rate into the neon 2s levels appears to be divided evenly between the $2s_2$ and $2s_3$ levels.[15] If we temporarily ignore any contributions by radiative cascade from higher levels and also any interactions among the 2s levels, the rate equation for the $2s_2$ and $2s_3$ neon levels will be similar to Eq. (2). The simplest treatment is to consider the two levels as a group with one A coefficient and one total de-excitation rate by reverse transfer. In this case the solution obtained for the population of the He 2^3S level will be similar to Eq. (3). It will be shown in the next section that A_2, the combined inverse lifetime of the $2s_2$ and $2s_3$ levels, is 1-2 orders of magnitude larger than Z'_{2r}, the reverse transfer rate out of these levels. As indicated in the rate diagram, the net effect is to make the resonant transfer process essentially unidirectional. Any reasonable contributions by radiative cascade from higher levels into the Ne 2s levels or from electron collisional excitation of neon ground state atoms to the $2s_2$ and $2s_3$ levels should then be expected to have little effect on the population of the 2^3S level. A form similar to Eq. (3) will thus be a valid solution for the population of the 2^3S level under the conditions $A_2 > Z'_{2r}$. Any redistribution of populations involving all the 2s levels would make the transfer process still more non-reciprocal so that the form of Eq. (3) with $A_2 > Z'_{2r}$ would correspond even more closely to the actual form of the solution for the 2^3S population. In obtaining relations of the form of Eq. (3) the currents are assumed to be sufficiently low so that only single step electron collisional processes need be considered. For very large currents Eq. (3) will no longer be valid.

The treatment of the neon levels is complicated by the helium induced population changes and also by the sheer number of interacting levels that must be considered. The lowest neon level, the metastable $1s_5$, appears to be somewhat more tractable than many of the higher levels. For this level there will be production terms involving electron collisional excitation of neon ground state atoms. There will also be direct excitation of the higher groups of neon levels with subsequent radiative cascade contributions to the $1s_5$ level. The direct excitation rates of the higher neon levels, however, will be substantially less than those of the 1s levels[12] so that these cascade contributions may be ignored. Helium atoms will also be supplying excited neon atoms at a net production rate of $E_1 + E_2$, where only the He 2^1S and 2^3S resonant transfer contributions are considered. The excited neon atoms will ultimately radiatively relax to the neon 1s levels, some fraction of which will become $1s_5$ atoms. The net production rate of neon $1s_5$ atoms due to the presence of helium will then be $k_1E_1 + k_2E_2$, where k_1 and k_2 are the respective fractional contributions arising from the He 2^1S and 2^3S resonant transfer collisions. A rigorous determination of k_1 and k_2 requires a simultaneous solution of a large number of neon level rate equations. Alternately k_1 and k_2 may be approximately determined by assuming that the discharge current is sufficiently large so that there is complete equilibrium among the various 1s levels at the electron temperature.[2] In this limit the k's will be identical and approximately equal to the ratio of the statistical weight of the $1s_5$ level to the total weight of the 1s levels (5/12). Similarly complete equilibrium may be assumed for direct

electron excitation rates among the 1s levels. In the equilibrium or high current limit stepwise excitation from the $1s_5$ levels to the Ne 2p levels need not be considered in the $1s_5$ level rate equation. Radiative relaxation from the 2p levels back to the 1s levels together with rapid equilibrium redistribution of the 1s level populations ensured that stepwise excitation is not a significant loss mechanism.[16]

Current independent loss mechanisms for the $1s_5$ level include diffusion and atomic collisional de-excitation. Relative rate estimates based on experimental data[17] show that diffusion is clearly a more important loss mechanism than atomic collisional de-excitation. If no current dependent loss mechanisms are included in the formulation of the neon $1s_5$ level rate equation the on axis population of this level the becomes

$$N(1s_5) = \frac{\alpha_3 N_e^0 N_n + k_1 E_1 + k_2 E_2}{D_3 \left\{\frac{2.405}{R}\right\}^2} \qquad (5)$$

where α_3 is the direct electron excitation rate constant, N_n is the concentration of ground state neon atoms and D_3 is the diffusion coefficient for this level. The omission of current dependent de-excitation terms greatly facilitates the evaluation of the population of this level, but it is these current dependent de-excitation terms that are responsible for achieving thermal equilibrium among the 1s levels in the high current limit. The exclusion of these terms, which would normally appear in the denominator of Eq. (5), implicitly assumes that low current conditions are valid and prevents the predicted 1s

level population from saturating. The first term of Eq. (5), the unsaturated neon term, will be particularly responsible for introducing large errors at higher values of current. The $1s_5$ level population thus determined will still be very approximately correct in the intermediate current region where both the high and low current approximations overlap.

III. DETERMINATION OF LEVEL POPULATIONS

The following evaluation for our particular system will be generalized sufficiently so that similar calculations may be used for pressure-diameter products in the optimum gain range. The approach used here will then be applicable in the range pd = 3-4 torr-mm for tube diameters of from 1-6mm. All pertinent quantities will be determined from a knowledge of pressure, diameter, current and gas mixture. In the evaluation of the cross sections, the spontaneous emission rates, the electron temperatures and the electron number densities, experimentally available data will be used whenever possible.

The values for electron temperature and electron number densities may be obtained from the measurements of Labuda and Gordon[6] but there will be some adjustment in this electron temperature depending on the gas mixture. For H-Ne ratios ranging from 5 to 1 and higher the adjustment will be small and may be estimated from the plots of Young's calculations.[18] There will also be some adjustments in the electron number density measurements, which were made at only one pressure-diameter product (3.6 torr-mm) for tube diameters ranging from 2-6mm. These results may be extended by use of the proportionality

$$N_e = \frac{pd}{T_e^{3/4}} \frac{I}{d^2} \qquad (6)$$

where T_e is the electron temperature and I is the tube current. This relation is based on ambipolar diffusion theory and assumes that, in the region of interest, the electron drift velocity is linearly proportional to E/p , the ratio of the electric field to the pressure.[20] For small tube diameters Labuda and Gordon's measurements yield a $1/d^2$ dependence on electron number density; however, for larger diameters the dependence is somewhat slower, indicating the shortcomings of ambipolar diffusion theory in this situation. In the optimum gain range the tube radius is several times the electron mean free path, as opposed to the conditions of ambipolar diffusion which require a tube radius that is very much larger than the electron mean free path. Equation 6, however, appears to have reasonable validity provided it is not used over very extended ranges. For our particular tube with a He^3 : Ne^{20} ratio of 9:1 at a total of 3.7 torr and a tube diameter of 1 mm, an electron temperature of 91,000° K and an on axis electron number density of $5.8 \cdot 10^{10}$ ($cm^{-3} mA^{-1}$) was obtained. This tube is normally operated at a current of about 5 mA.

The electron collisional excitation and de-excitation rates may be obtained by appropriately integrating the electron energy distribution over the energy dependent cross sections of the process in question. The integrals may be expressed as

$$\alpha N_e = N_e \frac{2}{m}^{1/2} \int_{E_1}^{\infty} \sigma(E) E^{1/2} f(E)\, dE \qquad (7)$$

where $\sigma(E)$ is the cross section, m is the electron mass and f(E) is the normalized electron energy distribution. The choice of energy

distribution is made particularly convenient by the relatively high average electron energy in the He-Ne discharge. For our discharge conditions the central portion of the distribution is indeed Maxwellian, but significant departures from a Maxwellian distribution are present for electron energies above 35 ev.[21] The cross sections under consideration peak well below 35 ev. and the electron temperature is sufficiently high so that most of the contribution to the rate integral is in the region before large departures from a Maxwellian occur. We will thus use the Maxwellian as a reasonable approximation to the actual energy distribution in the evaluation of the rate integrals.

The direct electron excitation cross sections may be approximated by an expression of the form xe^{-x}, which is known to have the desired functional characteristics.[22] In order to obtain a better fit with the available data, we will use a somewhat modified version of this cross section. The excitation cross section will then have the form

$$\sigma(E) = C_1 Q_0 \left\{\frac{E - C_2E_1}{C_3E_1}\right\} - \left\{\frac{E - C_2E_1}{C_3E_1}\right\} \qquad (8)$$

where Q_0 is the maximum value of the excitation cross section and E_1 is the threshold excitation energy. The above expression has a maximum at $E = (C_2 + C_3)E_1$ and also has the convenient characteristics that it may be integrated in closed form with a Maxwellian distribution. The cross section may then be substituted in Eq. (7) so that the resulting

excitation rate becomes

$$\alpha N_e = 6.69\cdot 10^7 \; E_1^{1/2} \left\{ \frac{E_1}{kT_e} \right\}^{3/2} \exp\left\{ 1 + \frac{C_2 - 1}{C_3} \right\} \left\{ \frac{C_1}{C_3} \; Q_0 \right\} \frac{e^{-\frac{E_1}{kT}}}{\left\{ \frac{C_3 E_1}{kT_e} + 1 \right\}^3}$$

$$N_e \left[C_3(1 - C_2) \left\{ \frac{C_3 E_1}{kT_e} + 1 \right\}^2 + C_3^2(2 - C_2) \left\{ \frac{C_3 E_1}{kT_e} + 1 \right\} + 2C_3^3 \right] \qquad (9)$$

where E_1 , kT_e are in ev. The cross section for the electron excitation of the He 2^3S level may be determined from the partly theoretical, partly experimental, compilation of Corrigan and von Engel.[8] For this level the values of the parameters in the simplified form of the excitation cross sections are $E = 19.8$ ev., $C_1 \simeq .7$, $C_2 \simeq .85$ and $C_3 \simeq .2$, with $Q_0 = 4.5\pi a_0^2 \cdot 10^{-2}$ (a_0 is the Bohr radius). For neon the location and magnitude of the maximum value of the total 1s level excitation cross section may be obtained from the appropriately scaled[8] data of Maier-Leibnitz.[12] If we assume that the higher energy falloff is similar to that of the helium 2 S level, we obtain $E_1 = 16.6$ ev., $C_1 \simeq .7$, $C_2 \simeq .85$ and $C_3 \simeq .2$, with $Q_0 \simeq 6\pi a_0^2 \cdot 10^{-2}$. The actual cross section of the Ne $1s_5$ level is taken to be 5/12 of this value, consistent with high current equilibrium considerations. The He 2^1S level has different excitation characteristics with a much broader maximum occuring at somewhat higher electron energies.[8] From the

experimental data of Dorrestein[23] $E_1 = 20.6$ ev , $C_1 \simeq C_2 \simeq 1$ and $C_3 \simeq .5$, with $Q_0 \simeq 1.8\pi a_0^2 \cdot 10^{-2}$.

The cross sections for stepwise excitation of optically allowed S to P states in helium do not appear to be experimentally available. Quantum mechanical distorted wave calculations have been derived for such transitions that may be approximated by a simplified relation[24] similar to the classical Thompson cross section for ionization.

$$\sigma(E) = 12\pi a_0^2 \frac{E_h^2}{E} \left\{ \frac{1}{E_n} - \frac{1}{E} \right\} f_{ik} \quad , \tag{10}$$

where E_n is the energy separation between excited states, f_{ik} is the oscillator strength and E_h is the ionization potential of the hydrogen atom. The oscillator strengths for the $2^1S - 2^1P$ ($E_n = -.602$ ev.) and the $2^3S - 2^3P$ ($E_n = 1.14$ ev) transitions are .376 and .539 respectively.[25] This cross section may be substituted in Eq. (7) and the resulting expression for the stepwise excitation rate becomes

$$\beta N_e = 1.31 \cdot 10^{-5} N_e \frac{f_{ik}}{(kT_e)^{3/2}} \left[\frac{\exp - \frac{E_n}{kT_e}}{\frac{E_n}{kT_e}} - \left\{ - E_i \left(- \frac{E_n}{kT_e} \right) \right\} \right] \tag{11}$$

where $- E_i(- E_n/kT_e)$ is the exponential integral and E_n, kT_e, are in ev. The Maxwellian velocity distribution assumption is more accurate in this case because the relatively low value of E_n will make the

central portion of the energy distribution more important.

The diffusion coefficients for 2^1S and 2^3S helium in unexcited helium may be obrained from the measurements of Phelps.[26] Phelps also measured the diffusion coefficients of $1s_5$ neon in unexcited helium and $1s_5$ neon in unexcited neon.[27] The values thus obtained were .52 (cm^2sec^{-1}), .43, .82 and .20 respectively, at 300°K and 760 torr. These coefficients are inversely proportional to pressure, but their temperature dependence will be more complicated. The diffusion of excited atoms in a gas is somewhat smaller than that of unexcited atoms in the gas due to the longer range interaction potential. Based on the Leonard-Jones potential considrable success has been achieved in the accurate calculation of diffusion coefficients of unexcited atoms over a wide range of temperatures.[28] The temperature dependence of the diffusion coefficient for excited atoms may be estimated from dimensional considerations that determine the temperature variation once the potential interaction for a given temperature is specified.[27] In the region of interest the potential interaction may be evaluated from Hirschfelder's tabulated integrals[28] and a slightly longer range of interaction potential may then be assumed for the diffusion of excited states. At constant density these considerations yield a temperature dependence for the diffusion coefficient that varies approximately as $T^{3/4}$ for both exicted helium and excited neon in the temperature range 300-500°K. The inverse diffusion coefficient for the helium and neon mixtures considered may be obtained by linearly weighting the individual inverse diffusion coefficients by their respective fractional partial pressures. Data does not appear to be available for the diffusion of excited helium

in neon but the actual value for this diffusion coefficient will be slightly smaller than the value of 1.06 ($cm^2 sec^{-1}$) for pure helium in neon.[27] For the particular high ratio He-Ne mixture considered here this contribution may be ignored.

The resonant transfer destruction cross sections for the 2^1S and 2^3S helium states involving collisions with neutral neon atoms have been measured experimentally.[29] The total destructive rates by resonant transfer, at 400°K, for the He 2^1S and He 2^3S states are

$$Z_{1r} = 2.4 \cdot 10^6 \, p_{Ne} \quad , \tag{12}$$

$$Z_{2r} = 1.6 \cdot 10^5 \, p_{Ne} \quad , \tag{13}$$

where p_{Ne} refers to the original filling pressure at 300°K. The reverse rates may be related to the forward rates and are given by[5]

$$Z'_{1r} = 9.4 \cdot 10^6 \, p_{He} \quad , \tag{14}$$

$$Z'_{2r} \simeq .51 \cdot 10^5 \, p_{He} \quad , \tag{15}$$

where there is some uncertainty in the reverse rate for the second process due to the interactions between the $2s_2$ and $2s_3$ levels. The lifetime of the Ne $3s_2$ level may be determined in part by experimental results[30] and in part by theoretical calculations.[31] These results yield $A_1 = 1.3 \cdot 10^{+7}$ as the inverse lifetime of this level. As measured by Bennett[32] the inverse lifetime of the Ne $2s_2$ and $2s_3$ levels are $1.0 \cdot 10^7$ and $.63 \cdot 10^7$ respectively. These lifetimes are 1-2 orders of magnitude greater than the reverse transfer rate, as determined from equation 15. In this situation reverse resonant transfer can clearly be neglected with respect to the A coefficients.

The parameters thus determined may be substituted in Eq. (3) to obtain the nonregenerative He 2^1S and 2^3S populations. The Ne $1s_5$ population may similarly be obtained from Eq. (5). For our particular system the populations of these levels are presented in Table 1.

Under regenerative conditions stimulated emission terms will become the dominant radiative loss mechanism. The net effect will be to replace the term A_1 in Eq. (3) by its much larger stimulated emission counterpart. In the limit the term $A_1/(Z'_{1r} + A_1)$ in Eqs. (3) and (4) will approach unity and the resonant transfer process will become unidirectional. The result of regenerative action will be to increase the net transfer rate into the Ne $3s_2$ level and decrease the population of both the $3s_2$ level and the He 2^1S level. The He 2^3S level will be essentially unaffected by 1.15 μ regenerative operation as the term $A_2/(Z'_{2r} + A_2)$ is approximately unity in any case. The regenerative populations are also presented in Table 1.

IV. EVALUATION OF FREQUENCY SHIFTING

The current dependent refractive index changes at a given wavelength may be determined from the relation[33]

$$\frac{\partial n}{\partial I} = \frac{e^2}{2\pi mc^2} \sum_{i=0}^{\infty} \sum_{k=i+1}^{\infty} \frac{\lambda^2 \lambda_{ik}^2}{\lambda^2 - \lambda_{ik}^2} \frac{\partial N_i}{\partial I} f_{ik} \quad , \qquad (16)$$

where λ is the wavelength, f_{ik} is the oscillator strength of the transition and N_i is the population of the lower level. In this expression the population of the upper level is assumed to be negligible in comparison with that of the lower level.

For the 1.15 μ transition the current dependent refractive index changes due to the 1.08 μ absorptive transition may be evaluated from Eq. (16) and substituted into the usual Fabry-Perot relation to obtain the following shift in the interferometer frequency

$$\frac{\partial f}{\partial I} = - 6.6 \cdot 10^{-7} \frac{\partial N(2^3S)}{\partial I} \qquad (17)$$

In this expression we have assumed that the plasma occupies the entire length of the cavity; this is nearly the case for our particular geometry. The current dependent population changes may be obtained from Table 1 and the results may be substituted in Eq. (17) to obtain the values plotted in Fig. 3. The experimental results are also indicated in Fig. 3 and the two curves appear to be surprisingly good agreement, considering the approximate nature of these calculations.

In a similar way the refractive index changes at 6328 Å due to the dispersive effects of the 6334 Å and 6402 Å absorptive transitions, may be evaluated to obtain

$$\frac{\partial f}{\partial f} = 5.1 \cdot 10^{-6} \frac{N(1s_5)}{\partial I} \qquad (18)$$

where minor contributions from the other 1s-2p transitions have been ignored; these contributions, in any case, cancel if equilibrium populations are assumed.

The results of Table 1 (regenerative) together with Eq. (18) may then be used to obtain the theoretical curves of Fig. 4. The curves indicate that the excitation dependence of the $1s_5$ level population

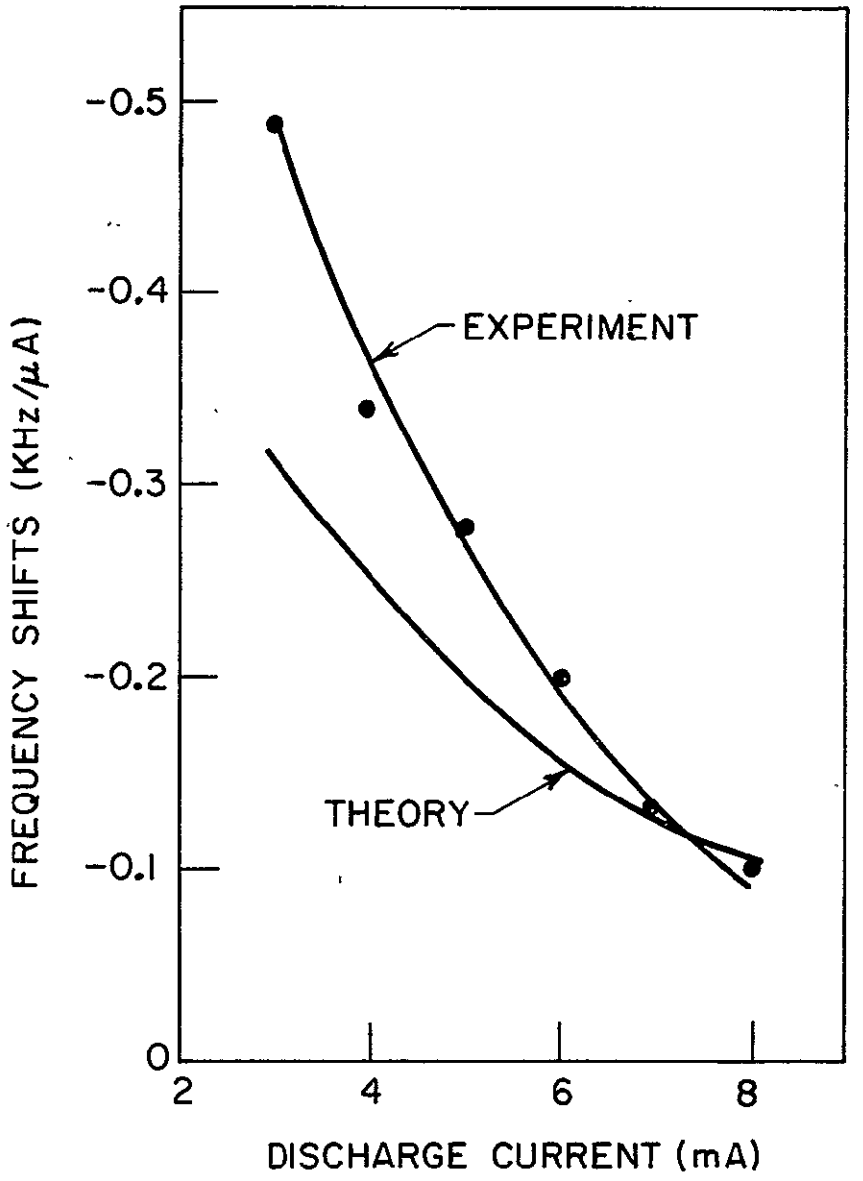


FIG. 3--Comparison of the theoretical and experimental values of the excitation induced off resonant dispersive shifts at 1.15 μ due to the 1.08 μ He 2^3S - 2^3P transition.

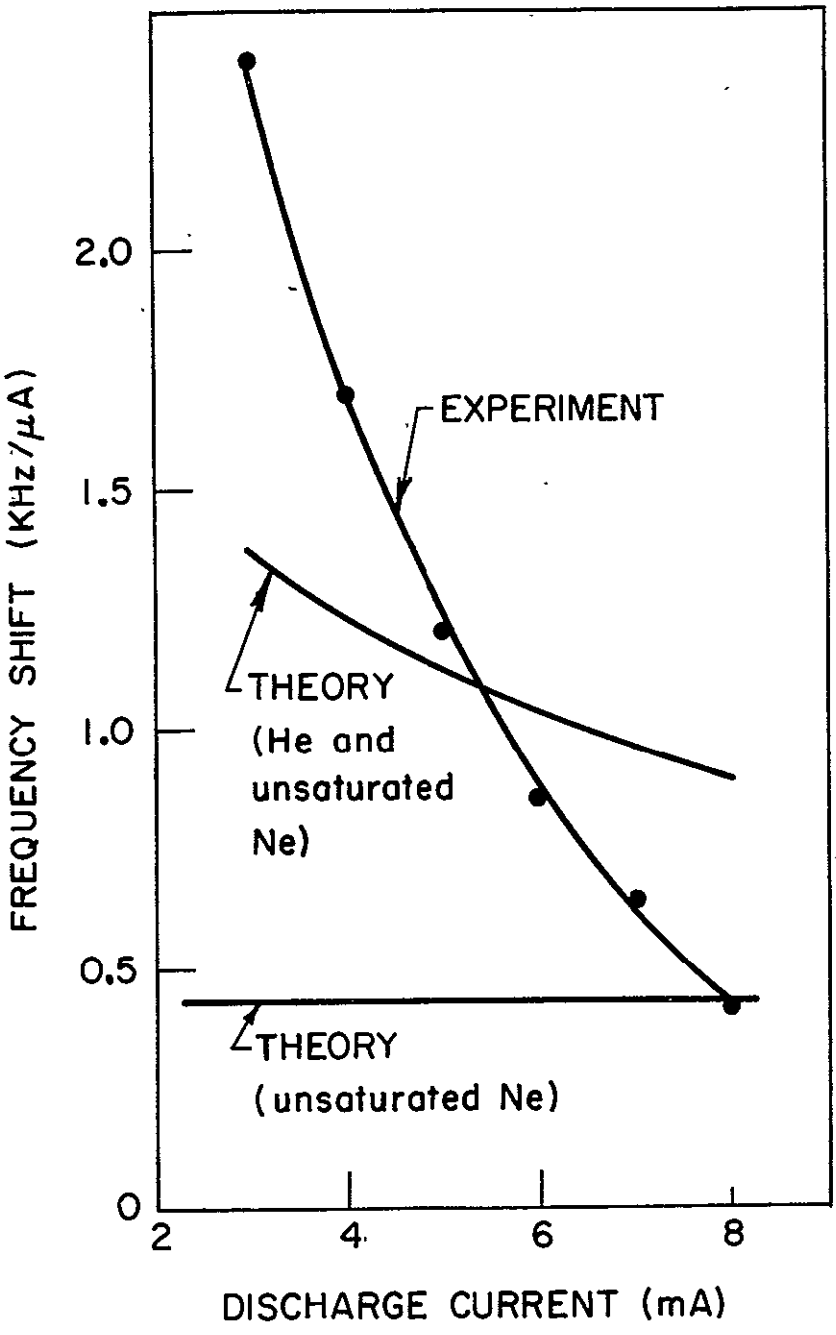


FIG. 4--Excitation induced frequency shifts at 6328 Å due to the Ne 6334 Å and 6402 Å transitions. The unsaturated neon value refers to the shifting effect in the absence of helium resonant transfer contributions to the Ne $1s_5$ level population.

and of the subsequent shifting effects is primarily determined by the helium in the He-Ne mixture. The experimental agreement with theory is somewhat poorer than that obtained at 1.15 μ ; however, qualitative considerations indicate that if other cascade contributions to the Ne $1s_5$ level population are considered and if $1s_5$ level saturation terms are included, the agreement will prove considerably.

Estimates of accuracy of the calculations for the He 2^1S population evaluation can be made by utilizing these techniques for comparison with other experimental data.[4,5] On this basis we conservatively estimate that accuracies within a factor of two can be obtained. Our own results indicate that the accuracy of the He 2^1S population calculations is probably at least this good. The Ne $1s_5$ level population is not as accurate, partly because of the formulation of Eq. (5) and partly because any determination of the opulation of this level is based on the previous determination of helium populations.

The experimental results, in conjunction with the population calculations, indicate that off resonance dispersive effects can be very important in He-Ne lasers. These effects are particularly significant at 6328 Å . If the constant pd scaling consideration[5]

$$\frac{N_e}{p} \propto \frac{I}{d} = \text{constant} \qquad (19)$$

is substituted in Eq. (3) together with the fact that both the diffusion rate and the resonant transfer rates are proportional to pressure at constant pd, it can be seen that the population of the He 2^3S level and the pushing effect at 1.15μ is inversely proportional to the tube radius. Similarly E_1 , E_2 and $\alpha_3 N_e^0 N_n$ in Eq. (5) will be proportional to $1/d^2$, so that the pushing effect at 6328 Å will also vary inversely with tube diameter.

V. CONCLUSIONS

If the lower state population is ignored the gain for both the 3.39 μ and 6328 Å transitions may be estimated from the calculated $3s_2$ populations. Gains thus obtained are only of order of magnitude accuracy and are too approximate for practical application. It appears that a determination of gain based on simple empirical relationships[7] is more convenient and more accurate at this time. The proposed model tends to overestimate the $3s_2$ population because the assumed Maxwellian velocity distribution does not decrease sharply enough at higher energies. The use of a more realistic velocity distribution, particularly for the He 2^1S level which has a slowly peaking excitation cross section, should considerable increase the usefulness of this model.

Order of magnitude calculations based on the $1s_5$ level population indicate that trapping of radiation from the $2p_4$ level to the $1s$ levels is negligible. Since both the 1.15 μ transition and the 6328 Å transition terminate on the $2p_4$ level this indicates that radiation trapping is not responsible for increasing the lower state population of the lasing transition.

The excitation induced shifting effects were most pronounced at 6328 Å . Residual plasma striations or perturbations in the discharge acting through off resonance dispersive mechanisms, may be responsible for limiting the spectral purity of presently available 6328 Å He-Ne lasers. For constant pressure-diameter product conditions such plasma effects will be substantially reduced by the use of larger bore tubes operated at higher currents.

ACKNOWLEDGEMENT

The author wishes to thank Professor A.E. Siegman and Dr. S. Koutsoyannis of Stanford University for much helpful and stimulating discussion.

REFERENCES

1. R. Arrathoon and A. E. Siegman, "Current Pushing of the Oscillation Frequency of a 6328 Å He-Ne Laser", Appl. Phys. Lett., 13, 197 (1968).

2. R. Ladenburg, "Dispersion in Electrically Excited Gases" Rev. Mod. Phys., 5, 243 (1933).

3. Yu. M. Kagan, R. Lyagushchenko and A. D. Khakhauen, "The Excitation of Inert Gases in the Positive Column of a Discharge at Medium Pressures I," Opt. Spectry, 14, 317 (1963).

4. A. D. White and E. I. Gordon, "Excitation Mechanisms and Current Dependence of Population Inversion in He-Ne Lasers", Appl. Phys. Lett., 3, 197 (1963).

5. E. I. Gordon and A. D. White "Similarity Laws for the Effects of Pressure and Discharge Diameter on Gain of He-Ne Lasers", Appl. Phys. Lett., 3, 199 (1963).

6. Labuda, E.F. and E. I. Gordon, "Microwave Determination of Average Electron Energy and Density in He-Ne Discharges", J. Appl. Phys., 35, 1647 (1964).

7. P. W. Smith, "On the Optimum Geometry of a 6328 Å Laser Oscillator", IEEE J. Quantum Electronics, Vol. QE-2, 77 (1966).

8. S.J.B. Corrigan and A. von Engel, "The Excitation of Helium by Electrons of Low Energy", Proc. Phys. Soc. Lond., 72, 786 (1958).

9. T. Holstein, "Imprisonment of Resonance Radiation in Gases, II", Phys. Rev., 83, 1159 (1951).

10. C.G.B. Garrett, Gas Lasers, (McGraw-Hill, New York 1967).

11. C.K.N. Patel, W.L. Faust, R.A. McFarlane, and C.G.B. Garrett, "Laser Action up to 57.355 μ in Gaseous Discharges (Ne, He-Ne)," Appl. Phys. Lett., 4, 18 (1964).

12. H.S.W. Massey and E.H.S. Burhop, Electronic and Ionic Impact Phenomena (Clarendon Press, Oxford, 1952).

13. H.S.W. Massey, "Excitation and Ionization of Atoms by Electron Impact" Handbuch der Physik (Springer-Verlag, Berlin, 1956), Vol. 36, 307.

14. L.F. Vellikok, A.E. Fotiadi and S.A. Fridrikhov, "Plasma-Optical Effects in Lasers at 0.63, 1.15 and 3.39 μ", Soviet Physics-Tech. Phys., Vol. 12, 811 (1967).

15. J.T. Massey, A.G. Schulz, B.F. Hochheimer and S.M. Cannon, "Resonant Energy Transfer Studies in a Helium-Neon Gas Discharge", J. Appl. Phys., 36, 658 (1965).

16. Yu. M. Kagan, R. Lyagushchenko, "The Excitation of Inert Gases in the Positive Column of a Discharge at Medium Pressures IV", Opt. Spectry, 17, 90 (1964).

17. A.V. Phelps, "Diffusion, De-excitation and Three Body Collision Coefficients for Excited Neon Atoms", Phys. Rev., 114, 1011 (1959).

18. R.T. Young, "Calculation of Average Electron Energies In He-Ne Discharges", J. Appl. Phys., 36, 2324 (1965).

19. James Dillon Cobine, Gaseous Conductors (Dover Publications Inc., New York, 1958).

20. J.C. Bowe, "Drift Velocity of Electrons in Nitrogen, Helium, Neon, Argon, Krypton, and Xenon", Phys. Rev., 117, 1411 (1960).

21. J.Y. Wada and Hans Heil, "Electron Energy Spectra in Neon, Xenon and Helium-Neon Laser Discharges", IEEE J. Quant. Elect., QE-1, 327 (1965).

22. C.E. Webb, "A New Technique for Meas. of Radial Dist. of Excit. Species in Plasmas and its Appl. to Cap. Discharges in Argon", Bell Tel. Labs, unpublished memorandum.

23. H.S.W. Massey, F.R.S. and B.L. Moiseiwitsch, "The Application of Variational Methods to Atomic Scattering Problems", Proceed. of Royal Soc., 227, 38 (1954).

24. Ya. B. Zel'dovich and Yu. P. Raizer, Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena, (Academic Press, New York, 1966).

25. W.L. Wiese, M.W. Smith and B.M. Glennon, Atomic Transition Probabilities (Hydrogen Through Neon), (National Bureau of Standards, Washington, D.C., 1966), Vol. 1.

26. A.V. Phelps, "Absorption Studies of Helium Metastable Atoms and Molecules", Phys. Rev., 99, 1307 (1955).

27. D.R. Bates, Atomic and Molecular Processes (Academic Press, New York 1962).

28. Hirschfelder, Curtiss, Bird, Molecular Theory of Gases and Liquids (John Wiley and Sons, Inc., U.S.A., 1965).

29. E.E. Benton, E.E. Ferguson, F.A. Matsen and W.W. Robertson, "Cross Sections for the De-excitation of Helium Metastable Atoms by Collisions with Atoms", Phys. Rev., 128, 206 (1962).

30. Th. Hansch and P. Toschek, "Measurement of Neon Atomic Level Parameters by Laser Differential Spectrometry", Phys. Lett., 20, 273 (1966).

31. Peter W. Murphy, "Transition Probabilities in the Spectra of Ne I, Ar I, and Kr I", J. Opt. Soc. America, 58, 1200 (1968).

32. W.R. Bennett, "Radiative Lifetimes and Collision Transfer Cross Sections of Excited Atomic States", Advances in Quantum Electronics (Columbia University Press, New York, 1961), pp. 28-43.

33. A.C.G. Mitchell and M.W. Zemansky, Resonance Radiation and Excited Atoms (Cambridge University Press, New York, 1934).

APPENDIX V

THE ANTENNA PROPERTIES OF OPTICAL HETERODYNE RECEIVERS

by

A. E. Siegman

Reprinted from APPLIED OPTICS, Vol. 5, page 1588, October 1966

The Antenna Properties of Optical Heterodyne Receivers

A. E. SIEGMAN

Abstract—An optical heterodyne receiver is, in effect, both a receiver and an antenna. As an antenna it has an effective aperture or capture cross section $A_R(\Omega)$ for plane wave signals arriving from any direction Ω. The wavefront alignment between signal and local-oscillator (LO) beams required for effective optical heterodyning may be summarized in the "antenna theorem" $\iint A_R(\Omega)d\Omega = [\overline{\eta^2}/\bar{\eta}^2]\lambda^2$ where the moments of the quantum efficiency η are evaluated over the photosensitive surface. Thus, an optical heterodyne having effective aperture A_R for signals arriving within a single main antenna lobe or field of view of solid angle Ω_R is limited by the constraint $A_R\Omega_R \approx \lambda^2$. Optical elements placed in the signal and/or LO beam paths can vary the trade-off between A_R and Ω_R but cannot change their product.

It is also noted that an optical heterodyne is an insensitive detector for thermal radiation, since a thermal source filling the receiver's field of view must have a temperature $T \approx [\ln(1+\bar{\eta})]^{-1} hf/k$ to be detected with S/N≈1. Optical heterodyning can be useful in practical situations, however, for detecting Doppler shifts in coherent light scattered by liquids, gases, or small particles. Another antenna theorem applicable to this problem says that in a scattering experiment the received power will be $\lesssim N\sigma\lambda/4\pi$ times the transmitted power, where N is the density of scatterers and σ is the total scattering cross section of a single scatterer. The equality sign is obtained only when a single aperture serves as both transmitting and receiving aperture, or when two separate apertures are optimally focused at short range onto a common volume.

Introduction

THE FUNDAMENTAL virtues and weaknesses of optical heterodyning as a coherent optical detection method are by now fairly well understood [1], [2]. The signal and noise properties of optical heterodyne and homodyne receivers have been debated [3], [4] and reasonably well verified experimentally [5]. The necessity for a stable single-frequency local oscillator has been pointed out [6]. The stringent alignment tolerances necessary to keep signal and local-oscillator wavefronts in phase over the photosensitive surface have also been pointed out [2], [7], [8], [9], and prescriptions for apparently relaxing these alignment tolerances have recently been offered [10], [11].

In discussing the directional characteristics of optical heterodyne detection, there is a simple and useful point of view based on antenna theory [2], [12] which does not seem to be as widely appreciated as it might be. Therefore, one main purpose of this paper is to point out that any optical heterodyne receiver is, in essence, both a *receiver* and an *antenna*. As an antenna, it can be characterized by the *effective aperture* or *capture cross section* that it presents to a signal plane wave arriving from any specified direction. Moreover, and more important, we will prove that, in common with any antenna at any frequency, the effective aperture integrated over all possible arrival directions (all solid angle) is essentially λ^2, the signal wavelength squared. As a consequence, there is an inescapable inverse trade-off between the directional tolerance or the angular field of view of an optical heterodyne receiver, and its effective aperture or capture area, with their product limited by the quantity λ^2. For optical wavelengths λ^2 is, of course, an uncomfortably small quantity. However, there does not seem to be any way to get around this limitation—relaxed directional alignment requirements will inevitably bring a penalty of reduced cross section or receiving aperture.

In addition to this main point, we also wish to make two secondary points concerning optical heterodyne characteristics. One point is simply to note that optical heterodyne reception is virtually useless, or at least is very insensitive, for detecting the incoherent thermal radiation emitted by any purely thermal source. The optical heterodyne cannot detect less than about one photon per resolving time, in a single spatial mode. Thermal sources at reasonable temperatures emit much less than this, at least at optical frequencies.

The other point has to do with the heterodyne reception of light scattered by a volume distribution of scatterers. This is a topic which has become of considerable interest recently in connection with optical heterodyne measurements of laser light scattered from small particles, gases, or liquids [13], [14], [15]. Optical heterodyning allows one to measure the Doppler shift and/or the Doppler broadening of this scattered light, and thereby to measure the flow rate, flow gradient, and/or various internal fluctuations such as critical opalescence in the scattering medium.

The main point to be made here is that there is another antenna therorem (possibly a novel one [16]) that seems to apply in this case, which limits the received power from the scattering medium to approximately $N\sigma\lambda/4\pi$ times the transmitted power, where N is the volume density of scattering centers and σ is the average scattering cross section of a single scatterer. This quantity appears to be a fundamental upper bound on the returned power in any elementary scattering experiment using coherent detection.

Manuscript received March 28, 1966. This work was supported by the U. S. Army Electronics Command, Contract DA-28-043-AMC-00446(E). The scattering analysis was originally developed as part of a consulting assignment supported by the Optics Department, Sylvania Electronic System—West, Mountain View, Calif.

The author is with the Department of Electrical Engineering, Stanford University, Stanford, Calif.

Fortunately, this quantity, although small, is not so small as to rule out many interesting experiments, at least in the laboratory [20].

The Antenna Theorem for Optical Heterodyning

As illustrated in Fig. 1, we consider a square-law photodevice which is illuminated by a local-oscillator (LO) wave of complex[1] scalar amplitude $\tilde{u}_0(x, y)$ and optical frequency ω_0, together with a signal wave of complex scalar amplitude $\tilde{u}_1(x, y)$ and optical frequency ω_1. For simplicity, the reference plane $z=0$ is taken as some plane near the photodevice, where the signal and LO waves have already been combined. The z direction is preferably taken along a main receiving lobe or direction of maximum sensitivity for the optical heterodyne. It is not at all necessary that the reference plane coincide with the photodevice surface, provided only that the variations in optical path length from the reference plane to the photodevice surface are small compared to the *difference* or *intermediate-frequency* wavelength corresponding to the difference frequency $(\omega_1-\omega_0)$. Transit time considerations in any practical photoelectric or photoconductive device will ensure that this condition is met.

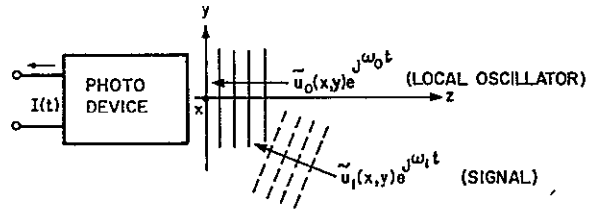


Fig. 1. General schematic of an optical heterodyne receiver.

For simplicity we also eliminate polarization effects by assuming that $\tilde{u}_0$ and $\tilde{u}_1$ refer to the same polarization. In general, if the quantum efficiency of the photodevice is isotropic in polarization, then any incoming signal can be resolved into polarization components "parallel" (in a general sense) and orthogonal to the local oscillator's polarization. The "parallel" component will then photomix with the local oscillator, and the orthogonal component will not. Still more generally, if the photoresponse (quantum efficiency) of the photodevice is anisotropic in polarization, then for any given LO polarization the signal will still be resolvable into components "parallel" and orthogonal (in a general sense) to the vector product of a photoresponse tensor times the LO polarization. We suppose here that $\tilde{u}_1$ contains only the "parallel" polarization, since the orthogonal component will elicit no response. Note that with an anisotropic quantum efficiency two optical waves of orthogonal spatial polarization *can* beat together [12], [17], [18]. Also, in either isotropic or anisotropic cases if a signal wave is randomly polarized with a uniform distribution, half of the signal intensity will always remain undetected.

[1] Any quantity having a superscript tilde will, in general, be complex.

The scalar amplitudes are assumed to be normalized so that the incremental photocurrent, produced at the photodevice's output terminals by the complex scalar intensity $\tilde{u}(x, y, t)$ falling on an incremental area dA on the reference plane, will be

$$dI(t) = \eta(x, y) \left| \tilde{u}(x, y, t) \right|^2 dA$$

or

$$I(t) = \int\int \eta(x, y) \left| \tilde{u}(x, y, t) \right|^2 dA,$$

where $\eta(x, y)$ is the quantum efficiency for light striking the point (x, y). In the elementary optical heterodyne case, the total scalar intensity is

$$\tilde{u}(x, y, t) = \tilde{u}_0(x, y)e^{j\omega_0 t} + \tilde{u}_1(x, y)e^{j\omega_1 t}.$$

Hence the total photocurrent is

$$I(t) = I_0 + I_1 + \tfrac{1}{2}\left[\tilde{I}_{10}e^{j(\omega_1-\omega_0)t} + \tilde{I}_{10}^{*}e^{-j(\omega_1-\omega_0)t}\right],$$

where the dc currents I_0 and I_1 induced by the local oscillator and the signal are given by

$$I_0 = \int\int \eta(x, y) \left| \tilde{u}_0(x, y) \right|^2 dA$$

$$I_1 = \int\int \eta(x, y) \left| \tilde{u}_1(x, y) \right|^2 dA,$$

and where the complex phasor amplitude of the difference-frequency photocurrent is given by

$$\tfrac{1}{2}\tilde{I}_{10} = \int\int \eta(x, y)\tilde{u}_1(x, y)\tilde{u}_0^{*}(x, y)dA.$$

If $\tilde{u}_0(x, y)$ and $\tilde{u}_1(x, y)$ have the same spatial variation (in magnitude and phase) over the reference plane, this reduces to the optimum photomixing situation

$$\left| \tilde{I}_{10} \right|^2 = 4I_0I_1$$

which gives maximum sensitivity [3], [4].

Suppose now that the incident signal wave is a uniform plane wave with wave vector $k_1=(k_x, k_y, k_z)$, so that

$$\tilde{u}_1(x, y, z) = \tilde{c}_1e^{-jk \cdot r}.$$

On the reference plane this gives

$$\tilde{u}_1(x, y, z = 0) = \tilde{c}_1e^{-j(k_xx+k_yy)}.$$

If the signal wave vector has magnitude $k_1=\left| k_1 \right| =\omega_1/c$, and if the direction of arrival of this wave is along spherical coordinates θ and ϕ in the coordinate system of Fig. 1 (where $\theta=0$ is along the z axis and ϕ is measured from the x axis in the x, y plane), then

$$k_x = k_1 \sin\theta \cos\phi$$

$$k_y = k_1 \sin\theta \sin\phi.$$

The difference-frequency photocurrent is then given by

$$\tfrac{1}{2}\tilde{I}_{10} = \tilde{c}_1 \int\int \eta(x, y)u_0^{*}(x, y)e^{-j(k_xx+k_yy)}dxdy.$$

Now we define a mean quantum efficiency $\bar{\eta}$ given by[2]

$$\bar{\eta} \equiv \frac{\iint \eta(x, y) \mid u_0(x, y) \mid^2 dxdy}{\iint \mid u_0(x, y) \mid^2 dxdy}.$$

Using this definition and the previous equation, we choose to write the magnitude squared of the IF photocurrent in the form

$$\mid I_{10} \mid^2 = 4I_0[\bar{\eta} \mid \bar{e}_1 \mid^2 A_R].$$

The product $\bar{\eta}\mid \bar{e}_1\mid^2$ in this expression is the dc photocurrent density, i.e., the dc photocurrent that would be produced by a plane wave of complex amplitude $\bar{e}_1$ falling on a photodevice of unit area and quantum efficiency $\bar{\eta}$. The whole quantity in the square brackets is the effective signal current that mixes with the LO dc current to produce an IF photocurrent. This effective signal current will be in general less than the actual signal dc current, which will be $\eta\mid \bar{e}_1\mid^2$ times the actual photocathode area. The quantity A_R has the dimensions of an area, and in fact may be interpreted as the effective receiver area or receiver aperture of the optical heterodyne receiver in receiving a plane wave signal by mixing with the LO wave.

This effective receiver aperture A_R is closely analogous to the effective aperture of an ordinary radio wave antenna. It is a function of the vector arrival direction Ω of the signal wave and, from the previous three equations, is given by

$$A_R(k_x, k_y) = A_R(\Omega) = \frac{\left|\iint \eta(x, y) u_0^*(x, y) e^{-j(k_x x + k_y y)} dxdy\right|^2}{\bar{\eta}^2 \iint \mid u_0(x, y) \mid^2 dxdy}.$$

From the form of this expression, the effective aperture $A_R(k_x, k_y)$ is proportional to the power spectral density in the space frequencies k_x, k_y of the spatial distribution $\eta(x, y)u_0^*(x, y)$, i.e., the LO complex amplitude weighted by the quantum efficiency.

We now wish to evaluate the integrated value of $A_R(\Omega)$ over all possible directions of arrival Ω. The differential of solid angle can be written

$$d\Omega = \sin\theta d\theta d\phi.$$

But we can also note that, from the relations above, we can write

$$dk_x dk_y = k_1^2 \cos\theta \sin\theta d\theta d\phi = k_1^2 \cos\theta d\Omega.$$

As mentioned earlier, we suppose that an optical heterodyne receiver will usually have only one (or at most a few) dominant (and reasonably narrow) main receiving lobe(s) pointed along the $\theta=0$ or z axis, so that over the range of angles for which $A_R(\Omega)$ is appreciable we can make the approximations $\cos\theta \approx 1$ and $d\Omega \approx k_1^{-2} dk_x dk_y$. We can then write

$$\iint_{\text{all solid angle}} A_R(\Omega) d\Omega \approx (1/k_1^2) \int\int_{-\infty}^{\infty} A_R(k_x, k_y) dk_x dk_y.$$

But, from the interpretation of $A_R(k_x, k_y)$ as a spatial-frequency power density we can easily find that[3]

$$\frac{1}{k_1^2} \iint A_R(k_x, k_y) dk_x dk_y = \left(\frac{2\pi}{k_1}\right)^2 \frac{\iint \eta^2(x, y) \mid u_0(x, y) \mid^2 dxdy}{\bar{\eta}^2 \iint \mid u_0(x, y) \mid^2 dxdy}$$

or

$$\iint A_R(\Omega) d\Omega = \frac{\overline{\eta^2}}{\bar{\eta}^2} \lambda^2,$$

where $\overline{\eta^2}$ is implicitly defined in the previous equation.

If the quantum efficiency is uniform over the photodevice, then the ratio $\overline{\eta^2}/\bar{\eta}^2$ is strictly unity; and in any reasonable case this ratio is not likely to be far from unity. We finally conclude, therefore, that the integrated effective aperture of an optical heterodyne receiver is given to a good approximation by the simple relation

$$\iint A_R(\Omega) d\Omega \approx \lambda^2$$

If the antenna has a single main receiving lobe that subtends a solid angular field of view of Ω_R steradians, with an effective aperture of A_R for sources inside this field of view and zero outside, then [2], [12]

$$A_R \Omega_R \approx \lambda^2.$$

These relationships are, of course, identical with the well-known antenna theorem for the integrated cross section or integrated gain of a single-mode antenna at any frequency. A major problem is the extremely small size of the wavelength at optical frequencies.

Examples

The simplest illustration of this antenna theorem is a photosurface of physical area A (diameter d) uniformly illuminated by a uniform plane-wave LO beam, as shown in

[2] We choose this definition for $\bar{\eta}$ because, with this definition, the minimum detectable signal power with an optimum spatial variation of $\bar{e}_1(x, y)$ will be just $\bar{\eta}^{-1}$ photons per reciprocal bandwidth (based on $S/N=1$, shot-noise-limited).

[3] We are basically using the Wiener-Khintchine relation for Fourier transform pairs.

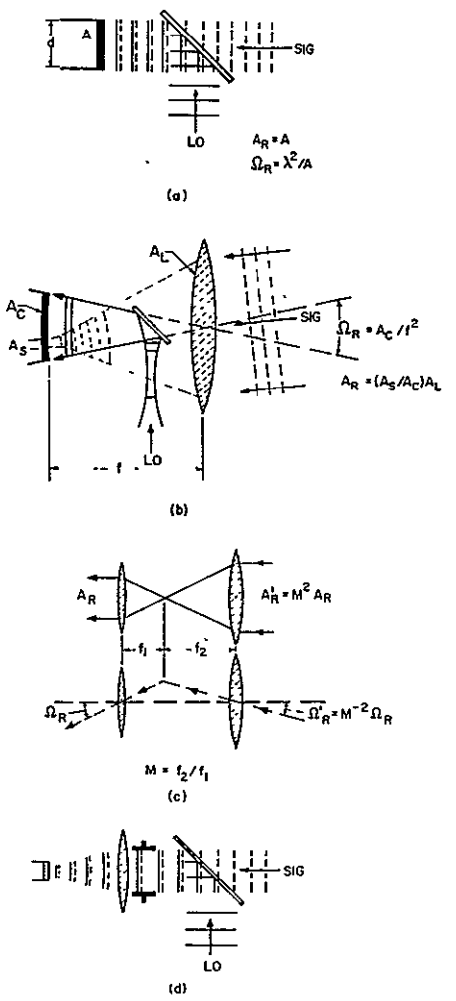


Fig. 2. Several sketches illustrating various concepts in optical heterodyne antenna theory.

Fig. 2(a). The effective aperture for a signal plane wave parallel to the LO plane wave is clearly just $A_R = A$. If the phase variation between signal and LO waves is not to exceed one wavelength across the aperture, the wavefront must not be tilted from parallelism by more than $\Delta\theta \approx \lambda/d$. Therefore the incoming signal wave must be confined to a cone of solid angle given by $\Omega_R \approx (\Delta\theta)^2 \approx \lambda^2/A$. Therefore, we have

$$A_R\Omega_R \approx A \times \left(\frac{\lambda^2}{A}\right) \approx \lambda^2,$$

in agreement with the theorem.

Figure 2(b) illustrates an arrangement that might at first seem to overcome the antenna theorem limitation. The signal wave is collected and focused by a lens of area A_L and focal length f onto a cathode of area A_C located along the focal surface of the lens. The signal spot size at the focal plane is A_S. The LO beam diverges as a spherical wave of proper curvature so as to uniformly illuminate the focal-plane cathode at normal incidence.

The matching of wavefronts in the overlap region of signal and LO beams can be made as ideal as desired, at least in principle, for a signal spot anywhere on the cathode surface. Therefore, the solid angular field of view is given by

$$\Omega_R = \frac{A_C}{f^2}.$$

At the same time it would appear that the effective aperture can be made indefinitely large simply by increasing the lens area A_L. The restriction in the present case comes about, however, because the lens area A_L and the focal spot size A_S are interrelated. For ideal optics the diffraction limit of the lens is

$$\frac{\lambda^2}{A_L} \approx \frac{A_S}{f^2}.$$

The LO dc current I_0 and the related shot noise are determined by the total LO intensity over the entire cathode. However, the signal light mixes only with the fraction A_S/A_C of the local-oscillator light lying within the focal spot. In effect, therefore, the conversion gain of the heterodyne process is reduced by this fraction. The effective aperture is, therefore, given by

$$A_R = \left(\frac{A_S}{A_C}\right) \times A_L.$$

The net result is

$$A_R\Omega_R \approx \left(\frac{A_C}{f^2}\right) \times \left(\frac{A_S A_L}{A_C}\right) = \frac{A_S A_L}{f^2} \approx \lambda^2,$$

which again agrees with the antenna theorem.[4]

We close this portion of the discussion with three general remarks.

1) First, as Fig. 2(c) illustrates, an M-power telescope ($M = f_2/f_1$) placed in the signal beam only, in front of an optical heterodyne, will magnify the apparent area of the cathode by M^2; but at the same time the angular sensitivity will be increased by M^2. Hence with such a telescope

$$A_R'\Omega_R' \approx (M^2 A_R) \times \left(\frac{1}{M^2}\Omega_R\right) \approx \lambda^2.$$

In general, optical elements of any sort placed in the signal beam before it enters the optical heterodyne will change both the effective aperture A_R and the angular field of view Ω_R, but always subject to the basic constraint that $A_R\Omega_R \approx \lambda^2$.

2) Optical elements of any sort placed in the signal and LO beams *after* they are combined will in general not alter

[4] References [10] and [11] give an experimental demonstration of essentially this scheme, in conjunction with an image dissector photodevice that passes only the photocurrent from the focal plane spot A_S. Their final system also automatically tracks the movements of the spot A_S. This system does not violate the antenna theorem, nonetheless, because its *instantaneous* field of view is still very narrow, so as to allow a relatively large aperture. The resulting critical alignment is made more convenient by the electronic and servo-controlled alignment method, but the alignment is still critical. Also a large waste of LO power is inherent in the scheme as used.

either A_R or Ω_R (provided that the total light intensity continues to strike the photosensitive element). Figure 2(d) illustrates this for a very simple example.

3) The integral expression given above for $A_R(k_x, k_y)$ in terms of $\eta(x, y)\tilde{u}_0(x, y)$ has exactly the form that would result from using Huygen's principle together with the Fraunhofer approximation to evaluate the intensity diffracted into a direction k_1 from a source amplitude distribution $\eta(x, y)\tilde{u}_0(x, y)$. Therefore, the following is a prescription for determining the antenna pattern of an optical heterodyne [12]. *Consider the complex LO amplitude distribution falling on the photodevice surface (weighted by the quantum efficiency distribution if necessary). Reverse the direction of propagation of this LO distribution and allow the reversed wavefront to propagate back out through any optical elements that an incident signal wave would traverse. The resulting far-field or Fraunhofer diffraction pattern will be the antenna pattern of the optical heterodyne receiver* [19].

Optical Heterodyne Detection of Scattered Light

As noted in the Introduction, one application where optical heterodyning has proven very fruitful is in making remarkably detailed measurements of the Doppler shift and/or the Doppler broadening of laser light scattered from gases, dust particles, and other scattering elements [13], [14], [15]. Therefore, we will next show that there appears to be a very simple and general expression for the maximum signal return to be expected in this type of experiment.

We consider an experimental situation such as that shown in Fig. 3, where the transmitter is a laser, and the receiver is an optical heterodyne with the usual field of view Ω_R and receiving aperture A_R. The transmitter and receiver beams intersect in a common volume which is presumed to contain a density N per unit volume of scatters. Each scatterer is assumed to have a total scattering cross section σ; for simplicity the scattering is assumed to be isotropic. Only single scattering is considered.

For simplicity we assume rectangular transmitter and receiver beams, as shown, so that the scattering volume common to the two beams is a parallelopiped with sides a, b, c. Then, the power density in the transmitter beam in the scattering region is $P_{\text{trans}}/ab \sin\theta$, and the total power scattered by an individual scatterer is $(\sigma/ab \sin\theta)P_{\text{trans}}$. Of this scattered power, a fraction $A_R/4\pi R^2$ will be intercepted by the optical heterodyne. The total number of scatterers in the common volume is $Nabc \sin\theta$, and the receiver field of view is equal to $\Omega_R = bc \sin\theta/R^2$. By combining all of these factors, we obtain

$$P_{\text{received}} = (Nabc\sin\theta)\left(\frac{A_R}{4\pi R^2}\right)\left(\frac{\sigma}{ab\sin\theta}\right)P_{\text{trans}}$$

$$= \frac{\lambda}{b\sin\theta}\,\frac{N\sigma\lambda}{4\pi}P_{\text{trans}}.$$

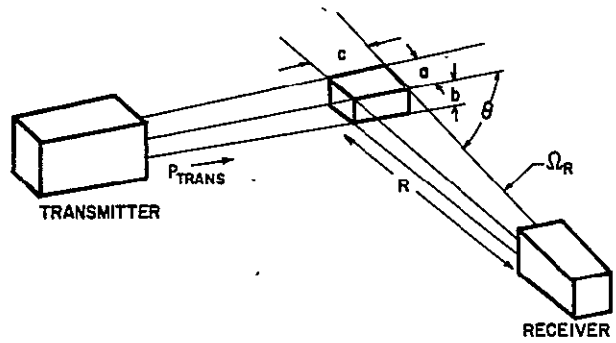


Fig. 3. Geometry for the analysis of heterodyne detection of scattered light.

If the transmitter and receiver are both focused on the scattering region to give a minimum value of b, the result with an $f/1$ lens will be $b \approx \lambda$. Moreover, the factor $(\sin\theta)^{-1}$ does not allow any large improvements, particularly in the focused case, because the region of beam overlap will be limited by the depth of focus. Therefore, by setting $\lambda/b \sin\theta \approx 1$ we are led to the simple limiting expression [20]

$$\frac{P_{\text{received}}}{P_{\text{trans}}} \lesssim \frac{N\sigma\lambda}{4\pi}.$$

This same result can be obtained directly by analyzing the situation in which transmitter and receiver are focused onto a common small volume.

This conclusion is reinforced if we consider as an alternative the case of a single transmitting and receiving aperture which we assume to be imbedded in the scattering medium. The geometry for this case is shown in Fig. 4. In this case the transmitted power density at a range r in the far field of the aperture (assuming the system to be focused at infinity) is $P_{\text{trans}}/r^2\Omega_R$. The power scattered by one scatterer is $(\sigma/r^2\Omega_R)P_{\text{trans}}$, and the fraction of this which is detected is $A_R/4\pi r^2$. The number of scatterers in a shell of thickness dr at range r is $N\Omega_R r^2 dr$. Therefore, we obtain for the received power from this shell

$$dP_{\text{received}} = (N\Omega_R r^2 dr)\left(\frac{A_R}{4\pi r^2}\right)\left(\frac{\sigma}{r^2\Omega_R}\right)P_{\text{trans}}$$

$$= \frac{N\sigma A_R}{4\pi}\,\frac{dr}{r^2}P_{\text{trans}}.$$

This expression holds only for ranges r in the far field of the aperture, beginning at about one "Rayleigh range" r_0 from the aperture. The "Rayleigh range" of an aperture is defined as the distance where the diffraction spread $r_0^2\Omega_R$ from the aperture is approximately equal to the original aperture size A_R, or

$$r_0 \approx \sqrt{\frac{A_R}{\Omega_R}} \approx \frac{A_R}{\lambda}.$$

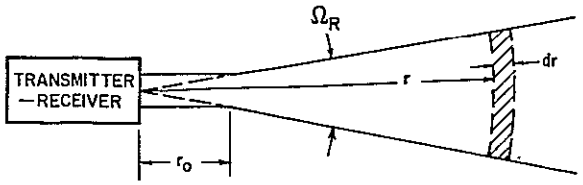


Fig. 4. Heterodyne detection of scattered light using the same aperture for transmitting and receiving.

Now, scatterers located closer to the aperture than about one Rayleigh range are "out of focus." That is, they scatter spherical waves which do not become sufficiently planar at the receiving aperture to be properly detected. Moreover, inside the Rayleigh range—that is, in the Fresnel region—the transmitted power density does not continue to rise as $1/r^2$, but rather varies between zero and the value at about one Rayleigh range.

It seems a reasonable approximation, therefore, to neglect those scatterers located inside the Rayleigh range, and to sum up the contributions only from shells located at and beyond the Rayleigh range. The result is

$$\frac{P_{\text{received}}}{P_{\text{trans}}} = \frac{N\sigma A_R}{4\pi}\int_{r=r_0}^{\infty}\frac{dr}{r^2} = \frac{N\sigma\lambda}{4\pi}.$$

This single-antenna result is the same as the limiting value of the two-antenna analysis just given, and is independent of the optical heterodyne's aperture or field of view. The same result is also obtained for a single antenna focused at any arbitrary range in the scattering volume [20].

This limiting formula, derived only by these approximate arguments, appears to be a general limitation on scattering experiments [19]. The expression given is useful in evaluating the minimum scattering density[5] $N\sigma$ that can be detected with a given transmitter power and optical heterodyne receiver. Fortunately, there are situations of much practical interest [13], [14], [15] in which very useful scattering measurements can be made taking advantage of the unique capabilities of the optical heterodyne receiver for detecting Doppler shifts and Doppler broadening of the scattered light.

Heterodyne Detection of Thermal Radiation

One question that seems to arise frequently is the utility of the optical heterodyne receiver for detecting optical radiation from incoherent thermal sources. This question is basic, for example, in determining whether optical heterodyne detection has applications in astronomy or radiometry. We will next give a simple argument to show that optical heterodyne detection is a relatively insensitive detection method for an incoherent thermal source.

Consider an optical heterodyne receiver of effective aperture A_R whose field of view Ω_R is filled by a blackbody thermal source at temperature T. It simplifies the derivation without altering the results if the source is assumed to be planar and located at a range R in the far-field region of the receiver. The total power ΔP emitted into a hemisphere, within frequency range Δf, from a small area ΔA on a blackbody is

$$\Delta P = \frac{2\pi}{\lambda^2}\frac{hf\Delta f}{e^{hf/kT}-1}\Delta A.$$

Since the power per steradian emitted normal to the surface is $1/\pi$ times this, while the solid angle subtended by the receiver aperture as seen from the blackbody is A_R/R^2, only a fraction $A_R/\pi R^2$ of the blackbody radiation is received by the optical heterodyne; and only $\frac{1}{2}$ of that on the average is properly polarized so as to be detected. Since emission from different points on a thermal source is incoherent, we can simply add the power contributed by each small area ΔA within the total area $R^2\Omega_R$ on the surface that lies within the field of view of the optical heterodyne receiver.

The net result is that the average signal power received by an optical heterodyne from a thermal source filling the field of view is

$$P_{\text{sig}} = \left(\frac{1}{2}\right)\left(\frac{A_R}{\pi R^2}\right)\left(\frac{2\pi}{\lambda^2}\frac{hfdf}{e^{hf/kT}-1}\right)(R^2\Omega_R) = \frac{hfdf}{e^{hf/kT}-1}.$$

But, the equivalent noise input to an optical heterodyne receiver, due to LO shot noise, is well known to be

$$P_{\text{noise}} = (1/\bar{\eta})hfdf.$$

Hence, the signal-to-noise ratio for optical heterodyne detection of the thermal source is

$$\frac{S}{N} \equiv \frac{P_{\text{sig}}}{P_{\text{noise}}} = \frac{\bar{\eta}}{e^{hf/kT}-1}.$$

For the thermal source to cause a signal-to-noise ratio of unity or better, the minimum source temperature must be

$$T \geq \frac{hf}{k}\frac{1}{\ln(1+\bar{\eta})} \approx \frac{hf}{k}\frac{1}{\bar{\eta}}, \qquad \bar{\eta} \ll 1.$$

For frequencies in the visible light range, the quantity hf/k is about 25 000°K. Therefore, it would require a source $1/\bar{\eta}$ times this hot, filling the receiver field of view, to cause the noise power level in the heterodyne IF channel to double. To put this in another way, an optical heterodyne receiver pointed directly at the sun (6000°K in the visible) would show virtually no additional signal.

Detection capabilities for thermal targets can, of course, be considerably increased by using chopping and time-integrating techniques, and other radiometric methods. Also, the relative sensitivity or the minimum detectable source temperature for an optical heterodyne improves directly with increasing wavelength as one goes

[5] In scattering from index variations in liquids, $N\sigma$ corresponds directly to the turbidity.

into the infrared. Nonetheless, it seems clear that the optical heterodyne method does not offer high sensitivity for detecting purely thermal radiation.[6]

Conclusions

The optical heterodyne receiver is usefully viewed as both a receiver and an antenna. As the latter, it shares, with all other antennas, an unavoidable reciprocal tradeoff between effective aperture A_R and angular field of view or angular alignment tolerance Ω_R; the product $A_R\Omega_R \approx \lambda^2$. The heterodyne receiver is relatively insensitive for detecting incoherent thermal sources. It is effective for detecting Doppler shifts, especially in scattered coherent laser light. There is a simple limiting expression which says that the scattered light received by an optical heterodyne cannot exceed $N\sigma\lambda/4\pi$ times the transmitted light, where N is the density of single scatterers and σ is the total scattering cross section of a single scatterer.

References

[1] S. Jacobs, "The optical heterodyne," *Electronics*, vol. 36(28), p. 29, July 12, 1963.

[2] A. E. Siegman, S. E. Harris, and B. J. McMurtry, "Optical heterodyning and optical demodulation at microwave frequencies," in *Optical Masers*, J. Fox, Ed. New York: Polytechnic Press and Wiley, 1963, p. 511.

[3] B. M. Oliver, "Signal-to-noise ratios in photoelectric mixing," *Proc. IRE (Correspondence)*, vol. 49, pp. 1960–1961, December 1961.

[6] The basic difficulty is, of course, that thermal radiation is distributed over all frequencies and all spatial modes; while an optical heterodyne can detect only a single spatial mode and only frequencies within its narrow receiving bandwidth. An optical heterodyne can detect at best $\overline{\eta}^{-1}$ photons per unit bandwidth in its single spatial mode; but a thermal source emits only $[\exp(hf/kT)-1]^{-1}$ photons per unit bandwidth into each separate spatial mode. By contrast, a photodetector used in incoherent fashion can receive energy over a spectral range very much broader than its modulation frequency bandwidth; and it can receive (but not resolve) a large number of spatial modes. Hence, incoherent detection is inherently much better suited to the detection of thermal radiation. Much the same point has been made by P. Connes, "High-resolution interferometric spectroscopy," in *Quantum Electronics and Coherent Light*, C. H. Townes and P. A. Miles, Eds. New York: Academic, 1964, p. 207.

[4] H. A. Haus, C. H. Townes, and B. M. Oliver, "Comments on 'Noise in photoelectric mixing,'" *Proc. IRE (Correspondence)*, vol. 50, pp. 1544–1545, June 1962.

[5] S. Jacobs and P. Rabinowitz, "Optical heterodyning with a CW gaseous laser," in *Quantum Electronics III*, P. Grivet and N. Bloembergen, Eds. New York: Columbia University Press 1964, p. 481.

[6] P. Rabinowitz, J. LaTourrette, and G. Gould, "AFC optical heterodyne detector," *Proc. IEEE (Correspondence)*, vol. 51, pp. 857–858, May 1963.

[7] V. J. Corcoran, "Directional characteristics in optical heterodyne detection processes," *J. Appl. Phys.*, vol. 36, p. 1819, June 1965.

[8] R. D. Kroger, "Motion sensing by optical heterodyne Doppler detection from diffuse surfaces," *Proc. IEEE (Correspondence)* vol. 53, pp. 211–212, February 1965.

[9] G. Gould et al., "Coherent detection of light scattered from a diffusely reflecting surface," *Appl. Opt.*, vol. 3, p. 648, May 1964.

[10] W. S. Read and D. L. Fried, "Optical heterodyning with noncritical angular alignment," *Proc. IEEE (Correspondence)*, vol. 51, p. 1787, December 1963.

[11] W. S. Read and R. G. Turner, "Tracking heterodyne detection," *Appl. Opt.*, vol. 4, p. 1570, December 1965.

[12] A. E. Siegman, "Detection and demodulation of laser beams," presented at the 1964 International Symposium, IEEE Microwave Theory and Techniques Group, New York, N. Y.

[13] H. Z. Cummins, N. Knable, and Y. Yeh, "Observation of diffusion broadening of Rayleigh scattered light," *Phys. Rev. Lett.*, vol. 12, p. 150, February 10, 1964.

[14] Y. Yeh and H. Z. Cummins, "Localized fluid flow measurements with an He-Ne laser spectrometer," *Appl. Phys. Lett.*, vol. 4, p. 176 May 15, 1964.

[15] J. W. Foreman, Jr., E. W. George, and R. D. Lewis, "Measurement of localized flow velocities in gases with a laser Doppler flow meter," *Appl. Phys. Lett.*, vol. 7, p. 77, August 5, 1965.

[16] A. E. Siegman, "Spatially coherent detection of scattered radiation," submitted for publication, *IEEE Trans. on Antennas and Propagation.*

[17] R. L. Smith, "Theory of photoelectric mixing at a metal surface," *Appl. Opt.*, vol. 3, p. 709, June 1964.

[18] A. J. Bahr, "The effect of polarization selectivity on optical mixing in photoelectric surfaces," *Proc. IEEE (Correspondence)*, vol. 53, p. 513, May 1965.

[19] This same point has been noted by R. E. Brooks, TRW Space Technology Labs, Redondo Beach, Calif., unpublished memorandum.

[20] *Note added in proof:* More recent calculations indicate that the limit seems to be given by $N\sigma\lambda/4$, i.e., a factor of π larger than stated in the text. This "theorem" has yet to be proved in general, but an exact calculation for a single transmit-receive antenna with a Gaussian beam pattern leads to exactly $N\sigma\lambda/4$.