Several easily implemented devices for doing ultrasensitive optical measurements with noisy lasers are presented. They are all-electronic noise cancellation circuits that largely eliminate excess laser intensity noise as a source of measurement error and are widely applicable. Shot-noise-limited optical measurements can now easily be made at baseband with noisy lasers. These circuits are especially useful in situations where strong intermodulation effects exist, such as current-tuned diode laser spectroscopy. These inexpensive devices (parts cost ~$10) can be optimized for particular applications such as wideband or differential measurements. Although they cannot eliminate phase noise effects, they can reduce amplitude noise by 55–70 dB or more, even in unattended operation, and usually achieve the shot-noise limit. With 1-Hz signal-to-noise ratios of 150–160 dB, they allow performance equal or superior to a complex heterodyne system in many cases, while using much simpler dual-beam or homodyne approaches. Although these devices are related to earlier differential and ratiometric techniques, their noise cancellation performance is much better. They work well at modulation frequencies from dc to several megahertz and should be extensible to ~100 MHz. The circuits work by subtracting photocurrents directly, with feedback applied outside the signal path to continuously adjust the subtraction for perfect balance; thus the excess noise and spurious modulation ideally cancel at all frequencies, leaving only the shot noise. The noise cancellation bandwidth is independent of the feedback bandwidth; it depends only on the speeds of the photodiodes and of the bipolar junction transistors used. Two noise-cancelled outputs are available; one is a high-pass filtered voltage proportional to the signal photocurrent and the other is a low-pass filtered voltage related to the log ratio of the signal and comparison photocurrents. For reasonable current densities, the noise floors of the outputs depend only on the shot noise of the signal beam. Four variations on the basic circuit are presented: low noise floor, high cancellation, differential high power, and ratio-only. Emphasis is placed on the detailed operation and design considerations, especially performance extension by compensation of the nonideal character of system components. Experience has shown that some applications advice is required by most users, so that is provided as well. © 1997 Optical Society of America

1. Introduction

Excess noise, spurious modulation, and power drift in lasers are common problems in optical measurements. The noise and spurious signals are usually worst at low modulation frequencies and can easily reach 50 dB above shot noise—even quite far from dc—which makes ultrasensitive optical measurements difficult. A great deal of work has been done on this problem, both to make lasers quieter and to circumvent the worst effects of the remaining noise.

Most precise optical measurement systems apply some sort of modulation to the beam to make their output signals periodic in time at a frequency as far from the low-frequency noise as necessary; these are then detected in bandwidths sufficiently narrow to exclude most of the excess noise. Typical examples are heterodyne interferometers,1 frequency modulation methods,2,3 beam-chopping systems with lock-in detection,4 and fast-scanning systems with signal averaging. (In this paper, excess noise refers both to true noise above the shot-noise level and to spurious modulation of the beam intensity, due, for example, to baseband mode beats or power supply ripple.)

Optical feedback stabilization of several kinds5,6 and all-electronic noise rejection schemes such as differential detection7–10 (subtraction) and normalization by analog division11 are used extensively; but when used unaided, they are limited by their noise floors, bandwidths, or finicky adjustment requirements to relatively low-performance applications. Of all the techniques listed, none results in measurements at the shot-noise limit except heterodyne interferometry.

Chopping techniques are too slow to escape the excess noise region, although as they move the mea-
measurements limited only by the shot noise of the signal in gas lasers, mode hopping, instability caused by adjustment of the balance. In addition to high levels of lidar and other measurements with better performance, and simpler apparatus than hitherto possible.

In this paper is presented design considerations, performance limitations, and several generalizations, including the important case of differential measurements. It concludes with detailed schematic diagrams and a discussion of the several versions, and a mathematical Appendix.

2. Laser Noise

Much of the work of designing an ultrasensitive optical apparatus involves circumventing the effects of laser noise. Laser noise arises in a great variety of ways, so an exhaustive list is difficult to give. A few of the more annoying types are baseband mode beats in gas lasers, mode hopping, instability caused by unintended optical feedback into the laser cavity, wiggle noise, and electrically or mechanically induced modulation from such sources as cooling water turbulence and power supply ripple.

Some laser applications, such as heating and pumping nonlinear elements or other lasers, require the beam itself to be stable in amplitude because it is driving the system. Others, such as the highest-resolution spectroscopy, require low noise in both phase and amplitude because both contribute to the spectral width of the laser line (the frequency modulation noise is generally far worse). In these cases, there is no alternative but to stabilize the laser as much as possible.

In most other measurements, photocurrent fluctuations resulting from laser amplitude noise are the primary nuisance. Both the signal photocurrent being measured and any static background that may be present are proportional to the laser power, and so intensity fluctuations translate directly into measurement noise. The laser-noise-induced fluctuations of the background appear as additive noise on top of the signal, while those of the signal itself appear as multiplicative noise, causing noise intermodulation, the impressing of noise sidebands on the desired signal.

In a measurement of a small shift on a large background, the fluctuations of the background are the worst problem, so experimenters usually aim to make zero background measurements, which minimize noise and drift from these fluctuations. In most such cases, the shot-noise contribution is not reduced by going to the zero background methods, so they would lose some of their attractiveness if truly quiet lasers were available. Exceptions to this rule include fluorescence, photochemical grating, and other dark-field measurements. These have other disadvantages, however, especially in their requirement for photomultiplier tubes, which have poor quantum efficiency and therefore lower SNR. Noise intermodulation is not addressed by zero background techniques, nor by most of the other available noise reduction methods.

A. Feedback Stabilization

Drift and low-frequency noise in lasers are often dealt with by using feedback stabilization techniques, by using external modulators or injection current modulation. These techniques work fairly well, but with the exception of current feedback with diode lasers, they tend to be complicated or expensive.

In measurements not involving squeezed light, the best-case noise floor of a photocurrent is set by its shot-noise current, given by

\[ i_n = \sqrt{2eB}, \]

where \( e \) is the electron charge and \( B \) is the measurement bandwidth.

All the noise suppression methods that follow rely on a comparison of one or more signal beams to a comparison beam. (The term comparison beam is better than the common term reference beam and than my earlier term sample beam, because it avoids confusion in interferometric and spectroscopic applications.)

In any technique that stabilizes one noisy signal by reference to another, the uncorrelated parts of the noise (i.e., the shot-noise contributions) ideally obey

\[ \frac{1}{\text{SNR}_{\text{out}}} = \frac{1}{\text{SNR}_{\text{signal}}} + \frac{1}{\text{SNR}_{\text{comp}}}. \]
where SNR\text{signal}, SNR\text{comp}, and SNR\text{out} are, respectively, the power signal-to-noise ratios of the signal, comparison photocurrents, and combined photocurrents. The effects of finite bandwidth and technical noise in the combining circuitry are ignored. Their power noise-to-signal ratios add, at least for noise frequencies well within their bandwidths. Thus such methods can at best bring the SNR of the resulting beam (or photocurrent) to the signal-to-shot-noise ratio of the comparison beam; if the two photocurrents are equal, then the noise of the output is 3 dB above that of the signal beam alone.

In feedback stabilizers, SNR\text{signal} is the signal-to-shot-noise ratio of the output beam with the stabilizer loop opened, and SNR\text{out} is that of the output beam with the stabilizer operating. Most users dislike losing half or more of their laser power in a stabilizer, so the comparison beam is usually chosen to be appreciably weaker than the main one; thus it contains relatively more shot noise, and this contributes to a high-noise floor.

Besides the noise floor problem, feedback stabilizers are intrinsically rather narrowband. For example, if 40-dB noise suppression is needed to reach the shot-noise floor, the feedback system must remove 99% of the noise current, so its closed-loop gain error must be less than 1%; assuming an open-loop unity gain bandwidth of 0.5 MHz and a 1-pole rolloff, this occurs at 5 kHz. Feedback stabilization is worthwhile for many kinds of optical systems but is incapable of reaching the shot-noise limit with noisy lasers.

B. All-Electronic Noise Suppression

All-electronic schemes have been known for some time as well and have been widely used. As can be seen from the block diagram in Fig. 1, these schemes differ from those discussed above in that no attempt is made to stabilize the laser beam itself, only the photocurrent at the output of the apparatus. The beam is sampled at the laser and detected to obtain a comparison photocurrent with the same fractional excess noise as that from the signal beam (which represents the output of the optical system). The two are then combined in some fashion to obtain an output current that, ideally, is completely free of excess noise.

Such methods rely on two important properties of most optical systems, namely, extremely wide temporal bandwidth and the use of highly linear photodetectors, such as photoconductive-mode photodiodes. The wide bandwidth guarantees that (if path delays are small) the optical system will not apply an gain variation or phase shift to the modulation of the beam—the instantaneous, fractional, excess amplitude noise of the comparison beam will be identical with that of the signal beam—so that if the cancellation is done properly, the noise suppression should be essentially perfect. The linearity of the photodiodes allows excellent cancellation performance with unmatched diodes, even if they are running at appreciably different current densities. With such fortunate circumstances, all-electronic noise cancellation should be extremely effective.

In the past, two all-electronic methods have been common, namely subtraction and division. In a subtractive noise rejection scheme, the comparison photocurrent is subtracted from the signal current. If the optical system is adjusted perfectly so that the two photocurrents are exactly equal, the excess noise and dc cancel, leaving only the signal and shot noise. Subtracters can have wide bandwidths, since the photocurrents can be subtracted directly (without prior conversion to voltages, which introduces poorly controlled phase shifts) and because the noise cancellation does not depend on feedback. The improvement is seldom more than 20 dB, because the adjustment required is finicky and because the intensity of the signal beam often varies somewhat during a measurement (e.g., from scanning), so that the currents cannot be exactly equal at all times. In addition, since the shot-noise currents of the signal and comparison photocurrents statistically are independent, they obey Eq. (2) and thus limit the system noise floor to 3 dB above the shot noise of the signal beam alone.

Dividers avoid the requirement for precise adjustment by dividing the signal current by the comparison current, hence canceling the fractional (rather than absolute) noise deviations. They have great attractions in theory, because dividing out the instantaneous intensity provides compensation for drift and noise intermodulation as well as removing additive excess noise. Dividers need not in principle suffer from the 3-dB additional noise problem of subtracters, since the comparison beam can be made stronger than the signal beam (and so relatively quieter). Unfortunately, dividers tend to be slow, so that the suppression bandwidth is limited and noisy, as suggested by the following example.

The Burr-Brown MPY-634 is neither the fastest nor the quietest divider available, but has a good compromise of noise, speed, and accuracy for this application. With a full-scale (10-V) denominator (the best case), it has a noise spectral density of approximately 1 \( \mu \text{V}/\sqrt{\text{Hz}} \) with a zero numerator and 2 \( \mu \text{V}/\sqrt{\text{Hz}} \) with a 10-V numerator; this is approximately 60 dB worse than the best operational amplifiers. Since its maximum input level is 10 V, its maximum SNR is 134 dB in a 1-Hz bandwidth. With appropriately chosen current-to-voltage con-
verter (transresistance amplifier) gain, and if the signal level is reduced by 3 dB to avoid clipping, this is equivalent to the signal-to-shot-noise ratio of a photocurrent of 8 \(\mu\)A. With a red helium neon laser and a silicon PIN diode (responsivity 0.3 A/W), that implies that a 27-\(\mu\)W laser beam can be quieted to 3 dB above the shot noise (granted a noiseless comparison beam), but a 3-mW beam, well within the linear operating regime of many photodiodes, can be quieted only to 20 dB above the shot noise, a poor performance.

Because of their use of feedback in the signal path and the frequency compensation problems alluded to above, dividers have poor high-frequency performance; swept sine measurements (0.2-V p-p, plus 7-V dc applied to both numerator and denominator, which were connected in parallel) on the unit mentioned above reveal 42-dB small-signal suppression at 100 Hz, deteriorating steadily above about 5 kHz, to 22 dB at 100 kHz, and virtually zero at 1 MHz.

Subtracters eliminate the additive excess noise, but can do nothing to suppress noise intermodulation; dividers in principle eliminate both. In high-accuracy (as opposed to merely high-dynamic-range) applications, where the signal is well above the additive noise floor and must be measured precisely, noise intermodulation limits the measurement accuracy to the SNR of the laser beam. For this reason, open-loop subtracters are most suitable for measurements of small changes in intensity with lasers that are already reasonably quiet.

For some purposes, then, dividers and open-loop subtracters can be very useful, but in general using them unaided does not allow shot-noise-limited measurements with noisy lasers.

3. Laser Noise Canceller

A. Principle

The open-loop subtracter can be improved substantially in many areas by the use of negative feedback outside the signal path, to continuously adjust the matching between the two photocurrents. By applying this idea, simple systems can be made that achieve the shot-noise limit without adjustments with most lasers; these systems are the subject of the remainder of this paper.

Feedback control requires an electronically variable current splitter that does not degrade the strict proportionality between the signal and comparison photocurrents and an electronic criterion for when the circuit is in balance. Figure 2 shows such a splitter. The comparison beam is made somewhat stronger than the signal beam, and the comparison photocurrent is split between two paths by using a bipolar transistor pair \(Q_1/Q_2\). The ratio of the collector currents of \(Q_1\) and \(Q_2\) is controlled by the difference \(\Delta V_{BE} = V_{BE2} - V_{BE1}\) in their base emitter voltages. In the Ebers–Moll model

\[
\frac{i_{C2}}{i_{C1}} = \exp\left(\frac{e\Delta V_{BE}}{kT}\right).
\] (3)

The BJT is unique in that this ratio does not depend on the value of \(i_{comp}\); its transconductance \(g_m = \partial i_C/\partial V_{BE}\) is proportional to its collector current, which makes BJT differential pairs highly linear as current splitters. Thus fluctuations in \(i_{comp}\) split in exactly the same way as the dc, so the differential pair will not degrade the noise cancellation, within the limits set by the BJT gain bandwidth product \(f_T\). Devices with good log conformity exhibit constant splitting ratios over several decades of collector current. To use this circuit, we make sure \(i_{comp}\) is somewhat larger than \(i_{signal}\) and adjust \(\Delta V_{BE}\) to dump the extra to ground by way of \(Q_1\). Used alone, this adjustment would still be finicky and manual, but it provides the electronic control needed.

The balance criterion is simple; since the excess noise splits just as the dc, ideal noise cancellation occurs when the total dc photocurrent into the summing junction of transresistance amplifier \(A_1\) is zero.

The combination of these two ideas yields the basic noise canceler, shown in Fig. 3, which uses negative feedback to keep the splitting ratio adjusted exactly. This circuit combines simplicity, freedom from adjustments, and good performance. Besides the differential pair \(Q_1/Q_2\) and transresistance amplifier \(A_1\), there is an integrating servo amplifier \(A_2\) that adjusts \(\Delta V_{BE}\) to force the dc output of \(A_1\) (and hence the dc current into \(A_1\)'s summing junction) to be zero, thus ensuring proper balance. Cascode transistor \(Q_3\) prevents the capacitance of the signal photodiode from loading the summing junction, which may allow as much as 40 times improvement in the amplifier bandwidth compared with a straightforward transresistance amplifier with the same feedback resistance. A small capacitor in parallel with \(R_1\) may help in controlling any high-frequency gain peaking in \(A_1\).

The feedback loop can be as fast or as slow as desired because the bandwidth of effective cancellation does not depend on the feedback bandwidth \(f_c\), only on the \(f_T\) of the transistors; the loop just tinkers with the adjustment. Because feedback is keeping the average voltage at the \(A_1\)'s output zero, \(A_1\)'s out-
output is a high-pass filtered version of the signal photocurrent, minus the excess noise; its corner frequency is \( f_c \).

Since the instantaneous excess noise fluctuations of the photocurrent are exactly proportional to their dc level, as discussed above, application of negative feedback to one of the transistor bases to keep the net photocurrent at zero results in essentially perfect noise cancellation out to very high frequencies, regardless of the bandwidth of the feedback loop. As an added benefit, the feedback voltage offers an alternative output; \( \Delta V_{BE} \) is related to the ratio of the comparison current to the signal current (within the feedback bandwidth) by Eq. (3),

\[
\Delta V_{BE} = -\frac{kT}{e} \ln \left( \frac{i_{comp}}{i_{signal}} - 1 \right). \tag{4}
\]

It is shown in Appendix A that if the emitter current of a differential pair has full shot noise, both collector currents have exactly full shot noise also, independent of the splitting ratio; thus the expected noise level of the output of \( A_2 \) is 3 dB above the shot noise of the signal current alone:

\[
V_{NA1} = R_1 \sqrt{4eB_i_{signal}}, \tag{5}
\]

where \( B \) is the bandwidth and \( R_1 \) is the feedback resistor of \( A_1 \). If the comparison beam contains signal information, as in a differential measurement, the SNR of the measurement can achieve the shot-noise level, but otherwise it is limited to 3 dB worse.

The noise at the output of \( A_2 \) is obtained by multiplying the total shot-noise current by the partial derivative of \( V_{A2} \) with respect to \( i_{signal} \):

\[
V_{NA2} = \frac{1}{\gamma} \left( \frac{4e}{\left( 1 - \frac{i_{signal}}{i_{comp}} \right)} \right),
\]

\[
= \frac{2kT}{\gamma \sqrt{e \gamma V_{A2} / kT}}, \tag{6}
\]

where \( \gamma \) is the voltage divider gain (0.025 here).

The voltage divider on the base of \( Q_2 \) performs three functions: It reduces the swing of \( A_2 \) so that the base-collector junction cannot be Zenered or forward biased enough to cause servo lockup. The choice of a 40:1 division ratio approximately cancels the factor of \( kT/e \) from Eq. (4), resulting in a scale factor of 2 V near null rather than 50 mV (1% change in \( i_{signal} \) gives a 20-mV change at \( A_2 \)). It also reduces the contribution of the input voltage noise of \( A_2 \) to the total circuit noise. Most opamps have at least several nanovolts per root hertz noise, which is considerably worse than the 1-nV/\( \sqrt{\text{Hz}} \) noise of good transistors; this will degrade the noise floor of the final circuit if the voltage divider is not used. The Thévenin resistance of the divider should be kept small.

If the feedback signal is used as the output, the performance is much like that of a divider in that intermodulation between noise and signal is suppressed (since \( \Delta V_{BE} \) depends only on the ratio of the two photocurrents). One important difference is that the new system does not get noisier as its loop bandwidth is approached, as dividers do; because the dc cancellation guarantees the cancellation of additive noise at all frequencies of interest, only the suppression of noise intermodulation declines. Another way of looking at this is that \( A_2 \) is integrating a signal (\( A_1 \)'s output) whose noise has been canceled at all frequencies, so the additive noise cancellation bandwidth of the log output is independent of the feedback bandwidth. This is a remarkable fact, because it means that the entire bandwidth of the log ratio output is useful for highly sensitive measurements rather than only 1% or so as with dividers and feedback systems. The cancellation of noise intermodulation does depend on the feedback bandwidth.

The log output is especially useful for situations in which the noise intermodulation is strong, such as current-tuned diode laser spectroscopy\(^{1,16}\); there the laser power may vary by as much as 2 or 3 to 1 during a scan, making small absorption peaks inconspicuous on a huge sloping background if this circuit (or some other background reduction technique) is not used. For applications such as this, \( Q_1 \) and \( Q_2 \) should be monolithically matched pairs to ensure that their temperatures are the same and that the dc offset voltage error in \( V_{A2} \) is small; if the log output is not needed, \( Q_1 \) and \( Q_2 \) may be replaced by discrete devices such as 2N3904's, which cost only a few cents.

There is nothing in the system that forces the feedback loop to be slow. The loop bandwidth \( f_c \) is found...
cant low-frequency noise and several peaks in the 1-2-MHz region. QI and Q2 were matched Motorola MRF904's, and D1 and D2 were a Hamamatsu type S1722-01. The signal and comparison beam powers were 5.6 and 7.2 mW, respectively, and the optical powers of 5.6 and 7.0 mW; output was taken from A1, and the feedback resistor Rf was 5.11 kΩ. The large ratio of $i_{C2}/i_{C1}$ (about 4) was chosen to reduce the transconductance of the differential pair, which reduces the extrinsic base resistance noise, as discussed below. The transresistance gain of the circuit is reduced by half here due to the 50-Ω source and load impedances. The expected shot-noise voltage spectral density $v_{ns}$ is 60.8 nV/√Hz from signal current shot noise alone.

In each plot of Fig. 4, the output noise with the canceller running is compared with the noise with the comparison beam blocked, so that $Q_2$ is turned off, and the circuit operates as an ordinary transresistance amplifier. Figure 4(a) is an averaged plot of the input-referred noise spectral density from 0 to 100 kHz, while Fig. 4(b) shows the broadband noise level. The flatband noise spectral density after cancellation is 88 nV/√Hz, which is in excellent agreement with the expected value of 60.8 √2 nV/√Hz ≈ 68.0 nV/√Hz; most of the remainder is contributed by the 29-nV/√Hz noise of the measurement system. This result represents a limiting dynamic range of 154 dB in 1 Hz, although over most of the usable range of $\Delta V_{BE}$, the base-spreading resistance noise will limit its dynamic range to approximately 150 dB in 1 Hz, as already noted. The resulting noise has Gaussian statistics, even far out in the wings of the Gaussian; considered as an imputed error in its variance, the deviation is less than 0.1 dB out to 7σ or more, even with a mode-hopping diode laser.

Figure 5 contains frequency response plots of the ultimate cancellation performance of the prototype. The beam was amplitude modulated by a few parts in 10³ before splitting by the spectrum analyzer's tracking generator driving a Pockels cell, with an analyzer to convert the resulting polarization shift to amplitude modulation. The data are from a spectrum analyzer; one sweep was taken with the canceller operating normally and a second sweep with the comparison beam blocked. The cancellation behavior depends on the ratio of the comparison and signal beam powers. The bandwidth of the amplifier $A_1$ was only 1.4 MHz and that of the feedback less than 100 Hz. The ultimate cancellation was nearly 60 dB to 100 kHz and deteriorated to 40 dB at approximately 1 MHz, primarily owing to the capacitance of the photodiodes (which seems to become worse at high light intensities) and the relatively poor high-frequency behavior of $Q_3$. By

\begin{equation}
 f_c = \frac{e}{2\pi kT} \left[ \frac{i_{signal}}{1 + \exp(e\Delta V_{BE}/kT)} \right] \gamma R_f \frac{RC}{RC'}
\end{equation}

B. Performance

Figure 4 shows the performance attainable with the basic noise canceller of Fig. 3 in a system with photon efficiency (excluding detector quantum efficiency) $\eta \approx 1$. $Q_1$ and $Q_2$ were MRF904's, selected for good differential splitter performance by using a curve tracer. These devices have good log conformity, high speed, and low noise, but have betas of around 30,
changing $Q_3$ to an MM4049, selected for good beta linearity, and reducing the beam power somewhat, the high-frequency situation can be improved a good deal, as shown in Fig. 6. The result in Fig. 6(b), where the cancellation is as great as 40 dB out to a modulation frequency of 10 MHz, is as good as a dominant-pole feedback stabilizer with a bandwidth of 1 GHz.

C. Effects of Nonideal Transistors

Real transistors do not follow the Ebers–Moll model exactly, and their collector currents are not exact replicas of their emitter currents. The most important sources of error are the emitter bulk (extrinsic) resistance $r_E$ and beta nonlinearity.

To achieve perfect cancellation, the collector current of $Q_1$ must be exactly proportional to the comparison photocurrent $i_{\text{comp}}$. The small-signal (differential) current gain of a transistor is

$$h_\rho = \frac{\partial i_C}{\partial i_B},$$

while the large-signal (average) current gain is

$$h_{FE} = \frac{i_C}{i_B}.$$

These two gains in general are not equal, which causes the fluctuations in the comparison photocurrent to split differently between the collector and base than the dc average; this degrades the attainable cancellation. In the absence of other effects, the lower limit imposed by beta nonlinearity on the uncanceled fraction of the noise with a given transistor is equal to the proportion of the emitter current of $Q_2$.
Fig. 7. Calculated limit to cancellation performance that is due to equal 0.5-Ω emitter bulk resistances $r_E$ in $Q_1/Q_2$ as a function of $\Delta V_{BE}$ and $i_{signal}$.

$$A_{min} = \frac{\Delta V_{BE}}{i_{comp}} \cdot \frac{\frac{\partial i_{C_2}}{\partial i_{comp}} - \frac{i_{C_2}}{i_{comp}}}{\frac{\partial i_{C_1}}{\partial i_{comp}} - \frac{i_{C_1}}{i_{comp}}}$$

(11)

In Appendix A, this is shown to be (to leading order in $eir_E/kT$):

$$A_{min} = \frac{e}{kT} \frac{(i_{signal} - i_{sample})}{i_{sample}} \times \frac{r_E}{r_{E_2}}$$

(12)

The limit depends on $\Delta V_{BE}$ and on collector current. Figure 7 shows the effects of $r_E$ as a function of collector current and $\Delta V_{BE}$ for small-signal transistors with $r_E \approx 0.5 \, \Omega$. Figure 8 shows the cancellation performance of the circuit of Fig. 5 at a frequency of 1 kHz for a variety of signal currents. Figure 9 shows the experimental curves for a temperature-controlled MAT-04 transistor array, showing the much more predictable behavior of this device, which is due to its better beta linearity and freedom from temperature drifts.

Besides these limits due to $r_E$ and beta nonlinearity, there is the Johnson noise contributed by the base spreading resistance $r_o$, typically $40-100 \, \Omega$ for small-signal devices, which contributes an ultimate noise floor of approximately $1 \, \text{nV}/\sqrt{\text{Hz}}$ on each base. Near null, this corresponds to a voltage SNR of $(2kT/e)/1.4 \, \text{nV}$ (about 150 dB) in 1 Hz, equivalent to the total system shot noise with a photocurrent of 800 μA (signal beam power of approximately 2.7 mW). This effect can be reduced by operating the circuit with a fairly small comparison beam, which forces $\Delta V_{BE}$ to be positive, reducing the transconductance of $Q_1$ and hence that of the stage; by choosing a tran-
The green Nd:YAG laser beam was sinusoidally modulated at 1 kHz by using an acousto-optical modulator, a variable attenuator, a Glan–Taylor prism to control the polarization, a Wollaston prism to split the beams, and a rotatable Glan–Thompson prism in the comparison beam to adjust the relative beam intensities. All were slightly misaligned to control etalon fringes. The interaction of $r_E$ degeneration and beta nonlinearity make the details of the cancellation performance difficult to predict a priori, but there is a clear trend toward better cancellation at lower currents.

Because the transistor type with low $r_b$ by paralleling devices to reduce the effective value of $r_b$, by reducing the transconductance of the differential pair by diode degeneration, as shown in Subsection 4.A; or by using the differential/high-dynamic-range version of the noise canceller, where $i_{C2}$ is a small fraction of the total signal photocurrent. The Thévenin resistance of the voltage divider used on the base of $Q_2$ also contributes noise, so care should be taken to keep it below $r_b$. Because the transconductance of a BJT increases with collector current, the shot-noise contribution of the diodes needs not be a limitation. In Appendix A, it is shown that with $N$ diodes in series with each emitter, the total noise current contributed by the diodes need not be a limitation. In Appendix A, it is shown that with $N$ diodes in series with each emitter, the total noise current contributed by the diodes need not be a limitation.

$$\langle i_{n2} \rangle^2 = 2eC_2 \left(1 - \frac{N}{N + 1} \frac{i_{C1}}{i_{C1} + i_{C2}} \right).$$

If the circuit is modified as shown in Fig. 10, and the comparison current made five times larger than $i_{signal}$ (so that 80% flows to ground), the shot noise in

![Fig. 8. Measured 1-kHz cancellation performance of the circuit of Fig. 3 as a function of the log ratio output voltage that reflects $i_{comp}/i_{signal}$. The green Nd:YAG laser beam was sinusoidally modulated at 1 kHz by using an acousto-optical modulator, a variable attenuator, a Glan–Taylor prism to control the polarization, a Wollaston prism to split the beams, and a rotatable Glan–Thompson prism in the comparison beam to adjust the relative beam intensities. All were slightly misaligned to control etalon fringes. The interaction of $r_E$ degeneration and beta nonlinearity make the details of the cancellation performance difficult to predict a priori, but there is a clear trend toward better cancellation at lower currents.](image)

![Fig. 9. Cancellation performance at 1 kHz of the circuit of Fig. 3 with a temperature-stabilized MAT-04 matched array used for $Q_1$ and $Q_2$. Excellent beta linearity of this device and elimination of temperature errors result in much more predictable cancellation performance. The deep minima near 0 V are due to symmetric cancellation of $r_E$ degeneration. The sharp deterioration with negative $V_{m}$ is due to the rapidly increasing current in $Q_1$, which increases its $r_E$ nonlinearity.](image)
the final split comparison current is still 4.8 dB less than that in the signal current, and the 3-dB penalty is reduced to 1.25 dB, a 1.75-dB improvement. A splitting ratio of 9:1 and 10 diodes per side reduces the penalty by another half decibel to 0.72 dB.

Figure 11 shows the improvement in performance obtained by using this circuit, with diode-connected MAT-04's used as the emitter diodes. (This is necessary, because real diodes do not obey the diode equation, whereas diode-connected transistors do.) The signal current was 0.189 mA (corresponding to a signal beam power of 0.61 mW) and the ratio $i_{\text{comp}}/i_{\text{signal}}$ was 5.0. The noise improvement is 1.7 dB, which is in excellent agreement with the expected value of 1.75 dB from Eq. (13).

The performance level reached by the basic (Fig. 3) noise canceler is the same as that of an ideal heterodyne system because the 3-dB extra noise resulting from the comparison beam shot noise is balanced by the factor of 2 decrease in bandwidth of homodyne or dual-beam measurements compared with heterodyne ones with the same temporal response. It is superior to that of an ideal chopped system because choppers exhibit 3-dB additional loss owing to a 50% duty cycle. The present variation achieves a noise level 1.7 dB better than the best heterodyne system.

B. Variation 2: High Spurious Suppression

Some lasers exhibit strong spurious signals that are spectrally narrow so that the power spectral density within the spur is very high; examples include power supply ripple and baseband mode beats in gas lasers. In such cases, enough of the spurs may remain after cancellation with the basic circuit to be bothersome. Since the effectiveness of the cancellation is limited by beta nonlinearity and emitter bulk resistance in the differential pair, we seek methods for compensating for their effects.

The emitter bulk resistance $r_E$ of a transistor appears in series with the emitter. Here it applies degenerative feedback to the emitters of $Q_1$ and $Q_2$, causing their transconductances to deviate from the predictions of the Ebers-Moll model. This degeneration can be compensated for by applying degenerative feedback to the bases so that the base-to-intrinsic-emitter bias is restored to the ideal value. The required bias is

$$\Delta V_{\text{comp}} = r_E i_C - r_E i_{\text{comp}},$$

(14)

where the second-order term in $i_E - i_C$ is neglected. We form the correction term as

$$\Delta V_{\text{comp}} = -i_C (r_E i_E + r_E i_{\text{comp}}) + i_{\text{comp}} r_E i_{\text{comp}},$$

(15)

which is implemented as shown in Fig. 12. The $i_{\text{comp}}$ term is obtained directly, the $i_{\text{comp}}$ term from a current mirror (for biasing reasons). Adjustment is accomplished by shunting part of the currents to ground via the potentiometers. As the $r_E$ term is small, the additional shot noise contributed by the current mirror is inconsequential.

Compensating for beta nonlinearity is more diffi-
The replacement of $Q_3$ with a Darlington pair of fast transistors or a field-effect transistor (FET) may help, but $Q_1$ and $Q_2$ cannot be replaced this way. FET differential pairs are ruled out because they do not obey Eq. (3). Darlinitons because of their poor transconductance linearity. The best course is to choose simple BJT devices with high beta and good linearity for $Q_1$ and $Q_2$. The Analog Devices MAT-04 transistor arrays used here are excellent in this regard, although their $f_T$'s are relatively low, especially for collector currents below 0.2 mA. The NEC UPA10x series of transistor arrays are claimed to have extremely high beta linearity, along with gigahertz $f_T$'s, and are under evaluation. If discrete devices are to be used, small-signal PNP devices generally have better beta linearity than NPN's, but significantly worse $r_E$, so better performance at low beam powers may sometimes be obtained by inverting the polarity of the entire circuit. Care must be taken with discretes, however, because even small temperature differences between transistors can limit cancellation performance at this level. Figure 13 shows the improvement in the cancellation performance when $r_E$ compensation is applied to a MAT-04 and demonstrates the possibility of one achieving as much as 28-dB improvement over the basic circuit of Fig. 3 at high current levels. A MPS64 Darlington transistor was used for $Q_2$ to avoid beta nonlinearity effects there; the flatness of the compensated curves suggests that the predominant limitation at this level is the beta linearity of $Q_2$. The improvement declines at lower current levels, as expected, but it is apparent that the low-current performance is not degraded significantly by the $r_E$ compensation circuitry. Because this method is capable of producing upward of 70 dB of cancellation (even with large signals), even minor departures from ideal conditions can degrade the overall system performance. Stray light, vignetting, photodiode response nonuniformity, and incidental etalon fringes can be serious problems. It also requires two trims and is mildly sensitive to the exact value of the photocurrents. Nevertheless, for high-dynamic-range, single-ended measurements with noisy lasers, it provides a level of performance otherwise unobtainable.

C. Variation 3: Ratio Only

In situations such as spectroscopy, in which only a normalized output is desired, it is possible to eliminate $A_1$ entirely to achieve a wider feedback bandwidth and somewhat better suppression on the logarithmic output, as shown in Fig. 14.\textsuperscript{15} The control voltage is now applied to the base of $Q_1$, since the elimination of $A_1$ inverts the sign of the loop gain. The loop bandwidth $f_c$ is as given in Eq. (7), with $R_E = R$ and the sign of $\Delta V_{BE}$ inverted, and the output noise level is given by Eq. (6) with the sign of $\Delta V_{BE}$ changed. Figure 15(a) shows the noise floor of the present version (averaged from 3 to 10 kHz) as a function of $i_{\text{signal}}$ for two values of $\Delta V_{BE}$. The noise floor is within 1 dB of the predicted value, and the limiting 1-Hz SNR near null is indeed approximately

![Graph](image)

**Fig. 13.** Cancellation performance of the high spurious suppression canceller of Fig. 12 (solid curve), under the same conditions as in Fig. 9, as a function of the strength of $i_{c_2}$. The data of Fig. 9 (dashed curve) are provided for comparison. The compensation circuit makes the behavior of the canceller stable over a wide range of signal and comparison currents. (a) $i_{\text{signal}} = 0.142$ mA. The curves are nearly identical because $r_E$ is not a significant limitation yet. (b) $i_{\text{signal}} = 0.665$ mA. The compensated device is several decibels better nearly everywhere and as much as 15 dB near null. (c) $i_{\text{signal}} = 1.26$ mA. The benefits of $r_E$ compensation become evident, with as much as 28-dB improvement. Cancellation is >60 dB almost everywhere.
150 dB, as predicted; the highest power signal corresponds to a noise voltage of 9 nV/√Hz and a scale factor of 220 mV, a 1-Hz SNR of 147 dB. Figure 15(b) shows 0–100-kHz noise floor corresponding to the rightmost point on the lower curve of Fig. 15(a). The noise is very flat and well behaved even in the low audio frequency range. Figure 15(c) shows the frequency response of the ratio-only version for two values of \( i_{\text{signal}} \). Even with beam powers of only a few hundred microwatts, bandwidths of a few megahertz are easily achieved. Figure 15(d) is a swept-sine plot of the cancellation performance of this variation, which shows performance similar to the others.

A disadvantage of the ratio-only variation is that it is no longer possible to obtain an uncanceled (I-V mode) signal simply by blocking the comparison beam; however, a switch can easily be added to change the feedback capacitor for a resistor and break the feedback loop.

Another approach, which is adopted in Figs. 18(c) and 19, is to block one of the beams, in this case the signal beam, and use another light source, such as a light-emitting diode or a flashlight, to provide the photocurrent. Because of the large variations in the gain of the log ratio output, it is vital to ensure that the replacement light power is the same as the original beam; this is done by one adjusting the replace-
ing only on the integrating capacitor and the voltage
than in Fig. 3 (for increased bandwidth, while still
the base of Q1, since the noise power density goes as iSignal •
properly, the total rms noise voltage at the base of Q1
damage to the transistor from excessive base-collector current when
achieve significantly better medium-frequency loop
gain—and hence better noise intermodulation
suppression—with lead-lag frequency compensation.

Figure 16 shows the intermodulation suppression of the log ratio canceller of Fig. 14. The experimental setup used a Pockels cell and a Glan-Taylor prism to modulate the laser beam by approximately 30% at 5 kHz. The resulting amplitude modulation had a severely distorted envelope that was due to the non-linearity of the Pockels cell–polarizer combination, so that its modulation spectrum was rich in harmonics of 5 kHz. The resulting beam was split by a Wollaston prism. One of the two resulting beams (the cancellation beam here) was further modulated by approximately 0.1% at 50 kHz with a liquid-crystal variable retarder and a film polarizer. The upper trace of Fig. 16 was taken with a flashlight to replace the signal beam, as discussed above, whereas the lower trace was taken with the canceler operating normally. The 50-kHz carrier is preserved, even increased slightly by the return of the sideband energy, but the very strong, harmonic-rich 5-kHz modulation is suppressed by more than 60 dB. (The small residual peaks at ±5 kHz, whose spectrum is different from the amplitude modulation spectrum of the upper trace, seem to be due to phase modulation of étalon fringes from the stray light, although straight feedthrough is not ruled out.) The noise intermodulation suppression of this technique is superior to any previous method. This feature makes possible accurate extinction measurements with lasers whose power varies greatly during a scan, such as current-tuned diode lasers. By the elimination of the large ramp signal that is due to the laser intensity change, this device makes possible such measurements as ultrasensitive tunable diode laser spectroscopy with a simple dual-beam approach. It is currently used in tunable diode laser measurements of the oxygen content of the Earth’s atmosphere.

Because the gain and dc level at the log ratio output depends on the temperatures of the differential pair Q1–Q2, some sort of temperature compensation or stabilization is required. It is possible to make efforts at temperature compensation of this output, for example by the use of forward-biased diodes connected in shunt with the bases of Q1 and Q2 (as is commonly done in integrated circuit operational transconductance amplifiers such as the LM1360024) or by the use of thermistors; however, the best way to do it is to temperature stabilize the circuit to eliminate the temperature variations themselves. Ovenizing the entire circuit is awkward, but adequate performance can be obtained by stabilizing the differential pair alone, an idea sometimes used in logarithmic amplifiers.25

Figure 17 shows a circuit fragment to control the temperature of the differential pair right on the die. Differential pair Q1 and Q2 are two diagonally opposite sections of a monolithic quad transistor, such as the Analog Devices MAT-04 or NEC UPA104. Of the other two sections, Q5 functions as a diode temperature transducer and Q4 as a heater. The feedback loop forces the VBE of Q5 to be 90% of its room-

1 February 1997 / Vol. 36, No. 4 / APPLIED OPTICS 915
Many of the most sensitive optical measurements are made in differential systems where the signal desired is the difference between the intensities of two beams. In such cases, the circuit of Fig. 19 is useful. It has two signal photodiodes; is; gl is made slightly larger than is; g2 so that the differential pair must always sink at least some current to keep the circuit in balance.

The two signal photocurrents is; gl and is; g2 are subtracted directly by putting the signal photodiodes in series (with cascode transistors as before), which eliminates the phase and gain error problem, and the (small) comparison current goes through the differential pair Q1/Q2. Because the current imbalance is assumed to be small, i comp can be kept small, and so the additional noise and error introduced are small as well. Figure 20 shows the improved performance of this version. Even with a total signal beam power of over 10 mW, 70 dB of low-frequency cancellation can be achieved without adjustments, and significantly more with occasional manual tweaking of the beam position on the photodiodes. This result should be compared with those of Fig. 8 for the single-ended model. High-frequency cancellation performance...


The differential noise canceller is especially suited to this situation. Because the two signal photocurrents can be large compared with the comparison current (provided the maximum imbalance in the signal currents is sufficiently low), the collector currents of the differential pair can be kept small, for good cancellation performance and low $r_b$ noise, while the signal currents can be made as large as necessary for the best SNR, within the limitations of the components used.

Figure 21 shows the noise floor of the present version, operating with a total signal beam power of 14 mW and exhibits a 1-Hz SNR of 160 dB, approximately 0.5 dB less than the theoretical value. This corresponds to a measurement uncertainty of 7 parts in $10^2$ in 1 s, even at the baseband. The effects of transistor nonlinearity and noise have been essentially eliminated here.

For single-ended measurements demanding such dynamic range, one of the signal inputs can be used as a second, higher-power comparison input that does not suffer from the nonideal behavior of $Q_1/Q_2$.

In a system demanding the highest dynamic range, it is vital that the current density in the photodiodes be kept sufficiently low. High-current densities lead to flattening of the potential gradient in the junction, causing nonlinearity resulting from intensity-dependent photocarrier recombination rates. Solar cells work well at very high current densities, but this technological sophistication seems slow to appear in signal photodiodes.

E. Performance Optimization in Operation

Whichever version is used, there are a few things to remember when using the noise canceller. The most important is that it removes the bulk of the laser amplitude noise. This (while a virtue) makes prominent some seldom-considered second-order effects, which can limit the performance of the experiment if care is not taken. These effects include the noise of the unpolarized spontaneous emission, spatial dependence of the laser noise, photodiode response non-uniformity, and etalon fringes.

The noise canceller cancels only the correlated part of the noise, so for best performance both photodiodes must see exactly the same mode spectrum. In practice, this means putting an efficient polarizer in front of the laser to eliminate spontaneous emission in the orthogonal polarization state — this will in general not be split in the same ratio as the laser light by the beam splitter — and making sure that none of the beams spills off the photodiodes or is otherwise vignetted.

In many lasers, the noise spectrum depends on the detector's position in the beam, so special care must be taken to avoid any vignetting (this effect is particularly strong in diode lasers, even single-longitudinal mode types). If the spatial dependence of the laser noise is strong, the nonuniformity (typically 1–5%) of the responsivity of the photodiodes can cause degraded cancellation; this can be improved by one moving the beams around on the photodiodes while watching the noise floor on a spectrum analyzer (this is usually unnecessary). This is often a problem in setups used for verifying the performance of the noise canceller itself; the acoustic delays in acousto-optic modulators make them unsuitable for high-frequency measurements at high cancellation efficiencies, and some Pockels cells exhibit enough retardation variation across their apertures to limit attainable cancellation to 30 dB or less.

Because of the high premium placed on preserving the correlations between beams, the cancellation per-
formance tends to degrade as the optical system becomes more complicated. Étalon fringes from incidental reflections (e.g., off the windows of metal-can photodiodes) are pervasive unless great care is taken. Particularly in tuned laser experiments, such as current-tuned diode laser spectroscopy, étalon fringes cause the spectra of the two beams to differ slightly and to change rapidly with scanning, vibration, and temperature. It is noted above that a seemingly innocent antireflection-coated photodiode window can cause measurement errors of 1 part in $10^{10}$ / °C and 3.8% per wave number due to étalon fringes alone. Errors introduced ahead of the beam sampler, which should seemingly be common mode (and hence canceled well) may sometimes not be, owing to the spatial sensitivity variations. Fringes often have a pronounced spatial structure, so they should be attended to even at points before the beam sampler.

Some optical components are worse than others; for the highest precision ratiometric measurements, components requiring near-normal incidence on one or more faces are disastrous. Cube beam splitters are particularly bad in this regard and should be replaced by Wollaston prisms or noncubic Brewster devices wherever possible. Wollaston prisms have plane-parallel faces, but because they deflect both the extraordinary and ordinary rays passing through them by an angle of ±5° to ±10°, a beam entering near-normal incidence leaves the prism at a large enough angle that back reflections are much less troublesome.

Photodiode windows can be removed conveniently and safely by holding the device in a vise and gently tapping the window with a curved blunt object, such as a ball-peen hammer or a pair of gas pliers. The glass is reduced to powder, which does not damage the diode itself and is easily removed. The size of the hole in the metal case, together with the large radius of curvature of the peen, ensures that the hammer does not hit the die. The competing technique of cutting the metal can with a lathe results in oil contamination and metal slivers, which can cause short circuits. Not all photodiodes are sufficiently well passivated to perform well after long exposure to ambient conditions, so well-passivated devices should be selected. The Hamamatsu diodes used in most of the circuits for the present paper exhibited large 1/f noise after approximately one day of exposure to ambient conditions after their windows were removed and so appear to be poor candidates for windowless operation.

Because of the coherent addition of the incidental reflections with the main beams, the intensity perturbations caused by weak interfering beams can be surprisingly strong, 1 μW of stray surface reflection can cause a ±6.4% intensity change in a 1-nW beam. Because étalon fringes are frequency dependent, they can also convert phase noise to amplitude noise, especially in multiple-longitudinal mode lasers and those (such as diode lasers) that exhibit mode hopping. The mode-hopping or mode-partition noise on the two detectors is partially decorrelated by the fringes because the relative amplitudes of the modes are changed. This extra amplitude noise will in general be different in the two beams and so will not cancel. Because of their relatively narrow line widths, fixed-tuned single-longitudinal mode lasers should be immune to this extra noise in most cases, but it may be a limiting factor with multiple-longitudinal mode units.

Besides respecting the integrity of the laser beams, there are other constraints on the optical system. Phase shifts that are due to path delays may be important at high modulation frequencies. When a perfect canceller operation is assumed, the lower limit on modulation feedthrough for a given path-length difference $\Delta z$ is

$$A_{\text{min}} = 2 \sin \left( \frac{\pi \Delta z_{\text{mod}}}{c} \right).$$

For example, if 40-dB cancellation is desired at 1-MHz modulation frequency, the path-length difference between the comparison and signal beams must be less than 0.01 rad at 1 MHz, or 48 cm. This limitation does not preclude the use of the noise canceller in long path-difference interferometers such as coherent lidars, provided that the (undelayed) local oscillator beam is significantly stronger than the delayed signal beam; in that case the delayed noise is not too large a fraction of the total noise and cancellation is effective.

Noise cancellers are quite tolerant of normal room light, but strong stray light and leaky photodiodes (such as large-area, room-temperature germanium units or longer wavelength devices with low shunt resistances, such as InAs diodes) are to be avoided, especially if the highest logarithmic output accuracy is required. Less leaky infrared photodiodes should work well, provided they can be used with at least half a volt or so of reverse bias to allow normal operation of the differential pair. PIN photodiodes should be used when possible, as they have lower capacitance for a given area. Excessive photodiode capacitance may cause the differential pair or the cascode transistor to oscillate, which will lead to poor results. Even without actual instability, capacitance will certainly degrade the high-frequency performance of the device, so when high-frequency operation is important, the smallest photodiodes that are sufficiently linear for the required current density should be used.

Because of the high SNR's they can achieve, noise cancellers are somewhat susceptible to pickup, capacitive loading on the summing junction or photodiodes, and power supply noise. They should be constructed with robust grounds (a ground-plane PC board is best, provided that the stray capacitance it adds to the summing junction of $A_1$ is not too large) and housed in a metal box, with the light entering through holes to aid in reducing pickup. Adequate supply by-pass capacitors (e.g., 10-μF solid tantalum in parallel with 0.1-μF monolithic ceramic from each supply to ground, on the board) and a quiet linear power supply should be used to prevent interfering signals from entering on the dc supplies. If switch-
improves somewhat less, as the component speed is still a limiting factor.

There are situations, especially with powerful lasers, in which excess noise must be reduced to levels below the ultimate noise floor of the basic noise canceller. The differential noise canceller is especially suited to this situation. Because the two signal photocurrents can be large compared with the comparison currents (provided the maximum imbalance in the signal currents is sufficiently low), the collector currents of the differential pair can be kept small, for good cancellation performance and low $r_{11}$ noise, while the signal currents can be made as large as necessary for the best SNR, within the limitations of the components used.

Figure 21 shows the noise floor of the present version, operating with a total signal beam power of 14 mW and exhibiting a 1-Hz SNR of 160 dB, approximately 0.5 dB less than the theoretical value. This corresponds to a measurement uncertainty of 7 parts in $10^5$ in 1 s, even at the baseband. The effects of transistor nonlinearity and noise have been essentially eliminated here.

For single-ended measurements demanding such dynamic range, one of the signal inputs can be used as a second, higher-power comparison input that does not suffer from the nonideal behavior of $Q_1/Q_2$.

In a system demanding the highest dynamic range, it is vital that the current density in the photodiodes be kept sufficiently low. High-current densities lead to flattening of the potential gradient in the junction, causing nonlinearity resulting from intensity-dependent photocarrier recombination rates. Solar cells work well at very high current densities, but this technological sophistication seems slow to appear in signal photodiodes.

E. Performance Optimization in Operation

Whichever version is used, there are a few things to remember when using the noise canceller. The most important is that it removes the bulk of the laser amplitude noise. This (while a virtue) makes prominent some seldom-considered second-order effects, which can limit the performance of the experiment if care is not taken. These effects include the noise of the unpolarized spontaneous emission, spatial dependence of the laser noise, photodiode response nonuniformity, and etalon fringes.

The noise canceller cancels only the correlated part of the noise, so for best performance both photodiodes must see exactly the same mode spectrum. In practice, this means putting an efficient polarizer in front of the laser to eliminate spontaneous emission in the orthogonal polarization state — this will in general not be split in the same ratio as the laser light by the beam splitter—and making sure that none of the beams spills off the photodiodes or is otherwise vignette.

In many lasers, the noise spectrum depends on the detector's position in the beam, so special care must be taken to avoid any vignetting (this effect is particularly strong in diode lasers, even single-longitudinal mode types). If the spatial dependence of the laser noise is strong, the nonuniformity (typically 1–5%) of the responsivity of the photodiodes can cause degraded cancellation; this can be improved by one moving the beams around on the photodiodes while watching the noise floor on a spectrum analyzer (this is usually unnecessary). This is often a problem in setups used for verifying the performance of the noise canceller itself; the acoustic delays in acousto-optic modulators make them unsuitable for high-frequency measurements at high cancellation efficiencies, and some Pockels cells exhibit enough retardation variation across their apertures to limit attainable cancellation to 30 dB or less.

Because of the high premium placed on preserving the correlations between beams, the cancellation per-
performance tends to degrade as the optical system becomes more complicated. Étalons fringes from incidental reflections (e.g., off the windows of metal-can photodiodes) are pervasive unless great care is taken. Particularly in tuned laser experiments, such as current-tuned diode laser spectroscopy, étalon fringes cause the spectra of the two beams to differ slightly and to change rapidly with scanning, vibration, and temperature. It is noted above that a seemingly innocent antireflection-coated photodiode window can cause measurement errors of 1 part in 10^3 °/° and 3.8% per wave number due to étalon fringes alone. Errors introduced ahead of the beam sampler, which should seemingly be common mode (and hence canceled well) may sometimes not be, owing to the spatial sensitivity variations. Fringes often have a pronounced spatial structure, so they should be attended to even at points before the beam sampler.

Some optical components are worse than others; for the highest precision ratiometric measurements, components requiring near-normal incidence on one or more faces are disastrous. Cube beam splitters are particularly bad in this regard and should be replaced by Wollaston prisms or noncubical Brewster devices wherever possible. Wollaston prisms have plane-parallel faces, but because they deflect both the extraordinary and ordinary rays passing through them by an angle of ±5° to ±10°, a beam entering near-normal incidence leaves the prism at a large enough angle that back reflections are much less troublesome.

Photodiode windows can be removed conveniently and safely by holding the device in a vise and gently tapping the window with a curved blunt object, such as a ball-peen hammer or a pair of gas pliers. The glass is reduced to powder, which does not damage the diode itself and is easily removed. The size of the hole in the metal case, together with the large radius of curvature of the peen, ensures that the hammer does not cancel. Because of their relatively narrow line widths, fixed-tuned single-longitudinal mode lasers should be immune to this extra noise in most cases, but it may be a limiting factor with multiple-longitudinal mode units.

Besides respecting the integrity of the laser beams, there are other constraints on the optical system. Phase shifts that are due to path delays may be important at high modulation frequencies. When a perfect canceller operation is assumed, the lower limit on modulation feedthrough for a given path-length difference Δz is

$$A_{min} = 2 \sin \left( \frac{\pi \Delta f \text{mod}}{c} \right).$$

For example, if 40-dB cancellation is desired at 1-MHz modulation frequency, the path-length difference between the comparison and signal beams must be less than 0.01 rad at 1 MHz, or 48 cm. This limitation does not preclude the use of the noise canceller in long-path-difference interferometers such as coherent lidars, provided that the (undelayed) local oscillator beam is significantly stronger than the delayed signal beam; in that case the delayed noise is not too large a fraction of the total noise and cancellation is effective.

Noise cancellers are quite tolerant of normal room light, but strong stray light and leaky photodiodes (such as large-area, room-temperature germanium units or longer wavelength devices with low shunt resistances, such as InAs diodes) are to be avoided, especially if the highest logarithmic output accuracy is required. Less leaky infrared photodiodes should work well, provided they can be used with at least half a volt or so of reverse bias to allow normal operation of the differential pair. PIN photodiodes should be used when possible, as they have lower capacitance for a given area. Excessive photodiode capacitance may cause the differential pair or the cascode transistor to oscillate, which will lead to poor results. Even without actual instability, capacitance will certainly degrade the high-frequency performance of the device, so when high-frequency operation is important, the smallest photodiodes that are sufficiently linear for the required current density should be used.

Because of the high SNR's they can achieve, noise cancelers are somewhat susceptible to pickup, capacitive loading on the summing junction or photodiodes, and power supply noise. They should be constructed with robust grounds (a ground-plane PC board is best, provided that the stray capacitance it adds to the summing junction of A, is not too large) and housed in a metal box, with the light entering through holes to aid in reducing pickup. Adequate supply by-pass capacitors (e.g., 10-μF solid tantalum in parallel with 0.1-μF monolithic ceramic from each supply to ground, on the board) and a quiet linear power supply should be used to prevent interfering signals from entering on the dc supplies.
ing supplies cannot be avoided, they should be further regulated on the card, with the use of linear voltage regulators or capacitance multipliers. The photodiodes should be mounted on the circuit board, close to transistors Q1–Q3; they should not be put on cables, lest the cable reflections, time delays, and capacitance reduce the cancellation or destabilize the circuit. Any of these rules may of course be broken by someone sufficiently familiar with noise cancellation and low-noise design in general, but experimenters attempting to reproduce the present results in their own systems for the first few times should observe them faithfully if they desire good performance.

5. Applications and Discussion

The laser noise canceller is really a family of devices for reducing or eliminating the effects of laser amplitude noise. It operates by applying feedback outside the signal path to adjust the ratio of two photocurrents with a bipolar transistor differential pair, so that when the two are subtracted, the dc component of the result is zero. Because of the near-ideal linearity and wide bandwidth of most optical systems, photodiodes, and transistors, cancellation of the dc implies cancellation of the correlated part of the noise fluctuations as well at all frequencies of interest.

The actual noise reduction achieved by the cancellers varies from 55 dB (in the basic version) to 70 dB or more (in the high spurious suppression and differential versions). These numbers reflect the realistic behavior of the canceller in unattended technological applications, and can be improved significantly by the continual adjustment that is possible in a laboratory setting. The cancellation degrades with increasing frequency but remains useful to several megahertz, and work is going on to extend this to 100 MHz or so. The system noise floor is improved to 3 dB above the shot noise of the signal beam in the basic version and even closer with the low-noise floor and differential versions, even with noisy lasers. This is accomplished in a single operation, using a small circuit card containing approximately $10 worth of commonly available parts.

As an added benefit, the noise canceller provides a second output related to the log ratio of the two photocurrents. The log ratio output has the same high suppression of the additive noise as the linear output and in addition has noise intermodulation suppression superior to any other technique.

The basic version of the noise canceller is limited to a dynamic range of approximately 150 dB in 1 Hz by the Johnson noise of the base spreading resistance of the bipolar transistors that are used. In applications where this is insufficient, the high-power differential version can be used to achieve 1-Hz dynamic ranges of 160 dB or higher (limited essentially by the photodiodes), equivalent to an uncertainty of 7 parts in $10^9$ in 1 s.

The variations presented here are flexible enough to cover most laser-based optical measurements, and have performance good enough to replace more complicated schemes for ultrasensitive measurements (such as heterodyne interferometry) in many cases. They are sufficiently simple, robust, and cheap to be used in real-world instruments, so that ultrasensitive measurements can be brought to cost-sensitive applications such as in-line sensors and environmental monitors. For those applications in which even more noise suppression is required, the noise cancellers can be combined with other techniques to reduce laser intensity noise, such as feedback stabilization.

Because the canceller reduces the excess noise by such a large factor, there are some subtle pitfalls in its use; these arise from a variety of seldom-noticed second-order effects. The most important of these effects are the dependence of the noise on polarization and position in the beam, photodiode nonuniformity, incidental étalon fringes, and the limitations of the bipolar transistors used in the device. The optical effects can be minimized by using a polarizer at the laser and eliminating vignetting and on-axis planar surfaces; the electronic ones by careful device selection or the use of one or more of the circuit variations presented here.

The caveats listed are not specific to this device; such problems will beset any feedback stabilization or all-electronic noise reduction system. They are more important here than elsewhere only because of the high degree of noise reduction attainable with the laser noise canceller.

Appendix A: Mathematical Results

A. Nonlinearity due to $r_e$

Here we derive Eq. (12) for the noise feedthrough due to $r_e$. Adding the effects of the degenerative feedback of the emitter bulk resistances into Eq. (3) yields

$$\frac{i_{c2}}{i_{c1}} = \exp \left( \frac{qV_{BE}}{kT} \right) \exp \left[ \frac{q}{kT} \left( i_{c1}r_{E1} - i_{c2}r_{E2} \right) \right].$$

(A1)

Assuming $ei_{c2}/kT \ll 1$, the second exponential may be replaced by the first two terms in its Maclaurin series. Differentiating implicitly and assuming that $i_{c1} + i_{c2} = i_{\text{comp}}$ and $i_{c2} = i_{\text{signal}}$ (i.e., that the current gain is infinite and the loop is in perfect balance), we obtain the partial derivatives to plug into Eq. (11), and we arrive at

$$A_{\text{min}} = \frac{q}{kT} \frac{r_{E1}(i_{\text{comp}} - i_{\text{signal}})}{i_{\text{comp}}} \frac{i_{\text{comp}}}{r_{E2}(i_{\text{comp}} - i_{\text{signal}}) - r_{E2}i_{\text{signal}}}. \quad \text{(A2)}$$

B. Shot Noise of $i_{c2}$

In this subsection, we derive an expression for the noise of the collector currents of a differential pair, perhaps with diode degeneration, which is fed by an emitter current source with full shot noise. If no diode degeneration is used, the collector currents have full shot noise also, regardless of the splitting ratio. We begin by introducing a simplified noise model of a BJT differential pair, which is as shown in Fig. 22; it contains only the emitter current shot-noise contribution. There are two terms in the col-
lektor noise current of $Q_2$: the split emitter current noise, which splits proportionally to the emitter conductances of the transistors; and the noise current owing to the shot-noise voltages on the emitters, in series with the two emitter resistances $r_e (r_e$ is the small-signal emitter resistance, equal to the reciprocal of the transconductance and is not the same as the emitter bulk resistance $r_E$). Thus the total instantaneous noise current at the collector of $Q_2$ is

$$i_{N2} = \frac{v_{N1} + v_{N2}}{r_{e1} + r_{e2}} + i_{NE} \left( \frac{1}{r_{e1}} + \frac{1}{r_{e2}} \right), \tag{A3}$$

where

$$\langle v_N^2 \rangle = (2 q i_C) r_e^2 = \frac{2 kT}{q_i C} \tag{A4}$$

in 1-Hz bandwidth. Assuming that the current gain is high (this assumption is really in the model already) and that the three noise contributions are uncorrelated, they combine to yield

$$\langle i_{N2}^2 \rangle = 2 q i_C r_e \tag{A5}$$

exactly full shot noise.

If there are $N$ ideal diodes (such as diode-connected transistors) in series with each emitter, the values of $r_e$ increase by a factor of $1 + N$ and the shot-noise voltages by $(1 + N)^{1/2}$. Putting these factors in, we obtain for the collector current noise

$$\langle i_{N2}^2 \rangle = 2 q i_C \left( 1 - \frac{N}{N + 1} i_{C1} + i_{C2} \right). \tag{A6}$$

References