REACHING THE

SHOT NOISE LIMIT

FOR $10

BY PHILLIP C. D. HOBBS

Laser-based optical measurements are a fundamental technology of our civilization. Their flexibility is amazing: interferometers maintain the accuracy of machine tools, serve as gyroscopes for navigation, do transform spectroscopy, and perform ranging for seismology and gravity wave detection;[1] lidar detects tornadoes, wind shear, and air pollution; other systems play compact disks and scan bar codes in supermarkets. Optical measurements in research include spectroscopy, Doppler velocimetry,[2,3] aerosol particle monitoring, ellipsometry, interferometry, fiber-based sensing, femtosecond pump/probe studies, and many, many more.

The limits of most of these measurements are set by noise and spurious signals of one kind or another. Perturbations caused by wind, sound, mechanical vibration, Johnson (thermal) noise, dark current noise, shot noise, excess laser noise, and extinction noise from aerosols are examples of true noise, in the sense that they are usually uncorrelated with the data (indeed, sometimes they are the data). Besides the noise, there are usually spurious optical signals, due, for example, to etalon fringes and laser self-modulation from front surface reflection, drift in light intensity, residual amplitude modulation from front surface reflection, drift in light intensity, or residual amplitude modulation (AM) and chirp. These are frequently correlated with the data and so introduce systematic error.

Generally speaking, these noise sources are more troublesome at low frequencies, with the power spectral density (PSD) of the noise in the detected signal showing a distribution that is very roughly proportional to

1/f (falling by 10 dB per decade of frequency). The proportionality is much better in textbooks than in real life.

The scientist or engineer trying to perform high-precision optical measurements seeks to minimize the effect of these nuisances on the optical system. Most of them can be alleviated by one means or another: wind, sound, and mechanical vibration can be reduced by baffles, boxes, and floating tables; electrical and optical spurious, by shielding and choosing an appropriate operating frequency; Johnson noise by increasing impedance levels or using avalanche devices. Most of the time, such techniques work well enough that it is primarily noise and spurious signals associated with the generation and detection of light that limit measurements. Unless the optical loss in the experiment is small, the lower limit of detection is set by the shot noise.

This article discusses ways to eliminate excess laser noise, by which I mean all unintended intensity modulation, except shot noise. It begins with a review of the problem and existing techniques for dealing with it, then introduces a new, low cost, all-electronic device that cancels *excess laser noise* very effectively, making shot noise limited measurements significantly cheaper and simpler than they have been up to now.

The shot noise limit

When a photodetector is illuminated by a laser beam, the resulting photocurrent is not perfectly steady. Because of the random detection times of the individual photons in the beam, and because each photon gives rise to the same size current pulse (one electron in the case of an ideal photodiode), the photocurrent is a Poisson process. Poisson processes have characteristic counting statistics; for a Poisson current of average value λ, the variance in the counted number of events N in a given time t, is equal to the mean,

$$\sigma_N^2 = <N^2-(<N>)^2> = <N> = \lambda t \qquad (1)$$

and so the photocurrent exhibits fluctuations, which are called shot noise. In engineering terms, a photocurrent i has an accompanying RMS noise current in a bandwidth B given by

$$i_n = \sqrt{2qiB} \qquad (2)$$

and a maximum (power) signal to noise ratio of

$$SNR = \frac{i}{2q\,B} \qquad (3)$$

The signal to noise ratio improves linearly as the photocurrent increases. The somewhat mysterious factors of two in Eqs. 2 and 3 are due to the 0.5 Hz equivalent noise bandwidth of a one second averaging window. A current obeying Eq. 2 is said to have full shot noise. Not all currents exhibit full shot noise; in metals, for instance, the Pauli exclusion principle forces the arrival times of electrons to be highly correlated, so that the intrinsic noise of a steady current is very small. To give an idea of how quiet "shot noise limited" is, a 5 mW H-Ne laser shining on a silicon photodiode (0.3 A/W) has shot noise equivalent to an absorption uncertainty of 1×10^{-8} in one second.

The derivation of Eq. 2 for the noise current depends on the assumption that the detection of a photon is a Poisson process. There is an entire field of optical research—squeezed states—devoted to this not being true. In quantum electrodynamics, shot noise results from the interference of a completely coherent beam with *vacuum fluctuations*, the zero-point motion of the electromagnetic field. This interference has the effect of suppressing the correlations between individual photons, thus changing a noiseless current into a Poisson one. There are various means of "squeezing" the vacuum state to reduce either the amplitude or phase noise of a beam, which (if used properly) permit sub-Poissonian noise levels in measurements.

Sub-Poissonian noise levels require the quantum efficiency of the entire optical system following the squeezer, including the photodetector, to be high. This is because the same mechanisms that cause absorption (or scattering) also emit (or scatter) vacuum fluctuations into the beam, thus increasing the detected noise proportionally.

Consider an optical system with a total quantum efficiency of 0.5—which is quite good. A 50% loss implies the emission of 50% of the full vacuum fluctuations into the beam, so that this system can beat the shot noise limit by no more than 3 dB.

Although squeezed light measurements are currently difficult, they may ultimately be very important in some technological applications; they will never be entirely general, because many optical measurements are intrinsically lossy (lidar, for example) or highly cost-sensitive. The shot noise limit is still a highly relevant concept.

Laser noise

Lasers exhibit noise of many kinds, and no type of laser is entirely free of excess noise. Generally speaking, there is a white noise region at high modulation frequencies and a gradually increasing noise density as the carrier is approached. This increased noise power spectral density at low modulation frequencies is often described as "1/f" noise.

The noise characteristics of a real laser are quite different from the usual textbook description, so it is useful to exhibit one for discussion. Figure 1 shows the actual noise spectrum of a commercial 8 mW He-Ne laser; the lower curve is one scan with an FFT spectrum analyzer, and the top curve is the envelope of 32767 scans. The noise spectrum has several features common to many kinds of laser. It has many discrete spurious signals at different frequencies that arise from switching power supplies, mode beats, acoustic resonances, power line pickup,

Noise spectrum of a commercial helium-neon laser, detected with a silicon PIN photodiode and a 20kV transimpedance amplifier. The lower trace is a single sweep with an FFT spectrum analyzer, the top trace is the envelope of 32767 sweeps, and the extra line at the bottom is the calculated shot noise. The noise spectrum is dominated by large discrete spurious signals, primarily baseband mode beats.

FIGURE 1

and so on; most of them move around somewhat, so that no frequency is really quiet. This is especially true of baseband mode beats, which are strong fourth- and higher-order effects arising from the heterodyning of one mode with an intermodulation product of two others, and which drift back and forth from DC to a few hundred kilohertz.

The beam was attenuated to 0.93 mW before measurement, and the measured DC photocurrent was 280 μA, corresponding to 5.6 V DC with the 20 kΩ transimpedance amplifier used. The calculated shot noise level is −134.4 dBV (dashed horizontal line at bottom). It is clear that this laser is very far from being shot noise limited, especially at low frequencies.

Noise reduction

Many of the most sophisticated optical measurement technologies consist mainly of ways to improve signal to noise ratios; just how this is done depends on the application. Jobs like heating and pumping other lasers or nonlinear optics require the beam itself to be stable, since it is driving the system (Nuss et al.[4] show a dramatic example). In such applications, noisy lasers must be stabilized, through improved design or external feedback/attenuator systems. These systems operate by taking a sample of their own output beams, and adjusting the laser operating current ("light control") or an external optical attenuator (such as a "Noise Eater"™) to keep the detected photocurrent from the sample beam constant. They are helpful, but none available today is capable even of approaching the shot noise because of their weak sample beams and poor
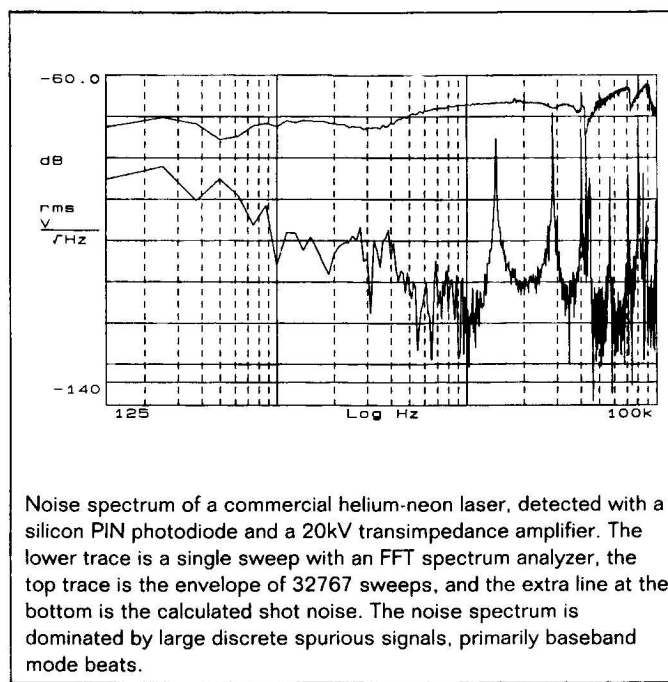
bandwidths.

The effectiveness of a feedback system is set by its gain error. If a laser beam that is, say, 40 dB above the shot noise at some modulation frequency f_m is fed into such a system, the system's gain error at f_m must be less than 1% if the output is to reach the shot noise. This usually occurs at about 1% of the 3 dB bandwidth of the system; thus, a feedback system with a bandwidth of 1 MHz would be limited to 10 kHz modulation frequencies.

Any system that relies on a sample beam as a comparison standard can at best attain the same signal-to-noise ratio as that beam; as Eq. 3 shows, weak sample beams have poorer signal to shot noise ratios. The output of even a perfectly constructed system exhibits a shot noise/signal ratio that is the sum of the shot noise/signal ratios of the sample and output beams. Therefore, any noise suppression system in which the sample beam is relatively weak will be noisy. Even if the sample and output beams are equal in intensity—meaning a power loss of more than half—the output beam's noise PSD will be at least 3 dB above the shot noise. Besides stabilizing the laser beam itself, there are many ways to design experiments to reduce the effect of noise and spurious signals, too many for detailed discussions here. A few examples can give a general flavor.

Chopping systems make the desired signal periodic at some low frequency, usually at a frequency well inside the l/f region, where the noise power spectral density is still high (though better than at DC). Such a small frequency shift can do little to improve the noise PSD, so look-in amplifiers are needed to narrow the detection bandwidth sufficiently to obtain adequate signal to noise ratios. Because the detection bandwidth must often be less than 1Hz, this method is very slow, especially when scanning is required. Signal averaging is a somewhat similar technique in which the detection bandwidth is narrowed by averaging repetitive scans, producing a "comb" frequency response made up of narrow spikes at the harmonic of the scan frequency; it often works better than chopping since more of the sensitive bandwidth is further away from DC.

To escape the (additive) excess noise entirely, a much larger frequency shift is usually needed; the best technique for this is heterodyne interferometry. Here, the two beams

in an interferometer are derived from the same laser, but one is frequency-shifted before they interfere; the desired beat signal is thereby translated up to a high frequency, outside the l/f region, where the background noise is low. This method improves the noise PSD (it usually reaches the shot noise), allowing quiet measurements to be made quickly, but is also complex and expensive compared to baseband techniques or choppers. Frequency-modulation techniques are another way to achieve much the same result; they are not as sensitive as heterodyne methods, but are often simpler experimentally, since not every measurement needs an interferometer.

None of these methods improves the noise *intermodulation*; since the detected signal is proportional to the laser power, laser noise sidebands are impressed upon the desired signal and these are not removed by frequency shifting.

All-electronic noise suppression

Measurement applications do not necessarily require quiet lasers, only quiet photocurrents. This somewhat subtle distinction is the basis for all-electronic noise suppression, whose history goes back at least 30 years.[5] Several techniques have been used, most of which are in the categories of subtracters and dividers.

In a subtractive system, a sample of the laser beam is split off and detected. Since the amplitude noise of the sample is coherent with that of the detected main signal beam at the far end of the optical path, subtracting the two photocurrents makes their noise components cancel. Figure 2 shows one kind of subtracter, with a transimpedance amplifier A1 to convert the difference current into a voltage. This simple system works fairly well; it is wideband, since it does not need a long (*i.e.*, slow) feedback loop in the main signal path. Its major drawback is that it requires finicky adjustment of the relative strengths of the sample and signal beams to obtain good cancellation performance. Besides being inconvenient, this is impossible if the beams' relative strengths change during the measurement (because of scanning, for example). The prevailing lore is that subtracters are often good for 20 dB improvement in the (electrical) signal-to-noise ratio (SNR), but seldom for much more. A variant on the simple subtracter is the symmetric differential detection scheme, in which both beams are treated equally; scanning Nomarski interferometry is one example.

A divider works similarly, except that instead of subtracting the two photocurrents, it divides them. This is done with an analog divider IC or occasionally with a voltage controlled amplifier (VCA). Dividers would seem to be the way to go, since the two beams are both proportional to the laser power, so dividing them should normalize out the
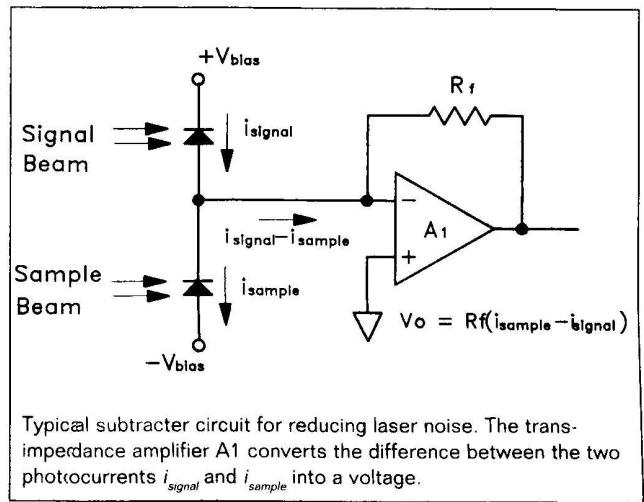

Typical subtracter circuit for reducing laser noise. The transimpedance amplifier A1 converts the difference between the two photocurrents i_{signal} and i_{sample} into a voltage.

FIGURE 2

Bipolar junction transistor (BJT) differential pair connected into the subtracter of Figure 2 to allow electronic adjustment. The ratio of the collector currents depends only on ΔV_{be}, and is not a function of the input current i_{sample}. The circuit's effective suppression bandwidth depends only on the current gain-bandwidth product f_r of the transistors.
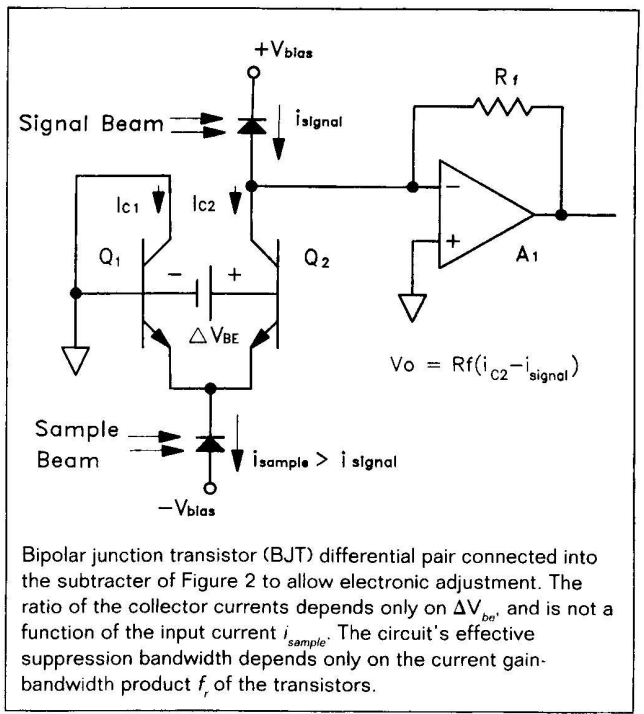
FIGURE 3

noise, without requiring fine adjustments. A more subtle advantage of division is that it eliminates the noise intermodulation.

Real dividers are not too useful in high dynamic range applications. For one thing, they are very noisy—about 60 dB noisier than a quiet operational amplifier (op amp). Dividers are usually made by putting an analog multiplier in the feedback loop of an op amp. The gain and speed of the multiplier varies with the denominator voltage, making the divider difficult to frequency-compensate (for stability). Hence, dividers are quite slow, with bandwidths of at most a few megahertz. As with laser stabilizers, the maximum

cancellation is set by the gain error, so that a system requiring 40 dB of cancellation to reach the shot noise is limited to a bandwidth of 10 kHz with a 1 MHz divider.

We have seen that all-electronic noise reduction schemes are quite useful, though limited. For technical reasons, they have not hitherto been capable of highly efficient noise suppression over reasonable bandwidths. It is worth a look at the basics of these sorts of techniques to see how their performance might be improved.

Optical systems are usually very linear and very wideband. An interference filter a nanometer wide in the visible has a temporal bandwidth of a terahertz; almost all optical components (including silicon photodiodes) are linear over a range of six or more decades of intensity. Thus, when an optical beam with amplitude noise sidebands passes through the system, the sidebands are treated just like the carrier— they suffer no differential gain or phase. If the optical system is minutely adjusted, so that the DC photocurrents cancel exactly, all the amplitude noise cancels identically at all frequencies of interest.

It is possible to replace the optical adjustment for subtracter balance with an electronic one. Figure 3 shows one circuit fragment that does this: a differential pair of bipolar junction transistors (BJTs). The emitter current i_{sample} is split into the two collector currents i_{c1} and i_{c2}, whose ratio is set by the difference Δ_{be} of the base-emitter voltages of the transistors,

$$\frac{i_{c1}}{i_{c2}} = \exp\left(\frac{q\Delta V_{be}}{kT}\right), \qquad (4)$$

which we set externally with a variable voltage source.

BJTs are unique in that this ratio does not depend on the value of i_{sample}; because of this crucial property, any fluctuations in i_{sample} split in exactly the same ratio as the DC, so the noise sidebands are still treated the same way as the carrier and the use of the differential pair will not degrade the effectiveness of the noise cancellation. BJT pairs are fast; transistors with current gain-bandwidth products $f_T > 5$GHz are available for about a dollar, so bandwidth is not such a worry as with the divider. To use this circuit, we make i_{sample}

Schematic diagram of a functioning noise canceller circuit (power supply connections omitted). It combines the electronically adjustable differential detector of Figure 3 with an integrating servo amplifier A2 to force the DC component of transimpedance amplifier A1's output to zero. This automatically adjusts the current splitting ratio, so that the subtractor is balanced and the noise cancels identically at all frequencies of interest, independent of the feedback loop bandwidth.

FIGURE 4

somewhat larger than i_{signal} and adjust ΔV_{be} to dump the extra to ground. The adjustment is still finicky and manual, but at least it is electronic.

Feedback noise canceller

With the possibility of excellent performance, an electronic balancing adjustment for the subtracter, and a simple electronic criterion for finding the perfect balance, all the basic elements of a better noise canceller are there; it remains only to add negative feedback.

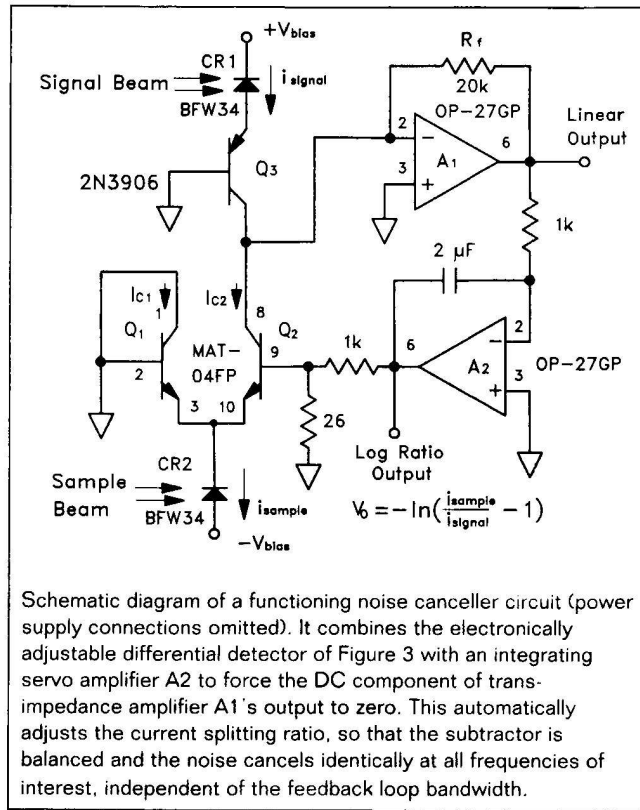Figure 4 is a schematic of a new device using these ideas.[6] It is complete and will work if constructed as drawn (the data shown here were taken with one of these). The parts cost less than $10, including the photodiodes, dominated by the $5 matched pair Q1/Q2.

The circuit is similar to that of Figure 3, with some additions. Transistors Q1 and Q2 are the differential BJT pair that perform the current splitting (Q3 is a "cascode" stage to prevent the capacitance of CR1 from loading the summing junction of A1 and causing instability; it is not fundamental to understanding the device). Integrating servo amplifier A2 senses the output voltage of A1 and adjusts the current splitting ratio of Q1/Q2 (via ΔV_{be}) to keep it at zero volts. The feedback loop can be as fast or as slow as desired, since the bandwidth of effective cancellation does not depend on the feedback bandwidth f_c, only on the f_T of the transistors.

The circuit has another output, from A2, which (as is easily shown from Eq. 4) is related to the log ratio of the signal and sample photocurrents:

$$V_{A2} = -\ln\left(\frac{i_{sample}}{i_{signal}} - 1\right) \qquad (5)$$

The factor of kT/q is eliminated by the choice of the voltage divider ratio. This output has the same advantages as a divider, with one nontrivial addition: since A2 is integrating a signal (A1's output) whose noise has been cancelled, the additive noise cancellation bandwidth of the log output is independent of the feedback bandwidth. Thus, the entire bandwidth is useful, rather than only 1% or so as with

dividers and feedback system; (the cancellation of noise intermodulation does depend on the feedback bandwidth). It is also intrinsically very much quieter than a divider.

This output is useful for situations in which the noise intermodulation is strong, such as current-tuned diode laser spectroscopy;[7] there, the laser power may vary by as much as 2 or 3 to 1 during a scan, making small absorption peaks very inconspicuous on a huge sloping background if this circuit is not used. Q1 and Q2 are a matched pair, to ensure that the offset voltage error in V_{A2} is small; if the log output is not needed, Q1 and Q2 may be replaced by discrete devices such as MPSA18s, which cost only a few cents.

As discussed above, we expect this system to have a noise floor 3 dB above the shot noise of the signal beam alone. This is because we are combining two currents—i_{signal} and i_{c1}—which have the same DC value and have full shot noise. The shot noise currents add in power, resulting in an RMS noise PSD twice that of i_{signal} alone.

It is not obvious, but if i_{sample} has full shot noise, both i_{c1} and i_{c2} have exactly full shot noise as well, regardless of the splitting ratio. If i_{sample} is much larger than i_{signal}, its relative shot noise will be much less. It is possible to redesign the differential pair to preserve this noise advantage, and thus eliminate the 3 dB penalty, by putting a series string of diodes in series with each emitter. The diodes provide feedback, reducing the relative shot noise of the split currents to essentially that of isample. Since this requires a large i_{sample} beam, it is most helpful when the optical system is lossy (so that taking relatively large sample beam need not significantly reduce the strength of the signal beam). Such a system can achieve the performance of an ideal heterodyne system.

Figure 5 shows the performance of the noise canceller of Figure 4 operating on the laser used earlier. The bottom trace is an average of 5 sweeps, the middle trace is the envelope of 1220 traces, and the top trace is the same as Figure 1, for comparison. It is easily seen that the excess noise is strongly suppressed; the larger discrete peaks are reduced by as much as 55 dB. Note that no adjustments are required other than setting i_{sample} a bit bigger than i_{signal}. The quieter regions between the noise peaks have been improved, too; the computed average noise voltage spectral
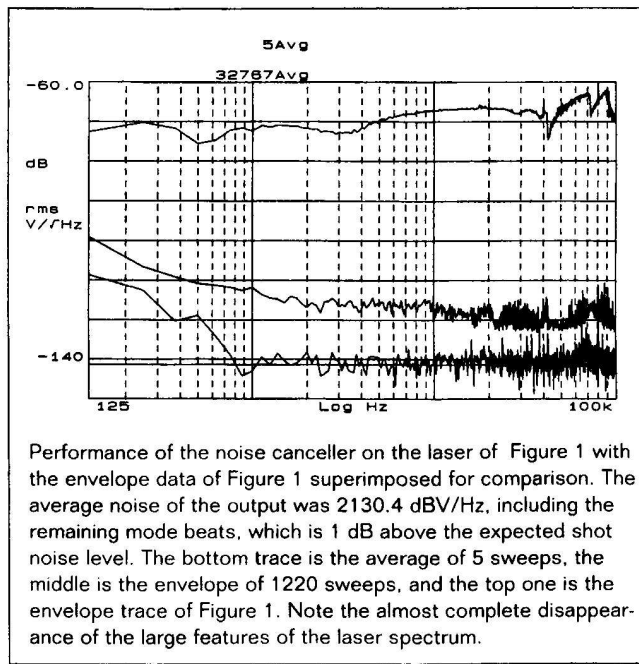

Performance of the noise canceller on the laser of Figure 1 with the envelope data of Figure 1 superimposed for comparison. The average noise of the output was 2130.4 dBV/Hz, including the remaining mode beats, which is 1 dB above the expected shot noise level. The bottom trace is the average of 5 sweeps, the middle is the envelope of 1220 sweeps, and the top one is the envelope trace of Figure 1. Note the almost complete disappearance of the large features of the laser spectrum.

FIGURE 5

density there is -130.4 dBV/Hz even with residual modebeats, which agrees well with the calculated value of -131.4 dBV/Hz from shot noise alone.

Thus, apart from a few very strong, very narrow spectral lines (which are greatly reduced), the noise of the output of A1 is limited only by the shot noise of the signal beam. Swept-sine measurements[6] of the ultimate suppression performance of the system in Figure 3 show it to be more than 55 dB near DC, and better than 40 dB out to 10 MHz (even though the op amps roll off near 4 MHz), with a feedback loop bandwidth of 10 Hz. The emitter-degenerated system has the same noise performance as an ideal heterodyne system, at a fraction of the cost and complexity. Even where an interferometer is necessary, in many cases a complex heterodyne system can be replaced by a much simpler homodyne approach.

There are a few things to remember when using the noise canceller. It can only cancel 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—since 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 beam spills off the photodiodes or is vignetted. In diode lasers, the noise spectrum depends on the detector's position in the beam, so special care must be taken to avoid vignetting. Etalon fringes from the windows of metal-can photodiodes are sometimes troublesome, too, since these affect each beam differently.

Phase shifts due to path delays may be important at high modulation frequencies. Assuming perfect canceller operation, the lower limit on modulation feedthrough for a given path length difference Δz is

$$A_{min} = 2 \sin(\frac{\pi f_m \Delta z}{c})$$ (6)

For example, if 40 dB cancellation is desired at 1 MHz modulation frequency, then the path length difference between the sample and signal beams must be less than 0.01 radian at 1 MHz, or 48 cm.

Applications

My colleagues and I have used this new technique for quantum-limited absorption spectroscopy,[7] coherent lidar, and other measurements, where the noise sources to be suppressed are the laser's own noise, incidental etalon fringes, self-modulation of diode and gas lasers due to optical feedback, and power variations from current tuning diode lasers. In every case, we have had results comparable to those shown above. The noise canceller is suitable for use in most systems where the limiting factor is amplitude noise in the laser beam, whether intrinsic or externally impressed.

Very many techniques exist for doing shot-noise limited optical measurements, but few are used outside the laboratory. The significance of this device is that it allows such measurements to be done routinely in cost-sensitive technological applications and in experimental setups where heterodyne systems are infeasible, or their added complexity prohibitive.

GREATER DEFINITION

BJT	Bipolar junction transistors
FFT	Fast Fourier transform
PSD	Power spectral density
RMS	Root mean square
SNR	Signal-to-noise ratio
VCA	Voltage controlled amplifier

References

1. Robert L. Forward, "Wideband laser-interferometer gravitational-radiation experiment", Phys. Rev. D **17**, 2, 1978, 379-390.
2. Yu. N. Dubnitshev, *et al.*, "Laser Doppler velocimeter as an optoelectronic data processing system", Appl. Opt. **14**, 1, 1975, 180-184.
3. B. Reuter and N. Talukder, "A new differential laser microanemometer", SPIE Proc. **236**, 1981, 226-230.
4. Martin C. Nuss, *et al.*, "Amplitude noise reduction of 50 dB in colliding-pulse mode-locking dye lasers," Optics Letters, **15**, 1990, 1026-1028.
5. E. Ingelstam, "Some quantitative measurements of path differences and gradients by means of phase contrast and new interferometric devices", J. Opt. Soc. Am., **47**, 1957, 536-544.
6. Philip C.D. Hobbs, "Shot noise limited optical measurements at baseband with noise lasers", in *Laser Noise*, Rajarshi Roy, ed., SPIE Proc. **1376**, 1991.
7. Kurt L. Haller and Philip C.D. Hobbs, "Double beam laser absorption spectroscopy: shot noise limited performance at baseband with a novel electronic noise canceller", in *Quantum Spectroscopy*, SPIE Proc., 1991.

Philip C.D. Hobbs is a research staff member at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y.