These articles, most of which are invited contributions by guest authors, are intended to serve the readers of THIS JOURNAL *by calling attention to new developments in the theory, design, or availability of chemical laboratory instrumentation, or by presenting useful insights and explanations of topics that are of practical importance to those who use, or teach the use of, modern instrumentation and instrumental techniques.*

XXXIX. Signal to Noise Optimization in Chemistry—Part One

THOMAS COOR, *Princeton Applied Research Corporation, Princeton, New Jersey 08540*

Thomas Coor graduated with a B.A. in Chemistry from Rice Institute and obtained a Ph.D. in Physics from Princeton University. He was associated with Brookhaven National Laboratory and Princeton University before joining in the founding of Princeton Applied Research Corporation in 1962. He has worked in the fields of Plasma and high energy physics and in electronic instrumentation.

INTRODUCTION

Chemistry, in common with the other disciplines of the physical sciences, in its pursuit of greater accuracy and finer small effect phenomena, is pushing ever closer to the fundamental limitations imposed on the observer by nature. The fundamental limitations arise from two facts: (1) all matter not at absolute zero has thermal fluctuations, (2) charge, light, and energy are quantized. As a result, all measurements made of physical quantities will have a certain irreducible level of noise associated with them. In addition to this "fundamental" noise in measurements, there is always noise arising from more prosaic sources; 60 Hz line pickup, building vibrations, variations in room temperature, radio stations, etc. These environmental disturbances can, in principle, be reduced to an arbitrarily small value, but in practice are very difficult to remove entirely from the picture.

In spite of his concern with noise as a limitation on detectability or accuracy in his work, the chemist is often unaware of simple techniques that will allow him to improve his results. He is generally even less aware of recent advances that have been made in signal processing, advances that allow vastly improved experimental signal-to-noise (S/N) ratios to be achieved. While much of the literature on the subject is couched in abstruse mathematics, most of the results are relatively simple to understand and can be applied easily to most experimental situations.

In this paper, we will first review the properties of the various sources of noise that confront the chemist as he makes sensitive physical measurements. Then consideration will be given to various ways of avoiding unnecessary noise in deriving an electrical signal that represents the quantity being determined. Next the theory of various kinds of simple information processing that the chemist can employ to enhance the S/N of electrical signals will be presented. And finally various types of commercial equipment useful in S/N enhancement will be surveyed. While the principles to be discussed are very general, we will restrict our attention to electrical signals and assume that all quantities of concern can be transformed into such by suitable transducers.

Although by no means an exhaustive list, the following represents typical areas where the chemist might well be interested in improving his S/N: (1) electron spin resonance, (2) nuclear magnetic resonance, (3) nuclear quadrupole resonance, (4) microcalorimetry, (5) differential thermal analysis, (6) all areas of analytical and research spectroscopy, (7) photochemistry, (8) chromatography, (9) electrochemistry. In a particular instrumental system there will usually be some form of a "probe"—be it an electrode, beam of light, thermocouple, or radio frequency pickup coil—inserted in some way into the medium under study. It is of course desirable that the phenomena of interest alone be monitored by the probe but this is never entirely the case. One must always contend with the fact that the signal representing the desired quantity will have noise in some form as contamination. This noise may be inherent in the physical process, arise from the "probing" action itself, or may be picked up from the environment.

A simple example will be illustrative. Consider a glass microelectrode being used to determine, say, the pH at a particular site in some small system. At equilibrium the Nernst equation specifies that the electrode potential relative to some reference will be:

$$E_{\mathrm{H}} = -2.303 \frac{RT}{nF} [p\mathrm{H}] + K$$

In reality, this equation can represent only the *average* potential at the electrode. Being a small electrode, the concentration of ions around the active region will fluctuate because there are only a finite number of them. This microscopic fluctuation in concentration will give rise to a corresponding fluctuation in electrode potential. The smaller the electrode size and the lower the concentration, the greater will be the relative fluctuation. However, even for large electrodes, the fluctuations or noise will place limitations on the ultimate attainable precision of measurement. What in fact we have described is the Brownian phenomena in an electrochemical framework.

To obtain a more precise measure of the true pH in our example of the glass microelectrode, we might try to take an average of many measurements of the fluctuating cell potential. One might expect that an arbitrary degree of precision to be achieved in this way, but this is not necessarily the case. One must also consider the "noise" properties of the probe and of the rest of the electronic system being used to make the measurement. The effects of glass membrane resistance, flicker noise effects in the glass electrodes and in amplifiers, extraneous signal pickup from the environment, temperature fluctuation "noise" in the constant temperature bath in which the measurements are being carried out; all these and more come into the picture and must be considered. Such effects will in general place severe limitations on the effectiveness of increasing the averaging time as a means of gaining greater precision in simple dc measurements. To understand why this is so, let us first consider the

(*Continued on page A534*)

fundamental sources of noise that arise in making electrical measurements and review their properties.

SOURCES OF NOISE AND THEIR PROPERTIES

Fundamental Noise

It was realized by Schottky (*1*) very soon after the advent of the vacuum tube amplifier that unlimited gain is not useful and that after a certain gain is achieved, random fluctuations or noise show up. These random fluctuations have been shown to arise from two basic sources: one is thermal in origin and its properties can be derived from thermodynamic and statistical mechanical arguments—the other arises from the fact that light and electrical charges are quantitized and current in transistors, vacuum tubes, and photocells flows by virtue of the uncorrelated transfer of *single electrons* across some void, hence is statistical in nature.

Johnson or Resistance Noise

Although intimately related to Brownian motion, the existence of thermal noise in resistors was realized comparatively late. This noise is usually called "Johnson noise" after its discoverer (*2*). Nyquist (*3*), a colleague of Johnson, first derived its properties.

It is instructive to derive the magnitude of Johnson noise in a resistor. The method here is related to, but not identical with, that of Nyquist, and will be given in outline form. Consider the idealized experiment pictured in Figure 1.

(1) Because of black-body radiation from the enclosure at temperature T_1, the dipole will be a source of average power, $\overline{P_A}$, which flows down the transmission line and is completely dissipated in R.

(2) The Second Law of Thermodynamics requires that resistor R must be a source of average power, $\overline{P_R}$, such that $\overline{P_R} = \overline{P_A}$ if $T_1 = T_2 = T$.

(3) Statistical Mechanics states that every mode of a system at equilibrium with

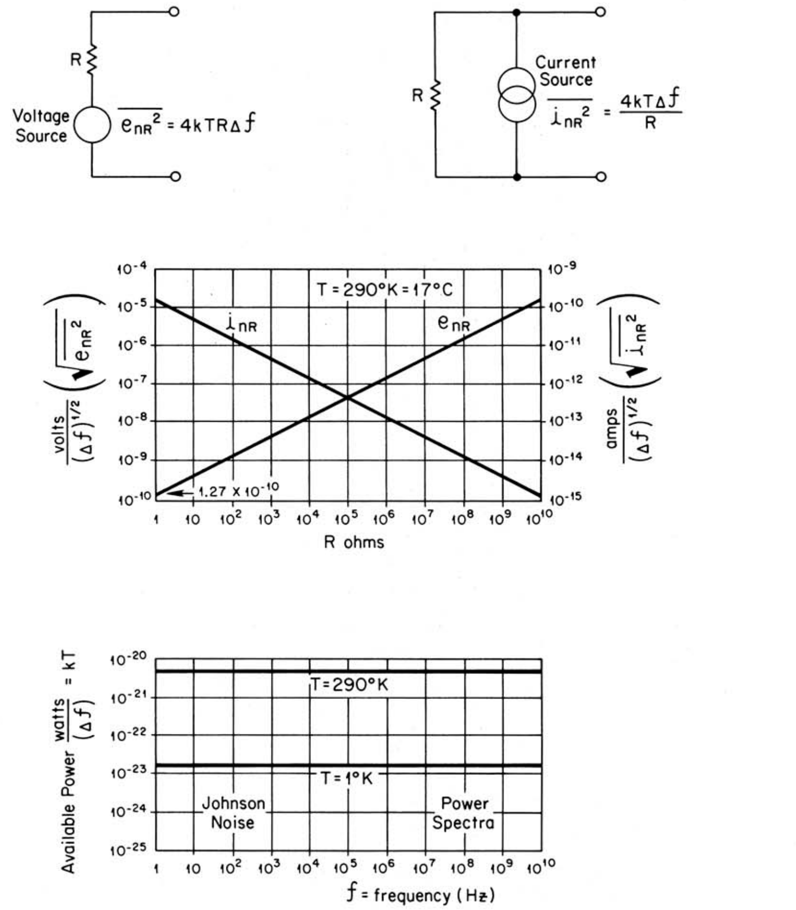

Figure 2. The two equivalent circuit representations of thermal noise in resistor *R* are shown together with the magnitude of the rms voltage and current per cycle as a function of *R*. Also illustrated in the flat or "white" power spectra of the noise at two different temperatures.

temperature T will have average total energy $= kT$, where k is the Boltzmann constant, but that the total energy of a given mode will fluctuate.

(4) Assume an instantaneous short at each end of the transmission line. Energy trapped on the transmission line results from both $\overline{P_R}$ and $\overline{P_A}$. Total energy trapped is:

$$(\overline{P_R} + \overline{P_A}) \times \begin{pmatrix}\text{time for signal} \\ \text{to traverse line}\end{pmatrix} = 2\overline{P_R}\,\frac{L}{C}$$

(5) Wave theory states that the number of standing waves, Δm, on the transmission line between frequency f and $f + \Delta f$, where Δf is the bandwidth, is

$$\Delta m = \frac{2L}{C}\,\Delta f$$

(6) The average total energy trapped on the transmission line having frequencies between f and $f + \Delta f$ will then be,

$$\Delta mkT = \frac{2L}{C}\,kT\Delta f$$

$$= 2\,\overline{P_R}\,\frac{L}{C}$$

(7) Solving for $\overline{P_R}$ we find

$$\overline{P_R} = kT\Delta f$$

(8) $\overline{P_R}$ results from an internal noise voltage source, e_{nR}^2, associated with R. Since the transmission line represents a matched (equal in resistance) load to the voltage source, e_{nR}^2, of internal resistance R,

$$\overline{P_R} = kT\Delta f = \frac{\overline{e_{nR}^2}}{4R}$$

$$\text{Hence, } \overline{e_{nR}^2} = 4kT\Delta f \qquad (1)$$

Similarly one can show that

$$\overline{i_{nR}^2} = \frac{4kT\Delta f}{R} \qquad (2)$$

(*Continued on page A536*)

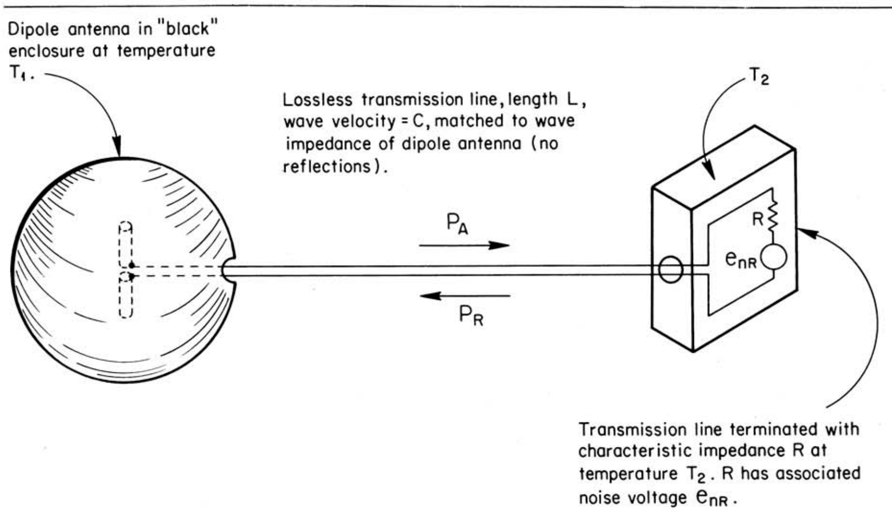

Figure 1. Gedanken experiment used to derive the relations for Johnson noise in resistance *R* at temperature *T*. Thermodynamic and statistical arguments lead immediately to the results.

where $\overline{i_{nR}^2}$ represents the noise current that will flow across the shorted terminals of a resistor R. The two equally valid ways of representing the noise of a resistor are shown in Figure 2.

Note that the eqns. (1) and (2) are functions only of the *frequency bandwidth*, Δf, and not the frequency, hence the magnitude of the average noise voltage or current depends only on the effective frequency bandwidth of the measuring instruments. This type of frequency spectrum, plotted on a power-per-unit-bandwidth basis, is known as a *white* spectrum, hence resistor or Johnson noise is called white noise.

We thus see that any dissipative element, whether it be the membrane resistance of a glass electrode or the load resistance of a photocell, gives rise to a real noise voltage or current and this noise will add to any signal that appears across the resistor. Purely reactive elements, capacitance and inductance, have no noise voltages associated with them *per se.* However, real inductors have winding resistance and real capacitors have dielectric loss, both of which give rise to the expected noise.

While discussing noise arising in resistors, it is well to note that apart from

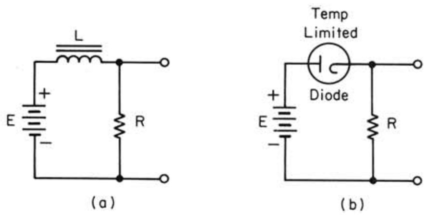

Figure 3. Circuit shown in (a) will exhibit only Johnson noise of *R* if components are perfect. Circuit (b) has additional shot noise associated with current that flows by virtue of independent event processes in temperature limited diode.

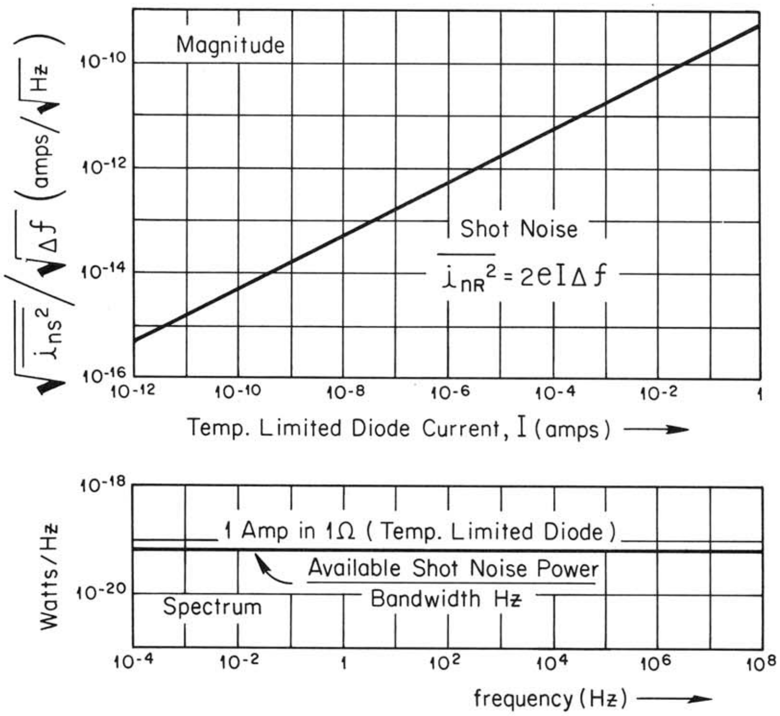

Figure 4. The magnitude and spectral characteristics of shot noise as a function of total diode current, *I*.

Johnson noise, which appears across any resistor whether or not is is carrying current, there are additional sources of noise in resistors that *are* carrying current. Wire wound resistors are in general quite good in this respect, microphonics being about the only source of additional detectable noise. On the other hand, composition, deposited carbon, and metal film resistors show voltage dependent noise, descending in that order. Good metal film resistors are almost—but not quite—as good as wire-wound resistors. This voltage dependent noise is not too well understood but seems to be due to the granular nature of the resistance elements. The spectral characteristics of this excess resistor noise is quite interesting, having a peculiar $1/f$ power spectrum, and has a flicker-like appearance on an oscilloscope or meter. This "$1/f$" or "flicker" noise is disastrous for low-frequency or dc measurements, because averaging for longer periods of time can result in no net increase in S/N!

Shot Noise

In the last section the statement was made that good resistors carrying current had little noise in excess of the Johnson noise. How can this statement be reconciled with the known fact that electrical conduction occurs *via* electrons and that charge is quantized? This is an important point that is often missed. Consider the circuit in Figure 3*a*. For frequencies where $R \gg 2\pi fL$, the noise at the output terminals is just the Johnson noise. However if a diode is introduced in the circuit, as shown in Figure 3*b*, the situation is vastly different, even though the current through the resistor is the same. In the diode (or photocell, transistor, triode, etc.) the passage of current is governed by a random single-event, electron emission process, while metallic conduction is a large scale correlated drift phenomenon with many discrete changes taking part at once. Hence the current through the single-event devices are subject to large statistical fluctuations while metallic conduction is not. The "shot noise" due to the random emission of electrons from cathodes or semiconductor junctions was predicted by Schottky (*1*) early in the era of electronics.

The derivation of the magnitude of this noise is straightforward (*4*) but only the result will be indicated here. The mean square shot noise current, i_{ns}^2, is given by:

$$\overline{i_{ns}^2} = 2Ie\Delta f \tag{3}$$

where e is the electron charge, 1.59×10^{-19} coulomb, I is the current that flows due to some single electron random emission process (e.g., temperature limited diode), and Δf is again the bandwidth at the measuring apparatus with which i_{ns}^2 is measured. We see that this noise is "white," i.e., has a flat spectrum. The magnitude of shot noise current is given in Figure 4.

All active amplifying devices used in electronics—the transistor, photocell, vacuum tube—exhibit shot noise. This noise

(*Continued on page A538*)

is an important factor in determining the overall performance of an amplifying and measuring system. However, the shot noise in a vacuum tube is much less than given by eqn. (3). Space charge effects which introduce a smoothing action on the random conduction process are responsible for this reduction.

Flicker Noise

The third source of noise which is important to our considerations is called *flicker-effect* noise and has been mentioned in conjunction with resistors. Flicker-effect noise is characterized by its unusual spectral characteristics and, for most electronic devices, dominates thermal or shot noise only for frequencies less than about 100 Hz. Flicker noise is especially troublesome for those trying to make small dc measurements, as we shall see.

Flicker noise is not completely understood but is found generally where electron conduction occurs in granular semiconducting material, or where cathode emission is governed by the diffusion of clusters of barium atoms to the cathode surface. Phenomenologically, flickering conduction can be pictured as resulting from a very large assemblage of series connected switch-resistor combinations randomly arranged in all possible ways between two terminals. The switches are then assumed to randomly open and close according to some statistical distribution. The particular distribution determines the power spectrum characteristics of the resulting noise. Almost all electronic elements exhibit flicker noise to some degree and their noise power spectra differ widely. Tubes and transistors typically show a power spectrum of the form $1/f^n$ where n is a constant near 1. The noise of several typical amplifying devices is shown in Figure 5.

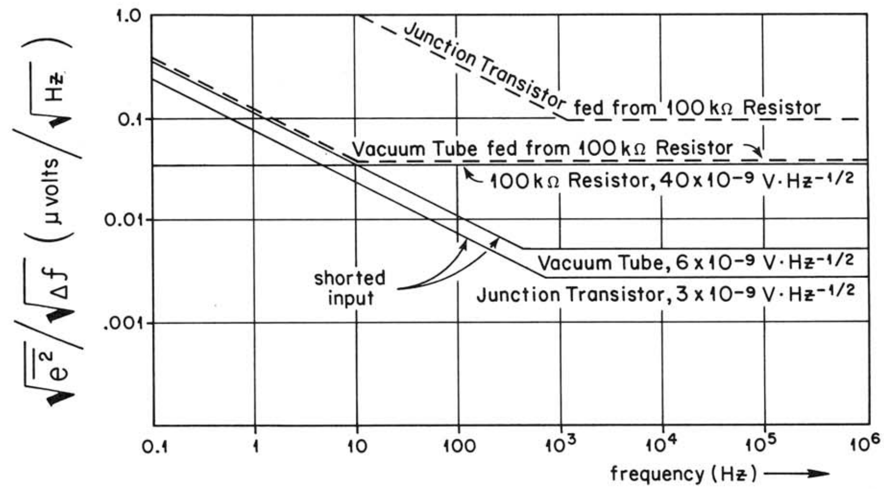

Figure 5. The noise voltage as a function of frequency of representative vacuum tubes and junction transistors under conditions of shorted input and with 100 kΩ source impedance (noise referred to input). Also shown is noise of 100 kΩ resistor at 290 °K. The effect of base current noise is immediately apparent.

The particularly disturbing factor about flicker noise is that the $1/f^n$ characteristic seems to hold down to as low a frequency as one cares to determine it. The long term drift in dc transistor or tube amplifiers would seem to be the manifestation of very low frequency flicker noise! Hence, dc electrical measurements are fraught with great difficulties. Even galvanometers exhibit flicker-noise-like effects. Moral: avoid dc measurements if S/N is a problem.

Environmental Noise

Environmental disturbances which

(*Continued on page A540*)

make their way into chemical measurements can, in principal, be reduced to an arbitrarily small value, but in practice, can be very annoying and difficult to eliminate entirely.

Sixty hertz (and higher harmonics) from the power line, radio stations, motor and switch sparks, corona, building vibrations, and room temperature fluctuations all fall into this category. One interesting and important characteristic of the sum of all of these disturbances is the apparent $1/f$ characteristic of the power spectra. Why this should be so is not entirely clear. Apparently the power spectra of temperature fluctuations, earth vibrations, and all the man-made fluctuations are determined by event disturbances not unlike those found in cathodes and other granular semiconductor processes. While not quantitative, Figure 6 illustrates the spectral characteristics of typical laboratory environmental noise. Apart from the power line frequencies, which are very sharp lines, the region between 10 and 10^6 Hz is comparatively quiet and is a good region to have information appear, as will be discussed later.

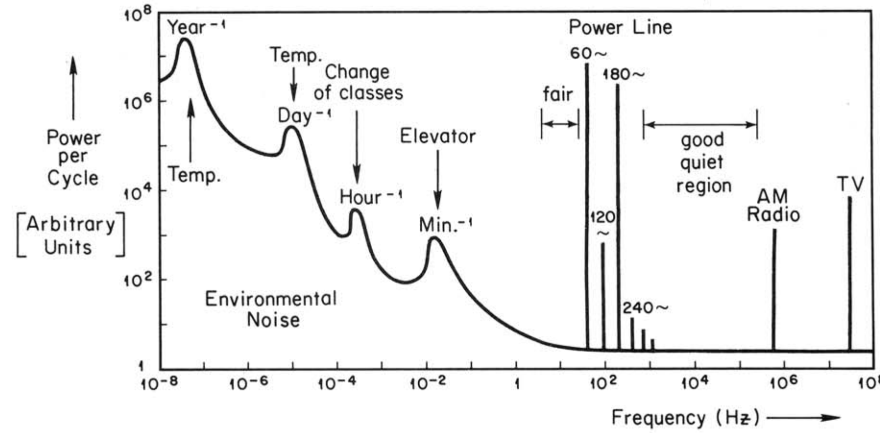

Figure 6. Pictorial representation of environmental noise in a typical location as function of frequency. Note the $1/f$ character at low frequencies.

THE SIGNAL MEASURING SYSTEM

Block Diagram and Noise Sources

Before considering the S/N properties of chemical instrumentation systems, let us make reference to Figure 7 in which a hypothetical system is shown in block diagram form, together with sources of noise. The chemical quantity under study, $Q(t)$, together with its inherent noise, is probed with a linear transducer of some type yielding an electrical signal $E(t)$. Generally, some provision is made to switch or move the transducer input from the object of study to some blank or standard. This allows a zero point to be established or a system calibration to be made. The signal $E(t)$ from the transducer is amplified, processed for noise reduction, and finally recorded in some fashion.

In designing such a system the characteristics of each element must be specified in such a way as to pass the information of interest, act on it in the way desired, and optimize the S/N performance of the entire system. The most important considerations in system specifications are discussed in the following paragraphs.

System Considerations: Nature of Signal

The first consideration in system specification is the nature of the information that is going to be handled by the system. The simplest and most common type of instruments are dc systems where the quantity of interest is steady or slowly varying with time, and the information "containing"

(Continued on page A542)

the magnitude of the quantity is handled as a dc signal.

Other systems carry this information as the amplitude of sinusoidal or square wave signals with fundamental frequency f_c. These ac signals can result from the modulation of the phenomena under study by some means, or can result from the nature of the transducer used. In the latter case a steady or dc input to the transducer is converted by some chopping or modulating action into an ac signal, as, for example, is done in a vibrating capacitor electrometer.

In still other systems, the information may be in the form of recurring waveforms, or in the amplitude of power spectra of signals from the material under study. Consideration of systems in which the information has these forms is beyond the scope of this paper and will not be dealt with here.

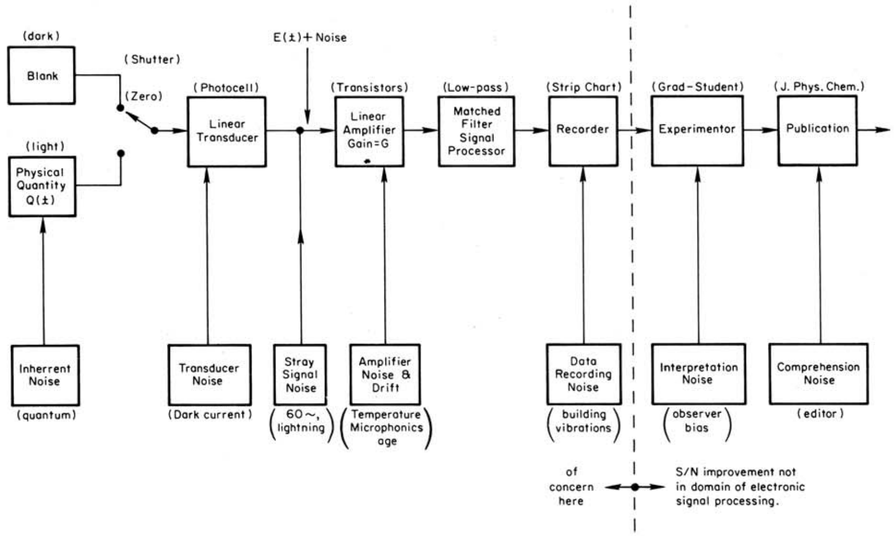

Figure 7. Block diagram of system required to record the changes with time of some quantity Q(*t*) (assumed to be light). Also shown are possible sources of noise.

Signal Frequency

In designing a chemical system, one often has the choice as to where the band of information of interest is made to lie. For example, in atomic absorption spectroscopy the changes in transmission of resonance radiation through the flame of a burner can be determined by a dc measurement. Alternatively, the resonance radiation can be chopped, or alternated between the flame path and a reference path at some frequency, say 400 Hz.

By chopping, the band of information is removed from very low frequencies (dc) and placed around the chopping frequency. This allows the use of ac amplifiers and thus avoids the inherent drift and flicker noise of dc equipment. In addition, it permits the use of some very powerful techniques for improving S/N. These techniques will be discussed later.

System Bandwidth

Having chosen the frequency at which the information is to appear, there is a further related question that must be considered—that of system bandwidth. If the quantity being chopped or modulated at frequency f_c has a steady or dc value, then the bandwidth of the information "channel" at f_c is very narrow. However, if the quantity is changing with time, the bandwidth Δf_c required to carry the information will be greater. For faithful reproduction of the original time-varying quan-

(*Continued on page A544*)

tity, every element of the system through which the information passes must have sufficient bandwidth to pass all components of the signal. However, as most noise sources have the property that the wider the bandwidth Δf of the system the greater will be the resultant noise, this should not be overdone. From the standpoint of S/N the system bandwidth should be only wide enough to faithfully reproduce the wanted time dependance of the original information. These considerations are equally true if the signals being processed are dc. Apart from $1/f$ effects, narrower bandwidths will in general yield less noise.

Dynamic Range

Dynamic range is a very important concept in information processing which is often neglected in system design. Any measuring system will have its own irreducible noise level and drift, which impose a lower limit on the level of signals that may be handled by the system. Likewise, the system will always overload and become non-linear for sufficiently large inputs. Often this overload point is dependent on whether the input is a signal at the chopping frequency f_c or is wide-band noise.

Dynamic range can be defined in two ways: (1) the ratio of the signal producing full scale output to the signal corresponding to the inherent system noise and drift, (2) the ratio of the noise level producing overload to the signal that corresponds to full scale output. Both of these definitions are valid and useful and will be referred to again. In general, definition (1) is more useful in quieter situations. Definition (2) serves better where the input S/N is less than unity. It is obvious that the dynamic range of a measuring system has great bearing on the precision with which measurements can be made and on the ability of a system to function properly in the presence of large quantities of noise.

Drift

Drift also is defined in two valid and useful ways: (1) *zero drift*—the variations with time of the output of a measuring system with zero input; (2) *gain drift*—the drift in the indicated value of a steady input with time. Zero drifts usually impose limitations on sensitivity or ultimate detectability, while gain drifts set limits on accuracy and precision.

Noise Figure

We have seen that any physical system providing some quantity to be measured will have a certain irreducible noise associated with it. What we wish to do, given this S_s/N_s from the source, is to transduce, amplify and then measure the characteristics of S_s with the least possible effect of N_s on the measurement. To do this, we should see that the transducer, amplifier, and other equipment through which the signals are processed, degrade the initial S_s/N_s as little as possible. Any noise these parts of the system inject into the information channel will produe a degraded S_o/N_o at the output. Hence they should be "low noise." One sees that a quantitative measure of the degradation of S/N in passing information through some subsystem is important. One possible measure is called the *Noise Figure* (NF) of the unit. It is defined as:

$$\text{NF} = 10 \log \frac{S_i/N_i}{S_o/N_o}$$

where S_i/N_i is the ratio of the available input-signal and noise *powers,* and S_o/N_o the ratio of the available output powers. It is expressed in dB. The NF of any network is a function of the nature of the source (the source temperature and impedance), the frequency of operation, and its gain. This complex subject has been extensively treated elsewhere (*5*) and only the important results will be given. In specifying the NF of any element, it is assumed that the signal source has an internal resistance at 20°C and that the available noise power N_2 is just the thermal noise of this impedance.

We have seen in Figure 5 that the noise of tubes and transistors is dependent on the source impedance to which they are connected. The excess noise with high source resistance is *in addition* to the Johnson noise generated by the source resistance, and results from a noise current (grid or base current) flowing from the device giving rise to increasing noise as the source resistance is raised. Hence the NF of a signal processor or transducer is a complicated function of source impedance, frequency, input device characteristics, etc. A perfect signal processing system will have a NF of 0 dB, while one that degrades the S/N of the signal source by a factor of two has a 3 dB NF. We shall be concerned with NF when we discuss amplifiers and other signal processing equipment.

Literature Cited

(1) W. Schottky, *Ann. Physik.* **57,** 54 (1918).
(2) J. B. Johnson, *Phys. Rev.* **32,** 97 (1928).
(3) H. Nyquist, *Phys. Rev.* **32,** 110 (1928).
(4) A. van der Ziel, "*Noise,*" Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1954.
(5) For a good treatment of NF and related subjects see: A. T. Starr, "*Radio and Radar Techniques,*" Pitman and Sons, Ltd., London, 1953, Appendix 9, "Noise."

To be concluded in the August issue

The Shimadzu QV-50 Spectrophotometer: Addendum

This instrument, imported by Cosmos Scientific Co., sells for $3250, including its power supply, rather than the lesser figure quoted in our April column. Accessories are available to permit measurements of circular dichroism, flow dichroism, and optical rotatory dispersion. *P. F. L. and G. W. E.*