The Fundamental Noise Limit of Linear Amplifiers*

H. HEFFNER†, FELLOW, IRE

Summary—**If the uncertainty principle of quantum mechanics is applied to the process of signal measurement, two theorems relating to amplifier noise performance can be deduced. First, it can be shown that it is impossible to construct a linear noiseless amplifier. Second, if the amplifier is characterized as having additive white Gaussian noise, it can be shown that the minimum possible noise temperature of any linear amplifier is**

$$T_n = \left[\ln \frac{2 - 1/G}{1 - 1/G}\right]^{-1} \frac{h\nu}{k}.$$

In the limit of high gain G this expression reduces to that previously derived for the ideal maser and parametric amplifier. It is shown that the minimum noise amplifier does not degrade the signal but rather allows the use of an inaccurate detector to make measurements on an incoming signal to the greatest accuracy consistent with the uncertainty principle.

Introduction

Since the advent of the maser, there have been a number of treatments of the noise figure or noise temperature of this and other potentially low noise devices such as the parametric amplifier.[1–4] Most of these have treated each specific device as a quantum system and have determined a limiting noise temperature arising because of amplified spontaneous emission. Although the details of the calculations differ, investigations of the minimum noise temperature due to this effect yield values of the order of $h\nu/k$ for both the maser and the parametric amplifier.

The maser and the parametric amplifiers are phase preserving amplifiers, or, in the terminology employed here, linear amplifiers. They have been characterized broadly as voltage amplifiers. There is another type of amplifier which does not preserve the signal phase which can be classed as a quantum counter. Weber[2,5] has proposed two forms of such amplifiers and has pointed out that they have no spontaneous emission noise. Thus these phase-insensitive counters have a zero-limiting noise temperature. This result opens the question of whether there are possible forms of linear amplifiers—linear in the sense of phase preserving—which also have a limiting noise temperature approaching zero.

In order to settle this question one needs to look for a general physical principle which will apply to all amplifiers regardless of the details of the specific amplification process. An appropriate one is the uncertainty principle of quantum mechanics which forms one of the basic postulates of quantum theory. Although it has resulted in important limitations on the accuracy of measurements possible in atomic systems, its corresponding limitations on the accuracy of signal measurements have until recently gone unnoticed. At the 1959 Quantum Electronics Conference, Serber and Townes[6] investigated the role of the uncertainty principle in maser noise and Friedburg[7] considered its implication on the noise figure of a general amplifier. Each of these papers is open to some criticism. First, the uncertainty principle as generally interpreted is a statement about the results obtained in a physical measurement. As such it can be applied to a signal detector but not to an amplifier. An amplifier is not a measuring instrument which produces a set of data. Amplification is rather a process, a transformation of the signal. The uncertainty principle can be applied to the measurement of the results of that processing but not directly to the processing itself. A second objection to both papers concerns their lack of rigor in problems of statistical averaging, particularly in regard to the phase of a signal. Finally, in the case of Friedburg, the wrong constant was used in the statement of the uncertainty principle. As implied before, Friedburg's conclusions strictly speaking apply only to a detector, not to an amplifier, and as such do not indicate what class of amplifier falls under the uncertainty limitation, nor how this limitation is affected by the gain.

This purpose of this paper is to develop in a simple but rigorous fashion the limitations on the noise performance of linear amplifiers which are implied by the uncertainty principle. First it will be shown that there is no such thing as a noiseless linear amplifier. Next, by characterizing the amplifier as adding white noise, the minimum possible noise temperature is derived. The resulting expression when the gain is large is exactly that derived for the limiting noise performance of the maser and the parametric amplifier.

* Received January 8, 1962; revised manuscript received, April 30, 1962.

† Stanford Electronics Laboratories, Stanford University, Calif.

[1] K. Shimoda, H. Takehashi, and C. H. Townes, "Fluctuations in amplification of quanta with application to maser amplifiers," *J. Phys. Soc. Japan*, vol. 12, pp. 686–700; June, 1957.

[2] J. Weber, "Maser noise considerations," *Phys. Rev.*, vol. 108, pp. 537–541; November, 1957.

[3] M. W. Muller, "Noise in a molecular amplifier," *Phys. Rev.*, vol. 106, pp. 8–12; April, 1957.

[4] W. H. Louisell, A. Yariv, and A. E. Siegman, "Quantum fluctuations and noise in parametric processes," *Phys. Rev.*, vol. 124, pp. 1646–1654; December, 1961.

[5] J. Weber, "Masers," *Rev. Mod. Phys.*, vol. 31, pp. 681–710; July, 1959. See also N. Bloembergen, "Solid-state infrared quantum counter," *Phys. Rev. Lett.*, vol. 2, pp. 84–85; February, 1959.

[6] R. Serber and C. H. Townes, "Amplification and Complementarity," in "Quantum Electronics," C. H. Townes, Ed.; Columbia University Press, New York, N. Y., pp. 233–255; 1960.

[7] H. Friedburg, "General amplifier noise limit," in "Quantum Electronics," C. H. Townes, Ed.; Columbia University Press, New York, N. Y., pp. 228–232; 1960.

The Uncertainty Principle

The uncertainty principle, first formulated by Heisenberg in 1927, claims that complete accuracy is impossible to obtain in the simultaneous measurements of certain physical quantities. In its most familiar form, it asserts the fundamental inaccuracy which must result in the simultaneous measurement of a particle's momentum p and position x. If we define the uncertainty in measurement to be the rms deviation from the mean in the distribution obtained from an ensemble of measurements, then the uncertainty in the measurement of momentum Δp and the uncertainty in the measurement of position Δx are related by

$$\Delta p \Delta x \geq h/4\pi. \tag{1}$$

This relation can be interpreted in the following way. The process of measuring cannot be divorced from the physical process being measured. Not only does the act of observing affect the system being observed, but it does this in a way which cannot be precisely predicted. It is this quality of unpredictability which formed the new content of the uncertainty principle.

In its most general form, the uncertainty principle applies to measurements of any two canonically conjugate quantities[8] such as, for example, the energy of a system and the precise time at which the system possesses this energy,

$$\Delta E \Delta t \geq h/4\pi. \tag{2}$$

Still another form of the principle applies to the measurement of the number of quanta in an oscillation and its phase

$$\Delta n \Delta \phi \geq \tfrac{1}{2}. \tag{3}$$

This latter statement of the principle may be made plausible by substituting the relations $E = nh\nu$ and $\phi = 2\pi\nu t$ into the preceding equation. We shall use this last formulation of the uncertainty principle to derive two basic theorems on amplifier noise performance.

The Unavoidable Noisiness of Linear Amplifiers

The first conclusion which emerges from the uncertainty principle is: *It is impossible to construct a noiseless linear amplifier*. We can prove this statement by postulating the existence of a noiseless linear amplifier and then showing that it violates the uncertainty principle.

Suppose we have a perfect linear amplifier by which we mean the following. If during any given interval we measure the number of photons n_2 produced at the output, we find that it is related to the number of input photons n_1 by a constant G, the gain of the amplifier.

$$n_2 = Gn_1. \tag{4}$$

Secondly, if we measure the phase ϕ_2 of the output, we find that it is equivalent to the input phase ϕ_1 with perhaps the inclusion of an additive phase shift θ.[9]

$$\phi_2 = \phi_1 + \theta. \tag{5}$$

Such an amplifier is linear in the sense that the phase is preserved and the output quanta are linearly related to the input quanta. It is perfect in that no noise is added. Note that frequency converters which derive their gain solely by the frequency conversion factor do not fall under this definition of a linear *amplifier* in that the photon gain G is unity.

Let us now attach to the perfect linear amplifier an ideal detector, ideal in the sense that it is capable of detecting the number n_2 of output photons and the output phase ϕ_2 within an uncertainty,

$$\Delta n_2 \Delta \phi_2 = \tfrac{1}{2}, \tag{6}$$

the minimum value allowed by the uncertainty principle. (See Fig. 1.) Thus we imagine that we make a measurement of the output photons and phase which together with the uncertainties introduced by the detector we write symbolically as $(n_2 \pm \Delta n_2)$ and $(\phi_2 \pm \Delta\phi_2)$.

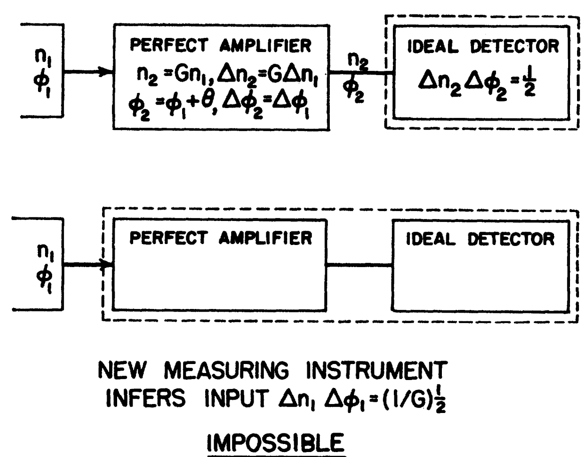

Fig. 1—Thought experiment to show the nonexistence of a perfect amplifier.

[8] In mathematical form, the uncertainty principle states that if the operators A and B which represent physical observables a and b satisfy the commutation relation

$$AB - BA = iC,$$

then the uncertainties in the measurement of a and b satisfy the relation

$$\Delta a \Delta b = \tfrac{1}{2}\langle C \rangle,$$

where the term uncertainty stands for the root mean square deviation from expectation value, *e.g.*,

$$\Delta a = [\langle A^2 \rangle - \langle A \rangle^2]^{1/2}.$$

[9] These are operationally valid constructs since we can prepare two signal sources, one of which has an accurately known photon output and the other of which has an accurately known phase. If each of these is applied in succession to the amplifier input and if in succession two detectors are applied to the amplifier output, one of which measures number of photons without giving phase information and another which measures phase without giving photon number information, we can determine the constants of the amplifier to arbitrary precision. Such a determination can, in fact, be made with a single signal source and a single detector if a sufficiently large signal level is used since the accuracy in the determination of phase and the relative accuracy in the simultaneous determination of photon number are each proportional to $1/n^{1/2}$.

We now change our point of view and look upon the combination of perfect amplifier and ideal detector as a single measuring instrument. The measurement of $(n_2 \pm \Delta n_2)$ and $(\phi_2 \pm \Delta\phi_2)$ when referred to the input of the amplifier implies an input number of photons $(n_1 \pm \Delta n_1) = 1/G(n_2 \pm \Delta n_2)$ and an input phase $(\phi_1 \pm \Delta\phi_1) = (\phi_2 \pm \Delta\phi_2) - \theta$. Thus the uncertainty in the measurement of input photons and phase is such that

$$\Delta n_1 \Delta\phi_1 = \frac{1}{G} \cdot \frac{1}{2}. \tag{7}$$

This conclusion is clearly impossible since it violates the uncertainty principle. Therefore our postulated perfect linear amplifier cannot exist. It must add some uncertainty, that is, noise.

Minimum Noise Temperature of a Linear Amplifier

We can pursue this argument even further to prove the following result. *The minimum noise temperature of a linear amplifier characterized by additive white Gaussian noise is*

$$T_n = \left[\ln \frac{2 - 1/G}{1 - 1/G}\right]^{-1} \frac{h\nu}{k}, \tag{8}$$

and the minimum mean square phase fluctuation is

$$\Delta\phi_a^2 = \frac{(G-1)h\nu B}{2P}. \tag{9}$$

Here h is Planck's constant, ν is the frequency, k is Boltzmann's constant, B is the bandwidth and P is the signal power.

The proof of this assertion employs the same conceptual scheme used previously, a linear amplifier (no longer considered noiseless) followed by an ideal detector. Again a measurement is made by the detector of the output phase ϕ_2 and photon number n_2. This measurement will include as before the same uncertainties introduced by the detection process but will also include the uncertainties which we have seen must be added by the amplifier. Let us label the uncertainties introduced by the detector Δn_d and $\Delta\phi_d$ and those added by the amplifier Δn_a and $\Delta\phi_a$. If we assume the processes which give rise to these two sets of uncertainties are uncorrelated, then the total uncertainties as actually measured, Δn_2 and $\Delta\phi_2$, can be obtained from

$$\Delta n_2^2 = \Delta n_a^2 + \Delta n_d^2$$
$$\Delta\phi_2^2 = \Delta\phi_a^2 + \Delta\phi_d^2. \tag{10}$$

These equations merely assert that the variances of two uncorrelated random processes add.

Let us again shift our point of view and look upon the combination of amplifier and detector as a single measuring instrument. The measured uncertainties Δn_2, $\Delta\phi_2$ now imply an uncertainty in the measurement of the input phase and photon number given by

$$\Delta n_1^2 = \frac{1}{G^2}(\Delta n_a^2 + \Delta n_d^2)$$
$$\Delta\phi_1^2 = \Delta\phi_a^2 + \Delta\phi_d^2. \tag{11}$$

Our proof proceeds by first demanding that the inferred input uncertainties be the least allowed by the uncertainty principle, *i.e.*, $\Delta n_1 \Delta\phi_1 = \frac{1}{2}$. This condition insures that the amplifier uncertainty is the smallest possible. We then characterize this amplifier uncertainty by white Gaussian noise and determine the noise temperature corresponding to this minimum uncertainty.

First, however, we must make sure that the detector is matched to the amplifier, for although the product of the uncertainties $\Delta n_d \Delta\phi_d$ is set, their ratio is not. We can assure the best detection performance by minimizing the product $\Delta n_1 \Delta\phi_1$ given by (11) with respect to the ratio $(\Delta n_d/\Delta\phi_d)$ while still demanding that

$$\Delta n_d \Delta\phi_d = \tfrac{1}{2}. \tag{12}$$

This process results in the relation

$$\frac{\Delta n_d}{\Delta\phi_d} = \frac{\Delta n_a}{\Delta\phi_a} \tag{13}$$

which simply states that the minimum over-all uncertainty comes about when the detector measures number and phase with the same relative uncertainties as those introduced by the amplifier.

We must make sure that the detector is matched to the amplifier in another sense. Let us assume that the amplifier has a bandwidth B. This characteristic implies that the detector should sample the output at intervals of $\tau = \frac{1}{2}B$. If the interval is made longer, the full information transmission capabilities of the amplifier are not being used, while if it is shorter, not only are some of the data redundant, but also, because of the fluctuations in the output, the uncertainties are greater than necessary. Thus the matching of the detector to the amplifier implies two things, first that the ratios of the uncertainties are made equal and second that the time interval over which the number of output photons and the phase are detected is one half of the reciprocal bandwidth.

Let us now multiply (11) together, introduce the conditions of (12) and (13), and demand that the uncertainty in the measurements referred to the input be the minimum allowed by the uncertainty principle, that is

$$\Delta n_1 \Delta\phi_1 = \tfrac{1}{2}. \tag{14}$$

The resulting equation is

$$\frac{G^2 - 1}{4} = \Delta n_a^2 \Delta\phi_a^2 + \Delta n_a \Delta\phi_a. \tag{15}$$

If we put this equation in the form

$$\left(\frac{\Delta\phi_a}{\Delta n_a}\right)^2 \Delta n_a^4 + \left(\frac{\Delta\phi_a}{\Delta n_a}\right) \Delta n_a^2 - \left(\frac{G^2 - 1}{4}\right) = 0 \tag{16}$$

we can solve for Δn_a^2 in terms of the ratio $(\Delta n_a/\Delta\phi_a)$ to give

$$\Delta n_a^2 = \frac{G-1}{2}\left(\frac{\Delta n_a}{\Delta\phi_a}\right). \tag{17}$$

Little more can be done unless we specify the nature of the amplifier uncertainty. Let us characterize the uncertainty as additive white noise. The statistical properties of a signal contaminated by white Gaussian noise have been extensively studied.[10] One result is that the probability density function for the output phase approaches a Gaussian distribution for large SNR's and has a variance given by

$$\Delta\phi_a^2 = \frac{\Delta P}{2P}\,. \tag{18}$$

Here ΔP is the noise power and P is the signal power. For large SNR's the variance in the power distribution is given by

$$\Delta P^2 = 2P\Delta P \tag{19}$$

so that (18) becomes

$$\Delta\phi_a^2 = \frac{\Delta P^2}{4P^2}. \tag{20}$$

Since the integration time of the amplifier is $\tau = \frac{1}{2}B$, (20) can be put in terms of the variance of the number of photons Δn_a^2 detected during this interval

$$\frac{\Delta n_a}{\Delta\phi_a} = \frac{2P\tau}{h\nu}\,. \tag{21}$$

This result inserted in (17) together with the relation

$$\Delta n_a^2 = \frac{\Delta P^2\tau^2}{(h\nu)^2} = \frac{2P\Delta P\tau^2}{(h\nu)^2} \tag{22}$$

gives for the minimum noise power introduced in the output of the amplifier

$$\Delta P = (G-1)h\nu B. \tag{23}$$

The effective noise temperature T_n is obtained by dividing by the amplifier gain to refer the noise power to the input and then determining what temperature is required for a black body to radiate the same power. That is, we must find the value T_n for which

$$\frac{h\nu B}{e^{h\nu/kT_n}-1} = \left(1-\frac{1}{G}\right)h\nu B. \tag{24}$$

The result is

$$T_n = \left[\ln\frac{2-1/G}{1-1/G}\right]^{-1}\frac{h\nu}{k}\,. \tag{8}$$

According to (19) and (23), the minimum mean square phase fluctuations are

$$\Delta\phi_a^2 = \frac{(G-1)h\nu B}{2P}\,. \tag{9}$$

These last two equations give relations for the best possible noise performance of any linear amplifier whose uncertainty is characterized by white Gaussian noise. It is interesting to note that in the limit of high gain, the minimum noise temperature becomes

$$T_n = \frac{1}{\ln 2}\frac{h\nu}{k}, \tag{25}$$

which is precisely the value obtained for the minimum noise temperature of the maser and the parametric amplifier derived from detailed consideration of the amplification mechanisms in the two cases. Thus we can say that both the maser and the parametric amplifier represent ideal amplifiers in so far as their ultimate noise performance is concerned.

We should also remark in passing, (8) and (9) indicate that the parametric up-converter which has power gain by virtue of the change in frequency but has unity photon gain $(G=1)$, possesses a limiting noise temperature and phase uncertainty of zero. This result is in agreement with the detailed calculations of Louisell, Yariv and Siegman.[4] This sort of amplifier which does not multiply the *number* of photons, however, does not improve the capability of detecting a signal. The accuracy of the detection process at the output of the up-converter is no better than if it were performed at the input. The change in photon frequency and hence energy is immaterial since the limiting detector uncertainty is dependent solely on the number of photons arriving, not upon their energy.

The Implications of the Minimum Noise Limit

At conventional communications frequencies, the minimum noise temperature given by (8) is entirely negligible. At optical frequencies though, it can amount to several tens of thousands of degrees. Such a number is misleading, however, for it implies that the insertion of even the best amplifier has seriously degraded our ability to detect a signal. We can show that such is not the case by recasting the results of the previous section in slightly different form. Eqs. (22) and (23) can be combined to give the uncertainty produced by the amplifier in the number of photons as

$$\Delta n_a^2 = (G-1)n_a, \tag{26}$$

[10] S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282–332, January, 1944; vol. 24, pp. 46–156, January, 1945; "Statistical properties of a sine-wave plus random noise," *Bell Sys. Tech. J.*, vol. 27, pp. 109–157, January, 1948.

and (9) can be rewritten in the form

$$\Delta\phi_a^2 = \frac{(G-1)}{4n_a} \tag{27}$$

to give the uncertainty in phase produced by the amplifier. If we refer these quantities to the input we have

$$\Delta n_{ai}^2 = \frac{1}{G^2}\Delta n_a^2 = (1 - 1/G)n_1 \tag{28}$$

and

$$\Delta\phi_a^2 = \frac{(1-1/G)}{4n_1}. \tag{29}$$

From (26) and (27) we see that if the amplifier gain is high, the output uncertainties introduced by the amplifier are considerably larger than those of even a poor matched detector. In this case, the total uncertainty Δn_1 in the inferred measurement of the input photon number and $\Delta\phi_1$ in the inferred measurement of input phase are closely given by (28) and (29), which relate to the amplifier alone. In the limit of high gain, the product of these uncertainties is

$$\Delta n_1 \Delta\phi_1 \cong \Delta n_{ai}\Delta\phi_a \cong \tfrac{1}{2}. \tag{30}$$

This is, of course, the minimum value allowed by the uncertainty principle. Thus the minimum noise amplifier allows us to use a poor detector, one which introduces uncertainties considerably larger than the minimum necessary, and still measure an incoming signal with an accuracy approaching the best allowed by the uncertainty principle. There still remains a question of what limitation is put on the rate of information transmission by this maximum allowable accuracy of detection. The answer to this question, however, must await the development of a quantum theory of communication.

Negative L and C in Solid-State Masers*

R. L. KYHL†, MEMBER, IRE, R. A. McFARLANE‡, MEMBER, IRE, AND M. W. P. STRANDBERG§, FELLOW, IRE

***Summary*—The analysis of solid-state cavity masers is extended to include the reactive component of the paramagnetic resonance. This reactance is inverted (in opposition to Foster's reactance theorem). A two-cavity network makes use of this negative frequency dependence of reactance to obtain a broad-band flat-topped amplifier response. In verification of this theory a ruby maser has been built which has a 95-Mc bandwidth at 14-db gain and operates at 9000 Mc and 1.5°K. This performance is comparable to that of published, tapered magnetic field traveling-wave masers. General network limitations on cavity maser amplifiers are derived. Broadbanding techniques that have been published for parametric amplifiers are essentially equivalent. The tuning of the broad-band amplifier is critical. The same performance can be achieved in a unilateral transmission maser by using circularly polarized cavities, but the problem of circuit design and tuning with the increased number of parameters has thus far prevented successful operation.**

* Received February 19, 1962; revised manuscript received, May 1, 1962. This work was supported in part by the U. S. Army Signal Corps under Contract DA36-039-sc-87376, the Air Force Office of Scientific Research, and the Office of Naval Research.

† Department of Electrical Engineering and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass.

‡ Bell Telephone Laboratories, Inc., Murray Hill, N. J. Formerly with Research Lab. of Electronics, M.I.T.

§ Fulbright Lecturer, University of Grenoble, France, 1961–1962; on leave from the Dept. of Physics and Research Lab. of Electronics, M.I.T.

I. Introduction

THE STIMULATED emission behavior of the active material in a solid-state maser can be characterized satisfactorily by its contribution to the complex electric or magnetic susceptibility of the material. (Beam masers are somewhat more complicated in this respect.) Typically this susceptibility shows a sharp resonance at a frequency corresponding to the quantum transition involved. The imaginary component of the susceptibility is, of course, responsible for the maser gain, but the real, or reactive, component must also be present. In narrow-band systems this reactance may be masked by the larger reactive effects of the microwave cavity or circuits. As larger gains and bandwidths are obtained with better substances and design configurations, the reactive component must be taken into consideration to obtain a correct analysis of the circuit behavior. When the population distribution between the quantum levels is inverted to achieve maser amplification, both components of the susceptibility reverse sign. The resulting frequency dependence of reactance corresponds to the situation that would obtain in conventional circuit analysis if the symbols