# Noise—confusion in more ways than one

## 1—Thermal noise and terminology

by K. L. Smith

*University of Kent at Canterbury*

Random fluctuations at the limit of sensitivity of radio receivers, amplifiers and other instruments that process tiny signals have become ever more significant in recent times. Because no single number exists for complete characterization of a system or device under all conditions, difficulties and confusion often arise when the quantities which have been suggested are misunderstood and used incorrectly. The ideas of the noise figure and noise temperature are discussed in this four-part article, with practical hints on how they can be measured with a good chance that the values obtained will bear some approximation to the truth. At first sight, the many special terms seem to have little in common: this first article looks at some of the basic ideas.

While sitting on the banks of our lake in the woods near the University Electronics Laboratory, the students' group discussion turned to noise. Not the airport or motorway sort—nothing could have been further from our thoughts, in the rather idyllic setting we are fortunate enough to have. Lest readers become incensed at this point with visions of lazy good-for-nothing students sitting about and are tempted to demand their taxes back from the Exchequer, I hasten to add that we were working—quite hard, trying to sort out a little order in that maze of a subject, electronic noise and its effect on the ultimate sensitivity of communications systems.

At one point in this discussion I asked, "What temperature would you observe using a microwave receiver whose horn aerial was pointing up at the clear sky, with the waveguide plumbing at around room temperature?" A quick reply was forthcoming, "Oh, about 300K". Such a high value for the received temperature of a system forming the subject of my question was incorrect and I said so. "What is it then?" was asked and I replied "Around 10K". "What, within a few degrees of absolute zero and the waveguide at 300K!" chorused back.

This serves to illustrate one of the first elementary fallacies found in the subject of noise in systems—that the temperatures everyone talks about must have something to do with physical temperatures. Following the line to show up this fallacy, I then pointed to another part of the sky and said, "Over there the same receiver could register an aerial temperature up to perhaps 10,000K". I was indicating in a direction towards the sun, of course.

The point is that the receiver measures the source effective temperature, not the actual physical temperature of the aerial components. To help drive the point home, the question was raised as to the temperature that would be seen if the horn "looked down" at a sheet of metal on the ground, which was therefore at the physical temperature of the earth (in other words, about 300K). The answer is the same as before, 10K.

The metal is so nearly a perfect reflector at centimetre wavelengths, that it is simply the sky that is seen. The sheet of metal hardly absorbs any r.f. signal power, therefore it hardly radiates any energy, even though it is at about 300K. The metal is anything but the "black body" of Max Planck, radiating at 300K. As it is such a poor absorber it is a very poor emitter indeed, its effective emission temperature probably being around a hundredth of a degree or less.

### The noise figure muddle

The debate around the lake came to an end at this point, and with parting reminiscences of black bodies and Kirchoff's radiation laws reminding us that we had studied a little physics in the dim and distant past, thoughts turned to how much we owe to thermodynamics for the ideas which have enabled us to realize the ultimate limits of small-signal reception and to build low-noise equipment for handling the tiny signals near these limits, so often part of modern radar, satellite and radio-astronomical systems.

When looking at noise problems in communications one finds that misconceptions and confusion abound in the subject. Even the professionals cannot agree. William Mumford and Elmer Scheibe[1] noted that no less than nine definitions of "noise figure" exist in the literature, so pity the poor student or junior engineer! Then we have all that talk about "temperatures". It is worth noting the range of quantities and concepts that exist around the subject of electronic noise, so that as these articles proceed we stand a good chance of stripping some of the fuzziness away from these ideas and show how they relate and how some of them have arisen.

A cursory glance into the literature always turns up the well-known quantity $F$, the noise figure. Or is it noise factor? It was H. T. Friis[2] who defined and named $F$ the noise figure. D. O. North defined what amounted to the same thing in a paper published at about the same time, only he called it noise factor. Therefore noise figure and noise factor have the same meaning. One meets $F$ as a simple ratio or as $10\log_{10}F$ (i.e. noise figure in decibels).

The concept of excess noise figure is met. It turns out to be simply $F-1$. The spot noise figure is the figure defined at one frequency, whereas the average noise figure, sometimes written $F$, is that effective over the whole bandwidth under discussion. A concept which may come into more general use is the operating noise figure, $F_{op}$ defined by Dr North as long ago as 1942[3].

Turning now to the idea of using absolute temperatures to discuss noise performance, one finds a bewildering array of signs and symbols, but paradoxically, working with temperatures is actually more straightforward than the confusing noise figure muddle. If the common use of noise temperatures had developed before noise figures as a concept, we may have been saved the duplication. Temperature is fundamental, being related to (thermal) energy and power. Noise temperatures are especially convenient when dealing with low-noise receivers and amplifiers, and the use of noise temperature is slowly taking over as a parameter in performance measurements.

Other than the concept of a general noise temperature (represented by $T_N$, $T_i$, etc.) one finds in particular the aerial temperature, $T_a$; the effective input noise temperature, $T_e$; a standard reference temperature, $T_o$ (we shall see that everyone has agreed to 290K for $T_o$, after some persuasion by the American IEEE). There is the excess noise temperature, which turns out to be $T_N-T_o$, or the number of degrees in excess of 290K, the standard room temperature. From this, a quantity known as the excess noise ratio is obtained,

$(T_N - T_o)/T_o$, which is written $t-1$, so that $t$ is the noise ratio, $T_N/T_o$. Finally, there is the concept of an operating noise temperature, $T_{op}$. Sometimes the system noise temperature, $T_{sys}$, is used instead of $T_{op}$ in some articles.

The fact that there has been this proliferation of concepts and quantities, shows that either there are hidden subtleties in the subject, or that some of the definitions are unsatisfactory, or both. Even so, there are one or two other definitions that are vital to an understanding of noise problems, but they are much more general and useful in other contexts as well. These are the ideas of the available power from a source and the related available power gain of an amplifier stage. (Or loss, if the gain is less than one, as in an attenuator.) Noise and signal bandwidths ($B_N$ and $B_s$) are also of importance.

## Thermal agitation

Ever since the fall of the old caloric theory centuries ago and the subsequent rise of the mechanical theory of heat as the energy of the random jostling of the molecules in substance, it was suspected that the warmth of an object might set the limit to the accuracy of measurements on it. This was found to be so in examples like the Brownian motion and the jumping around of the light spot of very sensitive galvanometers.

Thermal noise in electronic devices has received a great deal of attention since the theoretical discussion of H. Nyquist[4] and the experimental work of J. B. Johnson[5] was published in 1928. You can follow this up a little in Cathode Ray's articles, "Heads, Tails and Noise"[6] and, "More about Noise"[7] of some time ago. It is worth quoting a few opening remarks from J. B. Johnson's paper: "... a phenomenon has been described which is the result of spontaneous motion of the electricity in a conducting body. The electric charges in a conductor are found to be in a state of thermal agitation, in thermodynamic equilibrium with the heat motion of the atoms of the conductor. The manifestation of the phenomenon is a fluctuation of potential difference between the terminals of the conductor which can be measured by suitable instruments. . . ." The value of the mean of the voltage squared is found to be

$$\overline{v^2} = 4kTR(f_2 - f_1)$$

and the power

$$w = \frac{\overline{v^2}}{R} = 4kT(f_2 - f_1).$$

For an amplifier operated at room temperature and covering the approximate voice frequency range of 5kHz, this power is $0.82 \times 10^{-16}$ watt.

We have here the fundamental equation for the open-circuit noise e.m.f. derived experimentally by Johnson and theoretically by Nyquist

$$\overline{v^2} = 4kTRB \qquad \ldots . 1$$

in which $T$ is the absolute temperature of a resistor $R$, $B$ is the noise bandwidth in Hz, and $k$ is Boltzmann's constant (see Cathode Ray's discussion of "$k$" in *WW* November 1960[8]).

Dr Nyquist's derivation of equation 1 is now standard bookwork, see for instance, Robinson[9]. An interesting point is that equation 1 predicts a uniform output over the entire frequency spectrum. This gives rise to the expression white noise by analogy to white light (all frequencies present), Fig. 1. The presence of the thermal noise e.m.f. across a resistor means that power can be drawn by a load connected across it and the well-known equivalent circuit shown in Fig. 2 enables us to calculate the maximum, in this case thermal noise power, that can be drawn. From Fig. 2,

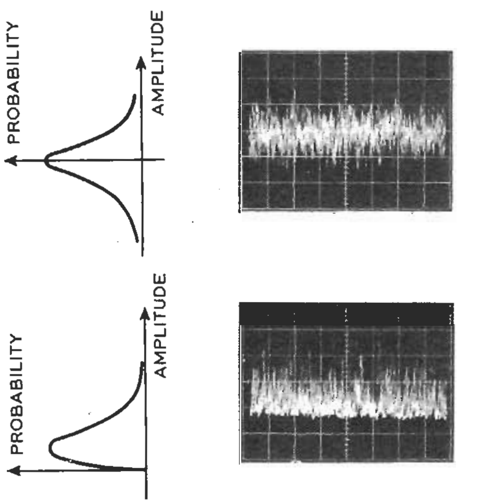



*Fig. 1. White noise has a Gaussian distribution (upper oscillogram). After band limiting, white noise typically shows a Rayleigh distribution (lower oscillogram).*

$$\overline{i^2} = \frac{\overline{v^2}}{(R+R_L)^2}$$

and the power dissipated in $R_L$ is $\overline{i^2}R_L$. If we now make $R_L = R$, we have the matched condition and the maximum power is drawn.

$$N = \frac{\overline{v^2}}{4R} = \frac{4kTRB}{4R} = kTB \quad \ldots\ldots\ldots 2$$

In this equation $N$ is the noise power under matched conditions. Of course, in thermal equilibrium, the two resistors feed this much power to each other so no net transfer of energy takes place. This balance is also true for every frequency band, as well as overall. Suppose for a moment that this was not so, and imagine a tuned circuit in parallel with the two resistors, which acts as a selective filter, then the resistor which had a higher output at one frequency band could continually supply power to the other even though the temperatures of the two are equal. Thermodynamics has something quite definite to say about the impossibility of doing that, so the balance is maintained across the whole spectrum. The $kTB$ expression is called the available noise power. Equation 2 shows that the available noise power is directly proportional to the bandwidth $B$ and absolute temperature $T$, but is independent of the value of $R$ and of the frequency.

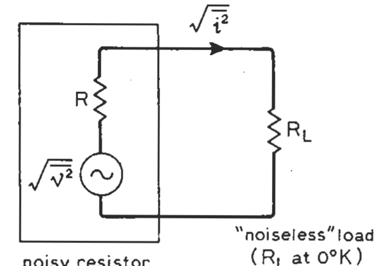



*Fig. 2. $R_L$ is absorbing power from R. Maximum power is absorbed when $R = R_L$, according to the well-known matching theorem.*

The above arguments about constancy with frequency begin to fail for frequencies around 1000GHz at room temperature. This frequency is much higher than the radio frequency spectrum in current use. The error is about 1%. But at low temperatures, say 1K, a 10% error already would exist at 10GHz. This is because the correct expression for thermal noise derived rigorously from thermodynamics is

$$\overline{v^2} = \frac{4RBhf}{\exp(hf/kT) - 1}$$

What all this means is that at very high frequencies and/or very low absolute temperatures we are running into quantum effects. If you glance into a mathematics textbook, exp $x$ or $e^x$ will usually be found expanded into a power series

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{2.3} + \ldots$$

If our $x$, which is $hf/kT$ is small the quantity $\exp(hf/kT)$ can be taken for all intents and purposes equal to the first two terms, that is $1 + (hf/kT)$. Putting this into the equation, cancelling, etc., immediately gives Nyquist's simple form for $\overline{v^2}$.

## Amplifiers and available gain

If amplifiers added no noise to signals they were processing, no problems would exist about gain and how much of it we could use. Also, it stands to reason that no amount of gain will put out a signal already buried in a large amount of noise. What is important is to amplify a weak signal already in noise so that it can resist further degradation by interference and while amplifying, to add the minimum amount of extra noise. Two requirements are needed for front-end amplifiers then: a high gain, and a low-noise performance. Even with the best amplifiers the output signal-to-noise ratio is worse than that at the input, because some noise will always be added by the circuit components.

There is some difficulty about what is meant by gain, which must be cleared up. Definitions of various gains abound in the electronics literature. There is the voltage gain, various power gains, current gain and so on. The usual requirement is to boost the power level of a weak signal so that it can operate fairly energetic transducers such as loudspeakers, pen recorders, etc.

Clearly, power gain is the idea we want. But, which power gain? If an amplifier has an extremely high input impedance, the input power is nearly zero. With a few watts output, a simple ratio of output power to input power is getting on for infinity! The best general definition to use is termed available power gain, $G_A$, and as can be seen from Fig. 3, is defined as $G_A =$

$$G_A = \frac{\text{available power from the output terminals}}{\text{available power from the input signal source}}$$

The point to remember now is that available power is the maximum that can be taken from a source or generator. In other words it is the matched power output into a load connected to a generator, as I mentioned earlier for noise power. It is at this stage that subtleties inviting confusion tend to crop up. The available power from the generator feeding some amplifier is, by definition, a fixed quantity depending only on the generator and its internal impedance.

It does not depend on the matching or otherwise at the input circuit. Similarly, the definition of output available power from an amplifier does not depend on the value of the load impedance. This is not to say that available power gain is independent of all matching conditions, because available output power is dependent on input matching. (In effect, available power from the source feeding the amplifier is a constant of the source, but actual power going into the input terminals does depend on matching.) Therefore, the definition of available power out includes the effects of any mismatching at the input, the conditions of which must be stated in the specifications of any particular case.

M. S. Gunston[10] wrote a critique on the use of available power gain and attempted to introduce a mismatch factor $M$, because of "errors" by using available gain ideas. I cannot agree and consider introducing more factors just adds complications. In the special case of complete input circuit matching, maximum available power gain is obtained. The idea of available power gain is versatile and useful. One or two cases are considered in the Appendix for readers who do not mind a few numbers. Whenever you see gain mentioned, I am talking of available power gain under the input conditions prevailing.

*To be continued.*

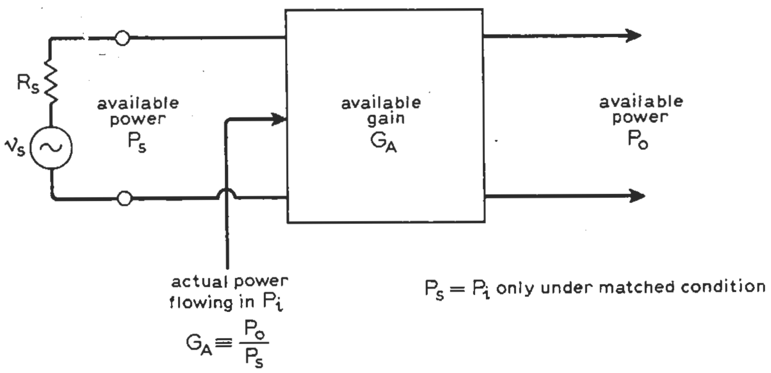



*Fig. 3. Available power gain has no direct connection with matching, but maximum available power gain is only obtained when $R_s$ matches the input resistance of the amplifier.*

## Appendix A

Available power

A very useful property resulting from the definition, is that available power from a source is unaltered if we put a network of reactances after it. On the other hand, a resistive network will change the available power, in practice always reducing it below the original value. You can see this by considering the examples shown in Fig. A1. The first shows a series reactance inserted after the generator. Available power from the generator on its own is $E_g^2/4R_g$. All we have to do is cope with $+jX$ and the available power is again drawn.

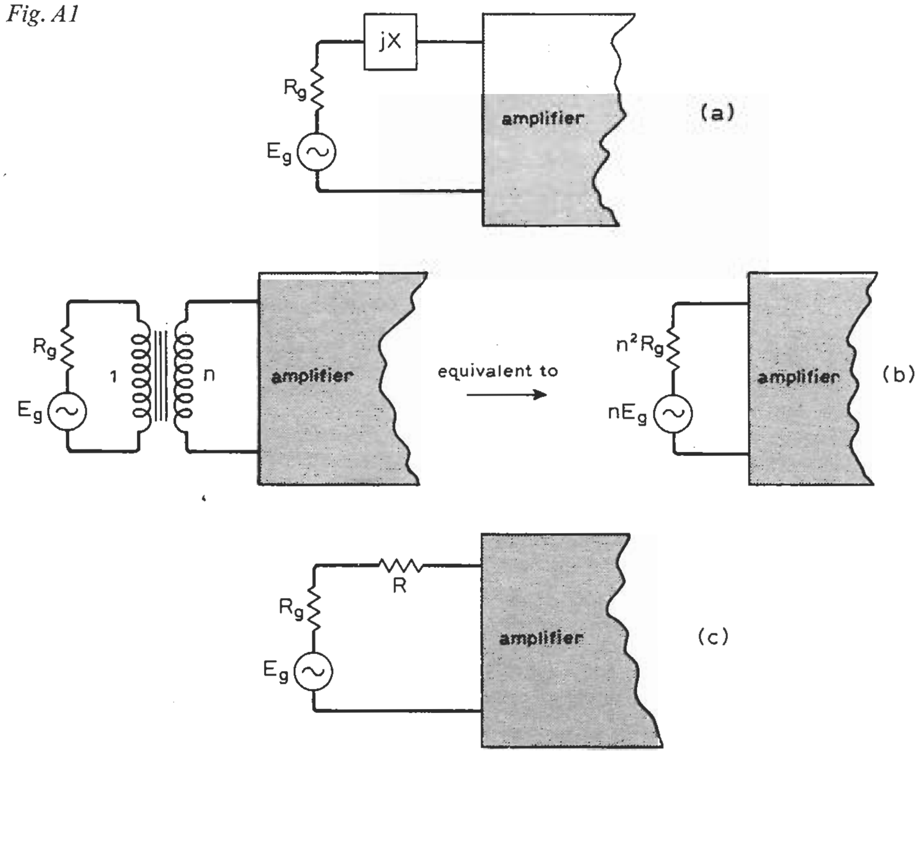
Fig. A1(b) shows the case of an ideal transformer with a step up ratio of 1:$n$ The output voltage from the secondary will be $nE_g$, and the effective source impedance will be $n^2R_g$ from transformer theory. Available power will be $(nE_g)^2/4n^2R_g$, which is as before. The last example shows that if a series resistance $R$ is placed after the generator, as in Fig. A1(c), available power is $E_g^2/4(R_g+R)$ which is obviously less than $E_g^2/4R_g$.

### Available power gain

Working from the definition for $G_A$ above, we can write down the two available powers, take the ratio and thereby obtain $G_A$. As an example, Fig. A2(a) shows a voltage amplifier with input resistance $R_{in}$ fed via a reactive network from a generator whose e.m.f. is $E_g$, with real internal impedance $R_g$. (Any imaginary part can be taken into the reactive network.) From network theory, any arrangement of reactances can be reduced to an equivalent transformer with a series resistance, for a shunt susceptance. Resulting from this you will see that the equivalent circuit shown in Fig. A2(b) can be drawn. Going one stage further in the simplification we arrive at Fig. A2(c), where the transformation of the generator voltage and internal resistance to $nE_g$ and $n^2R_g$ is shown. The magnitude of the voltage appearing across the terminals of the amplifier is

$$E_{in} = i_{in}R_{in} = \frac{nE_gR_{in}}{\sqrt{(n^2R_g + R_{in})^2 + X^2}}$$

*Fig. A1*



Available power output will be some constant of the amplifier, times $E_{in}^2$; or, what amounts to the same thing, it will be proportional to the square of the input terminal voltage.

$$\therefore P_{A(out)} = K \frac{E_g^2 R_{in}^2}{\left(nR_g + \frac{R_{in}}{n}\right)^2 + \frac{X^2}{n^2}}$$

Notice that I have deliberately rearranged $n$. As available input power is simply $E_g^2/4R_g$,

$$G_A = \frac{P_{A(out)}}{P_{A(in)}} = K' \frac{R_{in}^2 R_g}{\left(nR_g + \frac{R_{in}}{n}\right)^2 + \frac{X^2}{n^2}}$$

where $K'$ is some constant.

Straightaway, you can see that for *maximum* available power gain, $X$ should equal zero.

This means that any residual reactance should be tuned out at the front end. The term $[nR_g + (R_{in}/n)]^2$ is left in the denominator and for maximum available gain this should be minimized. If $n$, the transformer ratio, is very large the first term in the bracket dominates and $G_A$ is small. If $n$ is tiny, the second term becomes large and again $G_A$ is small. Somewhere between these extremes an optimum value for $n$ occurs to minimize the bracketed term, and this gives the largest $G_A$. By using calculus we can easily show there is a minimum when $n^2 = R_{in}/R_g$. This is the matching condition and gives the largest available power gain, as would be expected.

The whole idea of available gain is valid under any conditions, not just matching, as long as the conditions existing are given in any example. (In the example discussed here, these conditions would be the values of $R_{in}$, $R_g$, $X$ and $n$.)

Suppose now we consider a chain of amplifier stages with available gains $G_1$, $G_2$, $G_3$, etc., each under the source conditions offered by the preceding amplifier. Then we can see that $G_1 = P_1/P_s$ where $P_1$ is the output available power of amplifier number one, and $P_s$ is the available power from the generator. Similarly for $G_2$, $G_3$ etc. and so on. If we have $N$ amplifiers, the last one will have a gain $G_N = P_o/P_{N-1}$.

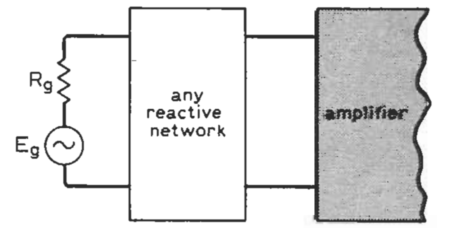
Overall available gain is clearly $G_A = P_o/P_s$ and this is

$$G_A = G_1\, G_2\, G_3 \, \ldots \, G_N =$$

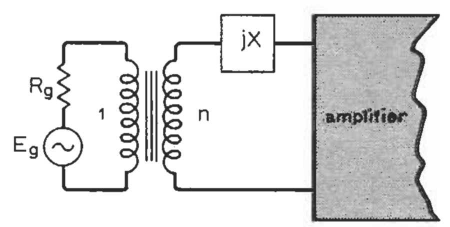
$$\frac{P_1}{P_s}, \frac{P_2}{P_1}, \frac{P_3}{P_2} \ldots \frac{P_o}{P_{N-1}} = \frac{P_o}{P_s}$$

So the overall gain of a series of stages is the product of the available gains, under the conditions prevailing at each input.

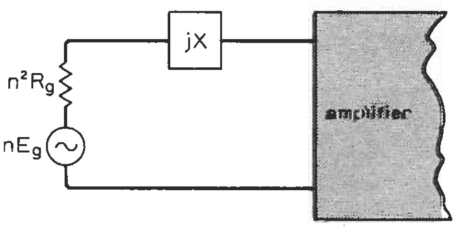
*Fig. A2*




(a)




(b)




(c)

## References


1. Mumford, W. W. and Scheibe, E. H. "Noise Performance Factors in Communication Systems", Horizon House.
2. Friis, H. T. "Noise figures radio receivers", *Proc. IRE*, vol. 32, 1944, p.419.
3. North, D. O. "Absolute sensitivity of radio receivers", *RCA, Review*, vol. 6, 1942, p.3.
4. Nyquist, H. "Thermal agitation of electric charges in conductors," *Phys. Rev.*, vol. 32, 1928, p.110.
5. Johnson, J. B. "Thermal agitation of electricity in conductors," *Phys. Rev.*, vol. 32, 1928, p.97.
6. "Cathode Ray". "Heads, tails and noise," *Wireless World*, vol. 62, 1956, p.235.
7. "Cathode Ray". "More about noise," *Wireless World*, vol. 62, 1956, p.267.
8. "Cathode Ray". "k—and why it is 1.38× 10 ⁻²³," *Wireless World*, vol. 66, 1960, p.567.
9. Robinson, F. N. H. "Noise in electrical circuits," chapter 4. Oxford University Press, 1962.
10. Gunston, M. A. R. "Errors and inconsistencies in the determination of system noise parameters," *Marconi Review*, 1969, p.142.
11. Mumford, W. W. "Broad-band microwave noise source," *Bell Systems Tech. Journal*, vol. 28, 1949, p.608.
12. Smith, K. L. "The ubiquitous phase sensitive detector," *Wireless World*, vol. 78, 1972, p.367.
13. Hosking, M. W. "Realm of microwaves, part 6 —microwave antennae," *Wireless World*, vol. 79, 1973, p.501.
14. Nielsen, E. G. "Behaviour of noise figure in junction transistors," *Proc. IRE*, vol. 45, 1957, p.957.
15. Walling, J. C. and Smith, F. W. "Solid-state masers and their use in satellite communication systems," *Philips Technical Review*, vol. 25, 1964, p.289.
16. Van der Ziel, A. "Solid State Electronics," Prentice-Hall.
17. Moxon, L. A. "Noise Factor," *Wireless World*, 1946, p.391, and 1947, p.11.
18. Greene, J. C. "The art of noisemanship," *Proc. IRE*, vol. 49, 1961, p.1223.
19. Cohn, Seymour B. "The noise figure muddle," *Microwave Journal*, vol. 2, March 1959, p.7.


## Black holes: radiation transformers?

The concept of a black hole may seem so bizarre that to complicate matters by asking what might happen if the black hole were electrically charged might seem to be a case of scientific masochism. Nevertheless, the question has been asked, and answered, by Ulrich Gerlach of Ohio State University. The answer bears on the relationship between electromagnetism and gravitation. Gerlach calculates that if the charged black hole is immersed in an electromagnetic field then any electromagnetic radiation which comes near it will be transformed into gravitational radiation. The black hole acts as a catalyst which transforms one type of radiation into another.

## The Psi particle

A new sub-atomic particle (or perhaps two similar ones) has been discovered independently at two US research laboratories (Brookhaven and Stanford). One team produced "Psi particles" by throwing protons at protons, the other by throwing electrons at protons. The Psi particle has a mass energy equivalent of 3GeV, which makes it about three times as heavy as a proton. Most particles of such a mass have very short lifetimes in isolation. The Psi article is a surprise in that it lives for about 10,000 times as long as would be expected before decaying into an electron and a positron. Physicists are trying hard to explain it. The one thing they seem sure of is that the Psi particle is *not* the elusive quark.

## New thermal imaging tube

A thermal camera capable of taking "recognizable images of human faces or hands" has been developed at SERL, Baldock. The tube has a germanium window, a target of the pyroelectric material triglycine sulphate which reflects low-energy electrons to an extent dependent on the incident heat and a fluorescent screen to turn the electron beam, after acceleration, into a visible image. Temperature differences of 1°K are detectable.

## Lasers detect paint-peeling masterpieces

If deterioration in oil paintings is detected early it is possible to take action to preserve the paintings. Laser holography has been shown to make early detection possible. The technique is the same as that used to measure minute changes in any object. A laser hologram is taken, the temperature of the painting is raised a little, and another hologram taken. Detached regions of paint, caused by its beginning to peel off the backing material, dissipate heat at a different rate from the rest and expand faster. Pronounced interference fringes appear when the holograms are superimposed, and where peeling has begun there are kinks in the fringes.

# Noise—confusion in more ways than one

## 2—Noise temperature and noise generators

by K. L. Smith

*University of Kent at Canterbury*

**In part 1, temperature was shown to play a large part in discussions about noise. In this part the noise temperature concept is discussed more fully, together with methods of measurement at low frequencies using a noise generator.**

If a resistor at room temperature is connected across the input terminals of an amplifier of bandwidth $B$(Hz), the available noise power $kT_0B$ is amplified by the gain $G_A$. This means that the output power from the amplifier is $G_AkT_0B$. The noise power added by the amplifier must also be taken into account. If this amplifier contribution is $P_{Na}$ at the output, it can be added to the above expression directly, because noise powers from different sources can simply be summed if they are unrelated. The total available output noise power $P_{NO}$ becomes $G_AkT_0B+P_{Na}$ as shown in Fig. 5(a).

This is the point at which we think up our first bit of convenient fiction. We imagine that the amplifier is completely noiseless and account for $P_{Na}$ by a (now fictitious) extra noise power available at the input terminals. So we write $P_{Na}=G_AkT_eB$. By this dodge we can replace a noisy amplifier by a noiseless equivalent, Fig. 5(b), whose output is

$$P_{NO}=G_AkT_0B+G_AkT_eB$$

or

$$P_{NO}=G_AkB(T_0+T_e).$$

The whole thing is equivalent to an input source resistor at a temperature of $T_0+T_e$ connected to a noise-free amplifier, where $T_0$ is the room temperature of the actual resistor at the input terminals (=290K) and $T_e$ is the effective input noise temperature of the amplifier. Like available gain, $T_e$ varies with input matching conditions, so there is not a unique $T_e$ for every system. It will depend on how the system is used. An amplifier with a low $T_e$ is better noisewise than one with higher temperatures, other things being equal. The idea of $T_e$ is a little abstract because it is not a physical temperature (the input of an amplifier with $T_e=4000$K would not be glowing white hot!).

One or two points arise at this stage. The first is that we are not limited to a source temperature of $T_0$ in every case. Thus the noise power output for a receiver whose effective input temperature is $T_e$ and connected to an aerial whose aerial temperature is $T_a$ is

$$P_{NO}=G_AkB(T_a+T_e).$$

Another point arising is to do with the bandwidth $B$—I have been assuming that we know all about it. $B$ is not the normal 3-dB bandwidth used by radio engineers, but is the noise power equivalent bandwidth and involves notions about the available gain-bandwidth product ($G_AB$). I have relegated these ideas to a brief discussion in Appendix B.

There is another very easily overlooked complication and that is the possibility of more than one channel allowing signals and/or noise to pass through the system. An obvious example is the superhetrodyne receiver with a response at the image frequency. I often wonder how many experimenters measure the noise performance of their v.h.f. converters, oblivious of the fact that they have a wide open channel at the image frequency. Incidentally, this "improves" the (erroneously) measured single-channel noise performance figures, so one should beware of excellent-looking figures on some manufacturers equipment specifications.

A useful concept in connection with the above arguments is that of the operational noise temperature, $T_{op}$. This is a measure of the overall system performance. A knowledge of $T_{op}$ enables the all important output signal to noise ratio to be calculated. As an example of how this idea arises, consider a superhet with a gain $G_s$ at the signal frequency and $G_i$ at the image frequency, as outlined in Fig. 6. The noise bandwidth is usually $B_{IF}$ for all channels, because it is set by the i.f. amplifier. The signal may occupy a bandwidth $B_s$($B_s<B_{IF}$ because if the i.f. is narrow it will limit $B_s$ to $B_{IF}$). The total available output noise power from this receiver will be

$$P_{NO}=k(T_s+T_e)B_{IF}G_s+k(T_i+T_e)B_{IF}G_i \quad (3)$$

where $T_s$ is the temperature of the aerial, signal generator etc., at the frequency of the signal channel, and $T_i$ is the same quantity but at the image frequency. If the temperature is constant over the two channels, then $T_s=T_i$.

The question arises, how do we handle $P_{NO}$ for signals to noise ratio purposes? The answer is that if the available output signal power is $P_{so}$, the signal-to-noise ratio is given directly by $P_{so}/P_{NO}$, a little thought shows this to be the important

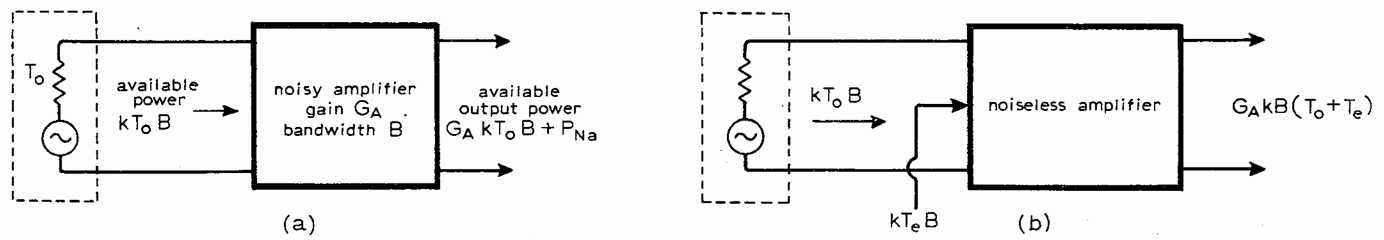



*Fig. 5. It is more convenient to replace a noisy real amplifier (a) with a noiseless one (b), and account for the noise by inventing a fictitious noise temperature $T_e$ at the input.*

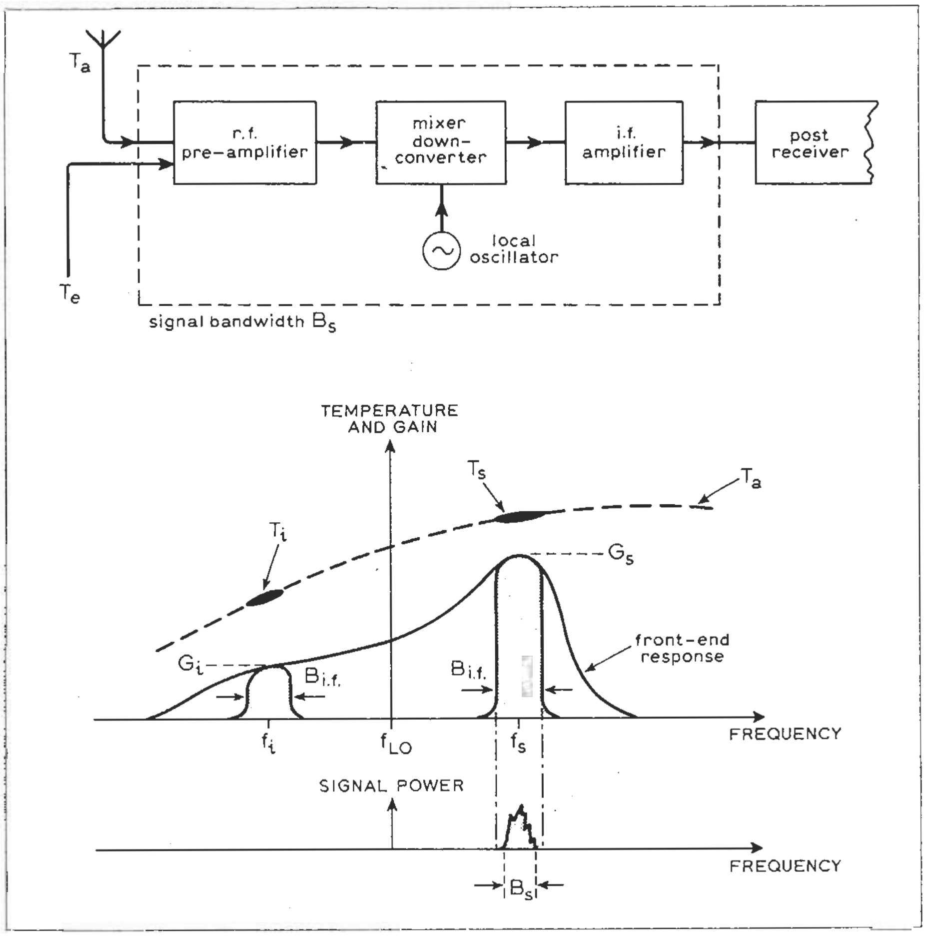



*Fig. 6. In a superhet receiver there are usually at least two channels through which noise can pass to the output. Unless signal information is also coming in via the image frequency $f_i$, it is always advantageous to reduce $G_i$ to the smallest possible value. The "shape factor" of the i.f. bandpass, $B_{IF}$ also has a significant effect on the noise performance.*

final parameter in any data link or communications system. The effect of the noise power is as though all of it is concentrated into the signal bandwidth $B_s$. Therefore we define another temperature, the operating noise temperature, $T_{op}$ as

$$P_{NO}/kB_sG_s.$$

Notice the particular gain bandwidth product used. You will be pleased to know this is about the limit of abstract thinking we need, so we will soon be back to more concrete things!) Substituting for $P_{NO}$, by using equation (3), and assuming for simplicity that $T_s = T_i$ and relabelling them $T_a$, the aerial temperature, operating noise temperature becomes

$$T_{op} = \frac{(T_a + T_e)B_{IF}(G_s + G_i)}{B_sG_s}$$

$$= \frac{B_{IF}}{B_s}(T_a + T_e)(1 + \frac{G_i}{G_s}).$$

This equation offers considerable meat to get one's teeth into. First, it illustrates the rationale of using temperatures in noise discussions. Awkward Boltzmann's constants cancel out and one is left with the various temperatures and parameters of the amplifier only. Clearly, the output signal to noise ratio degrades as $T_{op}$ becomes larger. The lowest $T_e$ should be the aim when designing the equipment and is achieved by noise matching and low noise components in the front end.

Care should be taken to understand the significance of $T_a$. For instance, the signal from a satellite is not enhanced when it is originating from the direction of the sun! ($T_a$ shoots up.) Significantly, simple but all too easily-forgotten pieces of work need to be attended to, such as making sure $B_{IF}$ is not greater than $B_s$. If the receiver bandwidth is twice as wide, say as that required to pass the signal, then $T_{op}$ is doubled. The noise coming in via the image channel increases $T_{op}$. If $G_i = G_s$ (as in microwave receivers) $T_{op}$ is again doubled. The receiver designer should reduce the bandwidth to the minimum ($B_s$) and filter out the image, (make $G_i = 0$) to obtain the minimum operating noise temperature. Then $T_{op} = T_a + T_e$.

There are certain wideband signals which are received with a sensitivity advantage if both channels are wide open. Radio astronomical signals are themselves wideband noise powers. This means that useful signal powers are received in both sidebands. In fact the wider the bandwidth of the radio astronomy receiver the more signal power will be received. There is a worsening of signal-to-noise ratio by a factor of two if a double-channel receiver is used to receive a single-channel signal.

If the gain of the first stage of an amplifier or receiver is high, intuition might suggest that noise power contributions by later stages are not significant. Although intuition is not very trustworthy sometimes, in this example it is all right, as the following argument shows.

If we consider the three stages with gains and effective temperatures as shown in Fig. 7 then the output noise power is

$$P_{NO} = G_1G_2G_3kB(T_i + T_e) \quad (4).$$

The noise output of the first stage is the noise power from the resistor times $G_1$ plus the contribution represented by $T_{e1}$.

Therefore the available noise output from stage one is $G_1kB(T_i + T_{e1})$. The output from the second stage is its own noise, represented by $T_{e2}$, plus the input from stage one multiplied by $G_2$. The output from stage two is

$$G_2kB[G_1(T_i + T_{e1}) + T_{e2}].$$

Similarly the output from stage three, which is the final output noise power, is

$$G_3kB\{G_2[G_1(T_i + T_{e1}) + T_{e2}] + T_{e3}\} \quad (5).$$

Equations (4) and (5) are both expressions for $P_{NO}$, therefore,

$$G_1G_2G_3kB(T_i + T_e) =$$
$$G_3kB\{G_2[G_1(T_i + T_{e1}) + T_{e2}] + T_{e3}\}.$$

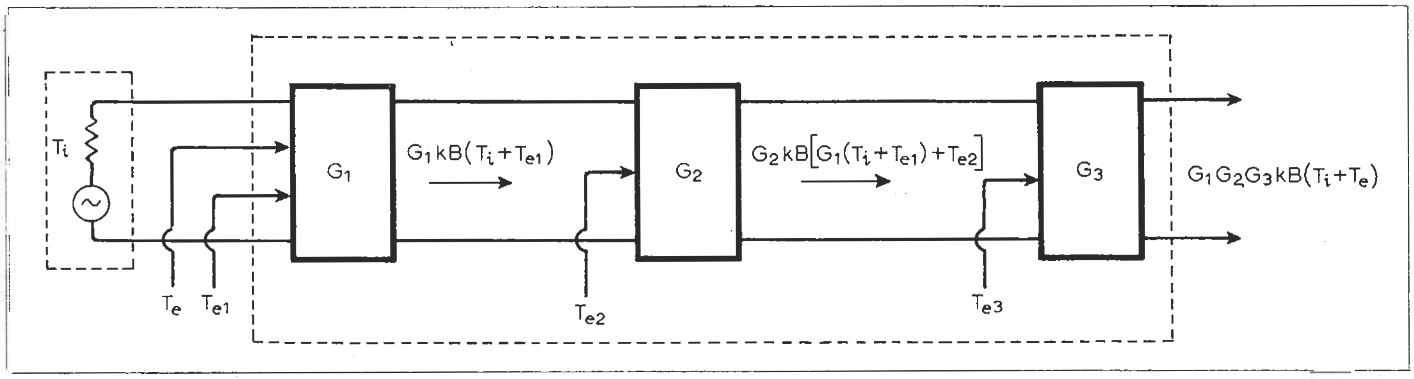



*Fig. 7. Overall noise temperature of a cascade of amplifier stages can be deduced as shown here.*

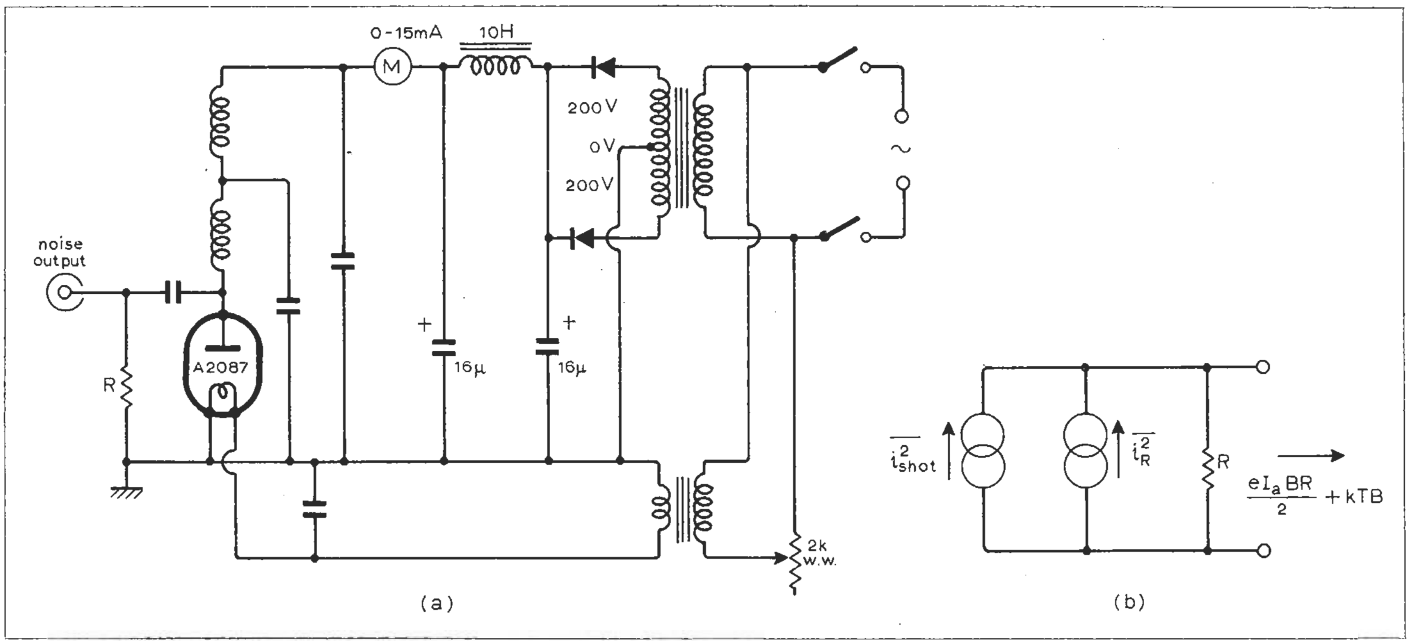



*Fig. 8. Still an extremely useful noise source for measurement purposes, the saturated thermionic diode is an absolute noise generator. (a) shows a typical circuit using an A2087, (b) is the equivalent circuit for calculation purposes.*

This cancels down to the final simple equation:

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1G_2}.$$

Notice that the term containing $T_i$ conveniently subtracts from both sides. This equation shows that if the first stage gain is, for example, 100 times and the effective noise temperature of the second stage is 300K, then the contribution to the overall $T_e$ by stage two is only 3K. Usually the third term can be neglected, unless $G_2$ is very small. The gain of stage three ($G_3$) has not entered into the picture. The argument can be extended to any number of stages. The equation is conveniently termed the cascading formula and in effect describes how the various noise temperatures throughout a chain of stages can all be referred to the front-end terminals as a single $T_e$, the system of stages is regarded from then on as noiseless.

## Measuring $T_e$

The way in which I have discussed the role of the absolute temperature in noise problems, shows the convenience of dividing the output noise power from a signal handling system into two parts. One part is the noise that comes in with the signal represented by $T_a$ and the other is that introduced by the local equipment, which accounts for the $T_e$ term. This means that all the various noise powers produced in the local equipment are lumped together under the title $T_e$ —whether they originate as thermal noise in the resistors, shot effect in the transistors or valves, flicker noise and so on.

If you have just built a receiver or a customer has ordered a system to be designed with a stated maximum $T_{op}$, it is essential to be able to make fairly accurate measurements of $T_e$, so that you know what you are talking about. The basis for any noise measurements involves generating accurately known noise powers. The standard noise generators are based on physical mechanisms including the saturated thermionic diode, the gas discharge tube and the noise generated in a reverse biased semiconductor diode. Sinewave signal generators can be used as standard power sources, but because they produce narrow band signals, their use in noise measurements involves difficulties interpreting what bandwidths mean and errors are very likely.

Before going on to the construction of noise sources, I will discuss a technique for measuring $T_e$. The following way for determining $T_e$ might be termed the ratio method. A noise source with an effective temperature $T_{hot}$ when it is fired, is coupled into the amplifier or receiver and the output $P_{NO(hot)}$ is noted on a power meter. The noise source is now switched off but still connected to the system. The temperature when the noise source is not fired can be labelled $T_{cold}$, with a corresponding output power from the system, $P_{NO(cold)}$. It is not necessary to know accurately the actual values of the output powers, only their ratio, $A$.

As an example, consider the superhet receiver for which equation (3) applies. Putting in the appropriate values for the "hot" and "cold" conditions, gives

$$P_{NO(hot)} = kB_{IF}(T_{hot} + T_e)(G_s + G_i)$$

and $P_{NO(cold)} = kB_{IF}(T_{cold} + T_e)(G_s + G_i)$.

Dividing them gives $A$

$$A = \frac{P_{NO(hot)}}{P_{NO(cold)}} = \frac{T_{hot} + T_e}{T_{cold} + T_e}.$$

From which we get

$$T = \frac{T_{hot} - A\ T_{cold}}{A - 1} \qquad (6).$$

All we require to know is $T_{hot}$, $T_{cold}$ and $A$. The bandwidths, gains and $k$ have cancelled. This straightforward result applies for any system whether there are more channels than two or any other complexities. For best results, the value of $A$ is often chosen to be two (the minimum error occurs near this value).

As usual, the assumptions made should be considered to avoid, or at least understand, errors that might creep in. $T_{cold}$ is usually taken to be $T_0$, but the temperature of the lab or workshop in which the measurements are made could very well differ by a few degrees from 290K, and there will be a corresponding error introduced. $T_{hot}$ must be known accurately for the particular noise source. The matching conditions of the noise source to the receiver or amplifier should duplicate the conditions that will apply in the operational system. It is not certain that the source impedance of the noise generator when it is fired will be the same as when it is cold. Any difference that does exist will introduce an error in $T_e$, but it is difficult to establish any such impedance shift.

The output meter should be a true square-law device with voltage or current. In other words it should be accurately linear as a function of power. Any overloading or non-linearity in the amplifier will introduce errors. For instance, the common f.m. receiver is non-linear for amplitude changes, and cannot be investigated by the above method. (The front end could be checked separately, but we are discussing a.m. noise, which would normally be eliminated in this kind of receiver anyway. In f.m. systems the more difficult f.m. noise has to be considered). Errors also arise at the higher frequencies, mainly because of the usual effects of the stray reactances.

## Sources of wideband noise, diode noise generators

One of the most useful noise generators for frequencies up to a few hundred megahertz is based on the temperature limited diode. The full shot-noise generated by a thermionic diode operated under these conditions can be calculated exactly, but involves fairly complex statistical ideas such as Campbell's theorem. A treatment can be found in reference 9 (see part 1). Pierce derived the shot noise equation very simply but his method lacks the rigour demanded by purists. It is an interesting derivation and I have included an outline of it in Appendix C. The full shot noise produced on a direct current $I_a$ in a bandwidth $B$, is

$$\overline{i^2}_{shot}=2eI_aB$$

where $e$ is the charge on an electron.

Because the diode is saturated, the effective source resistance of the shot noise generator is very high indeed. Fig. 8 shows a typical circuit for a diode noise generator with a source resistance, $R$. The equivalent circuit is also shown. The total available noise power from the generator is the sum of the noise power from the shot source and that from $R$, which is at the ambient temperature $T$. The two sources of noise power are not correlated, so that their outputs add directly as we have seen earlier. From Fig. 8(b) the available power from the two current generators is

$$P_N=\frac{\overline{i^2}_{tot}}{4G}=\frac{\overline{i^2}_{shot}+\overline{i^2}_R}{4G}$$

where the conductance $G$ is equal to $1/R$.

$$\therefore P_N=\frac{eI_aBR}{2}+kTB.$$

Excess noise temperature, $T_D$, for a saturated diode is obtained by equating the first term on the right hand side of this last equation to $kT_DB$, so that $T_D=eI_aR/2k$. The numerical values of the physical constants, $e$ and $k$ give the value 11,600 for the quotient $e/k$. Therefore $T_D=5800I_aR$. The total noise temperature of the fired source is $T_D$ plus the contribution from $R$

$$T_{hot}=5800I_aR+T \quad (7).$$

The cold temperature is simply $T$, because with the diode off, $I_a=0$ and no contribution is forthcoming from the shot noise term. From these considerations we know the values of $T_{hot}$ and $T_{cold}$ to use in equation (6). Putting in the quantities gives

$$T_e=\frac{5800I_aR+T-AT}{A-1}$$

which conveniently simplifies to

$$T_e=\frac{5800I_aR}{A-1}-T.$$

A number of authors have used the ideas of the noise ratio and excess noise ratio. I think we have enough detail from the preceding discussions to illustrate at this point, how these ideas are used. You may recall the definition involves the ratio of the temperature to 290K or the ratio of the excess temperature to 290K respectively. The ratios obtained are really noise power ratios, in which the bandwidth and Boltzmann's constant cancel. Being a power ratio, the results are often expressed in decibels. By dividing the equation above by 290K we obtain the noise ratio $t_e$

$$t_e=\frac{T_e}{290}=\frac{20I_aR}{A-1}-\frac{T}{290}$$

Often $T$ is taken equal to 290K (but see my earlier cautionary note); in that instance this equation becomes

$$t_e=\frac{T_e}{290}=\frac{20I_aR}{A-1}-1.$$

The *excess* noise ratio for a diode generator can be obtained from equation (7) by subtracting 290K from both sides, then dividing by 290K

$$\frac{T_{hot}}{290}-1=20I_aR+\frac{T}{290}-1$$

and again if $T=290$K

$$\frac{T_{hot}}{290}-1=20\,I_aR.$$

The diode noise source is very convenient because the temperature and noise ratios are directly proportional to $I_a$, and by just winding up the filament temperature, $I_a$ can be set to any convenient values on an accurate anode current meter. (With due care not to burn out the filament of course!)

Ordinary lumped-component circuitry begins to fail as the frequency of operation rises toward the GHz region. The diode noise generator is no exception and errors begin to affect the result when measuring at the frequencies in question. Another effect becomes important at the same time: transit time of the electrons across the cathode to anode space is significant in the hundreds of megahertz range and the shot noise equation begins to break down.

*To be continued*

## Appendix B

### Noise equivalent bandwidth

Perhaps you have noticed in the discussion so far, I have blandly assumed that $G_A$ is "the power gain", without any real attempt to discuss how this quantity varies with frequency. Most amplifiers, whether intended or not, are severely limited in their frequency response. This means that $G_A$ is a maximum somewhere near the centre of the band and drops off towards zero at both ends of the response, except for d.c. amplifiers. If you think of a constant distribution of energy over the frequency spectrum (white noise) then the bandpass function "weights" the contribution in each very small band at points across the response. The total output power is a sum of all these weighted contributions. This is the kind of reasoning we do when finding averages. Fig. B shows an example to make the point clear.

We can imagine $G_A$ to stay at its maximum value for a bandwidth $B$, then drop off sharply to zero at each side. If the width B of this fictitious rectangular bandpass curve is such that the output power is the same as from the actual response, then $B$ is defined as the "equivalent noise power bandwidth". What we have really said is that the area of the rectangular curve is made the same as the area of the actual curve. This gives us a clue about the mathematical approach to writing down the definition. If the available noise power is constant over the band then the available noise power in *any* small band $df$, is $Kdf$. $K$ is the constant level. Therefore the available output power is $G_A(f)Kdf$ and the total output power is

$$P_{No}=K\int_{bandpass}G_A(f)df.$$

By definition, the total output power is also

$$P_{No}=KBG_{A(max)} \quad (B1)$$

Equating these gives

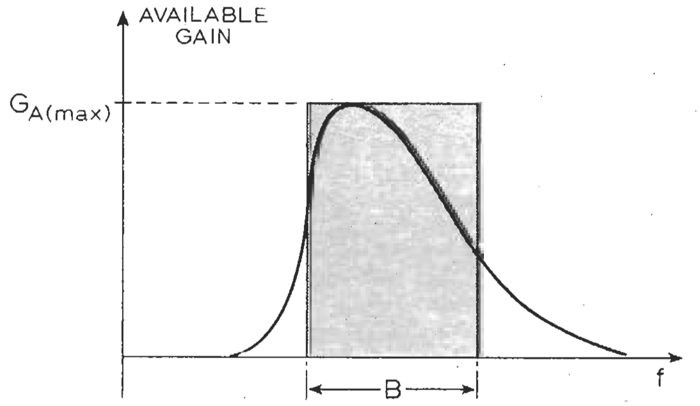
$$B=\frac{\int_{bandpass}G_A(f)df}{G_{A(max)}} \quad (B2)$$

This is alright if you can do the integration or look it up in tables, but if, as usual, no simple function exists for $G_A(f)$, then the integral would have to be solved numerically. Equation (B1) shows that the amount of noise power emanating from the output of a system is proportional to the gain-bandwidth product $BG_{A(max)}$.

Note that $B$ is not the ordinary "half-power" bandwidth; a simple example shows this to be true by relating the two bandwidths.

Consider the bandwidth to be limited by a series tuned circuit. The reactance at any frequency will be $X=\omega L-(1/\omega C)$. Using the equation for $G_A$ (p.110) available gain is

*Fig. B. Area of the shaded rectangle is equal to area under response curve. Height is $G_A$ (max) and width is definition of equivalent noise bandwidth, B.*



$$G_A = \frac{K'R_{in}^2 R_g n^2}{(n^2R_g+R_{in})^2+\left(\omega L-\frac{1}{\omega C}\right)^2}$$

which can be written

$$G_A = \frac{constant}{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}$$

where $R$ has been written for $n^2R_g+R_{in}$. From this, $G_{A(max)}$ is $constant/R^2$. At the 3-dB points $|X|=\pm R$ because $G_A$ is then equal to $\frac{1}{2}G_{A(max)}$. This condition enables us to write down the frequencies of the 3-dB down points. From $R=\omega C-(1/\omega L)$ and $-R=\omega C-(1/\omega\ L)$ we get two quadratic equations whose solutions are

$$\omega_1=\frac{R}{2C}\pm\left(\frac{R^2}{4C^2}+\frac{1}{LC}\right)^{\frac{1}{2}}$$

and $$\omega_2=-\frac{R}{2C}\pm\left(\frac{R^2}{4C^2}+\frac{1}{LC}\right)^{\frac{1}{2}}$$

Subtracting gives the frequency difference

$$B_{3dB}=f_1-f_2=\frac{\omega_1-\omega_2}{2\pi}=\frac{R}{2\pi C}.$$

Using equation (B2),

$$B=\frac{1}{2\pi}\int_0^\infty \frac{1}{1+\left(\frac{\omega C}{R}-\frac{1}{\omega LR}\right)^2}d\omega$$

The integral is a "do-able" one, and involves $\tan^{-1}$ type solutions. Carrying out this solution, $B$ is $R/4C$, which means that the relationship between $B$ and $B_{3dB}$ for a single tuned circuit is $B=\pi B_{3dB}/2$. Thus $B$ is somewhat wider than $B_{3dB}$. The Table shows a few relationships for other band-limiting filters.

| Circuit | Relationship |
|---|---|
| Two cascaded tuned circuits | $B=1.22B_{3dB}$ |
| Three cascaded tuned circuits | $B=1.16B_{3dB}$ |
| A staggered pair | $B=1.11B_{3dB}$ |
| A 4-pole Butterworth filter | $B=1.11B_{3dB}$ |
| A 6-pole Butterworth filter | $B=1.05B_{3dB}$ |

The noise bandwidth approaches the 3dB bandwidth more and more closely as the "shape factor" improves. For ordinary i.f. amplifiers with a number of tuned stages, there is very little error if you assume $B \doteq B_{3dB}$.

**Appendix C**
**Shot noise equation**

A simple but not very rigorous derivation of the shot noise current equation was ingeniously put forward in J. R. Pierce's paper, "Noise in Resistances and Electron Streams" published in the *Bell System Technical Journal*, volume 27 (1948). It goes something like this:

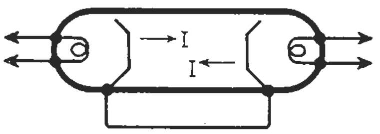



*Fig. C. Artificial double cathode "diode" used by J. R. Pierce to derive the shot noise equation.*

If a diode consisting of two emitting cathodes (Fig. C) has a potential $V$ between them, a current $I$ will pass equal to $I_V=I_0exp(eV/kT)$, where $I_0$ is the current that passes when $V=0$; that is, by the thermally energetic electrons "bridging the gap". Differentiating gives

$$\frac{dI_V}{dV}=\frac{1}{r_a}=\frac{I_o e}{kT}e^{eV/kT}$$

As the mean square noise current expected from a resistance $r_a$ is $\overline{i^2}=4kTB/r_a$ and the diode "resistance" at zero volts is $r_a=kT/I_0e$, it follows that $\overline{i^2}=4eIB$ after substituting for $r_a$. This is the total noise current produced by the special case of two cathodes exchanging current. Because noise powers add, then the mean square current of one cathode is half the value and therefore $\overline{i^2}=2eIB$, which is the shot noise equation.

*(To be continued)*

## Wireless World noise reducer

Next month's issue will contain the start of an article describing the *Wireless World* noise reducer, an add-on Dolby processor mainly for use with magnetic tape cassette machines. This constructional design, the only one of its kind, has been planned in close collaboration with Dolby Laboratories and will be available from *Wireless World* in kit form.

The unit includes a stereo Dolby B processor that is switchable for both encoding and decoding. This means that as well as decoding commercial Dolby B cassettes, encoded tapes can be prepared. For recording stereo broadcasts, a switched 19kHz pilot-tone filter is included. And should B-type encoding be adopted for f.m. transmissions, as in the USA, the unit will also decode those. There is another use of the processor. Because of the improved signal-to-noise ratio obtained with the unit, recordings can be made at a lower level that would otherwise be possible. Consequently some of the noise reduction can be traded for a lower distortion level at peak recorded levels.

The *Wireless World* Dolby processor can be aligned without using additional instrumentation. The circuit board has been designed to include the required alignment facilities—400Hz and 5kHz oscillators are constructed from components in the *WW* kit, together with a 1-kHz meter calibration oscillator. Full alignment and calibration instructions are included in the article, which starts in the May issue with a description of the Dolby system and its functioning.

## HF predictions

Predicted disturbed periods are March 23–28, April 4–10 and 19–25.

Seasonal trend and low solar activity combine to produce FOTs and LUFs which give a restricted choice of time and frequency for reliable day-to-day communication. The charts show that the restriction is severe when both ends of a circuit are in the northern hemisphere.