XII. NOISE IN ELECTRON DEVICES[*]

Prof. H. A. Haus
Prof. P. L. Penfield, Jr.

Prof. R. P. Rafuse
W. D. Rummler

RESEARCH OBJECTIVES

The research on noise in electron devices has two objectives:

(i) The study of specific devices, such as the parametric amplifier and the maser, to determine the physical sources of noise and the limitations they impose on the noise performance of the amplifier.

The parametric amplifier with coherent input signals at both the signal and idler frequency, and the degenerate parametric amplifier are the impetus for the second objective.

(ii) The determination of a measure of the optimum noise performance of multiterminal linear amplifiers. The optimum noise performance of a linear twoport amplifier is known to be expressible in terms of its minimum noise measure, a quantity that is characteristic of the amplifier. The optimum noise performance of a multiterminal pair amplifier excited by a multiterminal pair source is at present under study for the purpose of devoloping an analogous measure for the optimum noise performance of such an amplifier.

H. A. Haus, P. Penfield, Jr., R. P. Rafuse

A. SOLUTIONS TO THE PROBLEM OF THE OPTIMUM NOISE PERFORMANCE OF MULTITERMINAL AMPLIFIERS

In a previous report[1] we showed that under certain restrictions the stationary values of signal-to-noise ratio which can be obtained for a specified value of exchangeable signal power at a single-output terminal pair by imbedding an n-terminal pair source and an m-terminal pair amplifier in an n + m + 1 terminal pair lossless network are governed by two coupled matrix equations — Eqs. 1 and 2.

$$\overline{E_{na}E_{na}^{\dagger}}\,x_2 + 2\lambda(Z_a + Z_a^{\dagger})\,x_2 = 0 \tag{1}$$

$$\overline{E_s E_s^{\dagger}}\,x_1 + \mu\left[\overline{E_n E_n^{\dagger}} + 2\lambda(Z + Z^{\dagger})\right]x_1 = 0 \tag{2}$$

By a simpler derivation, it has been established that Eqs. 1 and 2 are valid regardless of the nature of the matrices characterizing the source and amplifier networks; however, for simplicity we shall assume here that $\overline{E_s E_s^{\dagger}}$ is positive definite, $Z + Z^{\dagger}$ is positive definite, and $Z_a + Z_a^{\dagger}$ is indefinite.

Components of the vectors x_1 and x_2 may be interpreted as complex voltage ratios. For instance, with all voltage sources in the source and amplifier short-circuited

except the source represented by E_{s1}, which is the open-circuit signal voltage at the first terminal pair of the source network, the open-circuit voltage at the output terminal pair is $x_{11}E_{s1}$, where x_{11} is the first component of the vector x_1. Similarly, the components of x_2 relate the open-circuit noise voltage at the output terminal pairs to the open-circuit voltages of the amplifier. Imbedding the source and amplifier networks in a given $n + m + 1$ terminal pair lossless network enables us to define a unique pair of vectors x_1 and x_2. The reverse is not true; it can be shown that there are an infinite number of lossless networks corresponding to a particular pair of vectors x_1 and x_2.

For Eq. 1 there are n eigenvalues λ_i and eigenvectors $x_2^{(i)}$. We may express the eigenvalue λ_i in terms of its eigenvectors as

$$\lambda_i = -\frac{x_2^{(i)\dagger}\overline{E_{na}E_{na}^{\dagger}}x_2^{(i)}}{2x_2^{(i)\dagger}\left(Z_a + Z_a^{\dagger}\right)x_2^{(i)}}. \tag{3}$$

For convenience later, we shall label the smallest positive eigenvalue as λ_1 and the remaining positive eigenvalues in ascending order, and we shall label the smallest negative eigenvalues λ_m and the remaining negative eigenvalues in descending order.

Equation 2, on the other hand, has n solutions for each of the m eigenvalues λ_i of Eq. 1; this gives us a total of $n \times m$ sets of solutions to Eqs. 1 and 2. Using a double-subscript notation for these eigenvalues and eigenvectors, we may express the eigenvalue μ_{ij} in terms of λ_i and the eigenvector $x_1^{(ij)}$ as

$$\frac{1}{\mu_{ij}} = \frac{x_1^{(ij)\dagger}\overline{E_nE_n^{\dagger}}x_1^{(ij)} + 2\lambda_i x_1^{(ij)\dagger}(Z+Z^{\dagger})x_1^{(ij)}}{x_1^{(ij)\dagger}\overline{E_sE_s^{\dagger}}x_1^{(ij)}}. \tag{4}$$

Here, we number our eigenvalues for a given λ_i by the order of the value of their reciprocals; the minimum value of $1/\mu_i$ (which may be negative) is $1/\mu_{i1}$ and the maximum is $1/\mu_{in}$.

We would now like to relate the quantities of interest, namely the signal-to-noise ratio at the output and the exchangeable signal power at the output, to the eigenvalues derived from Eqs. 1 and 2. For the optimal imbedding network corresponding to a set of values λ_i and μ_{ij} we write the signal-to-noise ratio as

$$\sigma_{ij} = \frac{x_1^{(ij)\dagger}\overline{E_sE_s^{\dagger}}x_1^{(ij)}}{x_1^{(ij)\dagger}\overline{E_nE_n^{\dagger}}x_1^{(ij)} + x_2^{(i)\dagger}\overline{E_{na}E_{na}^{\dagger}}x_2^{(i)}}, \tag{5}$$

and the exchangeable signal power as

$$p_{ij} = \frac{x_1^{(ij)\dagger} \overline{E_s E_s^\dagger} x_1^{(ij)}}{2x_1^{(ij)\dagger}(Z+Z^\dagger)x_1^{(ij)} + 2x_2^{(i)\dagger}\left(Z_a+Z_a^\dagger\right)x_2^{(i)}} \quad . \tag{6}$$

We now see why these quantities do not appear directly as eigenvalues. In Eq. 3 we see that λ_i is independent of the magnitude of $x_2^{(i)}$; similarly, μ_{ij} in Eq. 4 is independent of the magnitude of $x_1^{(ij)}$. On the other hand, both of the quantities defined in Eqs. 5 and 6 are dependent on the relative magnitude of $x_2^{(i)}$ and $x_1^{(ij)}$. We can explicitly demonstrate the consequences of this by considering the particular realization of the optimal network that places this fact in evidence.

Consider an imbedding of the m-terminal pair amplifier in an arbitrary m + 1 terminal pair lossless network. We know that we can always pick this network in such a manner that the mean-square voltage and the real part of the impedance at the output terminal pair are $x_2^{(i)'\dagger} \overline{E_{na} E_{na}^\dagger} x_2^{(i)'}$ and $1/2\, x_2^{(i)'\dagger}\left(Z_a+Z_a^\dagger\right)x_2^{(i)'}$, respectively, where $x_2^{(i)'}$ is a vector that is proportional to $x_2^{(i)}$. Similarly, we can losslessly reduce the source network to a one-terminal pair network with mean-square signal voltage
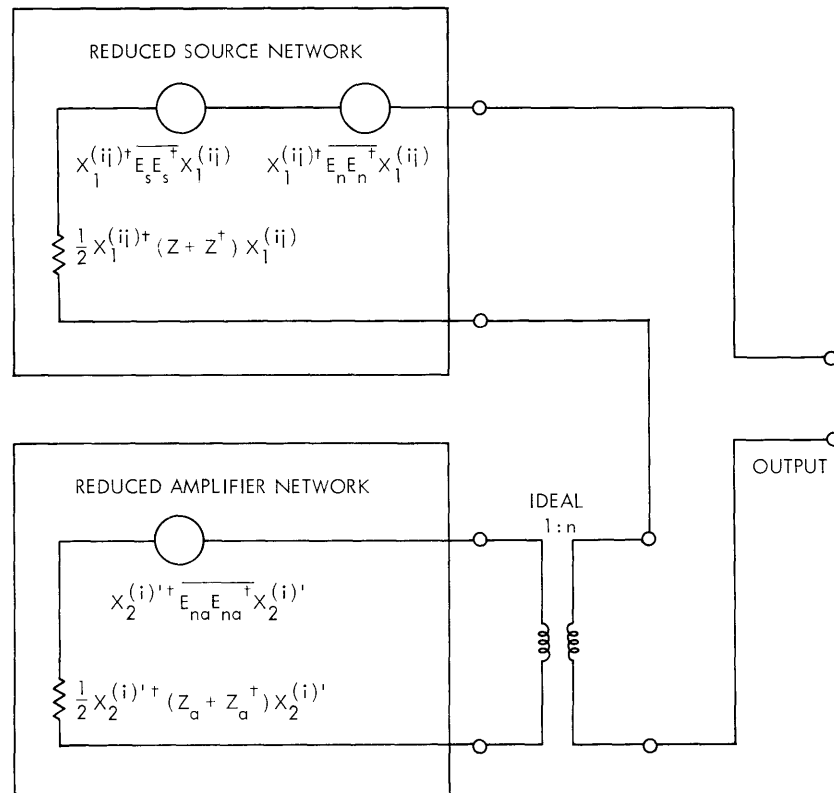

Fig. XII-1.   Realization of the optimal network.

$x_1^{(ij)\dagger} \overline{E_s E_s^\dagger} x_1^{(ij)}$, mean-square noise voltage $x_1^{(ij)\dagger} \overline{E_n E_n^\dagger} x_1^{(ij)}$, and impedance $1/2 \, x_1^{(ij)\dagger} (Z + Z^\dagger) x_1^{(ij)}$. Combining these two reduced networks as shown in Fig. XII-1, we see that the signal-to-noise ratio and the exchangeable signal power at the output are given by Eqs. 5 and 6, respectively, where $x_2^{(i)} = nx_1^{(i)'}$. This network is optimal in the sense of Eqs. 1 and 2 for any value of n. We see that varying the transformer ratio is equivalent to varying the ratio of $\left|x_2^{(i)}\right|$ to $\left|x_1^{(ij)}\right|$. Physically, then, when we vary n or $\left|x_2^{(i)}\right| / \left|x_1^{(ij)}\right|$, we are changing the amount of our use of the amplifier, since we are changing the exchangeable signal power at the output and, thereby, also changing the signal-to-noise ratio at the output. This interpretation is actually completely general and independent of the particular imbedding network that we are using. However, in general, variation of the ratio of $\left|x_2^{(i)}\right|$ to $\left|x_1^{(ij)}\right|$ can correspond to some complex variation of the imbedding network because of the multiple feedback loops that may exist between the source and amplifier networks. Figure XII-1, then, just gives a convenient way of visualizing the effects of varying the ratio of $\left|x_2^{(i)}\right|$ to $\left|x_1^{(ij)}\right|$.

We see then that setting the ratio of $\left|x_2^{(i)}\right|$ to $\left|x_1^{(ij)}\right|$ equal to zero or, equivalently, setting n=0 is the same as throwing away the amplifier. The values of σ_{ij} and p_{ij} for this limit are just the signal-to-noise ratio and the exchangeable power of the source network alone. We designate these two quantities as

$$\sigma_{ij, s} = \frac{x_1^{(ij)\dagger} \overline{E_s E_s^\dagger} x_1^{(ij)}}{x_1^{(ij)\dagger} \overline{E_n E_n^\dagger} x_1^{(ij)}} \tag{7}$$

$$p_{ij, s} = \frac{x_1^{(ij)\dagger} \overline{E_s E_s^\dagger} x_1^{(ij)}}{2x_1^{(ij)\dagger} (Z + Z^\dagger) x_1^{(ij)}} ,$$

respectively.

It follows that for a given pair of eigenvalues λ_i and μ_{ij} we may vary σ_{ij} and p_{ij} by varying the ratio of the magnitudes of the eigenvectors, that is, by varying only the transformer ratio in Fig. XII-1. Such a variation enables us to plot the stationary values of signal-to-noise ratio as a function of the exchangeable signal power at the output corresponding to this particular solution to Eqs. 1 and 2. For this purpose we need the relation

$$1/\sigma_{ij} = 1/\mu_{ij} - \lambda_i/p_{ij}, \tag{9}$$

which may be verified by using Eqs. 3-6. From Eq. 9 we see that it is more convenient to plot characteristic curves of $1/\sigma_{ij}$ as a function of $1/p_{ij}$. In the $1/\sigma_{ij} - 1/p_{ij}$ plane these characteristic curves are straight lines with slopes $-\lambda_i$ and intercepts $1/\mu_{ij}$ with the $1/\sigma_{ij}$ axis. Thus a set of eigenvalues λ_i and μ_{ij} merely determine a

characteristic line in the $1/\sigma_{ij}$ - $1/p_{ij}$ plane; moreover, we see that for the stated problem there are $n \times m$ such lines.

It must be pointed out, however, that only one-half of a characteristic line is realizable. Comparing Eqs. 5 and 7, we see that

$$\frac{1}{\sigma_{ij}} \geq \frac{1}{\sigma_{ij,\,s}} ,$$ (10)

and, using Eqs. 4, 7, 8, and 10 in Eq. 9, we find that either

$$\lambda_i > 0 \quad \text{and} \quad 1/p_{ij} \leq 1/p_{ij,\,s}$$

or (11)

$$\lambda_i < 0 \quad \text{and} \quad 1/p_{ij} \geq 1/p_{ij,\,s}.$$

The equality signs in Eqs. 10 and 11 will hold only when $\left|x_2^{(i)}\right| / \left|x_1^{(ij)}\right|$ is zero, that is, the realizable one-half of the characteristic line ends at the point determined by the source — the point whose coordinates are $1/\sigma_{ij,\,s}$ and $1/p_{ij,\,s}$. Hence, in the $1/\sigma_{ij}$ - $1/p_{ij}$ plane we can realize only the one-half of the characteristic line that is above and to the left of the source point for negative λ_i, and above and to the right of the source point for positive λ_i. We can only achieve a signal-to-noise ratio equal to μ_{ij} at infinite exchangeable signal power if $1/p_{ij} = 0$ satisfies one of the inequalities of Eq. 11.

In displaying the solution to the optimization problem in the $1/\sigma_{ij}$ - $1/p_{ij}$ plane, we shall show only those solutions for a given λ_i which correspond to $1/\mu_{i1}$, the minimum value of $1/\mu_i$, and $1/\mu_{in}$, the maximum value of $1/\mu_i$. We would like to find where the end points of these characteristic curves lie in the $1/\sigma_{ij}$ - $1/p_{ij}$ plane. If we consider how these end points change as λ_i changes, we find that they generate a closed curve. This is illustrated in Fig. XII-2, in which several of these characteristic curves are shown as they must appear for positive λ's and minimum $1/\mu_i$. With reference to Fig. XII-1, when we vary λ_i we are changing amplifiers and then reoptimizing the source for use with this new amplifier and thereby obtaining a new source point.
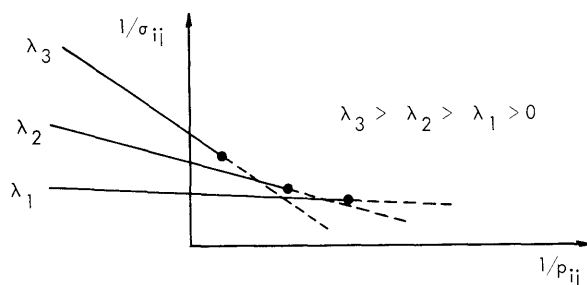

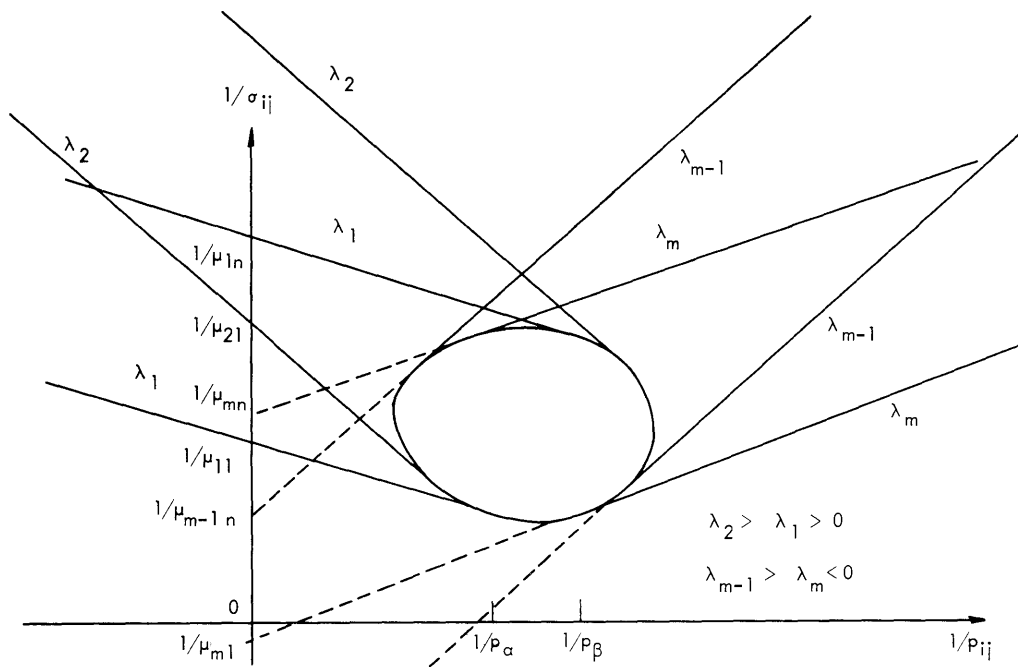Fig. XII-2. Characteristic curves for $\lambda_i > 0$, $1/\mu_{i1}$.

Fig. XII-3. Typical set of characteristic curves.

The closed curve generated is tangent to each of the characteristic lines. Physically, this curve represents the boundary of the region of values of the noise-to-signal ratio and the reciprocal of the exchangeable signal power that may be obtained at the output by imbedding only the source network, and not the amplifier, in an arbitrary lossless imbedding network.

In Fig. XII-3 we show a typical plot of characteristic curves for a positive definite source and an amplifier having both positive and negative eigenvalues. We have shown only those characteristics corresponding to extremal values of $1/\mu_i$ for four values of λ_i — the two smallest positive eigenvalues and the two smallest negative eigenvalues. (All solutions corresponding to intermediate eigenvalues of $1/\mu_i$ would give rise to characteristic curves that terminate at points inside the source region.) All of these characteristics may be interpreted. The line characterized by λ_1 and μ_{ij} is the true optimum-performance curve. It is the curve of the maximum signal-to-noise ratio as a function of exchangeable signal power for exchangeable signal power greater than p_α. The line characterized by λ_1 and μ_{1n} represents the worst way of increasing exchangeable signal power with the best reduction of the amplifier. The line characterized by λ_2 and μ_{21} is just a locus of points of stationary signal-to-noise ratio for fixed exchangeable signal powers. This same statement applies to all of the curves shown. The line characterized by λ_m and μ_{m1} represents the best way of reducing the

exchangeable signal power below p_β. Since λ_m is negative here, we are using the amplifier as a positive resistor; for that matter, we are using the least noisy positive resistor appearing in the cannonic form of the amplifier.

<div align="right">W. D. Rummler</div>

References

1. W. D. Rummler, Optimum noise performance of multiterminal amplifiers, Quarterly Progress Report No. 66, Research Laboratory of Electronics, M. I. T., July 15, 1962, pp. 71-77.