ANALYSIS OF SIGNALS AND NOISE IN LONGITUDINAL ELECTRON BEAMS

H. A. HAUS

TECHNICAL REPORT 306

AUGUST 18, 1955

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

The Research Laboratory of Electronics is an interdepartmental laboratory of the Department of Electrical Engineering and the Department of Physics.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army (Signal Corps), the Navy (Office of Naval Research), and the Air Force (Office of Scientific Research, Air Research and Development Command), under Signal Corps Contract DA36-039 SC-64637, Project 102B; Department of the Army Project 3-99-10-022.

This report is published by special permission from NOISE IN MICROWAVE DEVICES. To be published jointly by the Technology Press of M.I.T. and John Wiley & Sons, Inc., New York.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report 306 August 18, 1955

ANALYSIS OF SIGNALS AND NOISE IN LONGITUDINAL ELECTRON BEAMS

H. A. Haus

Abstract

The theory of signal propagation in longitudinal one-dimensional electron beams is reviewed. The kinetic power theorem is proven and used for the characterization of longitudinal-beam microwave amplifiers in terms of matrices of lossless networks.

The properties of noise in electron beams are studied. The two noise parameters, invariants with regard to lossless beam transformations, are derived from a simple theorem of matrix algebra. Equivalent noise impedances are defined. As a result, noise transformations in an electron beam can be handled by conventional impedance transformation methods.

The noise theory is then applied to derive the expression for the minimum noise figure of longitudinal-beam tubes. Applications to practical cases are discussed.

Preface

In the summer of 1955 a special course was given at Massachusetts Institute of Technology on "Noise in Electron Devices." Research workers from industrial research laboratories and from M. I. T. and other universities presented various topics on noise, including their own most recent work. The present report is based on notes used by the author for the course.

An attempt was made in the original notes to present a coherent story on noise in electron beams, starting from very simple fundamental notions. This attempt is continued in this report. The work done by the author and his friend and co-worker, F. N. H. Robinson, forms only a fraction of the material presented here. Yet, inclusion of other material was necessary in order to achieve clarity of presentation. It is the very inclusion of such material which, it is hoped, will make this report fill a need that cannot be met by short journal articles. It is hoped that in presenting the work of others proper credit has been given. If it has not, the author begs forgiveness for such unintentional omissions.

Grateful acknowledgment is given to Professor L. J. Chu, who made possible the results reported here through his own fundamental work. The major part of the author's work was carried out under Prof. Chu's personal guidance and supervision. All of the work had the benefit of his advice.

H. A. Haus

August 1, 1956

Table of Contents

1. ANALYSIS OF SIGNAL PROPAGATION ALONG ELECTRON BEAMS

1.1 INTRODUCTION

The power amplification of a conventional triode is based on the control of the current from the cathode by a potential applied to the grid. At low frequencies the control of the current can be achieved without expenditure of power if the grid current is reduced to zero by a proper bias. The electron flow can be analyzed mathematically like a time-independent stationary process. Any slow time variation of the process is represented by a simultaneous time variation of all parameters of the process. Such an analysis might be called "quasi-stationary." At frequencies at which the electron transit time through the tube is not negligibly small compared to a period of the applied rf grid voltage, the quasi-stationary analysis is inadequate. Transit-time effects cause induced grid currents. The existence of grid currents calls for a supply of rf power to the control grid. Successful attempts to minimize transit-time grid loading and other undesirable high-frequency effects have led to the modern microwave triode. Simultaneously, new principles of amplification have been recognized and put to use in a new class of amplifiers, microwave-beam amplifiers. In these amplifiers transit-time effects are used to advantage. In this report we shall deal with the noise performance of one subclass of microwave-beam amplifiers, longitudinal-beam amplifiers. They were chosen for attention partly because the noise in these amplifiers is by now fairly well understood, partly because all low-noise microwave-beam amplifiers that have been built to date are members of this class.

The analysis of longitudinal-beam amplifiers has some features in common with the analysis of the high-frequency triode and other related tubes. These we shall call "space-charge-control" tubes, referring to their basic principle of operation. The mathematical approach of section 1.2 is also used in the high-frequency analysis of noise in space-charge-control tubes.

One feature distinguishes microwave longitudinal-beam amplifiers from conventional space-charge-control tubes. In the latter tubes the applied rf fields act on the electron beam while it passes through the potential minimum in front of the cathode. The former tubes employ an electron beam formed in an electron gun that is free of applied rf fields. Examples of longitudinal-beam amplifiers are the traveling-wave tube (1), the klystron (2, 3, 4), the resistive-wall amplifier (5), the space-charge-wave amplifier (6), the double-stream amplifier (7), the rippled-wall and rippled-stream amplifier (8), the backward-wave amplifier (9), and so on. The designer of a low-noise, longitudinal-beam amplifier can take advantage of the fact that the electron beam is formed in a region free of applied rf fields. He can design structures surrounding the beam in front of the rf interaction region of the amplifier. Such a structure, if properly chosen, can reduce the noise output of the amplifier without affecting its gain. The theory of such noise-reducing schemes, and their limitations, will be the topic of this report. An expression will be derived for the minimum obtainable noise figure of a longitudinal-beam amplifier using a beam with a given noise. A logical definition of the "noisiness" of an electron beam follows from the expression for the minimum noise figure. It will be shown that noise-reducing structures preceding the amplifier are, in principle, sufficient for attaining the minimum noise figure, and that elaborate feedback and noise-cancellation schemes within the amplifier cannot lead to a lower minimum noise figure. Finally, it is shown that a traveling-wave tube with negligible loss in its rf structure preceded by conventional noise-reducing schemes (10) attains, in principle, the minimum noise figure.

1.2 ASSUMPTIONS

An exact analysis of the propagation of signals and noise along electron beams is extremely difficult. Certain approximations have to be made before the problem becomes amenable to a mathematical treatment.

In all longitudinal-beam amplifiers, amplification is obtained through an energy transfer from the electron motion to predominantly longitudinal electric fields. If a very large longitudinal magnetic focusing field confines the motion of the beam, the energy transferred by the electrons comes entirely from the kinetic energy associated with the longitudinal motion. In the analysis of longitudinal-beam amplifiers the assumption of an infinite magnetic focusing field is made quite often because it represents the physical facts adequately for many purposes, and leads to mathematical simplicity.

The electrons emitted at random from the cathode in the electron gun of a longitudinal-beam amplifier form a beam in which the velocity and the density of the electrons passing any reference cross section fluctuate statistically. Part of the noise output of an amplifier which employs the beam is caused by the currents induced in the amplifier structure by the fluctuations in the beam. Also, some of the electrons may be intercepted by the rf structure of the amplifier in a more-or-less random fashion if the beam is inadequately focused. This latter source of noise is commonly called "partition noise." In a well-designed amplifier the interception current can be kept to less than 0.5 per cent of the total beam current. Under such conditions the effect of partition noise is negligible compared to the noise induced in the rf structure. The present analysis will deal solely with the latter.

An electron beam consists of a large, but finite, number of electrons. The electrons in the beam interact by virtue of their Coulomb repulsion force. Contributions to the force on any particular electron come partly from the next neighbors, partly from electrons farther away. The forces upon any electron exerted by its next neighbors fluctuate rapidly with time. These forces are usually referred to as "short-range collisions." The forces exerted upon the electron by electrons farther away behave more regularly. These forces are the "long-range collisions." Up to the present time, all signal and noise analyses of electron beams have neglected the granular nature of the charge. (See reference 11 for an interesting discussion of this approximation.) The electron beam is treated as a "fluid," made up of an infinite number of infinitesimally small particles with an infinitesimal charge. The effect of the short-range collisions among the particles is neglected in the case of high-vacuum electron beams. An argument that this approximation is legitimate has been given by Mott Smith (12) for beams of reasonable length and current density. Experimental results corroborate the validity of this approximation (13).

The granular nature of the electron beam is the cause of noise. In the analysis of an electron beam as a "fluid" this effect is taken into account a posteriori in terms of appropriate noise input conditions at the cathode, or at some reference plane further on in the beam.

The small-signal theory is used for the analysis of noise in electron beams. The assumption is made that the excitation of the beam can be treated as a small perturbation of the time-average conditions in the beam. The approximations of small-signal theory are apparently good anywhere along the electron beam except at the potential minimum. Whether or not the small-signal assumption is applicable to the region of the potential minimum at high frequencies is not clear. But this question is academic, since no high-frequency analysis of the electron interaction in the potential minimum region exists today.[†]

Electrons emerging from the cathode have different velocities with a Maxwellian distribution. The analysis of an electron beam as a charged "fluid" still retains this picture. Parts of the fluid with higher velocities drift through parts with lower velocities without friction. Friction is neglected as soon as the effect of short-range collisions is disregarded. The single-velocity theory makes the assumption that a perturbation in the beam can be treated as if all electrons passing a beam cross section had the same velocity. This theory has been adopted almost exclusively in microwave work. Its justification has been discussed by various authors (14,15,16). The conclusion is that single-velocity theory yields results in good agreement with the more sophisticated multivelocity theory as long as the range of velocities possessed by the majority of the electrons is small compared to their average velocity. Such a situation prevails in a beam that has been accelerated to a few volts above the potential minimum. At the potential minimum in front of a space-charge-limited cathode this condition is obviously violated. The single-velocity theory applied to the potential minimum region cannot give better than qualitative answers. This difficulty is circumvented in the analysis of noise in electron beams in this report. The noise input conditions to the electron beam are stated at a cross section beyond the potential minimum, chosen so that the small-signal and single-velocity theories are applicable at, and beyond, the cross section. No specific values are assumed for the noise parameters at the reference cross section. The evaluation of the noise parameters is left to a detailed analysis of the potential minimum region which does not make the single-velocity assumption.[†]

[†] In the meantime, such an analysis has been carried out by P. K. Tien on a digital computer in a way that avoids the small signal assumption.

Finally, there is the assumption of the one-dimensional theory. Only a single spatial co-ordinate, the co-ordinate along the electron beam, is retained in the analysis. The one-dimensional theory in conjunction with all of the assumptions introduced above leads to a characterization of a perturbation in the beam in terms of two modulation parameters; for example, the velocity and current modulations. Strictly speaking, this assumption implies that the electron beam is of an infinite parallel-plane (or spherical) geometry, with the motion of the electrons confined to the longitudinal axis (in the direction of the radius vector). All parameters of the beam are assumed independent of the co-ordinates transverse to the beam. A more practical case which fits into the one-dimensional formalism is a freely drifting electron beam of finite cross section on which only two space-charge waves are excited (17, 18). Two waves can be described by two parameters; their respective amplitudes at a given cross section, for example. However, if the beam passes through transition regions in which its shape or time-average velocity is changed, cross-coupling among the different modes of the beam occurs and, in place of the original two waves, many other waves travel along the beam. Then, the one-dimensional formalism yields only approximate answers; the smaller the cross-coupling of modes in a transition region, the better the approximation should be. This happens in a thin beam, and the one-dimensional theory gives good approximate answers. Recently, the noise analysis has been extended to systems propagating any number of modes (19). The results are, however, rather complex, so that their presentation here does not seem warranted.

Before the noise in electron beams can be analyzed, the propagation of signals along an electron beam has to be understood. Section 1.2 and all of section 2 are devoted to this problem.

1.21 THE BASIC EQUATIONS

The first two equations of Maxwell give the electric field $\bar{E}(\bar{r}, t)$ and the magnetic field $\bar{H}(\bar{r}, t)$ produced by a given current distribution $\bar{J}(\bar{r}, t)$. The vectors $\bar{E}$, $\bar{H}$ and $\bar{J}$ are all functions of the radius vector $\bar{r}$, and the time t.

$$\nabla \times \bar{E}(\bar{r},t) = -\mu \frac{\delta}{\delta t} \bar{H}(\bar{r},t) \tag{1.1}$$

$$\nabla \times \bar{H}(\bar{r},t) = \bar{J}(\bar{r},t) + \epsilon \frac{\delta}{\delta t} \bar{E}(\bar{r},t) \tag{1.2}$$

The force equation gives the relation between the acceleration of charged particles and the field. If the motion of the particles is confined by an infinite magnetic focusing field in the z-direction, or, if the velocity is z-directed for other reasons in the absence of a time-average magnetic field, we have

$$\frac{\delta v(\bar{r},t)}{\partial t} + v(\bar{r},t)\,\frac{\partial v(\bar{r},t)}{\partial z} = \frac{d}{dt}v(\bar{r},t) = \frac{e}{m}E_z(\bar{r},t) \tag{1.3}$$

where e is the charge of the particle (for an electron it is a negative quantity) and m is its mass. Equation 1. 3 neglects the force upon the particle from rf magnetic fields, an approximation legitimate at nonrelativistic velocities. The quantity $v(\bar{r},t)$ is the z-component of the velocity. For the sake of brevity no subscript z is used. The continuity equation is, if the current is entirely z-directed,

$$\frac{\delta}{\delta z}J(\bar{r},t) = -\frac{\delta}{\delta t}\rho(\bar{r},t) \tag{1.4}$$

If the single-velocity assumption is made, the current density is given as the product of the velocity and space-charge density:

$$J(\bar{r},t) = v(\bar{r},t)\,\rho(\bar{r},t) \tag{1.5}$$

Under the small-signal assumption all quantities can be split into a time-average part, and a time-varying part which is much smaller in amplitude than the time-average part. In evaluating the time-dependent parts, cross products of the time-varying quantities can be neglected. The resulting equations for the time-varying quantities become linear, and thus a sinusoidal excitation of frequency ω causes all time-dependent quantities to vary at the same frequency. The superposition principle can be applied under the small-signal approximation. Complex notation can be used to represent the time-varying quantities. We can write

$$\begin{aligned}
\bar{E}(\bar{r},t) &= \bar{E}_o(\bar{r}) + \mathrm{Re}\,[\hat{E}(\bar{r})\,e^{j\omega t}] \\
\bar{H}(\bar{r},t) &= \bar{H}_o(\bar{r}) + \mathrm{Re}\,[\hat{H}(\bar{r})\,e^{j\omega t}] \\
J(\bar{r},t) &= J_o(\bar{r}) + \mathrm{Re}\,[J(\bar{r})\,e^{j\omega t}] \\
v(\bar{r},t) &= u(\bar{r}) + \mathrm{Re}\,[v(\bar{r})\,e^{j\omega t}] \\
\rho(\bar{r},t) &= \rho_o(\bar{r}) + \mathrm{Re}\,[\rho(\bar{r})\,e^{j\omega t}]
\end{aligned} \tag{1.6}$$

The circumflex is used to indicate complex vector quantities. These definitions introduced into Eqs. 1. 1 - 1. 5 lead to a separation between the time-average and time-varying parts. The time-dependent part of Maxwell's equations is

$$\nabla \times \hat{E}(\bar{r}) = -j\omega\mu\hat{H}(\bar{r}) \tag{1.7}$$

$$\nabla \times \hat{H}(\bar{r}) = \bar{a}_z J(\bar{r}) + j\omega\epsilon\hat{E}(\bar{r}) \tag{1.8}$$

where $\bar{a}_z$ is the unit vector in the z-direction. The current has been assumed to be entirely z-directed. The time-average and time-dependent parts of the force equation are under the same assumptions:

$$u(\bar{r}) \frac{\delta}{\delta z} u(\bar{r}) = \frac{e}{m} E_{oz}(\bar{r}) \tag{1. 9a}$$

$$j\omega v(\bar{r}) + \frac{\delta}{\delta z} [u(\bar{r})\, v(\bar{r})] = \frac{e}{m} E_z(\bar{r}) \tag{1. 9b}$$

The continuity equation gives

$$\frac{\delta}{\delta z} J_o(\bar{r}) = 0 \tag{1. 10a}$$

$$\frac{\delta}{\delta z} J(\bar{r}) = -j\omega\rho(\bar{r}) \tag{1. 10b}$$

and from Eq. 1. 5 we have

$$J_o(\bar{r}) = u(\bar{r})\, \rho_o(\bar{r}) \tag{1. 11a}$$

$$J(\bar{r}) = u(\bar{r})\, \rho(\bar{r}) + \rho_o(\bar{r})\, v(\bar{r}) \tag{1. 11b}$$

1. 22 THE INFINITE PARALLEL-PLANE BEAM

We shall now analyze the propagation of signals along an electron beam of infinite cross section. The time-average current density and velocity are constant throughout the cross section. The positive z-direction is picked as the direction of positive velocity and current. Thus, a convection current of negatively charged particles with a positive velocity is, by convention, negative. (See Figure 1. 1.)

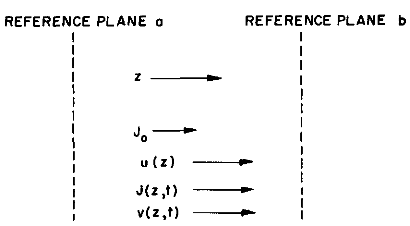

Fig. 1.1. Direction of positive current and velocity.

The time-average velocity can be found from an integration of the time-average part of the force equation, Eq. 1. 9a. Since the electron flow is assumed to depend merely upon the z-co-ordinate, we can replace the independent variable $\bar{r}$, which consists of all three co-ordinates, by z. We have, from Eq. 1. 9a,

$$\frac{m}{2e} \frac{d}{dz} [u(z)^2] = E_{oz}(z) \tag{1. 12}$$

The field, in turn, is determined by the space-charge distribution in the electron beam. From Gauss' law in its one-dimensional form, we have

$$\frac{d}{dz} E_{oz}(z) = \frac{\rho_o(z)}{\epsilon} \tag{1.13}$$

Since the time-average current density is independent of z (according to Eq. 1.10a) and is in turn related to the time-average velocity and space-charge density by Eq. 1.11a, we can express the space-charge density in terms of J_o and u(z). Once this is done, Eqs. 1.12 and 1.13 contain only two unknown functions, u(z) and $E_{oz}(z)$, and they can be solved subject to appropriate boundary conditions. The variety of such boundary conditions is great. It is conceptually possible to construct an arbitrary dc potential distribution with the aid of infinitely permeable grids, open-circuited for rf, to which arbitrary dc potentials are applied.

Let us assume that the time-average velocity and space-charge density have been found in the way described above. Let us further assume that the rf excitation applied to the electron beam is independent of the transverse co-ordinates x and y.

In this case the curl of $\hat{H}(\bar{r})$ in Eq. 1.8 must be zero. We thus have, from Eq. 1.8,

$$E_x(z) = E_y(z) = 0$$

and

$$J(z) + j\omega\epsilon E_z(z) = 0 \tag{1.14}$$

Equation 1.14 shows that the sum of the convection and displacement current densities is zero in an infinite parallel-plane beam. It can also be observed that there is no transverse electric field in an infinite parallel-plane electron beam. Under an excitation uniform in the transverse direction the motion of the electrons is entirely longitudinal. No magnetic focusing field is required for the confinement of the motion. Omitting, from now on, any explicit indication of the z-dependence of the process, we obtain from Eqs. 1.14 and 1.9b,

$$j\omega v + \frac{d}{dz}[uv] = -\frac{e}{m}\frac{1}{j\omega\epsilon} J \tag{1.15}$$

Equations 1.10b and 1.11b lead to the expression

$$j\omega J + u\frac{d}{dz} J = j\omega\rho_o v \tag{1.16}$$

Equations 1.15 and 1.16 are put into a symmetrical form by introducing the new dependent variable

$$V = \frac{m}{e} uv \tag{1.17}$$

(The quantity V was introduced by L. J. Chu as the "kinetic voltage modulation"(21). Its particular significance will become apparent later.) Further, we write

$$\frac{e}{m} \frac{\rho_o}{\epsilon} = \omega_p^2$$

ω_p has the dimensions of frequency and is, in general, a function of distance. It is commonly called the plasma frequency, because an electron plasma of uniform density ρ_o oscillates at this frequency. With this new notation we can write Eqs. 1.15 and 1.16 as

$$\left(j\frac{\omega}{u} + \frac{d}{dz}\right)V = -\frac{1}{j\omega\epsilon} J \tag{1.15a}$$

$$\left(j\frac{\omega}{u} + \frac{d}{dz}\right)J = j\omega\epsilon \frac{\omega_p^2}{u^2} V \tag{1.16a}$$

Physical reasoning leads us to a change of the dependent variables in Eqs. 1.15a and 1.16a which improves their appearance. An electron beam is a system of charged particles that interact through their space-charge repulsion forces. The sytem moves with the time-average velocity u. The mere fact that the electrons move with a finite time-average velocity implies that any perturbation applied to the electron beam at some point, let us say z = 0, arrives time-delayed at a later point z > 0. If space-charge forces have time to act upon the electrons during their travel between the points z = 0 and z, the velocity of the electrons gets modified. This is an effect over and above the natural time delay. We acknowledge this time delay by introducing a transformation of the dependent variables in which the time delay is brought out explicitly:

$$V = Ue^{-j\theta} \tag{1.18}$$

$$J = Qe^{-j\theta} \tag{1.19}$$

where θ is the transit angle between the reference plane z = 0 and the point z.

$$\theta = \omega \int_o^z \frac{dz}{u(z)} \tag{1.20}$$

In terms of these new variables, Eqs. 1.15a and 1.16a assume the more attractive form

$$\frac{d}{dz} U = -\frac{1}{j\omega\epsilon} Q \qquad (1.21)$$

$$\frac{d}{dz} Q = j\omega\epsilon \frac{\omega_p^2}{u^2} U \qquad (1.22)$$

Equations 1.21 and 1.22 are special forms of the equations formulated by Llewellyn (22). They have the appearance of transmission-line equations with a pure imaginary impedance per unit length $Z = 1/j\omega\epsilon$ and a pure imaginary admittance per unit length $Y = -j\omega\epsilon\ \omega_p^2/u^2$. U and Q play the role of voltage and current on the analog transmission line (see Fig. 1.2). Similar equations have been obtained for a spherical flow (24) and one-dimensional flow of any general geometry (25). The fact that the impedance and admittance per unit length of the analog transmission line are purely imaginary implies that the transmission line is lossless. Along such a transmission line the power must be independent of distance. The time-average power flow along a transmission line is given by one-half of the real part of the complex product of the voltage U and current Q on the transmission line, Re $[UQ^*]/2$. Thus, we have for the analog power

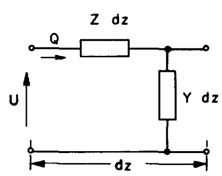

Fig. 1.2. Analog transmission line of an electron beam.

$$\frac{1}{2} \mathrm{Re}[U(z_1)\, Q(z_1)^*] = \frac{1}{2} \mathrm{Re}[U(z_2)\, Q(z_2)^*]$$

where z_1 and z_2 are the positions of two reference planes along the transmission line. According to the definitions of Eqs. 1.18 and 1.19 we have $UQ^* = VJ^*$, and thus

$$\frac{1}{2} \mathrm{Re}\, [V(z_1)\, J(z_1)^*] = \frac{1}{2} \mathrm{Re}\, [V(z_2)\, J(z_2)^*]$$

L. J. Chu (21) defined the quantity

$$\frac{1}{2} \mathrm{Re}\, [V(z)\, J(z)^*] = \mathrm{Re}\,(S_k) \qquad (1.23)$$

as the "real kinetic power density" of the electron beam. Since one can easily check that the kinetic power density has the dimension of power per unit area, the name seems, at least, partly justified. Later on we shall see that there are even more compelling reasons for the name. For the moment it is sufficient to note that the kinetic power density is independent of distance in any system of infinite parallel-plane geometry, regardless of the distribution of the time-average potential. In such a system the magnetic field $H(z, t) = 0$ and, correspondingly, there is no flow of electromagnetic power in an infinite parallel-plane beam. In section 2. 2 we shall see that this fact is responsible for the conservation of kinetic power.

The "transmission-line equations" (1. 21 and 1. 22) lead to differential equations of second order with two solutions. These solutions can be used to satisfy arbitrary boundary conditions. If the kinetic voltage and current-density modulations are known at any cross section of the beam, they are known everywhere. From the mathematical point of view the cross section at which the initial conditions are given is arbitrary.

However, since the electromagnetic power flow in an infinite parallel-plane geometry is zero, an excitation is not transmitted through the beam by electromagnetic radiation, but is transported along the beam by the electrons. The excitation propagates in the direction of motion of the electrons. An excitation in the region of an electron beam is thus most naturally given in terms of the boundary conditions at the input to the region.

1. 23 THE BEAM IN A DRIFT REGION

An electron beam that flows between two electrodes that are a finite distance apart, both at the same potential, causes a potential depression between the electrodes with a potential minimum situated half way between them. The closer together the two electrodes are, the smaller the potential depression. The potential depression is negligibly small; thus the time-average fields are negligible when the spacing between the electrodes is infinitesimal. An electron beam that drifts freely with no time-average forces acting upon it can be realized by a system of electrodes, all at equal time-average potential, spaced very closely together and open-circuited for rf.

An electron beam between two equipotential electrodes a finite distance apart, neutralized by heavy positive ions, acts in the same way. The ions cannot follow the rf changes of the field and thus do not affect them. However, all time-average fields are eliminated, since the ions can follow them, although sluggishly, until they fill the potential minima and compensate the charge of the electrons with their own positive charge. The two methods of realizing a drift region are artifices used to adapt the model of an infinite parallel-plane beam to represent a more physical situation: a finite longitudinal beam surrounded by a perfectly conducting cylindrical wall and confined by a large, ideally infinite, longitudinal magnetic field. In this latter case the time-average electric fields produced by space charges are entirely radial. Therefore the time-average velocity of the electrons is independent of z. The main features of propagation of signals along such a finite beam are contained in the model of the infinite parallel-plane beam in a drift region. This accounts for the importance attributed to the problem of the infinite parallel-plane beam in a drift region.

In the absence of time-average fields, the time-average velocity of the electrons in the infinite parallel-plane beam cannot change. Correspondingly, the dc charge density ρ_o, and thus the plasma frequency ω_p, are independent of distance. Equations 1. 21 and 1. 22 can be solved very easily in terms of two arbitrary constants:

$$U = U_+ e^{j\beta_p z} + U_- e^{-j\beta_p z}$$

$$Q = \beta_p \omega \epsilon (U_+ e^{j\beta_p z} - U_- e^{-j\beta_p z}) \quad , \quad \text{with } \beta_p = \frac{\omega_p}{u}$$

The kinetic voltage and the current density become

$$V = (U_+ e^{j\beta_p z} + U_- e^{-j\beta_p z})\, e^{-j\beta_e z} \tag{1.24}$$

$$J = \beta_p \omega \epsilon (U_+ e^{j\beta_p z} - U_- e^{-j\beta_p z})\, e^{-j\beta_e z} \quad , \quad \text{with } \beta_e z = \omega z/u \tag{1.25}$$

If we single out one part of the beam of cross-sectional area F, the time-average current flow through this area is $I_o = u F \rho_o$, and the time-varying current is $i = FJ$. We define the characteristic admittance of the beam by

$$Y_o = F \beta_p \omega \epsilon = \frac{F \frac{e}{m} \rho_o}{u} \frac{\omega}{\omega_p} = \frac{I_o}{\frac{m}{e} u^2} \frac{\beta_e}{\beta_p} \tag{1.26}$$

Note that both I_o and e are negative quantities. Thus, the last expression above is positive. The characteristic impedance of the beam Z_o is defined as the inverse of Eq. 1.26, $Z_o = 1/Y_o$. With the aid of these definitions we can write the solutions for the kinetic voltage and rf current in the beam of cross section F in the form:

$$V = U_+ e^{j\beta_p z} + U_- e^{-j\beta_p z})\, e^{-j\beta_e z} \tag{1.24a}$$

$$i = Y_o (U_+ e^{j\beta_p z} - U_- e^{-j\beta_p z})\, e^{-j\beta_e z} \tag{1.25a}$$

According to Eqs. 1.24a and 1.25a, two wave solutions exist in the beam. Their propagation constants are, respectively, $(\beta_e + \beta_p)$ and $(\beta_e - \beta_p)$. The wave with the propagation constant $\beta_e + \beta_p$ has a phase velocity smaller than the time-average beam velocity, u. The wave with the propagation constant $\beta_e - \beta_p$ travels with a phase velocity larger than the beam velocity provided that $\beta_e > \beta_p$. The case of $\beta_e < \beta_p$ never occurs in practice. The reason for this will be discussed in section 2.1. Thus, we shall call the wave with the propagation constant $\beta_e - \beta_p$ simply the "fast wave," implying that the inequality $\beta_e > \beta_p$ is satisfied.

$$\frac{1}{2} \mathrm{Re}\, [F V J^*] = \frac{1}{2} \mathrm{Re}\, [V i^*] = \frac{1}{2} Y_o (|U_+|^2 - |U_-|^2) \tag{1.27}$$

According to Eq. 1.27 the real kinetic power of the two waves is additive (orthogonality of power flow) – the fast wave carrying positive power, the slow wave carrying negative power.

The kinetic power carried by a beam in a drift tube, Eq. 1.27, can be written in a more elegant form by defining normalized wave amplitudes a_1 and a_2 as

$$a_1 = (2Z_o)^{-1/2} U_+ \qquad a_2 = -(2Z_o)^{-1/2} U_- \tag{1.28}$$

With the aid of this definition we can write the kinetic power in the form

$$\frac{1}{2} \mathrm{Re}\,[V i^*] = |a_1|^2 - |a_2|^2 \tag{1.29}$$

We shall find the use of normalized wave amplitudes very convenient later on.

The resemblance of Eqs. 1.24a and 1.25a to transmission-line solutions is apparent. This result is not surprising, since Eqs. 1.21 and 1.22, from which Eqs. 1.24a and 1.25a have been derived, have the form of transmission-line equations. It is important, however, to note the difference between transmission-line solutions and the solutions of a beam in a drift region. Both the voltage and current modulations are multiplied by a factor $e^{-j\beta_e z}$, which does not appear in the transmission-line solutions. Instead of two waves, one forward and one backward, we have a fast wave and a slow wave, both with phase velocities in the direction of the flow of the beam.

It is natural to suppose that techniques of conventional transmission-line theory can be applied to the beam problem. Equations 1.24a and 1.25a show that this is possible if proper precautions are taken. As is common practice in transmission-line analysis, we define an impedance at any cross section z by

$$Z = \frac{V}{i} = Z_o \frac{1+\Gamma}{1-\Gamma} \tag{1.30}$$

where

$$\Gamma = \frac{U_-}{U_+} e^{-2j\beta_p z} \tag{1.31}$$

Equation 1.30 is formally identical with the well-known relation between the impedance and the reflection coefficient on a transmission line. The reflection coefficient is defined, as usual, as the ratio of the amplitudes of the voltage in the wave-carrying negative power (-) and the wave-carrying positive power (+). The role of the (-) wave is played, in the beam problem, by the slow wave, whose kinetic power is negative, quite analogous to the negative electromagnetic power carried by the reflected wave of transmission-line theory. Equation 1.31 shows one important difference between the transmission-line problem and the beam problem. The angle of the reflection coefficient Γ decreases with increasing z, whereas the opposite is true for the reflection coefficient in transmission-line theory. The role of the wavelength is played in the beam problem by the quantity $\lambda_p = 2\pi/\beta_p$, the so-called plasma wavelength.

The bilinear relation between Z and Γ as shown in Eq. 1.30 is conveniently represented in the plane of complex Γ, the Smith chart of transmission-line theory (27). Motion in the positive z direction along the electron beam corresponds to clockwise rotation in the Γ-plane at constant Γ. A shift by half a plasma wavelength along the electron leaves Γ unchanged.

1.24 THE GENERAL SOLUTIONS OF LLEWELLYN'S EQUATIONS

Equations 1.21 and 1.22 can be solved for conditions other than those of a drift region. An important case is the one of an electron beam traveling between two completely permeable grids at potentials V_{oa} and V_{ob}, under the influence of its own space charge. The details of the solution are rather tedious and not within the scope of this discussion. The details are given in references 22 and 23. We list here only the results in the form of a table (Table I). Certain changes of notation have been made to conform with our notation. In particular, the change in the sign convention for the current should be noted.

Table I

1. Time-Average Solutions

Relation between potential and velocity:

$$\left|\frac{e}{m}\right| V_o = \frac{1}{2} u^2 \quad , \quad \text{where } |e/m| = 1.76 \times 10^{11} \text{ in rationalized mks units}$$

Definition of space charge factor:

$$\xi = 3\left(1 - \frac{T_o}{T}\right)$$

where T is the transit time of the electrons between planes a and b, T_o is the transit time between the same planes in the absence of space charge, with the potentials at the cross sections unchanged.

Relation between the space charge factor ξ, distance between reference cross sections d, transit time T, and initial and final velocities u_a and u_b:

$$d = \left(1 - \frac{\xi}{3}\right)(u_a + u_b)\frac{T}{2}$$

Current density:

$$J_o = \frac{\epsilon m}{e}(u_a + u_b)\frac{2\xi}{T^2}$$

Ratio of the actual current density to the maximum possible current density J_{max}:

$$\frac{|J_o|}{|J|_{max}} = \frac{9\xi}{4}\left(1 - \frac{\xi}{3}\right)^2$$

Maximum current density:

$$|J|_{max} = \frac{2.33 \times 10^{-6}[V_{oa}^{1/2} + V_{ob}^{1/2}]^3}{d^2} \text{ amps/unit area}$$

2. RF Solutions

$$V_b = AV_a + BJ_a$$

$$J_b = CV_a + DJ_a \qquad (1.32)$$

V_a and V_b are the kinetic voltage modulations at the reference cross sections a and b, respectively; J_a and J_b are the corresponding current-density modulations. In the equations above the assumption is made that the two grids at the cross sections a and b are rf open-circuited; the solutions obtained by Llewellyn and Peterson are more general (22, 23).

The coefficients A to D are given by:

$$A = \frac{1}{u_a}[u_a - \xi(u_a + u_b)]e^{-j\theta} \qquad (1.33)$$

$$B = \frac{-T^2}{2\epsilon}(u_a + u_b)(1 - \xi)\frac{e^{-j\theta}}{j\theta} \qquad (1.34)$$

$$C = \frac{2\epsilon\xi}{u_a T^2} \frac{u_a + u_b}{u_b} j\theta e^{-j\theta} \qquad (1.35)$$

$$D = \frac{1}{u_b}[u_b - \xi(u_a + u_b)]e^{-j\theta} \qquad (1.36)$$

2. MATRIX REPRESENTATION OF MICROWAVE AMPLIFIERS

The problem of interaction between an electron beam and electromagnetic fields is solvable in closed form only under the assumption of small-signal theory. Once this assumption is made, the differential equations of the system are linear. The solutions are then linear functions of the excitation of the system on its boundaries. It is convenient to write linear relations among sets of variables in matrix form. In section 2.1 we derive a basic relation of small-signal theory that will suggest a convenient matrix representation of an amplifier. Sections 2.2 and 2.3 are devoted to a study of the restrictions imposed on the matrices.

2.1 KINETIC POWER THEOREM

The definition of kinetic power density was introduced by Eq. 1.23. The significance of the kinetic power concept is studied in greater detail in this section.

Amplification of electromagnetic energy in an electron tube occurs at the expense of the kinetic energy of the electrons. The flow of kinetic energy into a longitudinal-beam microwave amplifier minus the flow of the kinetic energy out of the tube is equal to the electromagnetic power delivered to the rf structure surrounding the beam. Unfortunately, difficulties are encountered in attempting to make use of this simple statement.

The small-signal theory linearizes the equations of the electron beam and thus facilitates a solution. But small-signal theory neglects squares and cross products of the amplitudes of the excitation. Energy and power relations involve squares and cross products of the small-signal amplitudes which are of the same order of magnitude as the terms neglected in the small-signal approximation. Thus, it seems that a discussion of energy and power associated with an electron beam is bound to be inconsistent if it is based on small-signal assumptions.

A closer look at the problem is less discouraging. An identity analogous to the Poynting theorem can be derived for the longitudinal beam of Fig. 2.1, starting from the small-signal equations (Eqs. 1.7 and 1.8). These equations hold for a beam whose electrons are confined to an entirely longitudinal motion.

We take a scalar product of Eq. 1.7 with $\hat{H}(r)^*$, and of the complex conjugate of Eq. 1.8 with $E(r)$. Subsequent subtraction of the two equations gives

$$-\nabla \cdot [\hat{E}(\bar{r}) \times \hat{H}(\bar{r})^*] = E_z(\bar{r})\, J(\bar{r})^* + j\omega\,[\mu \hat{H}(\bar{r}) \cdot \hat{H}(\bar{r})^* - \epsilon \hat{E}(\bar{r}) \cdot \hat{E}(\bar{r})^*] \qquad (2.1)$$

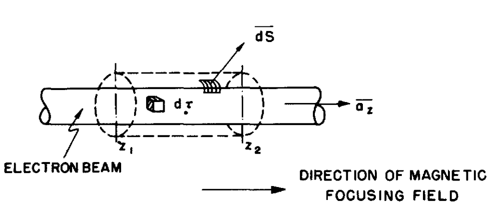

Fig. 2.1. Volume of integration in Eq. 2.5.

Equation 2.1 looks like the conventional Poynting theorem. It differs from it in establishing an identity among the approximate small-signal solutions of Maxwell's equations.

Through the use of the force equation (Eq. 1.9b), the continuity equation (Eq. 1.10b), and the relation between current density, charge density, and velocity (Eq. 1.11b), we find

$$E_z(\bar{r})\, J(\bar{r})^* = \frac{m}{e}\left\{ j\,\omega \rho_o(\bar{r})\, |v(\bar{r})|^2 + \frac{\delta}{\delta z}\, [u(\bar{r})\, v(\bar{r})\, J(\bar{r})^*] \right\} \tag{2.2}$$

We define the complex kinetic power density according to L. J. Chu (compare Eq. 1.23):

$$\hat{S}_k(\bar{r}) = \frac{1}{2}\,\frac{m}{e}\, u(\bar{r})\, v(\bar{r})\, J(\bar{r})^*\, \bar{a}_z = \frac{1}{2} V(\bar{r})\, J(\bar{r})^*\, \bar{a}_z \tag{2.3}$$

where $V(\bar{r})$ is the kinetic voltage modulation defined by Eq. 1.17, and $\bar{a}_z$ is the unit vector in the z-direction. The definition of Eq. 2.3 introduced into Eq. 2.2 and that, in turn, applied to Eq. 2.1 leads to an alternate form of the small-signal Poynting theorem. Noting that

$$\frac{1}{2}\,\frac{\delta}{\delta z}\, [V(r)\, J(r)^*] = \nabla \cdot \hat{S}_k(r)$$

we find

$$- \nabla \cdot \left[\frac{1}{2} \hat{E}(\bar{r}) \times \hat{H}(\bar{r})^* + \hat{S}_k(\bar{r}) \right] = \frac{1}{2} j\omega \left[\mu \hat{H}(\bar{r}) \cdot \hat{H}(\bar{r})^* - \epsilon \hat{E}(\bar{r}) \cdot \hat{E}(\bar{r})^* + \frac{m}{e} \rho_o(\bar{r})\, |v(\bar{r})|^2 \right] \tag{2.4}$$

$\hat{S}_k$ is a complex vector in the direction of the flow, i.e., the z-direction, with the dimension of power density. We shall call it the "complex kinetic power density." Integration of Eq. 2.4 over the volume τ enclosed by the surface S, shown in Fig. 2.1 gives,

$$- \oint \left[\frac{1}{2} \hat{E}(\bar{r}) \times \hat{H}(\bar{r})^* + \hat{S}_k(\bar{r}) \right] \cdot d\bar{S} = \frac{1}{2} j\omega \int \left[\mu \hat{H}(\bar{r}) \cdot \hat{H}(\bar{r})^* - \epsilon \hat{E}(\bar{r}) \cdot \hat{E}(\bar{r})^* + \frac{m}{e} \rho_o(\bar{r})\, |v(\bar{r})|^2 \right] d\tau \tag{2.5}$$

The real part of Eq. 2.5 is

$$\mathrm{Re} \oint \left[\frac{1}{2} \hat{E}(\bar{r}) \times \hat{H}(\bar{r})^* + \hat{S}_k(\bar{r}) \right] \cdot d\bar{S} = 0 \tag{2.6}$$

Since the small-signal amplitudes of the electric and magnetic fields have been found by neglecting terms involving squares and cross products of the small-signal amplitudes, the integral $\mathrm{Re}\,[\oint E(r) \times H(r)^* \cdot dS]/2$ cannot give the electromagnetic power flow through the surface S exactly. However, it is clear that the integral gives the electromagnetic power flow correctly within second order of the small-signal amplitudes and neglects only terms of higher order. Such an approximation is legitimate, provided that the applied rf fields are very small perturbations of the time-average conditions in the beam and the rf structure surrounding the beam is not resonant at a multiple of the fundamental frequency. By means of Eq. 2.6 we can identify the electromagnetic power delivered by the beam in the volume τ by computing the net real kinetic power flow $-\mathrm{Re}\,[\oint \hat{S}_k(\bar{r}) \cdot \overline{dS}]$ into the volume on a small-signal basis. This is the content of the "Kinetic Power Theorem" first formulated by L. J. Chu (21).

The usefulness of the kinetic power theorem stems from its generality. It is applicable to electron flows of arbitrary geometry as long as the motion of the electrons is confined to one direction; in our case, the z-direction. Thus, for example, the electron motion in a freely-drifting, thin, longitudinal beam is governed by the kinetic power theorem. If the beam is surrounded by a perfectly conducting cylinder, no electromagnetic power can be extracted from the electron beam. A detailed analysis shows that such a thin electron beam propagates two space-charge waves whose field and current density are approximately uniform throughout the cross section of the beam, not unlike the waves propagating along an infinite parallel-plane beam, as found in section 1.22. These waves have propagation constants $\beta_e + \beta_q$ and $\beta_e - \beta_q$, where $\beta_e = \omega/u$ is the beam propagation constant as before, and β_q is the so-called reduced plasma propagation constant. It is related by a factor of less than unity to the plasma propagation constant β_p, computed from the space-charge density ρ_o and time-average velocity u, $\beta_p = (e\rho_o/m\epsilon u^2)^{1/2}$. The factor β_q/β_p is often referred to as the plasma frequency reduction factor. It is a function of the frequency of operation, ω, and geometry.

A complete analogy can be established between the propagation of the two space-charge waves along a thin beam and the waves in an infinite parallel-plane electron beam. Equations 1.24a and 1.25a apply to the propagation of space-charge waves along a thin beam if we replace the plasma propagation constant β_p by β_q, and the characteristic admittance Y_o of Eq. 1.26 by

$$Y_o = \left| \frac{I_o}{\frac{m}{e}u^2} \right| \frac{\beta_e}{\beta_q} \tag{2.7}$$

Instead of Eqs. 1.24a and 1.25a for the kinetic voltage and current in the thin electron beam, we have

$$V(z) = [U_+ e^{j\beta_q z} + U_- e^{-j\beta_q z}]e^{-j\beta_e z} \tag{2.8}$$

$$i(z) = Y_o\,[U_+ e^{j\beta_q z} - U_- e^{-j\beta_q z}]e^{-j\beta_e z} \tag{2.9}$$

In practice, β_q is always smaller than β_e, so that here at least the name "fast wave" is justified for the wave with the propagation constant $\beta_e - \beta_q$ (compare the statement in section 1.22). U_+ and U_- in the equations above are the amplitudes of the fast and slow waves, respectively, at z = 0. If we introduce normalized wave amplitudes a_1 and a_2 according to Eq. 1.28,

$$a_1 = (2Z_o)^{-1/2} U_+ \quad \text{and} \quad a_2 = -(2Z_o)^{1/2} U_- \tag{1.28}$$

we can write the real part of the kinetic power in a particularly simple form:

$$\frac{1}{2} \text{Re} \, [V(z) \, i(z)^*] = |a_1|^2 - |a_2|^2 \tag{2.10}$$

The real part of the kinetic power is independent of distance. This is to be expected on the basis of the kinetic power theorem if no electromagnetic power is extracted from the beam.

An excitation in an infinite parallel-plane beam is not accompanied by an rf magnetic field. On the other hand, an excitation in a thin beam is always associated with a finite rf magnetic field and thus causes, in general, both an electromagnetic and a kinetic power flow. It has been shown (28), however, that the electromagnetic power flow associated with the fast or slow wave is smaller in magnitude than the real kinetic power flow of the wave by a factor β_q/β_e, usually a small number.

Electromagnetic power can be extracted from a thin beam if it flows through a structure other than a drift tube. The helix of a traveling-wave tube is an example of a structure whose fields may impart to, or extract from, the beam electromagnetic power. In this instance it is convenient to adapt Eq. 2.6 for a thin beam; then the kinetic voltage $V(\bar{r})$ and the current density modulation $J(\bar{r})$ are independent of the transverse coordinates. The integration in Eq. 2.6 can be carried over the cross section of the beam with the result (see Fig. 2.1):

$$\frac{1}{2} \text{Re} \left[\oint \hat{E}(r) \times \hat{H}(r)^* \cdot \overline{dS} \right] = \frac{1}{2} \text{Re} \, [V(z_1) \, i(z_1)^* - V(z_2) \, i(z_2)^*] \tag{2.6a}$$

where i(z) is the rf current modulation in the beam at the cross section z. According to Eq. 2.6a any time-average electromagnetic power extracted from the electron beam between two cross sections z_1 and z_2 is balanced by a decrease in the real kinetic power. On the other hand, if electromagnetic power is fed into the beam, and the integral on the left is negative, then the real kinetic power between the two cross sections z_1 and z_2 must increase correspondingly.

Now that the role of the kinetic power flow is known, we can attempt to obtain a physical understanding of its meaning. No claim to rigor will be made in the following discussion. Equation 2.10 shows that the real kinetic power flow associated with the fast wave is positive. This follows from the fact that the kinetic voltage, and the current in the fast wave are in phase, as is shown in Eqs. 2.8 and 2.9. Thus, if we view at a particular cross section z an electron beam propagating a fast wave only, we find that the kinetic voltage reaches its maximum at the same instant of time as the current modulation. A positive value of the kinetic voltage corresponds, according to the definition of Eq. 1.17, to a negative value of the velocity modulation v, since the electron charge e is negative. Thus, at the instant of time when the kinetic voltage is a maximum, the total velocity of the electrons passing the cross section z reaches the minimum value, $u - |v|$. The electrons passing the cross section at this instant of time travel more slowly than they would travel in the absence of an excitation. Simultaneously, the current reaches its maximum instantaneous value, $I_0 + |i|$. An excess of positive current over the current I_0 in the absence of an excitation in a beam of negative charge occurs when there is a deficiency of negative particles. Thus, when the current swings into its maximum, the number of electrons passing the cross section is less than it would be in the absence of an excitation. Conversely, an excess velocity of the electrons occurring when the kinetic voltage modulation swings negatively is accompanied by an excess of particle current. Thus, the number of electrons that passes the cross section with a velocity higher than u is larger than that passing the cross section with a velocity less than u. We may therefore conclude that the electron beam carries, on the average, electrons with a higher kinetic energy in the presence of a fast wave than it carries in the absence of an excitation.

Conversely, we find that in the slow wave the kinetic voltage and current modulations are 180° out of phase. Thus, if only the slow wave is excited, the number of electrons passing a given cross section with a velocity higher than u is, on the average, smaller than the number of electrons with a velocity less than u. On the average, the beam transports less kinetic energy when it propagates a slow wave than it would carry in the absence of such an excitation. This interpretation of the kinetic power flow, although not quite rigorous in view of the limitations of small-signal theory, gives a useful physical picture. According to this picture, a negative kinetic power flow does not signify a transport of energy in the negative z-direction, but rather a transport of a lack of kinetic energy in the positive z-direction.

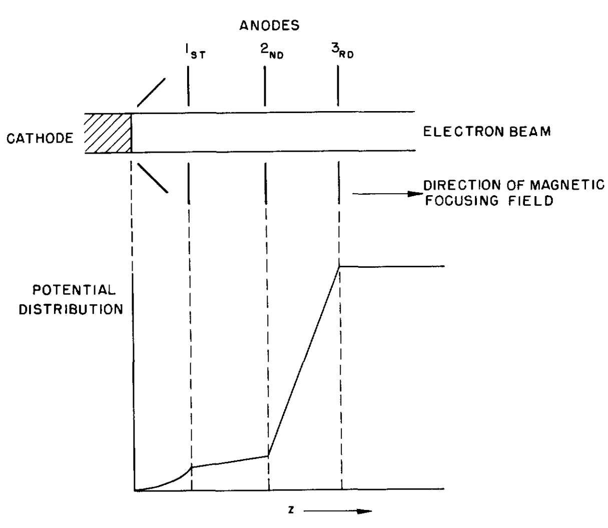

Fig. 2.2. Multielectrode gun and its potential distribution on the beam axis.

2.2 MATRIX REPRESENTATION OF BEAM TRANSDUCERS

The electron beam of a longitudinal beam amplifier is formed in an electron gun in which it is accelerated to anode potential. Following the anode there may be some accelerating or decelerating regions like those used in modern low-noise amplifiers (Fig. 2.2). These regions are termed "beam transducers" (27). No exact analysis exists for an accelerated beam of finite diameter confined by a large magnetic field. Instead, the one-dimensional analysis is used, with a simple substitution of the reduced plasma frequency ω_q for the plasma frequency $\omega_p = (e\rho_0/m\epsilon)^{1/2}$. However, some general statements which are based on less restrictive premises can be made about the nature of the accelerating regions.

If the motion of the electrons is predominantly longitudinal, and if the excitation is approximately uniform across the cross section of the beam, the one-dimensional representation can be used. The beam excitation at one cross section a, given in terms of the kinetic voltage and current modulations V_a and i_a determines uniquely the excitations V_b and i_b at a cross section b further down the beam. Linear relations must exist among these quantities under the small-signal assumption. Defining the column matrices†

$$w_a = \begin{Bmatrix} V_a \\ i_a \end{Bmatrix} \qquad w_b = \begin{Bmatrix} V_b \\ i_b \end{Bmatrix} \tag{2.11}$$

we can write the linear relations in the form

$$w_b = K w_a \tag{2.12}$$

The K matrix is sometimes called the "matrix of generalized circuit parameters" (29) or the (ABCD) matrix. The K matrix is usually written as the following array of complex scalars (compare Eqs. 2.11 and 2.12 with Eqs. 1.32 to 1.36):

$$K = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

If no rf electromagnetic power is extracted from the beam in the region between the cross sections a and b, as is true for the beam transducer of Fig. 2.2, the real part of the kinetic power must be conserved in accordance with Eq. 2.6a.

$$(V_a^* i_a + i_a^* V_a) - (V_b^* i_b + i_b^* V_b) = 0 \tag{2.13}$$

It is expedient to write Eq. 2.13 in matrix form. For this purpose we introduce the permutation matrix

$$R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{2.14}$$

† The operations and theorems of matrix algebra which we use here can be found in many texts on matrices or applied mathematics. See, for example, F. B. Hildebrand, "Methods of Applied Mathematics" (Prentice-Hall, New York, 1952).

According to the rules of matrix multiplication we find that

$$RR = I \tag{2.15}$$

where I is the identity matrix. Equation 2.15 can also be written in the form

$$R = R^{-1}$$

indicating that R is equal to its own inverse. Further, we define by A^+ the Hermitian (complex) conjugate of the matrix A. The Hermitian conjugate of a matrix A is obtained by taking the complex conjugate of all elements of A and then transposing it.

$$(A^+)_{ij} = A^*_{ji}$$

In particular, the Hermitian conjugate of a column matrix is a row matrix. Referring to definition 2.11 we have, for example,

$$w_a^+ = [V_a^* \ , \ i_a^*]$$

With the aid of these definitions we can write Eq. 2.13 in matrix form:

$$w_a^+ R w_a - w_b^+ R w_b = 0 \tag{2.16}$$

The vector w_b in Eq. 2.16 can be expressed in terms of w_a through Eq. 2.12. For this purpose we note only that the Hermitian conjugate of a product of two matrices A and B is equal to the product in reverse order of the Hermitian conjugates of the matrix factors:

$$(AB)^+ = B^+A^+ \tag{2.17}$$

From Eq. 2.16, with the aid of Eq. 2.12, we obtain

$$w_a^+ (R - K^+RK) w_a = 0 \tag{2.18}$$

Equation 2.18 has to be satisfied for an arbitrary choice of the vector w_a, that is, an arbitrary choice of the boundary conditions. This is possible if and only if

$$K^+RK = R \tag{2.19}$$

This condition is the restriction imposed upon the K matrix by the requirement for the conservation of the real kinetic power, Eq. 2.13.

For the analysis of noise in the electron beam it will be convenient to use another form of Eq. 2.19. In order to obtain that alternate form we shall make use of some additional definitions of matrix algebra.

A matrix A is termed nonsingular if its determinant det(A) is not equal to zero. A nonsingular matrix A always has an inverse, A^{-1}. Further, it follows from the properties of matrix multiplication and the multiplication of determinants that the determinant of a matrix product is equal to the product of the determinants of the matrix factors.

$$\det(AB) = \det(A)\det(B) \tag{2.20}$$

With the aid of Eq. 2.20 we obtain from Eq. 2.19

$$\det(K^{+})\det(K) = |\det(K)|^2 = 1 \tag{2.21}$$

By virtue of Eq. 2.21, the determinant of K is finite; correspondingly, K is nonsingular. The matrix K has an inverse. Multiplying Eq. 2.19 from the left by $K^{-1}R$ we have

$$K^{+}RKK^{-1}R = RK^{-1}R$$

or, with the aid of Eq. 2.15,

$$K^{+} = RK^{-1}R \tag{2.22}$$

This is the equation that we shall use in the analysis of noise in electron beams.

A drift region is a simple example of a lossless beam transducer. With the aid of Eqs. 2.8 and 2.9 it is easy to show that the kinetic voltage and the current modulations, V_a and i_a, at the plane a, transform into corresponding modulations, V_b and i_b, at the plane b, through the following equations:

$$\begin{aligned} V_b &= [V_a \cos\theta_q + i_a j Z_o \sin\theta_q]e^{-j\theta} \\ i_b &= [i_a j Y_o \sin\theta_q + i_a \cos\theta_q]e^{-j\theta} \end{aligned} \tag{2.23}$$

where $\theta_q = \omega_q(z_b - z_a)/u$ is the transit angle measured in terms of the plasma period, $T = 2\pi/\omega_q$; and $\theta = \omega(z_b - z_a)/u$ is the conventional transit angle. We shall refer to θ_q as the "plasma transit angle." Equations 2.23 show that the K matrix of the transformation of voltage and current by a drift region is

$$K = \begin{bmatrix} \cos\theta_q & jZ_o\sin\theta_q \\ jY_o\sin\theta_q & \cos\theta_q \end{bmatrix} e^{-j\theta} \tag{2.24}$$

Simple matrix manipulations show that K as given by Eq. 2.24 indeed satisfies the condition of conservation of the real part of kinetic power, Eq. 2.19.

A relation analogous to Eqs. 2.23 is given in Table 2 of section 1.24, for the voltage-current transformation by an accelerated electron beam. It is not difficult to confirm that the matrix K, whose coefficients are given by Eqs. 1.33 to 1.36, satisfies the condition of power conservation Eq. 2.19.

We earlier discussed another set of parameters that is also able to describe the excitation of a beam: the normalized amplitudes of the fast and the slow waves a_1 and a_2. Consider a lossless beam transducer that extends from cross section a to cross section b. Imagine that the transducer is preceded and followed by drift regions of characteristic impedance Z_{oa} and Z_{ob}, respectively. Then the normalized amplitudes a_1 and a_2 of Eq. 1.28 can be found uniquely in terms of the voltage V_a and current i_a by the use of Eqs. 2.8 and 2.9, in which we set z = 0, thus choosing an appropriate origin of the co-ordinate system in the input drift region.

$$\begin{aligned} a_1 &= \frac{1}{2}\left[(2Y_{oa})^{-1/2} i_a + (2Z_{oa})^{-1/2} V_a\right] \\ a_2 &= \frac{1}{2}\left[(2Y_{oa})^{-1/2} i_a - (2Z_{oa})^{-1/2} V_a\right] \end{aligned} \tag{2.25}$$

Similarly, we can choose the origin of the co-ordinate system in the output drift region to coincide with the cross section b. In order to avoid confusion we denote the normalized amplitudes of the fast and slow waves in the output drift region by b_1 and b_2, respectively. Thus, we have

$$\begin{aligned} b_1 &= \frac{1}{2}\left[(2Y_{ob})^{-1/2} i_b + (2Z_{ob})^{-1/2} V_b\right] \\ b_2 &= \frac{1}{2}\left[(2Y_{ob})^{-1/2} i_b - (2Z_{ob})^{-1/2} V_b\right] \end{aligned} \tag{2.26}$$

The linear relations among the kinetic voltage and current at reference cross sections a and b, respectively, summarized in Eq. 2.12, imply corresponding linear relations among the normalized wave amplitudes. Introducing the column matrices

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{2.27}$$

we conclude that the column matrices a and b are related by a matrix M of second order

$$b = Ma \tag{2.28}$$

The real kinetic power carried by the waves at cross section a is given by Eq. 1.29. Equation 1.29 can be written in matrix form if we introduce the "parity" matrix

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.29}$$

It should be noted that the P matrix is its own inverse.

$$PP = I \tag{2.30}$$

The real kinetic power at cross section a can be written as

$$\frac{1}{2} \operatorname{Re} (V_a i_a^*) = a^+ P a \tag{2.31}$$

The real kinetic power at cross section a has to equal that at cross section b if the transducer is lossless:

$$a^+ P a = b^+ P b$$

Using Eq. 2.28 in this relation, we obtain

$$a^+ (P - M^+ P M) a = 0 \tag{2.32}$$

Equation 2.32 can be satisfied for an arbitrary choice of a if and only if

$$M^+ P M = P \tag{2.33}$$

Equation 2.33 is the condition imposed upon the matrix M by the conservation of the real kinetic power analogous to the relation satisfied by the K matrix, Eq. 2.19. In many cases it is convenient to write the transformation between the wave amplitudes and the current and kinetic voltage, Eqs. 2.25 and 2.26, in matrix form. A normalization matrix N has to be defined for this purpose.

$$N = \begin{bmatrix} (8Z_o)^{-1/2} & 0 \\ 0 & (8Y_o)^{-1/2} \end{bmatrix} \tag{2.34}$$

The subscripts a and b will be applied to the N matrix to indicate whether it is referred to the beam at plane a or at plane b. It is easy to show that the matrix form of Eq. 2.25 is

$$a = (I + PR)N_a w_a \tag{2.35}$$

where I is, as usual, the unit matrix; and P is the parity matrix defined by Eq. 2.29. Equation 2.26 becomes

$$b = (I + PR)N_b w_b \tag{2.36}$$

The wave formalism is applicable even when the transducer under consideration is not preceded or followed by a drift region. Then Eqs. 2.35 and 2.36 are the definitions of quantities a and b, which have but mathematical significance. The admittance Y_0 in definition 2.34 is then arbitrary but is conveniently chosen to correspond to the characteristic admittance of a drift region with a time-average voltage and a current density equal to those existing at the reference cross section.

2.3 MATRIX REPRESENTATION OF LONGITUDINAL-BEAM AMPLIFIERS

The kinetic power concept formulated in Eq. 2.6a is common to all one-dimensional electron beam systems. (To remind the reader – a one-dimensional beam is one whose excitation can be characterized in terms of only two parameters.) With the aid of Eq. 2.6a an interesting formalism can be developed for all longitudinal-beam microwave amplifiers. Consider, for example, a traveling-wave tube as shown in Fig. 2.3. The excitation of the fast wave and the slow wave in the beam at the gun end we denote, as usual, by a_1 and a_2. The excitation of the same set of waves at the collector end we denote by b_1 and b_2. The normalized amplitude of the incident wave at the input of the amplifier we denote by a_3; the incident wave at the output of the amplifiers, by a_4. The reflected waves at the input and output are denoted by b_3 and b_4. The arrows in Fig. 2.3 indicate whether a particular wave carries power into or out of the amplifier. The wave amplitudes a_1 to a_4 can be adjusted by means external to the amplifier. Thus, we could, for example, excite the beam before it enters the

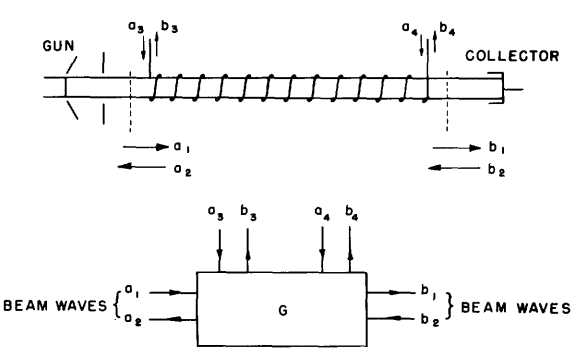

Fig. 2.3. Schematic of a traveling-wave tube.

amplifier. In this way a_1 and a_2 could be adjusted arbitrarily. Further, the output transmission line of the amplifier could be terminated in a matched load, which corresponds to choosing $a_4 = 0$. Power could be fed through an attenuator matched to the input transmission line of the amplifier. The wave a_3 would thus be fixed in amplitude and phase. This example shows that the wave amplitudes a are under our control. The wave amplitudes b_1 to b_4 must then be related to the quantities a_1 to a_4 by linear relations (small-signal theory.)

The amplifier can be characterized in terms of a four-by-four matrix G, so that

$$b = Ga \tag{2.37}$$

where b is the column matrix $\begin{Bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{Bmatrix}$ and a is the column matrix $\begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{Bmatrix}$

Chu's kinetic power theorem, Eq. 2.6a, imposes some interesting conditions upon the matrix elements of G. Let us assume, first, that no ohmic loss occurs within the amplifier structure. Accordingly, the difference between the kinetic power at the input of the amplifier and that at the output of the amplifier must be equal to the electromagnetic power delivered to the circuit. The latter is the difference between the power flowing out in the output transmission line and the power flowing in, in the input transmission line. The electromagnetic power fed to the amplifier in the input transmission line is

$$|a_3|^2 - |b_3|^2$$

The electromagnetic power leaving the amplifier in the output transmission line is

$$|b_4|^2 - |a_4|^2$$

(For a discussion of the magnitude of the electromagnetic power carried by an electron beam, see reference 28 and the remark on p. 18) Thus we must have, according to Eqs. 2.6a and 2.10

$$|b_4|^2 - |a_4|^2 - |a_3|^2 + |b_3|^2 = |a_1|^2 - |a_2|^2 - |b_1|^2 + |b_2|^2$$

or (2.38)

$$|b_1|^2 - |b_2|^2 + |b_3|^2 + |b_4|^2 = |a_1|^2 - |a_2|^2 + |a_3|^2 + |a_4|^2$$

Introducing a parity matrix P

$$P = \mathrm{diag}\,(1, -1, 1, 1) \tag{2.39}$$

we can write Eq. 2.38 in a more elegant form:

$$b^{+}Pb = a^{+}Pa \tag{2.40}$$

By the use of Eq. 2.37 we can express Eq. 2.40 as

$$a^{+}(G^{+}PG - P)a = 0 \tag{2.41}$$

The matrix equation (2.41) has to be satisfied for an arbitrary choice of the a matrix. This is possible if and only if

$$G^{+}PG = P \tag{2.42}$$

The matrix equation (2.42) contains scalar equations of the form

$$|G_{13}|^2 - |G_{23}|^2 + |G_{33}|^2 + |G_{43}|^2 = 1 \tag{2.43}$$

and

$$G_{13}G_{14}^{*} - G_{23}G_{24}^{*} + G_{33}G_{34}^{*} + G_{43}G_{44}^{*} = 0 \tag{2.44}$$

Without any loss of generality we can assume that the input source is matched to the input transmission line and the load to the output transmission line. Any mismatch between the terminations and the amplifier can be taken into account by a proper choice of the matrix elements G_{ij} which represent the amplifier from the circuit point of view as a two-terminal-pair device. These elements are G_{33}, G_{34}, G_{43}, and G_{44}. The term $|G_{43}|^2$ is the power gain; that is, the ratio of the output power, $|b_4|^2$, over the available input power, $|a_3|^2$, with the load matched to the output transmission line, $a_4 = 0$.

Equation 2.43 has an obvious interpretation. If the power gain of the amplifier is appreciably greater than unity, $|G_{43}|^2 >> 1$, Eq. 2.43 implies that $|G_{23}|^2$, the only term preceded by a minus sign, must also be appreciably greater than unity. In other words, the input wave a_3 must couple strongly to the slow mode in the beam leaving the amplifier, if the amplifier has an appreciable gain. The electromagnetic power gain is obtained at the expense of kinetic power in the beam, which is correspondingly large, and negative, when the beam leaves the amplifier.

If the amplifier contains elements with ohmic loss, this reasoning has to be modified. Because of added complications, it is believed that a discussion of such structures is not warranted. The qualitative results of such an investigation agree in essence with those for lossless structures. Reference 19 gives the details.

3. NOISE IN ELECTRON BEAMS

The preceding sections were devoted to the analysis of electron beams and the interaction of electron beams with rf structures under steady-state excitation at a frequency ω. The matrix equation (2.12) showed that a steady-state, sinusoidal modulation of the electron beam is completely determined by the knowledge of the kinetic voltage and current modulations at one reference cross section of the beam. This result can be used as the starting point of the analysis of noise in electron beams.

Noise is a statistical process which must be analyzed by statistical methods. Here we shall make use of the extension of Fourier integral theory to the harmonic analysis of random functions (43).

We assume that the noise process in the electron beam is stationary and has no hidden periodic components. An observation of the noise process at the reference cross section a of the beam gives the kinetic noise voltage $V_a(t)$ and the current modulation $i_a(t)$ as functions of time. The subsequent analysis will be devoted to finding the statistical properties of the kinetic voltage and current modulations at some other reference cross section b in terms of the statistical properties of $V_a(t)$ and $i_a(t)$.

The kinetic voltage modulation, $V_a(t)$, is not a periodic function of time and therefore cannot be analyzed by Fourier series methods. Fourier integral methods are not adequate, either, for the analysis of $V_a(t)$, since it is intuitively obvious that the integral

$$\int_{-\infty}^{\infty} |V_a(t)|^2 dt$$

carried over an infinite interval does not converge. We can choose, however, a function $V_{aT}(t)$ defined by

$$V_{aT}(t) = \begin{cases} V_a(t) & -T < t < T \\ 0 & T < |t| \end{cases} \tag{3.1}$$

The function $V_{aT}(t)$ satisfies the requirement of convergence of the integral

$$\int_{-\infty}^{\infty} |V_{aT}(t)|^2 dt$$

as long as T is chosen finite. It represents accurately the random function $V_a(t)$ over the interval 2T. We can form the Fourier transform

$$V_{aT}(\omega) = \int_{-\infty}^{\infty} V_{aT}(t) e^{-j\omega t} dt$$

In a similar way we can define a function $i_{aT}(t)$ with the corresponding Fourier transform $i_{aT}(\omega)$. The function $i_{aT}(t)$ represents the random noise current $i_a(t)$ over the finite time interval 2T.

A linear beam transducer relates the Fourier transforms $V_{aT}(\omega)$ and $i_{aT}(\omega)$ of kinetic voltage and current modulations applied to its input cross section to the corresponding Fourier transforms at its output b.

$$V_{bT}(\omega) = A V_{aT}(\omega) + B i_{aT}(\omega) \tag{3.2}$$

$$i_{bT}(\omega) = C V_{aT}(\omega) + D i_{aT}(\omega) \tag{3.3}$$

The coefficients A to D are, in general, functions of frequency; $V_{bT}(\omega)$ and $i_{bT}(\omega)$ are the Fourier transforms of the time functions $V_{bT}(t)$ and $i_{aT}(t)$. These, in turn, give the output of the beam transducer produced when the modulations $V_{aT}(t)$ and $i_{aT}(t)$ are applied to the input over a finite period of time 2T. It may be expected that $V_{bT}(t)$ and $i_{bT}(t)$ will resemble the true noise output of the transducer $V_b(t)$ and $i_b(t)$ over a portion of the period 2T. This portion encompasses the response of the transducer to the noise input $V_{aT}(t)$ and $i_{aT}(t)$ over the time during which it is possible to neglect the transients in the transducer set up at t = -T. The squares of the absolute values of Eqs. 3.2 and 3.3 are

$$|V_{bT}(\omega)|^2 = |A|^2 |V_{aT}(\omega)|^2 + |B|^2 |i_{aT}(\omega)|^2 + AB^* V_{aT}(\omega) i_{aT}(\omega)^* + A^* B V_{aT}(\omega)^* i_{aT}(\omega) \tag{3.4}$$

$$|i_{bT}(\omega)|^2 = |C|^2 V_{aT}(\omega) + |D|^2 i_{aT}(\omega) + CD^* V_{aT}(\omega) i_{aT}(\omega)^* + C^* D V_{aT}(\omega)^* i_{aT}(\omega) \tag{3.5}$$

The functions $V_{aT}(t)$ and $i_{aT}(t)$ do not represent the random functions $V_a(t)$ and $i_a(t)$ exactly as long as the interval 2T is held finite. If, however, the interval T is allowed to go to infinity, we see from the definition of 3.1 that $V_{aT}(t)$ becomes indistinguishable from $V_a(t)$. The same statement can be made concerning $i_{aT}(t)$ and $i_a(t)$, $V_{bT}(t)$ and $V_b(t)$, as well as $i_{bT}(t)$ and $i_b(t)$. Generalized harmonic analysis proves that in the limit $T \rightarrow \infty$ the following quantities approach finite limits

$$\lim_{T\to\infty} \frac{\pi}{T} \overline{|V_{aT}(\omega)|^2} = \Phi_a(\omega) \tag{3.6}$$

$$\lim_{T\to\infty} \frac{\pi}{T} \overline{|i_{aT}(\omega)|^2} = \Psi_a(\omega) \tag{3.7}$$

$$\lim_{T\to\infty} \frac{\pi}{T} \overline{V_{aT}(\omega) i_{aT}(\omega)^*} = \left[\lim_{T\to\infty} \frac{\pi}{T} \overline{V_{aT}(\omega)^* i_{aT}(\omega)}\right]^* = \Theta_a(\omega) \tag{3.8}$$

The bar over the quantities on the left side of Eqs. 3.6, 3.7, and 3.8 indicates an ensemble average. To obtain such an average, consider a set (ensemble) of statistical processes of identical statistical character. In our particular case we can imagine that measurements are performed on a large number of identical electron beams. The average over such a set of measurements is then the ensemble average.

The quantity Φ_a is the self power density spectrum (SPDS) of the kinetic noise voltage modulation, Ψ_a is the SPDS of the noise current modulation at the reference cross section a. The cross power density spectrum (CPDS) between the kinetic voltage and current modulations is Θ_a. The frequency dependence of these quantities will be henceforth implied and the parentheses (ω) will be omitted. If we take the limit $T \to \infty$ of Eqs. 3.4 and 3.5 multiplied by π/T we find, after taking an ensemble average,

$$\Phi_b = |A|^2 \Phi_a + |B|^2 \Psi_a + AB^* \Theta_a + A^* B \Theta_a^* \quad (3.9)$$

$$\Psi_b = |C|^2 \Phi_a + |D|^2 \Psi_a + CD^* \Theta_a + C^* D \Theta_a^* \quad (3.10)$$

Multiplication of Eq. 3.2 by the complex conjugate of Eq. 3.3, transition to $T \to \infty$ of the resultant equation multipled by π/T, and an ensemble average lead to the relation

$$\Theta_b = AC^* \Phi_a + BD^* \Psi_a + AD^* \Theta_a + A^* D \Theta_a^* \quad (3.11)$$

Equations 3.9, 3.10, and 3.11 give the self- and cross-power density spectra of the kinetic voltage and current of the noise at cross section b in terms of the corresponding quantities at cross section a. The three quantities, Φ_a, Ψ_a, Θ_a, the last of them complex, characterize the noise in an electron beam sufficiently for most practical purposes. Thus, a noise process in an electron beam is specified by four real parameters.

The SPDS's are related by a factor of $4\pi\Delta f$ to the more commonly used quantities, the "mean-square fluctuations within a frequency band Δf." Thus, for example, the SPDS of pure shot noise in a beam with a direct current I_o is: $\Psi = eI_o/2\pi$, whereas the mean-square fluctuations of the current within the frequency band Δf are known to be: $\overline{|i^2|} = 2eI_o\Delta f$. The reason for the deviation from conventional engineering use of the definition of the SPDS lies in the simplicity of the resulting relation between the mean square of the current fluctuations and the frequency integral of Ψ. We have

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} i^2(t)\, dt = \int_{-\infty}^{\infty} \Psi(\omega)\, d\omega$$

The SPDS of the noise current Ψ at a cross section z of the electron beam is a measurable quantity. A cavity with a short gap at the position z has a power output proportional to the value $\Psi(\omega_o)$ at the resonant frequency ω_o of the cavity. The classical experiment by Cutler and Quate (13) was a measurement of Ψ as a function of distance in a drifting beam performed in the described way.

Equations 3.9, 3.10, and 3.11 can be obtained in an alternate way by matrix methods. The advantage of such an approach lies partly in its elegance, partly in the fact that general theorems of matrix algebra can be applied to the noise problem. We define the column matrices

$$w_{aT}(\omega) = \begin{bmatrix} V_{aT}(\omega) \\ i_{aT}(\omega) \end{bmatrix} \quad \text{and} \quad w_{bT}(\omega) = \begin{bmatrix} V_{bT}(\omega) \\ i_{bT}(\omega) \end{bmatrix} \tag{3.12}$$

Equations 3.2 and 3.3 can be cast in matrix form.

$$w_{bT}(\omega) = K w_{aT}(\omega) \tag{3.13}$$

where

$$K = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

We define the matrices

$$W_a = \lim_{T \to \infty} \frac{\pi}{T} \overline{w_{aT}(\omega)\, w_{aT}(\omega)^+}$$

and

$$W_b = \lim_{T \to \infty} \frac{\pi}{T} \overline{w_{bT}(\omega)\, w_{bT}(\omega)^+} \tag{3.14}$$

A study of the definitions of Eq. 3.14, and the definitions of Eqs. 3.6, 3.7, 3.8, and 3.12, shows that the matrix W_a is composed of the SPDS's and the CPDS's as follows:

$$W_a = \begin{bmatrix} \Phi_a & \Theta_a \\ \Theta_a^* & \Psi_a \end{bmatrix} \tag{3.15}$$

A similar expression holds for W_b. The relation between the matrices W_a and W_b can be found by multiplying by Eq. 3.13 by π/T and its Hermitian conjugate and by transition to the limit $T \to \infty$. The result is

$$W_b = K W_a K^+ \tag{3.16}$$

Equations 3.9 to 3.11 are contained in the matrix equation (3.16), as can easily be demonstrated by matrix multiplication and the aid of Eq. 3.15. Although the matrix equation (3.16) is a short-hand expression for four equations, it contains only three distinct relations; the 12 element of the matrix equation (3.16) is equal to the complex conjugate of the 21 element. This fact is a direct consequence of the Hermitian character of the W matrices, namely.

$$W_b = W_b^+ \quad \text{and} \quad W_a = W_a^+$$

3.1 TRANSFORMATION OF NOISE BY LOSSLESS BEAM TRANSDUCERS

An important class of beam transducers is "lossless", i.e., conserves the real part of the kinetic power. It is therefore of practical interest to devote special attention to noise transformations by means of such transducers.

The K matrix of a lossless transducer satisfies Eq. 2.22. The transformation of Eq. 3.16 can then be written in the alternate form:

$$W_b = KW_aRK^{-1}R$$

Multiplication of this equation from the right by R and the condition on the R matrix (Eq. 2.15) lead to the transformation

$$W_bR = KW_aRK^{-1} \tag{3.17}$$

Equation 3.17 shows that the noise matrix WR undergoes a <u>similarity transformation</u> when the beam is passed through a lossless transducer. A similarity transformation leaves the trace and the determinant of a 2×2 matrix invariant. The WR matrix written out explicitly has the form

$$WR = \begin{bmatrix} \Theta & \Phi \\ \Psi & \Theta^* \end{bmatrix} \tag{3.18}$$

The trace of the WR matrix is

$$\mathrm{Tr}(WR) = \Theta + \Theta^* = 2\,\mathrm{Re}(\Theta) \tag{3.19}$$

The determinant is

$$\det(WR) = |\Theta|^2 - \Phi\Psi \tag{3.20}$$

The physical meaning of the first invariant, the trace, is not hard to grasp. According to Eq. 3.8, Θ is proportional to the kinetic power carried in a narrow frequency band around the frequency ω. The real part of the kinetic power in a particular frequency band has to be conserved in a transition through a linear lossless transducer, hence the invariance of the trace of WR.

The meaning of the second invariant is less self-evident. Indeed, conventional network theory has no analog to this invariant. In order to demonstrate this we consider, first, a random voltage V(t) applied to a two-terminal network characterized by the admittance function $Y(\omega)$. Again, an equivalent Fourier transform of the voltage V(t) based on a sample of length 2T can be constructed. Denote the Fourier transform by $V_T(\omega)$. The Fourier transform of the current flowing into the network under the influence of $V_T(\omega)$ is $i_T(\omega) = Y(\omega)\, V_T(\omega)$. Forming the CPDS Θ between the voltage and current we obtain

$$
\begin{aligned}
\Theta(\omega) &= \lim_{T\to\infty} \frac{\pi}{T} \overline{V_T(\omega)\, i_T(\omega)^*} \\
&= \lim_{T\to\infty} \frac{\pi}{T}\, Y(\omega)^* \overline{|V_T(\omega)|^2} \\
&= Y(\omega)^* \lim_{T\to\infty} \frac{\pi}{T} \overline{|V_T(\omega)|^2} \\
&= Y(\omega)^*\, \Phi(\omega) \qquad (3.21)
\end{aligned}
$$

where, as before, we denote the self power density spectrum of the voltage by $\Phi(\omega)$. In a similar way we obtain the self power density spectrum of the current, $\Psi(\omega)$,

$$
\begin{aligned}
\Psi(\omega) &= \lim_{T\to\infty} \frac{\pi}{T} \overline{i_T(\omega)\, i_T(\omega)^*} \\
&= |Y(\omega)|^2\, \Phi(\omega) \qquad (3.22)
\end{aligned}
$$

Introducing Eqs. 3. 21 and 3. 22 into Eq. 3. 20 we find that for this process

$$\det(WR) = 0 \qquad (3.23)$$

Consider, next, a cascade of linear two terminal-pair transducers of conventional network theory. The admittance seen across any terminal pair within the cascade is determined uniquely by the admittance connected to the end of the cascade. Forming the matrix W of the SPDS's and CPDS's of the noise voltage across, and the noise current into, any terminal pair of the cascade we find, again, that det(WR) = 0. Thus, noise propagating along a cascade of lossless two terminal-pair networks has only one invariant parameter, namely, $Re(\Theta)$, which is proportional to the time average power of the noise within a narrow frequency band around the frequency ω fed into the termination of the cascade. Since a conventional transmission line can be considered as the limit of a cascade of an infinite number of infinitesimal two terminal-pair networks, the above reasoning applies as well to noise propagation along conventional transmission lines.

The close analogy between lossless beam transducers and lossless, two terminal-pair networks makes one wonder whether or not the determinant of the matrix WR for a noise process in an electron beam can have a finite value. But, there is a basic physical difference between a drift region and a transmission line. An electromagnetic wave of a given frequency incident upon a transmission-line termination is accompanied by a reflected wave, the phase and amplitude of which depend upon the termination. The voltage and current in the incident and reflected waves combine to satisfy the boundary condition imposed by the termination. The admittance at any cross section of the transmission line is determined by the terminating admittance. In an electron beam the role of the incident and reflected waves is played by the fast and slow wave. Both of these waves have a group velocity equal in magnitude and direction to the time-average velocity of the electron beam. Thus, these waves can be excited only at the entry plane into a drift region, the plane passed first by the electrons. The phase relation between the fast and slow waves is determined by the method of excitation prior to the entry of the electron beam into the drift region. One could imagine a method of excitation by which the kinetic voltage and current are not put into a definite ratio to each other. We shall return to the question of noise excitation later. At this point it is sufficient to state that the kinetic noise voltage and noise current in the electron beam may or may not have a definite ratio to each other, and therefore we must allow for the possibility that $|\Theta|^2 \neq \Phi\,\Psi$. It is known from statistical theory that the inequality

$$|\Theta|^2 \leq \Phi\Psi \tag{3.24}$$

holds. With the aid of this inequality and Eq. 3.20 we find, for a noise process in an electron beam,

$$\det(WR) \leq 0 \tag{3.25}$$

According to Eq. 3.25, noise in an electron beam may well possess two invariants with regard to lossless beam transformations. Let us choose distinct symbols for the two invariants of a noise process in an electron beam, for they will prove of great importance. Since Re(Θ) plays the role of power carried by noise propagation along a transmission line we use the Greek letter Π to refer to it.

$$\mathrm{Re}(\Theta) \equiv \Pi \tag{3.26}$$

The imaginary part of Θ we denote by

$$\mathrm{Im}(\Theta) \equiv \Lambda \tag{3.27}$$

The determinant of WR can be written, according to Eqs. 3.20, 3.26, and 3.27, as

$$\det(WR) = \Pi^2 - (\Phi\Psi - \Lambda^2) \tag{3.28}$$

Since Π in itself is an invariant, we must conclude that the term $\Phi\Psi - \Lambda^2$ must also be an invariant with regard to lossless transformations. We introduce here a symbol for it. We set

$$S = (\Phi\Psi - \Lambda^2)^{1/2} \tag{3.29}$$

The inequality (3.25) assures that S is always real. We shall choose S positive by definition. We shall see later that the invariant S is more easily explained in terms of physical quantities than the invariant det(WR) itself.

The accelerating regions of a multielectrode electron gun as shown in Fig. 2.2 are lossless beam transducers. The noise in any of the regions is determined by the noise at its input plane, the plane first passed by the electrons. The parameters S and Π are invariant with regard to lossless transformations and may be traced back to the input of the first region. Where, then, should the input plane of the first drift region be chosen, and what determines the values of the two noise invariants? These two questions are intimately connected. The first region of the multielectrode gun is formed by a space-charge-limited diode. A space-charge-limited diode is a lossless transducer, provided that the small-signal, single-velocity approximations are applicable. These approximations hold as long as the range of velocities possessed by the majority of the electrons is small compared to the average velocity of the electrons. The single-velocity approximation will hold at potentials as low as a few volts above cathode potential. The input plane for the first region can be picked in front of the cathode beyond the potential minimum at a plane a few volts above the cathode potential. The values of S and Π at this plane are conserved throughout the multielectrode gun under the assumption that no electromagnetic power is extracted from the beam on the way. The parameters S and Π are thus entirely functions of the conditions in the potential minimum-cathode region in which the single-velocity assumption does not hold.

A word of caution is in order. If the potentials applied to the successive electrodes of the gun differ widely, a strong electrostatic lens may result which may cause crossovers of the electron paths. It is believed that a deviation from laminar electron flow is accompanied by an enhancement of noise over the values predicted by one-dimensional, single-velocity theory.

Little is now known about the effect of the region of the potential minimum upon the noise. Pierce (32) made the assumption that the effect of the potential minimum region upon the noise is negligible. Then, it is quite reasonable to suppose that the noise at the input reference plane, beyond which the single-velocity approximation is legitimate, consists of a current modulation of full shot-noise value, and an equivalent velocity modulation of the Rack value (33). The SPDS of the shot noise of a beam with a direct current I_o is

$$\Psi = \frac{e I_o}{2\pi}$$

The SPDS of the kinetic voltage corresponding to the Rack equivalent velocity is

$$\Phi = \left(\frac{m}{e} u\right)^2 \frac{1}{4\pi\Delta f} \overline{(\delta v)^2} = \left(1 - \frac{\pi}{4}\right) \frac{mkT_c}{eI_o} u^2$$

where u is the average velocity of the electrons at the reference plane, T_c is the cathode temperature, and the value of the mean-square velocity fluctuations $\overline{(\delta v)^2}$ is taken from reference 33. If the reference plane is taken exactly at the potential minimum, a choice not very convincing in view of the warnings given above, we can take the value $2kT_c/m$ for u^2. Further, Pierce assumed that the velocity and current are uncorrelated at the first input plane. Under this assumption the Π parameter is equal to zero. For the S parameter we obtain

$$S = (\Phi\Psi)^{1/2} = \left(1 - \frac{\pi}{4}\right)^{1/2} \frac{kT_c}{\pi} \tag{3.30}$$

3.2 AN INTERPRETATION OF THE S-PARAMETER

A drift region is a lossless beam transducer with a K matrix given by Eq. 2.23. The matrix equation (3.16) carried out explicitly for a drift region gives, with the aid of a simple trigonometric identity,

$$\Phi_b = \frac{1}{2}(\Phi_a + Z_o^2\Psi_a) + \frac{1}{2}(\Phi_a - Z_o^2\Psi_a)\cos 2\Theta_q + Z_o\Lambda_a \sin 2\Theta_q \tag{3.31}$$

$$\Psi_b = \frac{1}{2}(Y_o^2\Phi_a + \Psi_a) + \frac{1}{2}(\Psi_a - Y_o^2\Phi_a)\cos 2\Theta_q - Y_o\Lambda_a \sin 2\Theta_q \tag{3.32}$$

$$\Theta_b = \Pi_a + \frac{1}{2} j\,[Z_o\Psi_a - Y_o\Phi_a]\sin 2\Theta_q + j\Lambda_a \cos 2\Theta_q \tag{3.33}$$

Equation 3.33, split into its real and imaginary parts, and use of the definitions of Eqs. 3.26 and 3.27 lead to

$$\Pi_b = \Pi_a \tag{3.33a}$$

$$\Lambda_b = \frac{1}{2}[Z_o\Psi_a - Y_o\Phi_a]\sin 2\Theta_q + \Lambda_a \cos 2\Theta_q \tag{3.33b}$$

Equation 3.33a is simply an expression of the conservation of Re(Θ) through a transformation by a section of a drift region which is a lossless transducer. Equations 3.31 and 3.32 show that the SPDS of the kinetic noise voltage and current modulations have the form of standing waves as functions of Θ_q, the plasma transit angle. The maxima of the kinetic voltage and current modulations lie angles $\Delta \Theta_q = 90^\circ$ apart. (See Fig. 3.1.) The maximum of the current SPDS has the value

$$\Psi_{max} = \frac{1}{2}(Y_o^2 \Phi_a + \Psi_a) + \frac{1}{2}[(Y_o^2 \Phi_a - \Psi_a)^2 + 4 Y_o^2 \Lambda_a^2]^{1/2} \tag{3.34}$$

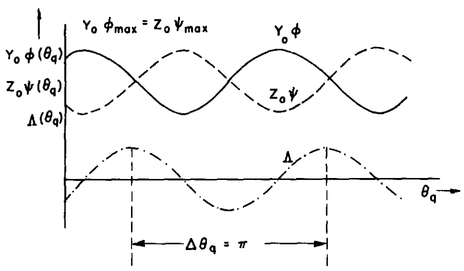

Fig. 3.1. Plot of the power density spectra as functions of plasma transit angle in a drift region.

The minimum is

$$\Psi_{min} = \frac{1}{2}(Y_o^2 \Phi_a + \Psi_a) - \frac{1}{2}[(Y_o^2 \Phi_a - \Psi_a)^2 + 4 Y_o^2 \Lambda_a^2]^{1/2} \tag{3.35}$$

From Eqs. 3.34 and 3.35 we find that the product of the maximum and the minimum is

$$\Psi_{max}\Psi_{min} = Y_o^2(\Phi_a \Psi_a - \Lambda_a^2) = Y_o^2 S^2 \tag{3.36}$$

Equation 3.36 shows that the product of the maximum and the minimum of the current SPDS are proportional to the square of the invariant S. The proportionality factor is the square of the characteristic admittance of the beam.†

The theoretically predicted conservation of the parameter S under lossless transformations can be checked experimentally. The noise standing-wave ratio in a drifting beam can be varied by adjustments of the voltages on the electrodes of a multielectrode gun preceding the drift region. If the potential of the drifting beam is left unchanged in the process, the characteristic admittance of the beam, Y_0, is not changed. Theory then predicts that the product $\Psi_{max}\Psi_{min}$ has to stay invariant. Experiments were performed (20) which checked the theory satisfactorily. At standing-wave ratios Ψ_{max}/Ψ_{min} greater than 10 db, the experimentally observed product $\Psi_{max}\Psi_{min}$ exceeded that observed at lower standing-wave ratios. This effect was ascribed to the existence of higher order space-charge modes which are not accounted for in the one-dimensional theory (34).

3.3 THE EQUIVALENT NOISE ADMITTANCE

Two noise processes with the same power density spectra, that is, the same WR matrix, are transformed by a transducer with a given K matrix in the same way. If a noise process in an electron beam has a WR matrix with det(WR) = 0, it is always possible to find an analog noise process with the same WR matrix for which an

---

†The invariance of the product $\overline{|i_{max}|^2}\;\overline{|i_{min}|^2}$ as first proved by Pierce is a special case of the proof of the invariance of S under lossless beam transformations.

admittance $Y(\omega)$ can be defined according to Eqs. 3.21 and 3.22

$$Y(\omega) = \frac{\Psi}{\Theta} \tag{3.37}$$

We may think of the analog noise process as propagating along transmission lines interconnected with two terminal-pair networks so chosen that their K matrices are identical with those of the electron beam drift regions and beam transducers.† The noise propagating along the transmission lines, henceforth called the analog noise process, permits the definition of an admittance according to Eq. 3.37. Once the admittance of the analog noise process is defined at any reference terminal pair, the admittance at any other terminal pair is uniquely determined according to the well-known laws of transformations of admittances by two terminal-pair networks. Many properties of the analog process can be derived from the knowledge of its admittance. Thus assume, for example, that the admittance $Y(\omega)$ of the analog noise is given at a reference cross section of one of the transmission lines. A particular value of Y is associated with a particular current standing wave ratio on the transmission line:

$$\left|\frac{i_{max}}{i_{min}}\right| = \frac{1 + \left|\frac{Y_o - Y}{Y_o + Y}\right|}{1 - \left|\frac{Y_o - Y}{Y_o + Y}\right|}$$

Since a noise process is defined by power density spectra rather than by Fourier amplitudes, this relation can be conveniently modified so that it yields the ratio of the maximum to the minimum of the current SPDS along the line. We have

$$\frac{\Psi_{max}}{\Psi_{min}} = \left[\frac{1 + \left|\frac{Y_o - Y}{Y_o + Y}\right|}{1 - \left|\frac{Y_o - Y}{Y_o + Y}\right|}\right]^2 \tag{3.38}$$

Equation 3.38 is an example of one of the many uses of the admittance concept for noise propagation along transmission lines. The admittance concept of the analog noise process can be applied directly to the corresponding noise process in the electron beam. The vast body of knowledge concerning admittance and impedance transformations can be brought to bear directly on the analysis of noise processes in electron beams for which an analog transmission-line noise process exists.

---

† In general, nonreciprocal networks may be needed to accomplish this.

A noise process in an electron beam with det(WR) $\neq$ 0 does not possess a network theory analog. The admittance concept is, therefore, not directly applicable. However, if the noise is transformed by lossless transducers only it is still possible to define an analog admittance, as we shall now show.

The WR matrix of a general noise process can be written with the aid of Eq. 3.18 and the definitions of Eqs. 3.26 and 3.27 in the form

$$WR = \begin{pmatrix} \Pi + j\Lambda & \Phi \\ \Psi & \Pi - j\Lambda \end{pmatrix} \tag{3.39}$$

Let us split the WR matrix of a general noise process with det(WR) $\neq$ 0 into two parts as shown below.

$$WR = W'R - I\Delta\Pi \tag{3.40}$$

where I is the identity matrix and $\Delta\Pi$ is a real scalar given in terms of the invariants S and Π of the WR matrix,

$$\Delta\Pi = S - \Pi \tag{3.41}$$

According to Eqs. 3.39, 3.40, and 3.41 the explicit form of the W'R matrix is

$$W'R = \begin{pmatrix} S + j\Lambda & \Phi \\ \Psi & S - j\Lambda \end{pmatrix} \tag{3.42}$$

The determinant of the W'R matrix is, from Eq. 3.29,

$$\det(W'R) = 0 \tag{3.43}$$

Let us now study the transformation of the WR and W'R matrices by lossless transducers. Denoting the values of the WR and W'R matrices at the input to the transducer by the subscript a and at the output by the subscript b, we have, corresponding to Eq. 3.40,

$$W_aR = W_a'R - I\Delta\Pi$$

$$W_bR = W_b'R - I\Delta\Pi \tag{3.44}$$

where the scalar $\Delta\Pi$ is the same in both equations because of the choice of Eq. 3.41 and the invariance of S and Π with regard to lossless transformations. The lossless transformation (Eq. 3.17) applied to Eq. 3.44 gives

$$W_b R = W_b' R - I\Delta\Pi = KW_a RK^{-1} = KW_a' RK^{-1} - I\Delta\Pi$$

or

$$W_b' R = KW_a' RK^{-1}$$

and

$$W_b R = KW_a RK^{-1} \tag{3.45}$$

Comparison of Eqs. 3.39 and 3.42 shows that the W'R matrices have the same off-diagonal elements and the same imaginary parts of their diagonal elements, as the WR matrices. According to Eqs. 3.45 the W'R matrix transforms in the same way as the WR matrix. Thus, the transformation of the off-diagonal elements, Φ and Ψ, and the transformation of the imaginary parts of the diagonal elements, $j\Lambda$, of the WR matrix can be studied with the aid of the transformation of the W'R matrix. The latter satisfies the condition of Eq. 3.43 and thus has an analog noise process with the admittance

$$Y'(\omega) = \frac{\Psi'}{\Theta'} = \frac{\Psi}{S + j\Lambda} \tag{3.37a}$$

This admittance can be used to describe the current SPDS of the W'R matrix, as shown earlier. But, this is the same as the current SPDS of the WR matrix, and thus the latter is also described by the admittance of Eq. 3.37a. It follows that the admittance concept can be applied to any noise process, even those with $\det(WR) \neq 0$, if only transformations by lossless transducers are studied.†

Transformations that extract power from the electron beam do not conserve S and Π. A split of the WR matrix according to Eq. 3.40 is not independent of the reference cross section. In general, no analog process with a W'R matrix which is such that $\det(W'R) = 0$ can be found whose Φ, Ψ, and Λ would transform identically with those of the WR matrix.

---

†The admittance concept was first used in noise transformations by S. Bloom and R. Peter (35).

3.4 ALTERNATE REPRESENTATION OF NOISE

In many computations it is more convenient to use the wave formalism rather than the voltage-current representation of an excitation in an electron beam. For this purpose it is necessary to describe a noise process in terms of normalized wave amplitudes. The Fourier transforms of the kinetic voltage and current in an electron beam of a noise process viewed during a time 2T were written in the form of a column matrix in Eq. 3.12. Normalized wave amplitudes can be assigned to these Fourier transforms according to Eqs. 2.35 and 2.36. We have

$$a_T(\omega) = (I + PR) N_a w_{aT}(\omega) \tag{3.46}$$

$$b_T(\omega) = (I + PR) N_b w_{bT}(\omega) \tag{3.47}$$

The formation of the noise matrices W_a and W_b suggests similar noise matrices formed of the normalized wave amplitudes. We define

$$A = \lim_{T\to\infty} \frac{2\pi}{T} \overline{a_T(\omega)\, a_T(\omega)^+} = 2 \lim_{T\to\infty} \frac{\pi}{T} (I + PR)\, N_a \overline{w_{aT}(\omega)\, w_{aT}(\omega)^+} N_a (I + PR)^+$$
$$= 2(I + PR)\, N_a W_a N_a (I + PR)^+ \tag{3.48}$$

where we have made use of the fact that N_a is a Hermitian matrix. The factor of 2 was introduced for reasons of normalization. In order to explain the normalization involved in Eq. 3.48 we give an example. Imagine that a lossless transmission line is terminated at its two ends into matching resistors at a temperature T. The resistors exchange noise power over the transmission line. Denote by a_1 the normalized amplitude of the wave traveling from left to right, by a_2 the normalized amplitude of the wave from right to left. The values of the elements of the matrix A as defined by Eq. 3.48 would then be

$$A_{12} = A_{21} = 0 \,, \quad \text{since the two waves are uncorrelated,}$$

and

$$A_{11} = A_{22} = \frac{kT}{4\pi} \tag{3.49}$$

according to the Nyquist formula, which in its more conventional form gives the mean power carried in any one of the waves within the frequency band Δf as $kT\Delta f$. The definition of the A matrix thus differs by a normalization factor of $4\pi\Delta f$ from the conventional engineering use. (Compare with the normalization of the current SPDS discussed on page 31.)

A similar "noise wave matrix" can be defined at reference cross section b by

$$B = \lim_{T \to \infty} \frac{2\pi}{T} \overline{b_T(\omega)\, b_T(\omega)^+} = 2(I + PR) N_b W_b N_b (I + PR)^+ \tag{3.50}$$

A beam transducer between the reference cross sections a and b is conveniently characterized in terms of the transformation of Eq. 2.24. This transformation establishes a relation between the noise wave matrices A and B. We have

$$B = \lim_{T \to \infty} \frac{2\pi}{T} \overline{b_T(\omega)\, b_T(\omega)^+} = \lim_{T \to \infty} \frac{2\pi}{T} M\, \overline{a_T(\omega)\, a_T(\omega)^+}\, M^+ = MAM^+$$

Thus

$$B = MAM^+ \tag{3.51}$$

If, in particular, the transducer M is lossless, we must have according to Eq. 2.29

$$M^+PM = P \tag{2.29}$$

or

$$M^+ = PM^{-1}P \tag{3.52}$$

Introducing Eq. 3.52 into Eq. 3.51, and multiplying from the right by P (which is equal to its own inverse), we find

$$BP = MAPM^{-1} \tag{3.53}$$

Equation 3.53 is analogous to the transformation of the WR matrix by lossless transducers, Eq. 3.17. Equation 3.53 leads to the conclusion that the transformation by a lossless transducer must leave the trace and the determinant of the AP matrix equal to that of the BP matrix. Naturally, the invariance of the trace and determinant of the AP matrix must be related to the invariance of the same quantities pertaining to the WR matrix. In order to find how these are related, let us study more carefully the transformation of Eq. 3.48. Multiplying Eq. 3.48 from the right by P, and noting that RR = I according to Eq. 2.15, we can obtain

$$AP = 2(I + PR) N_a W_a RR N_a RR (I + PR) P \tag{3.54}$$

Now, it is easy to prove, by virtue of the definitions of Eqs. 2. 14 and 2. 34, that

$$8RN_aR = N_a^{-1}$$

Further, the following relation is easily proven with the aid of the definitions of Eqs. 2. 14 and 2. 29.

$$R(I+PR)^{+}P = 2(I+PR)^{-1}$$

Thus, we find that Eq. 3. 54 can be written as

$$AP = \frac{1}{2}[(I+PR)N_a]W_aR[(I+PR)N_a]^{-1} \tag{3. 55}$$

The AP matrix and the WR matrix are related by a similarity transformation. Accordingly, there is a direct relation between the traces and the determinants of the two matrices.

$$\text{Trace}(AP) = A_{11} - A_{22} = \frac{1}{2}\text{Trace}(WR) = \Pi \tag{3. 56}$$

$$\det(AP) = |A_{12}|^2 - A_{11}A_{22} = \frac{1}{4}\det(WR) = \frac{1}{4}(\Pi^2 - S^2) \tag{3. 57}$$

The trace of the AP matrix is equal to the real part of the CPDS of the noise process. The inverse of Eq. 3. 55, which gives the matrix W_aR in terms of AP, is also of interest. It is obtained from Eq. 3. 55 by premultiplication by $[(I + PR)N_a]^{-1}$, and postmultiplication by $[(I + PR)N_a]$.

$$W_aR = 2[(I+PR)N_a]^{-1}AP[(I+PR)N_a] \tag{3. 58}$$

The relations between the elements of the WR matrix, as given in Eq. 3. 39, and the elements of the AP matrix are written out in detail below. The set of Eqs. 3. 59 to 3. 61 is obtained by carrying out the matrix operations in Eq. 3. 55; the set of Eqs. 3. 62 to 3. 65 follows from Eq. 3. 58.

$$A_{11} = \frac{1}{4}(Y_o\Phi + Z_o\Psi) + \frac{1}{2}\Pi \tag{3. 59}$$

$$-A_{12} = \frac{1}{4}(Y_o\Phi - Z_o\Psi) - \frac{1}{2}j\Lambda \tag{3. 60}$$

$$-A_{22} = -\frac{1}{4}(Y_o\Phi + Z_o\Psi) + \frac{1}{2}\Pi \tag{3.61}$$

$$\Pi = A_{11} - A_{22} \tag{3.62}$$

$$\Lambda = j(A_{12}^* - A_{12}) = -j(A_{12} - A_{21}) \tag{3.63}$$

$$\Phi = Z_o(A_{11} + A_{22} - A_{12} - A_{12}^*) \tag{3.64}$$

$$\Psi = Y_o(A_{11} + A_{22} + A_{12} + A_{12}^*) \tag{3.65}$$

The subscript a of the elements of the WR matrix has been omitted in these equations, thus emphasizing their applicability at any reference plane.

The transformation of the A matrix by a section of a drift region is much simpler than that of the WR matrix. A section of a drift region of length Θ_q, measured in terms of the plasma transit angle and transit angle Θ, has the M matrix:

$$M = \begin{pmatrix} e^{-j(\Theta-\Theta_q)} & 0 \\ 0 & e^{-j(\Theta+\Theta_q)} \end{pmatrix} \tag{3.66}$$

The expression for M, Eq. 3.66, applied to the transformation of Eq. 3.51, gives (written in detail):

$$\begin{matrix} B_{11} = A_{11} \\ B_{22} = A_{22} \end{matrix} \qquad B_{12} = A_{12}e^{2j\Theta_q} \tag{3.67}$$

The transformation of Eq. 3.67 is easy to comprehend. Along a drift region the fast and the slow waves change in phase only. This accounts for the invariance of the diagonal elements A_{11} and A_{22}. Within a drift region of plasma transit angle Θ_q the fast and the slow wave get out of phase by an angle $2j\Theta_q$, which accounts for the relationship between B_{12} and A_{12}.

A slightly different interpretation of Eq. 3.67 is possible. Let us assume that the reference plane a in a drift region is continuously varied. A movement of the reference plane forward by a plasma transit angle Θ_q corresponds to a change of the argument of A_{12} by an amount $2\Theta_q$. The reference plane can be so chosen that A_{12} is real and positive. Equation 3.65 shows that this reference plane coincides with the maximum of the current SPDS in the drift region. On the other hand, if the argument of A_{12} is not equal to zero, Eqs. 3.65 and 3.67 indicate that the plane of the maximum of the SPDS lies a plasma transit angle equal to $\Delta\Theta_q = \arg(A_{12})/2$ in front of the reference plane. "In front" means against the direction of the time-average velocity u of the beam.

If the off-diagonal element of the A matrix, A_{12}, is equal to zero, the current SPDS is independent of distance, according to Eqs. 3.65 and 3.67. The fast wave and the slow wave in the beam are uncorrelated. The A matrix is diagonal. Equation 3.60 shows that this is the case when

$$\Phi = Z_o^2 \Psi \quad \text{and} \quad \Lambda = 0 \tag{3.68}$$

The following question will prove of interest: Given a general noise process at a reference plane a with some noise wave matrix A, is it possible to pass this noise process through a lossless beam transducer so that the noise matrix B at its output b is diagonal? A general proof of the possibility of such a transformation and how it is achieved is given in reference 19. Here we shall answer the question by simple physical reasoning which will help towards an understanding of noise transformations.

We have shown in section 3. 3 that it is possible to find an equivalent admittance for an arbitrary noise process as long as transformations of noise by lossless beam transducers are considered. The transformation of the current SPDS of the analog noise process with the admittance $Y'(\omega)$ of Eq. 3.37a is the same as that of the general noise process. The question above can be recast into the terminology of admittance transformations. The requirement that the noise wave matrix be diagonal at the cross section b is tantamount to (see Eq. 3. 68)

$$\Psi_b = Y_o^2 \Phi_b \quad \text{and} \quad \Lambda_b = 0$$

The admittance of the analog noise process at cross section b must, according to Eq. 3. 37a and the definition of Eq. 3.29, be

$$Y_b' = Y_o$$

The equivalent admittance at reference cross section a is

$$Y_a' = \frac{\Psi_a}{S + j\Lambda_a}$$

Thus, the original question may be formulated as follows: Is there a lossless transducer that transforms the admittance Y_o at its output into Y_a' at the input? The admittance at the output of the transducer, Y_b', is positive real; the admittance at the input has a positive real part. But, it is always possible to find a lossless transducer that transforms any admittance with a positive real part into any other admittance with a positive real part.

Returning to the discussion of the original noise process we can conclude that it is always possible to find a lossless beam transducer, which, inserted between the reference cross sections a and b, transforms an arbitrary noise wave matrix into a diagonal noise wave matrix. Whether such a beam transducer is physically realizable is an entirely different problem. We know that a great variety of noise transformations can be achieved with a multielectrode gun (24) as shown in Fig. 2. 2. In the subsequent discussion we shall postulate that we can always find the beam transducer that is theoretically required.

When the noise matrix A is brought into a diagonal form B by a lossless beam transducer, the trace and the determinant have to be conserved. According to Eqs. 3.56 and 3.57, we must have

$$B_{11} - B_{22} = \Pi$$

$$B_{11}B_{22} = \frac{1}{4}(S^2 - \Pi^2)$$

with the result that

$$B_{11} = \frac{1}{2}(S + \Pi) \qquad B_{22} = \frac{1}{2}(S - \Pi) \tag{3. 69}$$

If M is the particular transducer that brings the noise wave matrix A into the diagonal form B, we have

$$MAM^+ = B \tag{3. 70}$$

with B diagonal. Premultiplying this equation by M^{-1}, and postmultiplying it by $(M^+)^{-1}$, we obtain

$$A = (M^{-1})B(M^{-1})^+ \tag{3. 71}$$

Since M satisfies the condition of power conservation

$$M^+PM = P \tag{2. 33}$$

we find, after premultiplying Eq. 2. 33 by $(M^+)^{-1}$, and postmultiplying by M^{-1},

$$(M^{-1})^+P(M^{-1}) = P \tag{3. 72}$$

Equation 3.72 shows that the matrix M^{-1} is the matrix of a lossless transducer. Using this knowledge we can interpret Eq. 3.71 as follows: Any noise process with the matrix A can be represented by a diagonal noise process B followed by a lossless transformer with the matrix M^{-1}. The matrix M^{-1} is the inverse of the matrix M required to diagonalize A by the operation of Eq. 3.70.

4. THE MINIMUM OBTAINABLE NOISE FIGURE

4.1 THE NOISE FIGURE EXPRESSION

The generally adopted measure of the sensitivity of an amplifier is the noise figure (36, 37). Here we shall deal exclusively with the spot-noise figure. The spot-noise figure can be evaluated as the ratio of the total noise output of the amplifier N, within the frequency band Δf, over the hypothetical noise output N_0 which the amplifier would have if no additional noise were introduced by the amplifier

$$F = \frac{N}{N_0}$$

The frequency band Δf is picked small enough so that the amplifier characteristics can be assumed to be constant within the band Δf. The noise output N of the amplifier can be ascribed to two sources: the contribution N_i caused by the noise internal to the amplifier, and the noise N_0 from the input circuit which would be present even if the amplifier were noise-free. The two contributions N_0 and N_i are, in general, uncorrelated. Thus, the total noise output N can be written as the sum of N_0 and N_i. The noise figure can be written in the form

$$F = 1 + \frac{N_i}{N_0} \tag{4.1}$$

Figure 4.1 shows a schematic of a microwave longitudinal-beam amplifier. The excitation at the input reference plane on the transmission line feeding the amplifier can be expressed in terms of the normalized amplitudes of the incident and reflected waves, a_3 and b_3. The excitation on the output transmission line is given in terms of the incident and reflected wave amplitudes a_4 and b_4. The electron beam enters the amplifier at the gun end with an excitation of the fast and slow waves a_1 and a_2, respectively. The excitation in the beam leaving the amplifier at the collector end is characterized by the normalized amplitudes of the fast and slow waves, b_1 and b_2. The amplifier can be described by the 4 × 4 matrix G which gives the linear relation between the column matrices a and b, each formed of the four normalized wave amplitudes.

$$b = Ga \tag{2.37}$$

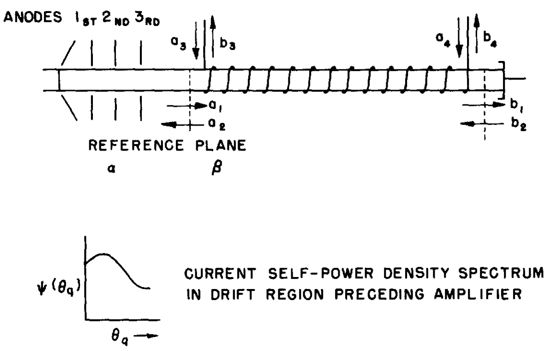

Fig. 4.1. Schematic of microwave longitudinal-beam amplifier with multi-electrode gun.

Without loss of generality we can assume that the input source is matched to the input transmission line and that the output load is matched to the output transmission line. For the noise leaving the amplifier with the output transmission line matched to the load, and thus a_4 equal to zero,† we have

$$N = \lim_{T \to \infty} \frac{2\pi}{T} \overline{b_{4T} b_{4T}^*} 4\pi\Delta f$$

The fourth element of the matrix relation (Eq. 2. 37) gives the amplitude b_4.

$$b_4 = G_{41} a_1 + G_{42} a_2 + G_{43} a_3 \tag{4. 2}$$

The noise power N_o is found by disregarding the noise contribution of the electron beam, i. e., by setting $a_1 = a_2 = 0$.

$$N_o = \lim_{T \to \infty} \frac{2\pi}{T} |G_{43}|^2 \overline{a_{3T} a_{3T}^*} 4\pi\Delta f = |G_{43}|^2 A_{33} 4\pi\Delta f \tag{4. 3}$$

The noise power carried by the incident wave is given by the Nyquist formula (compare Eq. 3. 49):

$$A_{33} 4\pi\Delta f = kT\Delta f \tag{4. 4}$$

where k is Boltzmann's constant, and T is the temperature of the input circuit. The noise power N_i caused by the noise internal to the amplifier comes from the noise in the electron beam. We have

$$N_i = \lim_{T \to \infty} \frac{2\pi}{T} \overline{(G_{41} a_{1T} + G_{42} a_{2T})(G_{41} a_{1T} + G_{42} a_{2T})^*} 4\pi\Delta f$$

$$= (|G_{41}|^2 A_{11} + |G_{42}|^2 A_{22} + G_{41} G_{42}^* A_{12} + G_{41}^* G_{42} A_{12}^*) 4\pi\Delta f \tag{4. 5}$$

Combining Eqs. 4. 1, 4. 3, 4. 4, and 4. 5 we find the noise figure of the amplifier.

$$F = 1 + \frac{4\pi}{kT} \frac{1}{|G_{43}|^2} [|G_{41}|^2 A_{11} + |G_{42}|^2 A_{22} + G_{41} G_{42}^* A_{12} + G_{41}^* G_{42} A_{12}^*] \tag{4. 6}$$

4. 2 MINIMIZATION OF THE NOISE FIGURE

Let us imagine that the beam passes through a multielectrode gun as shown in Fig. 4. 1 before it enters the amplifier. The multielectrode gun is a lossless transducer

---

† Note the remarks on the normalization of Eq. 3. 48.

by which the noise can be adjusted at the reference cross section at the gun end of the amplifier. In many amplifiers the beam passes through a drift region before it enters the amplifier. We shall assume this to be the case, merely because the adjustment of the standing wave of the current SPDS in this drift region gives a good visual representation of the noise transformations by the multielectrode gun. We shall assume that the multielectrode gun is versatile enough to be able to give an arbitrary standing-wave ratio of the current SPDS within the drift region. Stated in mathematical language this assumption is equivalent to the requirement that A_{11}, A_{22}, and A_{12} at the input reference cross section can be adjusted arbitrarily, subject to the condition of conservation of S and Π by lossless transformations. We have, from Eqs. 3. 56 and 3. 57,

$$\begin{aligned} A_{11} - A_{22} &= \Pi = \text{constant} \\ (A_{11} + A_{22})^2 - 4|A_{12}|^2 &= S^2 = \text{constant} \end{aligned} \tag{4. 7}$$

The requirement of the invariance of S and Π leaves two of the four noise parameters adjustable. One of these parameters is the argument of A_{12} which does not enter into the conditions of Eq. 4. 7. The choice of arg (A_{12}) which minimizes the expression of Eq. 4. 6 is, obviously,

$$\arg (A_{12}) = \arg (G_{42}) - \arg (G_{41}) + \pi \tag{4. 8}$$

This condition is achieved when the maximum of the current SPDS in the drift region preceding the amplifier lies a plasma transit angle,

$$\Delta \Theta_q = \frac{1}{2}\left[\arg (G_{42}) - \arg (G_{41}) + \pi\right] = \frac{1}{2}\left[\arg\left(\frac{G_{42}}{G_{41}}\right) + \pi\right] \tag{4. 9}$$

in front of the reference cross section in the beam. The first minimization leads to the noise figure expression

$$F = 1 + \frac{4\pi}{kT} \frac{1}{|G_{43}|^2} \left(|G_{41}|^2 A_{11} + |G_{42}|^2 A_{22} - 2\,|G_{41} G_{42}|\,|A_{12}|\right) \tag{4. 10}$$

This expression allows one further minimization. We can express A_{11} and A_{22} in terms of $|A_{12}|$ with the aid of Eq. 4. 7.

$$\begin{aligned} A_{11} &= \frac{1}{2}\left[(S^2 + 4|A_{12}|^2)^{1/2} + \Pi\right] \\ A_{22} &= \frac{1}{2}\left[(S^2 + 4|A_{12}|^2)^{1/2} - \Pi\right] \end{aligned} \tag{4. 11}$$

These expressions can be introduced into the noise figure expression (Eq. 4.10), which becomes a function of a single variable, $|A_{12}|$. Minimization with regard to this variable is achieved when

$$|A_{12}| = \left| \frac{G_{41} G_{42}}{|G_{41}|^2 - |G_{42}|^2} \right| S \tag{4.12}$$

The minimum noise figure is

$$F_{min} = 1 + \frac{2\pi}{kT} |D| \left[S - \frac{D}{|D|} \Pi \right] \tag{4.13}$$

where

$$D = \frac{|G_{42}|^2 - |G_{41}|^2}{|G_{43}|^2}$$

The minimization leads to a duplicity of sign of a square root. The duplicity is resolved by noting that the contribution of the beam noise to the output must be positive, regardless of the sign of Π.

The standing-wave ratio of the current SPDS required for the minimization of the noise figure is (see Eq. 3.65) from Eqs. 4.11 and 4.12,

$$\frac{\Psi_{max}}{\Psi_{min}} = \frac{A_{11} + A_{22} + 2|A_{12}|}{A_{11} + A_{22} - 2|A_{12}|} = \left(\frac{|G_{41}| + |G_{42}|}{|G_{41}| - |G_{42}|} \right)^2 \tag{4.14}$$

The minimum noise figure of the amplifier obtainable by means of a lossless transducer in the beam depends, according to Eq. 4.13, partly on the two noise invariants S and Π, partly on the amplifier structure through the constant D. In practice, it is often necessary to find an appropriate multielectrode gun that will give the best possible noise figure when used in a given amplifier. In view of our limited theoretical knowledge of the noise input conditions beyond the potential minimum in the electron gun, a semiempirical approach to this problem may be used. First of all, the parameters G_{41}/G_{43}, and G_{42}/G_{43} can be computed for the given amplifier. For a traveling-wave tube these parameters are, in Pierce's (1) notation,

$$\frac{G_{41}}{G_{43}} = jC(\delta_2 + \delta_3)\left(\frac{Z_o}{K}\right)^{1/2} - (\delta_2 \delta_3 - 4QC) \frac{2V_o C^2}{I_o} (KZ_o)^{-1/2}$$

$$\frac{G_{42}}{G_{43}} = -jC(\delta_2 + \delta_3)\left(\frac{Z_o}{K}\right)^{1/2} - (\delta_2 \delta_3 - 4QC) \frac{2V_o C^2}{I_o} (KZ_o)^{-1/2}$$

where K is the helix impedance and Z_0 is the characteristic impedance of the beam as defined by Eq. 2.7. The δ's are the incremental propagation constants that can be found from a solution of the determinantal equation for the propagation constants of the traveling-wave tube. C and QC are parameters defined by Pierce to characterize the strength of coupling between the circuit and the beam, and the space charge in the beam. The values for G_{41}/G_{43} and G_{42}/G_{43} computed from the equations given above can be introduced into Eqs. 4.9 and 4.14 to find the position of the current standing-wave maximum and the noise current standing-wave ratio for the optimum noise figure. Once this is found, a multielectrode gun (see Fig. 2.2) can be tested for the standing wave of the noise current SPDS. This can be done by a sliding-cavity beam tester which is now in use in several laboratories (13,20,46). The standing-wave ratio can be adjusted to the desired value (Eq. 4.14) by changes in the electrode potentials. If the level of the product $\Psi_{max}\Psi_{min}$ as given by Eq. 3.36 does not change in the process we can assume that the gun performs as ideally predicted by the theory. Should the level rise, the gun design is suspect and must be changed. Experience shows that gun designs with the smoothest possible potential distribution - as shown, for example, in Fig. 2.2 - perform closest to the idealized theory. When the proper current standing wave is achieved, the gun can be built into the amplifier. Small final adjustments in the potentials of the electrodes may be necessary in order to achieve the best performance.

The adjustments described above do not require the knowledge of Π. The value of Π serves only to determine the final value of the optimized noise figure.

Deviations of the noise behavior from that predicted by the idealized theory are ascribed, today, to two effects. First, there is the lens action of the various electrodes in the multielectrode gun. The lens action may result in cross-overs of the electron trajectories. Such cross-overs are believed to be harmful to low noise figures, since they may cause a transformation of the random motion in the transverse direction (which is always present under physically realizable magnetic focusing fields) into the longitudinal direction. Secondly, an electron beam propagates, aside from the dominant space-charge waves included in the one-dimensional theory (see remark on p. 38), higher order space-charge waves. These higher order space-charge waves are excited by the input noise and are picked up by a sliding cavity beam tester. Thus, a sliding cavity beam tester observes not only the standing wave of noise current set up by the dominant space-charge waves, but also the noise current pattern of the higher order space-charge waves. Their effect is particularly pronounced at the standing-wave minima of the noise current in the dominant space-charge waves and may influence a reading of the standing-wave ratio of the noise current in the dominant space-charge waves. It has been found experimentally, on a number of low-noise, multielectrode guns that the product $\Psi_{max}\Psi_{min}$ is conserved through noise transformations that give standing-wave ratios of noise current of less than 10 db. These findings have been taken as the experimental proof that the effect of the higher order space-charge waves upon a reading of the noise current standing-wave ratio is negligible for current standing-wave ratios of less than 10 db. The value of the standing-wave ratio of noise current to be used in the computations of the noise figure of an amplifier is that of the dominant space-charge waves. This follows directly from the fact that the theory is based on a two-wave picture and can, therefore, account only for noise carried in two waves. Since the coupling of slow wave structures to the higher order space-charge waves is weak, the noise figure of a traveling-wave tube is not believed to be strongly affected by the existence of the higher order space-charge waves.

4.3 MAGNITUDE OF THE PARAMETER D FOR A LOSSLESS AMPLIFIER

The subsequent analysis will be limited to lossless amplifier structures. (The effect of loss in the amplifier is treated in general terms in reference 19.) It may be stated here that the minimum noise figure of an amplifier with loss can only be higher than or, at best, equal to, that of a lossless amplifier.

Equation 2.42 imposes upon the elements of the G matrix a condition that has a profound influence upon the characteristic constant D. From Eq. 2.42 it follows that

$$|\det(G)|^2 = 1$$

Thus, the G matrix has a reciprocal, G^{-1}. Premultiplying Eq. 2.42 by GP, and postmultiplying it by $(PG)^{-1}$, we get

$$GPG^+ PG(PG)^{-1} = GPP(PG)^{-1}$$

and since the P matrix is its own reciprocal, we have

$$GPG^+ = P \qquad (4.15)$$

Superficially, Eq. 4.15 looks like Eq. 2.42. But it implies relations among the rows of the G matrix - as, for example,

$$|G_{41}|^2 - |G_{42}|^2 + |G_{43}|^2 + |G_{44}|^2 = 1 \qquad (4.16)$$

From Eq. 4.16 we find for the amplifier constant D (see Eq. 4.13),

$$D = \frac{|G_{42}|^2 - |G_{41}|^2}{|G_{43}|^2} = 1 - \frac{1 - |G_{44}|^2}{|G_{43}|^2} \qquad (4.17)$$

Since $|G_{43}|^2$ is the power gain of the amplifier, it must always be greater than unity. Thus, D is always positive. Two classes of amplifiers have to be distinguished. One class has $|G_{44}| < 1$; the other has $|G_{44}| > 1$. Let us recall the significance of the matrix element G_{44}. The fourth element of the matrix relation (Eq. 2.37) reads

$$b_4 = G_{41}a_1 + G_{42}a_2 + G_{43}a_3 + G_{44}a_4$$

With no excitation in the electron beam, $a_1 = a_2 = 0$, and with the input transmission line matched to a passive load, $a_3 = 0$, we find that the reflection coefficient measured in the output transmission line is $G_{44} = b_4/a_4$. A reflection coefficient of magnitude less than unity is caused by a passive load, a reflection coefficient of magnitude greater than unity is produced by an impedance with a negative real part. For $|G_{44}| > 1$ the output impedance of the amplifier with the input transmission line matched, has a negative real part. The amplifier is only conditionally stable with regard to end-loading.

Conversely, if $|G_{44}| < 1$, the output impedance of the amplifier has a positive real part. In this case it is always possible to find a lossless two terminal-pair network which, inserted between the amplifier and the output transmission line, matches the output impedance of the amplifier to the line. When this is done, the output power for a given input power is also maximized. The quantity $|G_{43}|^2$ becomes the "available power gain", G_{av}, of the amplifier, i. e., the ratio of the available output power over the available input power. The element G_{44} is reduced to zero. The parameter D becomes

$$D = 1 - \frac{1}{G_{av}} \tag{4.18}$$

The class of amplifiers with $|G_{44}| > 1$ cannot be matched to the output transmission line. The quantity $|G_{43}|^2$ still has the meaning of power gain, G, the ratio of the output power over the available input power. For this class of amplifiers,

$$D \geq 1 - \frac{1}{G_{av}} \tag{4.19}$$

Comparison of Eqs. 4.18 and 4.19 shows that amplifiers with an output impedance with a negative real part are, in general, less desirable from the point of view of noise performance.

A third possibility should not be excluded, namely, amplifiers with $G_{44} = 1$. For these amplifiers Eq. 4.17 gives directly,

$$D = 1 \tag{4.20}$$

independent of the gain.

Since D is always positive, the minimum obtainable noise figure as given by Eq. 4.13 depends only upon the difference of the noise invariants, $S - \Pi$. (This result does not depend upon the assumption of zero loss, as shown in general in reference 19.)

The lowest possible noise figure with a given gain is achieved with amplifiers that have a $|G_{44}| < 1$ and have been subsequently matched to the output transmission line. The minimum noise figure of this class of amplifiers has the form

$$F_{min} = 1 + \left(1 - \frac{1}{G_{av}}\right)\frac{2\pi}{kT}(S - \Pi) \tag{4.21}$$

Equation 4.21 shows that the noise figure of an amplifier can, in general, be reduced to unity at a corresponding sacrifice of gain. Such a result is not surprising. Indeed, one can almost always decouple the microwave structure from the electron beam, thus preventing any beam noise from entering the structure, with a resulting noise figure of unity. Naturally, the gain is then reduced to unity also. The dependence of the minimum noise figure upon the gain may suggest another scheme for achieving amplification with a small noise figure. Assume that a set of n electron guns is available, all with the same lowest possible value of $S - \Pi$. Then, construct n amplifiers using these guns, each of the amplifiers with a low gain and low noise figure corresponding to Eq. 4.21. Is it then possible to achieve a noise figure lower than that given by Eq. 4.21 at some large available gain, G_{av}, by cascading the n amplifiers? To find an answer to this question let us assume that all n amplifiers have the same available gain as the first amplifier, g_1. The gain of the cascade is

$$G_{av} = g_1^n \tag{4.22}$$

The noise figure of the cascade can be found from the well-known formula for the noise figure of a cascade of amplifiers:

$$F = F_1 + \frac{F_2 - 1}{g_1} + \dots \frac{F_n - 1}{g_{n-1}} \tag{4.23}$$

where the subscripts refer to the order of the arrangement of the amplifiers. In our case we have assumed that all of the amplifiers have the same gain, and according to Eq. 4.21 they have the same noise figure. For the case of n such amplifiers we find, from Eq. 4.23,

$$F = 1 + (F_1 - 1)\frac{1 - \frac{1}{g_1^n}}{1 - \frac{1}{g_1}}$$

where F_1 and g_1 refer to the noise figure and gain of the first amplifier. Introducing Eq. 4.21 for the noise figure of the first amplifier and using Eq. 4.22, we find

$$F = 1 + \left(1 - \frac{1}{G_{av}}\right)\frac{2\pi}{kT}(S - \Pi) \qquad (4.24)$$

The over-all noise figure is identical with the noise figure obtainable with a single amplifier at a corresponding gain, G_{av}. Thus, the cascading scheme cannot lead to any improvement over the noise figure of Eq. 4.21.

4.4 APPLICATIONS

The application of the preceding analysis to a series of practical cases is interesting. In principle, a lossless traveling-wave tube, as shown schematically in Fig. 4.1, can be matched to the input and output transmission lines. Then, no reflection occurs when power is fed into the output with the input transmission line terminated into a matched resistor, $G_{44} = 0$. Equation 4.21 for the minimum noise figure applies directly. Thus, the lossless traveling-wave tube is a microwave-beam amplifier that theoretically achieves the lowest possible noise figure (Eq. 4.21). If we assume, following Pierce, that S is given by Eq. 3.30, and $\Pi = 0$, we find that the limiting noise figure is

$$F_{min} = 1 + \left(1 - \frac{1}{G_{av}}\right)(4 - \pi)^{1/2}\frac{T_c}{T} \qquad (4.25)$$

This expression was obtained by various authors in the limit of large gain (42, 35, 40). With the choice of $T = 300°K$, $T_c = 1200°K$, and $G_{av} \rightarrow \infty$, we have

$$F_{min} \approx 6 \text{ db}$$

Next let us consider a klystron amplifier. In order to be consistent with the assumptions of the preceding theory, we must postulate that the cavities of the klystron are lossless. If, in addition, the cavity gaps are very short, the beam loading admittance is theoretically zero. The output cavity viewed from the output transmission line looks like a reactive termination, thus $G_{44} = 1$. According to Eq. 4.20, $D = 1$, and the minimum noise figure is (32)

$$F_{min} = 1 + \frac{2\pi}{kT}(S - \Pi) \qquad (4.26)$$

The formalism that led to Eq. 4.21 for the minimum noise figure was developed under the assumption of a lossless microwave structure. It is not hard to modify it so that it is applicable to several practical cases with loss. One case of practical interest is the traveling-wave tube with a severed helix. Somewhere between the input and output, the helix of the traveling-wave tube is interrupted by a lossy section so that regenerative feedback from reflections from the output is prevented. One can represent this effect by assuming that the traveling-wave tube has two lossless helices, one end of each terminated in matched loads (Fig. 4.2). The traveling-wave tube is now characterized by a lossless 6×6 G matrix, relating the column vectors a and b, each of sixth order. The numbering of the waves is shown in Fig. 4.2. The fast and slow waves in the beam at the input end are still a_1 and a_2. The input and output of the traveling-wave tube are denoted by the subscripts 3 and 4. The two new terminal pairs that are terminated in matched resistances are terminals 5 and 6. The 4,4 element of the matrix relation (Eq. 4.15), with G a matrix of sixth order and P = diag (1, -1, 1, 1, 1, 1), reads,

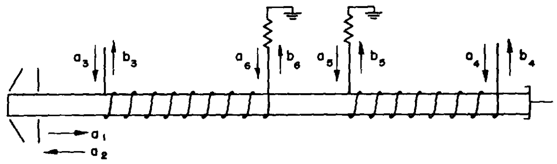

Fig. 4.2. Schematic of traveling-wave tube with severed helix.

$$|G_{41}|^2 - |G_{42}|^2 + |G_{43}|^2 + |G_{44}|^2 + |G_{45}|^2 + |G_{46}|^2 = 1 \tag{4.27}$$

If the second helix is matched to the output, we have $G_{44} = 0$. Further, a wave traveling in a direction opposite to the flow of the beam cannot excite the beam. Hence, the wave a_6 cannot couple to the output, $G_{46} = 0$. The quantity $|G_{43}|^2$ is the available gain of the over-all amplifier, G_{av}; the term $|G_{45}|^2$ would be the available gain, G'_{av}, if only the output helix section were used as an amplifier. If the contribution to the noise by the matching resistor 5 is neglected, Eq. 4.17 for the parameter D still holds unchanged. These conditions introduced into Eq. 4.17 give

$$D = 1 - \frac{1}{G_{av}} + \frac{G'_{av}}{G_{av}} \tag{4.28}$$

Comparison of Eq. 4.28 with Eq. 4.18 shows that D is not necessarily much larger than the D of a lossless traveling-wave tube. If the gain in the first section of the severed helix is appreciable, G'_{av}/G_{av} can be kept small. The noise figure then gets established essentially in the first section of the amplifier. The over-all traveling-wave tube acts quite like a cascade of two amplifiers in which the noise figure of the second amplifier does not affect the over-all noise figure if the gain in the first amplifier is sufficiently large.

Finally, if we consider a backward wave amplifier as an example, we find that its formalism is identical with that of a traveling-wave tube. A backward-wave amplifier with a lossless microwave structure can be described by a lossless 4×4 G matrix. The input terminals 3 are now situated at the collector end; the output terminals 4 are at the gun end. Aside from this difference in physical appearance, the mathematical formalism is identical with that of a traveling-wave tube, and the minimum noise figure expression (Eq. 4. 21) applies to the backward-wave amplifier as well.

4. 5 AN ALTERNATE DERIVATION OF THE MINIMUM NOISE FIGURE

An amplifier has electromagnetic power gain at the expense of kinetic power in the electron beam. Electromagnetic power can be extracted from the electron beam only if the negative kinetic power content in the electron beam is increased. In other words, an amplifier has to couple to the slow wave in the electron beam which carries negative kinetic power. This phenomenon was recognized when the magnitude of the term $|G_{23}|^2$ in Eq. 2. 43 was discussed.

Noise is carried in the beam both by the fast and by the slow wave. As long as there is partial correlation between the two waves, the noise in one can be used to cancel part of the noise in the other. We know that a lossless beam transducer can be used to achieve this partial cancellation until both noise waves are uncorrelated. This is equivalent to saying that the noise matrix B beyond the transducer is diagonal.

Let us now assume that the noise waves at the input reference plane of the amplifier of Fig. 4. 1 are uncorrelated; the B matrix is diagonal. We may then ask, What is the best possible amplifier which gives the lowest noise figure? In answering this question we shall gain a better understanding of the noise parameter $S - \Pi$.

The noise figure expression (Eq. 4.6) can be applied directly to the problem by merely identifying the elements of the A matrix with those of the diagonal B matrix. We have

$$F = 1 + \frac{4\pi}{kT} \frac{1}{|G_{43}|^2} [|G_{41}|^2 B_{11} + |G_{42}|^2 B_{22}] \qquad (4. 29)$$

Next, we may ask, How can we optimize Eq. 4. 29 by a proper choice of the amplifier? Since both B_{11} and B_{22} are positive, Eq. 4.29 is clearly optimized if $|G_{41}|^2$ and $|G_{42}|^2$ are selected as small as possible and $|G_{43}|^2$ is selected as large as possible. Condition 4.16 shows that this is the case when $G_{41} = G_{44} = 0$, and

$$\frac{|G_{42}|^2}{|G_{43}|^2} = 1 - \frac{1}{|G_{43}|^2} = 1 - \frac{1}{G_{av}} \qquad (4. 30)$$

But, according to Eq. 3.69, $B_{22} = (S - \Pi)/2$. Combining Eqs. 3. 69, 4. 30, and 4. 29, we find expression 4. 21 for the minimum noise figure.

This simple minimization teaches a useful lesson. As long as there is correlation between the slow wave and the fast wave there is always a hope of canceling the noise in one wave with that of the other. Such a scheme fails in the absence of correlation. The best that can be done from the point of view of noise figure is to couple to the slow wave only. Such a coupling is necessary for the operation of the amplifier. The fast wave should not couple to the output ($G_{41} = 0$), since it carries positive kinetic power and thus cannot be used as a source of electromagnetic power. Coupling to the fast wave would only introduce additional noise.

4.6 CONCLUSIONS

Based on the assumptions of the one-dimensional, single-velocity, small-signal theory, the preceding investigation has shown that there is a lower limit to the noise figure for microwave-beam amplifiers with large gain. This lower limit is a function solely of the noise process in the region of the potential minimum. The value of the basic noise parameter $S-\Pi$, established in the region of the potential minimum, gives the minimum noise figure at large gain, F_{min}.

$$F_{min} = 1 + \frac{2\pi}{kT}(S - \Pi)$$

It was also shown that a traveling-wave tube with a microwave structure of zero (i.e., small) loss achieves the theoretical minimum noise figure solely with the aid of a conventional multielectrode gun.

Future work on low-noise microwave tubes will, therefore, have to be concerned with two major questions: (a) What is the lowest possible value of the noise parameter, $S-\Pi$? (b) How can the minimum noise figure given by Eq. 4.24 be lowered by schemes that fall outside the realm of validity of the assumption of one-dimensional, single-velocity theory?

The first question has been approached theoretically by P. K. Tien (44) in an ingenious and elaborate digital computation. He analyzed the motion of the electrons emitted at random from the cathode between the cathode and the potential minimum. Mutual repulsion forces among the electrons were taken into account. Then, essentially the parameters S and Π were evaluated slightly beyond the potential minimum by an averaging process analogous to that of Rack. Although there may still be some doubt as to whether or not it is legitimate to compute averages at a reference plane at which the single velocity approximation does not hold, Tien's results seem to establish rather convincingly that the noise parameter Π is approximately zero at all frequencies and that S is a function of frequency, as shown in Fig. 4.3. It is noteworthy that S is small at low frequencies, rises at somewhat higher frequencies, but takes a decided dip when the frequency of operation is somewhat below the plasma frequency at the potential minimum. If this dip actually exists, in spite of unavoidable irregularities in the cathode, it would prove of great practical importance.

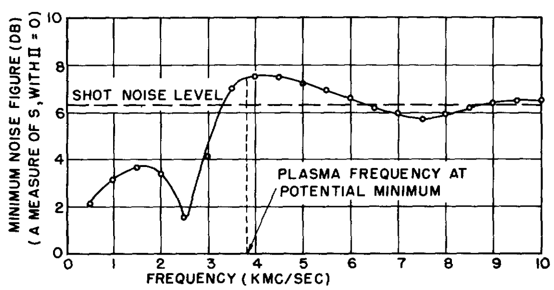

Fig. 4.3. Tien's computed minimum noise figure, $F_{min} = 1 + (2\pi/kT)S$, versus frequency.

Undoubtedly, more theoretical and experimental work will be concentrated around this intriguing question in the future.

The second possibility mentioned above is concerned with a means of circumventing the lower limit on the microwave tube noise figure as established in Eq. 4.24. This would have to be accomplished by schemes for which the single-velocity, one-dimensional assumptions do not hold. One might try to influence the emission from the cathode, or potential minimum, so that the noise parameter $S-\Pi$ would be reduced. Or one might look for microwave tubes other than those of the longitudinal-beam type, which would possess a limit on noise figure lower than that of the longitudinal-beam tube. Among these, the transverse field tube has been considered (45), particularly because it seems to allow a reduction of noise by beam collimation, a scheme that promises to be simpler than are noise reduction schemes that influence the cathode emission in longitudinal tubes.

Finally, we note that the one-dimensional assumptions used in the present report imply that the beam can propagate only two waves, the dominant space-charge waves. If, on the other hand, other waves are, or are made to be, important, the present theory is not applicable. The more elaborate theory of reference 19 would have to be used. It is conceivable that the value of the lower limit of the noise figure could be influenced by a proper choice of geometry.

References

1. J. R. Pierce, "Traveling-Wave Tubes", D. Van Nostrand Co., New York (1950).
2. D. R. Hamilton, J. K. Knipp, J. B. H. Kuper, "Klystrons and Microwave Triodes", Rad. Lab. Ser., Vol. 7, McGraw-Hill Book Co., Inc., New York (1948).
3. J. C. Slater, "Microwave Electronics", D. Van Nostrand Co., New York, (1950).
4. W. Kleen, "Einführung in die Mikrowellen Elektronik", S. Hirzel, Zürich (1952).
5. C. K. Birdsall, J. R. Whinnery, "Waves in an electron stream with general admittance walls", J. Appl. Phys., 24, 314-323 (1953).
6. L. M. Field, P. K. Tien, D. A. Watkins, "Amplification by acceleration and deceleration of a single-velocity stream", Proc. I. R. E., 39, 194 (1951).
7. A. V. Haeff, "The electron-wave tube - a novel method of generation and amplification of microwave energy", Proc. I. R. E., 37, 4-10 (1949).
8. C. K. Birdsall, "Rippled wall and rippled stream amplifiers", Proc. I. R. E., 42, 1628-1636 (1954).
9. R. Kompfner, N. T. Williams, "Backward-wave tubes", Proc. I. R. E., 41, 1602-1611 (1953).
10. D. A. Watkins, "Noise reduction in beam type amplifiers", Technical Report 31, Electronics Research Laboratory, Stanford University (1951).
11. G. Ecker, "Gesamtheiten mit kollektiver Wechselwirkung", parts I and II, Zeitsch. f. Physik, 140, 274-307 (1955).
12. H. M. Mott-Smith, "Change of electron temperature in an electron beam", J. Appl. Phys., 24, 249-255, (1953).
13. C. C. Cutler, C. F. Quate, "Experimental verification of space-charge and transit time reduction of noise in electron beams", Phys. Rev., 80, 875-878 (1950).
14. D. A. Watkins, "The effect of velocity distribution in a modulated electron stream", J. Appl. Phys., 23, 568-573 (1952).
15. A. M. Clogston, and L. R. Walker, unpublished memos, Bell Telephone Co.; also, "The dispersion formula for plasma waves", Appl. Phys., 25, 131-132 (1954).
16. H. A. Haus, "Propagation of noise and signals along electron beams at microwave frequencies", Sc. D. Thesis, M. I. T. (1954).
17. W. C. Hahn, "Small signal theory of velocity-modulated electron beams", G. E. Rev., 42, 258-270 (1939).

18. S. Ramo, "Space charge and field waves in an electron beam", Phys. Rev., 56, 276-283 (1939).

19. F. N. H. Robinson, H. A. Haus, "Analysis of noise in electron beams", Journal of Electronics, Ser. 1, IV, 373 (1956).

20. L. D. Smullin, C. Fried, "Microwave noise measurements on electron beams", Trans. I. R. E., PG-ED, ED-1, No. 4, 168-183 (1954).

21. L. J. Chu, "A kinetic power theorem", paper delivered at the I. R. E. conference, PG-ED, Durham, N. H., June (1951).

22. F. B. Llewellyn, "Electron Inertia Effects", Cambridge University Press (1941).

23. F. B. Llewellyn, L. C. Peterson, "Vacuum-tube networks", Proc. I. R. E., 32, 144-166 (1944).

24. R. W. Peter, "Low noise traveling wave tube", RCA Review, 13, 344-368 (1952).

25. B. Agdur, H. A. Haus, "A one-dimensional electron beam as a four-terminal network", Quarterly Progress Report, April 15, 1954, Research Laboratory of Electronics, M. I. T., pp. 30-32.

26. J. Müller, "Untersuchungen über Elektronenströmungen", Zeitsch. f. Angew. Math. u. Physik, 5, part I, 203-232; part II, 409-419 (1954).

27. S. Bloom, R. W. Peter, "Transmission-line analog of a modulated electron beam", RCA Rev., 15, 95-112 (1954).

28. W. H. Louisell, J. R. Pierce, "Power flow in electron beam devices", Proc. I. R. E., 43, 425-427 (1955).

29. E. A. Guillemin, "Communication Networks", Vol. II, John Wiley & Sons, Inc., New York (1935).

30. J. R. Pierce, "A theorem concerning noise in electron streams", J. Appl. Phys., 25, 931-933 (1954).

31. H. W. König, "Korrelationsverhältnisse beim Schroteffekt", Archiv d. El. Üb., 9, 109-116 (1955).

32. H. A. Haus, "Noise in one-dimensional electron beams", J. Appl. Phys., 26, 560-570 (1955).

33. A. J. Rack, "Effect of space charge and transit time on the shot noise in diodes", Bell System Tech. J., 17, 592-619 (1938).

34. H. E. Rowe, "Shot noise in electron beams at microwave frequencies", Sc. D. Thesis, M. I. T. (1952); also Technical Report 239, Research Laboratory of Electronics, M. I. T. (1952).

35. S. Bloom, R. W. Peter, "A minimum noise figure for the traveling-wave tube", RCA Rev., 15, 252-267 (1954).

36. H. T. Friis, "Noise figure of radio receivers", Proc. I. R. E., 32, 419-422 (1944).

37. K. Fränz, "Messung der Empfängerempfindlichkeit bei kurzen elektrischen Wellen", Zeitsch. Elektr. Elektroak., 59, 105-112 (1942).

38. H. Nyquist, "Thermal agitation of electric charge in conductors", Phys. Rev., 32, 110-113 (1928).

39. H. A. Haus, F. N. H. Robinson, "The minimum noise figure of microwave beam amplifiers", Proc. I. R. E., 43, 981-991 (1955).

40. F. N. H. Robinson, "Microwave shot noise in electron beams and the minimum noise factor of traveling wave tubes and klystrons", J. Brit. I. R. E., 14, 79-86 (1954).

41. C. E. Muehe, Jr., "Noise figure of traveling-wave tubes", M. S. Thesis, M. I. T. (1952); also, Technical Report 240, Research Laboratory of Electronics, M. I. T. (1952).

42. J. R. Pierce, W. E. Danielson, "Minimum noise figure of traveling-wave tubes with uniform helices", Appl. Phys., 25, 1163-1165 (1954).

43. Y. W. Lee, "Application of statistical methods to communication problems", M. I. T., Research Laboratory of Electronics, Technical Report 181 (1950).

44. P. K. Tien, "A dip in the minimum noise figure of beam-type microwave amplifiers", Proc. I. R. E., 44, 938 (1956).

45. G. Wade, K. Amo, D. A. Watkins, "Noise in transverse-field traveling-wave tubes", J. Appl. Phys., 25, 1514-1520 (1954).

46. Quarterly Progress Report No. 3, Research and development on microwave generators, mixing devices and amplifiers, Radio Corporation of America, Aug. 1955.