Waves in Electron Streams and Circuits

By J. R. PIERCE

(*Manuscript Received Jan. 9, 1951*)

This paper reviews some of the assumptions made and some of the general problems involved in analyzing the behavior of electron streams coupled to circuits. It explains why a wave approach is used. The propagation constant of the wave is obtained in terms of the properties of the electron stream and the impedance of the circuit. Some general properties of waves are discussed. The importance of fitting boundary conditions in the solution of an actual problem is discussed, and examples, including that of "backward-gaining" waves, are discussed.

INTRODUCTION

Of recent years, a good deal of work has appeared concerning small linear perturbations of uniform clouds of electrons and ions.[*1–4] A number of questions can be raised concerning the physical interpretations of such mathematical labors.

First of all, for there to be a very direct physical interpretation, the unperturbed state must exist at some time or place and then be modified in the manner described by the perturbation. This condition is satisfied, for instance, in the case of an electron stream of moderate current shot into a long metal tube and confined by a longitudinal magnetic field. However, if the current is made large enough, the uniform flow becomes unstable[5, 6] and the method of perturbations can be used only to establish such instability and not to determine what form the flow will assume. I feel some misgivings about drawing physical interpretations from perturbations of uniform d-c. plasmas and infinitely extending clouds of charge unless these unperturbed states can be shown to exist physically, or unless the results can be shown

* A few late references only are given; others are quoted in those cited.

1 D. Bohm and E. P. Gross, Theory of Plasma Oscillations: A. Origin of Medium-Like Behavior, *Phys. Rev.*, Vol. 75, pp. 1851–1864 (1949); B. Excitation and Damping of Oscillations, *Phys. Rev.*, Vol. 75, pp. 1864–1876 (1949). Effects of Plasma Boundaries in Plasma Oscillations, *Phys. Rev.*, Vol. 79, pp. 992–1001 (1950).

2 J. A. Roberts, "Wave Amplification by Interaction with a Stream of Electrons," *Phys. Rev.*, Vol. 76, pp. 340–344 (1949).

3 V. A. Bailey, "The Growth of Circularly Polarized Waves in the Sun's Atmosphere and Their Escape into Space," *Phys. Rev.*, Vol. 78, pp. 428–443 (1950).

4 "Traveling Wave Tubes," J. R. Pierce, Van Nostrand, 1950.

5 A. V. Haeff, "Space-Charge Effects in Electron Beams," *Proc. I.R.E.*, Vol. 27, pp. 586–602 (1939).

6 J. R. Pierce, "Limiting Stable Current in Electron Beams in the Presence of Ions," *Jour. App. Phys.*, Vol. 15, pp. 721–726 (1944); and "Note on Stability of Electron Flow in the Presence of Positive Ions," *Jour. App. Phys.*, Vol. 21, p. 1063, Oct. 1950.

to be approximations to those which would be obtained for more realistic but mathematically more refractory situations.

Other misinterpretations have arisen through combining non-relativistic equations of motion with Maxwell's equations and then attaching significance to terms of the order $(v/c)^2$.[7]

Finally, granting that all else is well, it is unsafe to draw conclusions from the examination of particular solutions of differential equations. In a very simple example, it is impossible to determine the gain of an amplifier tube which uses an electron stream simply by examining various "waves" which can travel on the stream. In solving a physical problem, one must not only solve the differential equation involved but he must satisfy the appropriate boundary conditions as well.

In all, such confusion as there has been concerning waves in clouds of electrons and ions seems to have arisen not through lack of mathematical ambitiousness but rather through simple errors in physical interpretation.

The following material concerns itself with some particular types of "waves" and with the importance and consequences of fitting boundary conditions. The work treats a very easy case, simplified and abstracted from a physically realizable system. The case was made so simple in order to avoid painful mathematics which might obscure the actual points to be made. The purpose is to explore this simple case thoroughly, avoiding basic misunderstandings. If it is objected that matters so simple should not be treated at such length, because no one could misunderstand them anyway, I can only reply that I did misunderstand some of the matters recounted herein.

I. Why Are Waves Introduced?

We will consider the case of a narrow or thin beam of electrons across which we can assume that the electric field is constant.† In our calculations we assume that all electrons in a given very small region have the same velocity, thus neglecting the thermal velocity distribution.‡ We assume that the flow is a smoothed-out jelly of charge,‡ with the charge per unit mass characteristic of electrons; thus, we neglect individual interactions between electrons, and consider only a sort of average effect.

We will write the quantities involved in the following forms

$$\text{velocity} = v + u_0$$

[7] L. R. Walker, "Note on Wave Amplification by Interaction with a Stream of Electrons," *Phys. Rev.*, Vol. 76, pp. 1721–1722 (1949).

† This is in itself a drastic abstraction. No attempt will be made to justify it here, beyond saying that it is useful in considering the problems that follow.

‡ Other drastic approximations for which no justification will be given.

Here u_0 is a constant component and v is a small fluctuating or a-c. component

$$\text{charge density} = \rho + \rho_0$$

where again ρ_0 is the average or d-c. component, which will of course be negative, and ρ is the a-c. component

convection current density

$$i - I_0$$

Here I_0 is the average or a c. current density and, as the electrons are assumed to move in the $+z$ direction, the current density in the $+z$ direction is taken as $-I_0$. In other work I have used i and I_0 as current rather than as current density; I hope that this will cause no confusion.

It is assumed that there is no average field. It is assumed that there is an a-c. field in the z direction only, and this is called E.

We have two equations to work with. One is

$$\frac{d(v + u_0)}{dt} = -\frac{e}{m} E$$

Here e/m, the charge-to-mass ratio of the electron, is taken as a positive quantity. The time derivative is that moving with an electron. We can instead take derivatives at a fixed point

$$\frac{d(v + u_0)}{dt} = \frac{\partial(v + u_0)}{\partial t} + (v + u_0)\frac{\partial(v + u_0)}{\partial z}$$

which gives

$$\begin{aligned} &\frac{\partial v}{\partial t} + u_0 \frac{\partial v}{\partial z} \\ &+ \frac{\partial u_0}{\partial t} + (v + u_0)\frac{\partial u_0}{\partial z} \\ &+ v\frac{\partial v}{\partial z} = -\frac{e}{m} E \end{aligned} \tag{1.1}$$

The terms on the second line are zero because $\partial u_0/\partial t = 0$, $\partial u_0/\partial z = 0$. Further, let us consider a series of solutions of (1.1) for fields in which E has the same form in time and space, but varies in magnitude. As E is made smaller and smaller, v will become smaller and smaller, and the term $v\partial v/\partial z$, which is a product of two a-c. quantities, will become relatively smaller than the other two terms involving v. In our small signal theory we neglect the term $v\partial v/\partial z$, and write

$$\frac{\partial v}{\partial t} + u_0 \frac{\partial v}{\partial z} = -\frac{e}{m} E \tag{1.2}$$

We note, then, that this approximates the true equation for small values of E and v only.

We have another equation

$$\frac{\partial}{\partial z}(i - I_0) = -\frac{\partial}{\partial t}(\rho + \rho_0) \tag{1.3}$$

This is the equation of continuity, or of conservation of charge. If we integrate it over a small distance Δz we obtain

$$(i - I_0)_{z+\Delta z} - (i - I_0)_z = -\frac{\partial}{\partial t}[(\rho + \rho_0)\Delta z]$$

The quantity $(\rho + \rho_0)\Delta z$ is the charge per unit cross section in the distance Δz. Thus, the right-hand side is the rate at which charge in the distance Δz decreases. The quantity on the left is obviously the rate at which charge per unit cross section is flowing out of the space Δz long.

If we carry out the operations in (1.3) we obtain

$$\frac{\partial i}{\partial z} - \frac{\partial I_0}{\partial z} = -\frac{\partial \rho}{\partial t} - \frac{\partial \rho_0}{\partial t}$$

As $\partial I_0/\partial z = 0$, $\partial \rho_0/\partial t = 0$

$$\frac{\partial i}{\partial z} = -\frac{\partial \rho}{\partial t} \tag{1.4}$$

We need to add that the convection current is given by

$$\begin{aligned} i - I_0 &= (\rho + \rho_0)(v + u_0) \\ i - I_0 &= \rho_0 v + \rho u_0 + \rho_0 u_0 + \rho v \end{aligned} \tag{1.5}$$

The term $\rho_0 u_0$ is a constant term and is to be identified with $-I_0$

$$-I_0 = \rho_0 u_0 \tag{1.6}$$

The term ρv is a product of a-c. quantities. Suppose we solve all our equations neglecting ρv. Then, the error caused by this approximation will be less as ρ and v are less, that is, at small signal levels. Thus, we write

$$i = \rho u_0 + v\rho_0 \tag{1.7}$$

We now have three approximate equations, which are good approximations at small signal levels

$$\frac{\partial v}{\partial t} + u_0 \frac{\partial v}{\partial z} = -\frac{e}{m} E \tag{1.2}$$

$$\frac{\partial i}{\partial z} = -\frac{\partial \rho}{\partial t} \tag{1.4}$$

$$i = \rho u_0 + v\rho_0 \tag{1.7}$$

We can eliminate ρ and v from these equations and obtain an equation relating i and E. To do this we solve (1.7) for v

$$v = \frac{1}{\rho_0} i - \frac{u_0}{\rho_0} \rho$$

differentiate

$$\frac{\partial v}{\partial t} = \frac{1}{\rho_0} \frac{\partial i}{\partial t} - \frac{u_0}{\rho_0} \frac{\partial \rho}{\partial t}$$

use (1.4)

$$\frac{\partial v}{\partial t} = \frac{1}{\rho_0} \frac{\partial i}{\partial t} + \frac{u_0}{\rho_0} \frac{\partial i}{\partial z}$$

differentiate (1.2) with respect to t,

$$\frac{\partial}{\partial t}\left(\frac{\partial v}{\partial t}\right) + u_0 \frac{\partial}{\partial z}\left(\frac{\partial v}{\partial t}\right) = -\frac{e}{m} \frac{\partial E}{\partial t}$$

and substitute for $\partial v/\partial t$, obtaining

$$\frac{\partial^2 i}{\partial t^2} + 2u_0 \frac{\partial^2 i}{\partial z \partial t} + u_0^2 \frac{\partial^2 i}{\partial z^2} = -\frac{e}{m} \rho_0 \frac{\partial E}{\partial t} \tag{1.8}$$

This is an equation relating i and its derivatives with E. It is a linear equation; that is, i and its derivatives, and E appear to the first power only. This is because we have neglected non-linear terms, saying that *at low levels* they are small compared with the linear terms.

Now, the electron flow interacts with surroundings of some sort, or, we shall say, with a circuit. Let us consider as an example of a circuit a transmission line with a distributed capacitance C per unit length and a distributed inductance L per unit length, which will transmit a slow wave. Suppose that the electron stream flows along very close to the line. Then if the current σi of the electron stream, where σ is the area of electron flow, changes with distance, a current J will flow into the line per unit length as shown in Fig. 1.1 where

$$J = -\sigma \frac{\partial i}{\partial z} \tag{1.9}$$

If V is the voltage on the line and I is the current in the line we write

$$\frac{\partial I}{\partial z} = -C \frac{\partial V}{\partial t} - \sigma \frac{\partial i}{\partial z} \tag{1.10}$$

$$\frac{\partial V}{\partial z} = -L \frac{\partial I}{\partial t} \tag{1.11}$$

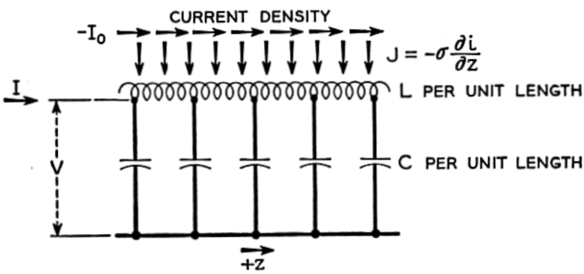

Fig. 1.1—A transmission line with an electron beam very close to it.

We can eliminate I by differentiating

$$\frac{\partial^2 I}{\partial z \partial t} = -C \frac{\partial^2 V}{\partial t^2} - \sigma \frac{\partial^2 i}{\partial z \partial t}$$

$$\frac{\partial^2 V}{\partial z^2} = -L \frac{\partial^2 I}{\partial z \partial t} \tag{1.12}$$

$$\sigma \frac{\sigma^2 i}{\partial z \partial t} = \frac{1}{L} \frac{\partial^2 V}{\partial z^2} - C \frac{\partial^2 V}{\partial t^2}$$

We can further identify the field acting on the electrons as

$$E = -\frac{\partial V}{\partial z} \tag{1.13}$$

In (1.8), let us replace E by means of (1.13), and let us differentiate with respect to z and again with respect to t. We obtain

$$\frac{\partial^2}{\partial t^2}\left(\frac{\partial^2 i}{\partial z \partial t}\right) + 2u_0 \frac{\partial^2}{\partial z \partial t}\left(\frac{\partial^2 i}{\partial z \partial t}\right) + u_0^2 \frac{\partial^2}{\partial z^2}\left(\frac{\partial^2 i}{\partial z \partial t}\right) = \frac{e}{m} \rho_0 \frac{\partial^4 V}{\partial z^2 \partial t^2}$$

We can substitute for $\partial^2 i/\partial z \partial t$ from (1.12) and obtain

$$
\begin{aligned}
u_0^2 \frac{\partial^4 V}{\partial z^4} + 2u_0 \frac{\partial^4 V}{\partial z^3 \partial t} + \left(1 - u_0^2 LC - \sigma \frac{e}{m} L\rho_0\right) \frac{\partial^4 V}{\partial z^2 \partial t^2} \\
- 2u_0 LC \frac{\partial^4 V}{\partial z \partial t^3} - LC \frac{\partial^4 V}{\partial t^4} = 0
\end{aligned}
\tag{1.14}
$$

Thus, we have obtained a linear partial differential equation in V, z and t.

So far, nothing has been said about waves or wavelike behavior. We might solve (1.14) for any boundary conditions on V and its derivatives that we chose, by any means, as by using a differential analyzer or a digital computer. There is, however, a well-established technique for dealing with linear partial differential equations with constant coefficients, such as (1.14) is. It is known that they have solutions of the form

$$V = Ae^{j\omega t}e^{-j\beta z} \tag{1.15}$$

As (1.14) is an entirely real equation, if (1.15) is a solution, the real part of (1.15) is also a solution, i.e.,

$$\text{Re}\,(Ae^{j\omega t}e^{-j\beta z})$$

is a solution. Hence, we may regard the real part of the complex V as the true physical solution.

If we substitute (1.15) into (1.14) we obtain

$$
\begin{aligned}
u_0^2 \beta^4 - 2u_0 \omega \beta^3 + \left(\frac{1}{L} - u_0^2 LC - \sigma \frac{e}{m} L\beta\right) \omega^2 \beta^2 \\
+ 2u_0 LC\omega^3 \beta - LC\omega^4 = 0
\end{aligned}
\tag{1.16}
$$

Now (1.16) is an algebraic equation in ω and β. How are we to interpret it?

Suppose we are interested in devices driven from sinusoidal generators, such as amplifiers.* This means that ω is real, and that it is the radian frequency of the applied signal. We may then regard (1.16) as an equation in β, and, as it is a fourth degree equation, there will in general be four roots. We may regard these as pertaining to four waves, whose voltages vary as

$$V_1 = A_1 e^{j(\omega t - \beta_1 z)}$$

$$V_2 = A_2 e^{j(\omega t - \beta_2 z)}$$

$$V_3 = A_3 e^{j(\omega t - \beta_3 z)}$$

$$V_4 = A_4 e^{j(\omega t - \beta_4 z)}$$

* We might, on the other hand, be interested in devices with an imposed spatial pattern, as in a magnetron oscillator. In this case we might assume β as a given, real quantity and solve for real or complex values of ω.

Each of these four components is a solution of the *differential equation*. The solution of an actual *physical problem* will be the sum of the four components, or, if we like, the real part of that sum, and the amplitude factors $A_1 - A_4$, which are in general complex, will depend on the particular physical problem which is solved.

What has been the purpose of this argument? First of all, it is intended to indicate how the waves get into the picture. The differential equations for a long beam of constant average velocity u_0 and charge density ρ_0 were linearized by neglecting terms in which the products of a-c. quantities appeared. By this means a linear partial differential equation with constant coefficients which relates i and E was found. This was combined with the linear partial differential equation for a uniform transmission-line circuit, and an overall partial differential equation for V was obtained, linear and with constant coefficients. Such an equation could be solved by any means, but it is known to have wave-type solutions, and the solution of the original physical problem must be a sum of all such solutions.

In general, we will not expect so simple a relation between i and V or E as (1.12), that for a simple transmission line. Further, for broad electron streams the electronic behavior cannot be expressed so simply as it has been in (1.8). Nonetheless, we will find wave solutions in which all quantities vary with time and distance as

$$e^{j\omega t}e^{-j\beta z}$$

as long as

(1) the d-c. beam properties (the undisturbed electron flow) and the circuit properties do not vary with z.

(2) the signal amplitude is low enough so that terms involving products of a-c. quantities can be neglected.

When this is so, the solution of a physical problem can be expressed as the sum, or the real part of the sum, of such wave solutions, taken with the proper amplitudes.†

II. The Component Waves

Once we are convinced that the solution of our problem can be expressed as the sum of a number of waves which are solutions of a linear partial differential equation, it is simplest to use this fact directly in finding certain properties of the waves of which the solution is to be made up.

Let us, for instance, let E in (1.8) contain the factor

$$e^{j\omega t}\,e^{-j\beta z}$$

† An additional overall condition is that the electron flow has no velocity distribution.

In other words, let E in (1.8) be one of the wave components of a solution. Then (1.8) becomes

$$(-\omega^2 + 2u_0\omega\beta - u_0^2\beta^2)i = -j\omega \frac{e}{m} \rho_0 E$$

$$i = \frac{-j\omega\epsilon\left(-\frac{e}{m}\frac{\rho_0}{\epsilon}\right)}{u_0^2\left(\frac{\omega}{u_0} - \beta\right)^2} E \tag{2.1}$$

Here ϵ, the dielectric constant of vacuum, has been introduced for reasons which will become apparent later. It is further of interest to introduce other simple parameters.

$$-\frac{e}{m}\frac{\rho_0}{\epsilon} = \omega_p^2 \tag{2.2}$$

$$\frac{\omega_p}{u_0} = \beta_p \tag{2.3}$$

$$\frac{\omega}{u_0} = \beta_0 \tag{2.4}$$

The quantity ω_p is called the *plasma frequency* (a radian frequency). ω_p^2 is positive because ρ_0 is negative. β_0 would be the phase constant of a wave traveling with the electron velocity. While β_p would be the phase constant of a wave traveling with a phase velocity equal to the electron velocity, and having a frequency ω_p, we may merely regard β_p as a convenient parameter which increases as the beam current is increased. In terms of β_p and β_0

$$i = \frac{-\beta_p^2}{(\beta_0 - \beta)^2}(j\omega\epsilon E) \tag{2.5}$$

This may seem a strange form in which to write the equation. It will perhaps seem less strange, however, if we recall that the current density I in a dielectric medium is given by

$$I = j\omega\epsilon E$$

Thus, we see that for real values of β the electron convection current density i is that which would correspond to a negative dielectric constant or a negative capacitance. Its magnitude depends on β_p^2, which is proportional to the d-c. beam current density; and the magnitude becomes very large when the phase velocity of the wave approaches the velocity of the electrons, that is when β approaches β_0.

Suppose we consider a beam of area σ. We can write the total electron convection current I_e in the form

$$I_e = \sigma_i = Y_e E \tag{2.6}$$

$$Y_e = \frac{-j\omega\epsilon\sigma\beta_p^2}{(\beta_0 - \beta)^2} \tag{2.7}$$

We will call Y_e the electronic admittance; it is measured in mho meters.

Later we will deal with waves in which the electron stream transfers power to the circuit, and it is interesting to see under what conditions this can take place. Let the amplitude of the wave under consideration vary with distance as

$$e^{(\alpha_1 - j\beta_1)z}$$

We may take the complex nature of the propagation constant into account by substituting in (2.7)

$$\begin{aligned} -j\beta &= \alpha_1 - j\beta_1 \\ \beta &= j\alpha_1 + \beta_1 \end{aligned} \tag{2.8}$$

This leads to

$$\begin{aligned} Y_e &= \frac{-j\omega\epsilon\sigma\beta_p^2}{(\beta_0 - \beta_1 - j\alpha_1)^2} \\ Y_e &= \frac{\omega\epsilon\sigma\beta_p^2[2\alpha_1(\beta_0 - \beta_1) - j(\beta_0 - \beta_1)^2]}{[(\beta_0 - \beta_1)^2 + \alpha_1^2]} \end{aligned} \tag{2.9}$$

The electron stream can transfer energy to the circuit only if the real part of Y_e is negative (a negative conductance). For a wave which increases in the direction of electron flow (the $+z$ direction), α_1 is positive and the electronic conductance will be negative if $\beta_1 > \beta_0$; that is, if the electron velocity is greater than the phase velocity of the wave.[8]

For a wave which decreases in the $+z$ direction, the conductance will be negative if the electron velocity is smaller than the phase velocity of the wave.

Let us now consider the interaction of our thin electron stream with the circuit. Here there is some possibility of confusion. In (1.12) the field caused by impressing a current on a circuit was calculated. This may be likened to the voltage along an impedance Z caused by an impressed current I.

[8] This is indicated by very elementary arguments (J. R. Pierce and L. M. Field, "Traveling Wave Tubes," *Proc. I.R.E.*, Vol. 35, pp. 108–111, Feb. 1947). It is easy to forget, however, and was recently pointed out to me, to my consternation, by Dr. L. J. Chu.

Figure 2.1 will help to make this clear. Here the impressed current I flows to the right and back through the circuit of impedance Z. The voltage will increase to the right and hence the field will be directed to the left.

In general, for an impressed current I we will write the field produced as

$$E = -Z(\omega, \beta)I \tag{2.10}$$

Here $Z(\omega, \beta)$ is a circuit impedance per unit length, which is usually a function of ω and β. In terms of an admittance, the relation connecting impressed current and field is

$$I = -Y(\omega, \beta)E \tag{2.11}$$

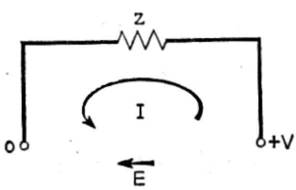

Fig. 2.1—The voltage and field produced by a current impressed on an impedance Z.

This can also be made clearer by means of an illustration. Suppose that the impressed current density in a very broad beam is i and the "circuit" is merely free space. Then

$$\rho = \frac{\beta}{\omega} i$$

and from Poisson's equation

$$\frac{\partial E}{\partial z} = -j\beta E = \frac{\rho}{\epsilon} = \frac{\beta}{\omega}\frac{i}{\epsilon}$$

$$i = -j\omega\epsilon E$$

But, the admittance of a unit cube is just $j\omega\epsilon$, and the current through this admittance is $j\omega\epsilon E$.

Thus, when we have calculated the field caused by an impressed convection current, the admittance is the negative of the field divided by the convection current.

In (2.5), i, or rather, σi, where σ is the area of the beam, may be regarded as the impressed current. If $Y(\omega, \beta)$ is the circuit admittance, one way of writing the condition for a natural mode of propagation of stream and circuit is

$$\sigma i = -Y(\omega, \beta)E \tag{2.12}$$

Another way of putting this is to say

$$Y_e + Y(\omega, \beta) = 0 \tag{2.13}$$

From (2.13) and (2.7) we obtain

$$\beta = \beta_0 \pm \beta_p \sqrt{\frac{j\omega\epsilon\sigma}{Y(\omega, \beta)}} \tag{2.14}$$

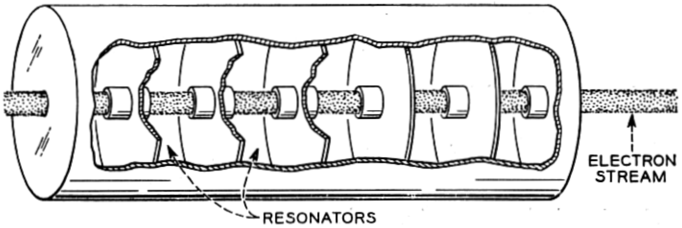

Fig. 2.2—An electron stream passing through a series of resonators, as in a multiresonator klystron.

Suppose, for instance, that the circuit admittance is capacitive and is equal to that for a longitudinal electric field in vacuum of area equal to the beam area σ. Then

$$Y(\omega, \beta) = j\omega\epsilon\sigma \text{ mho meter}$$

and we have two unattenuated waves

$$\beta = \beta_0 \pm \beta_p$$

We see that whenever (1) the circuit admittance is inductive or (2) the circuit admittance has a dissipative component, β will be complex, and there will be increasing and decreasing waves. Either of these conditions can be achieved, for instance, by surrounding the electron stream by a succession of essentially uncoupled resonators, tuned to be inductive, or with dissipation, as shown in Fig. 2.2. This is merely a continuous multi-resonator klystron.

In a transmission-line type of circuit such as we have considered and such as is used in the traveling-wave tube, for instance, the circuit admittance depends strongly on the phase constant β, and in solving (2.14) for β we must take cognizance of this fact.

We can, for instance, derive the circuit admittance from (1.12). We can use

$$E = j\beta V$$

and rewrite (1.12) as

$$\sigma i = \frac{-j}{\omega L \beta^2} (\omega^2 LC - \beta^2) E$$

Now, if the impressed current σi is zero, β must have a value β_1 such that

$$\beta_1 = +\sqrt{\omega^2 LC}$$

Also, the characteristic impedance of the line, K, is

$$K = +\sqrt{L/C}$$

In terms of these quantities

$$\sigma i = \frac{-j}{K \beta_1 \beta^2} (\beta_1^2 - \beta^2) E$$

and the circuit admittance $Y(\omega, \beta)$ is

$$Y(\omega, \beta) = -\frac{\sigma i}{E}$$

$$Y(\omega, \beta) = \frac{j}{K \beta_1 \beta^2} (\beta_1^2 - \beta^2) \quad \text{mho meter} \tag{2.15}$$

Here K and β_1 are positive quantities. We note that this admittance is capacitive for $\beta < \beta_1$, that is, for waves with a phase velocity greater than the natural phase velocity of the circuit, and inductive for $\beta > \beta_1$, that is for waves with a phase velocity less than the natural phase velocity of the circuit. This is easily explained. For small values of β the wavelength of the impressed current is long, so that current flows into and out of the circuit at widely separated points. Between such points the long section of series inductance has a higher impedance than the shunt capacitance to ground; the capacitive effect predominates and the circuit impedance is capacitive. However, for large values of β current flows into and out of the circuit at points close together. The short section of series inductance between such points provides a path of lower impedance than that through the capacitances and ground; the inductive impedance predominates and the circuit is inductive. Thus, for fast waves (β small) the circuit is capacitive and for slow waves (β large) the circuit is inductive.

We can, then, immediately make one observation. For a lossless circuit, any increasing or decreasing wave must have a phase velocity less than the natural phase velocity of the circuit.

We can make another observation as well; if the circuit has loss, $Y(\omega, \beta)$ will have a real component, and from (2.14) all the waves must have an imaginary component of β, that is, they must be increasing or decreasing.

If we like, we can combine (2.15) with (2.14). Doing this directly, we obtain

$$(\beta - \beta_0)^2 = \omega\epsilon\sigma K\beta_1\beta^2 \frac{\beta_p^2}{(\beta_1^2 - \beta^2)} \tag{2.16}$$

Unless the electron velocity is near the wave velocity (β_0 near to β_1) we will expect two sorts of solutions: one sort, for which β is near to β_0 corresponding to "space-charge" waves; and the other, for which β is near to $\pm\beta_1$, corresponding to "circuit" waves. If β_0 is not near to β_1, we can easily obtain approximate values of β for these two types of wave.

To obtain β for the space-charge waves we put $\beta = \beta_0$ on the right-hand side of (2.16) and obtain

$$\beta = \beta_0 \pm \beta_p \sqrt{\frac{\omega\epsilon\sigma K\beta_1\beta_0^2}{\beta_1^2 - \beta_0^2}} \tag{2.17}$$

If (2.17) gives a value of β differing by a small fraction from β_0, then (2.17) is to be trusted.

To obtain β for the forward circuit wave we put $\beta = \beta_1$ on the left of (2.16) and in the numerator on the right. This gives for the forward wave

$$\beta = \beta_1\left(1 - \frac{\omega\epsilon\sigma K\beta_p^2\beta_1}{(\beta_1 - \beta_0)^2}\right)^{1/2} \tag{2.18}$$

To obtain the backward wave, we put $\beta = \beta_1$ on the left of (2.16) and in the numerator on the right, and obtain

$$\beta = -\beta_1\left(1 + \frac{\omega\epsilon\sigma K\beta_p^2\beta}{(\beta_1 - \beta_0)^2}\right)^{1/2} \tag{2.19}$$

Again, (2.18) and (2.19) are to be trusted as long as β as given by (2.18) differs by a small fraction only from β_1.

We see that according to (2.19) the space-charge waves are unattenuated (real β) for $\beta_0 < \beta_1$, that is, for electrons traveling faster than the circuit phase velocity, while there are increasing and decreasing waves for $\beta_0 > \beta_1$, that is, for electrons traveling more slowly than the circuit phase velocity. We see from (2.18) and (2.19) that the circuit waves are unattenuated (for lossless circuits), and travel a little more slowly than in the absence of electrons.

Further, we see that (2.17) and (2.18) are not to be trusted when β_0 is close to β_1, that is, when the electron velocity is near to the circuit phase velocity. As a simple example, let

$$\beta_0 = \beta_1 \tag{2.20}$$

It will turn out that β will be very nearly equal to β_0. Hence,

$$\beta = \beta_0 + j\delta - j\beta = -j\beta_0 + \delta \tag{2.21}$$

Then, from (2.16) we have

$$j\delta = \pm\beta_p \sqrt{\frac{\omega\epsilon\sigma K\beta_0(\beta_0 + j\delta)}{\delta(-2j\beta_0 + \delta)}}$$

If we neglect δ with respect to β_0 in the sums inside the radical we obtain the equation

$$\delta^3 = -j\beta_p^2\beta_1\omega\epsilon\sigma K \tag{2.22}$$

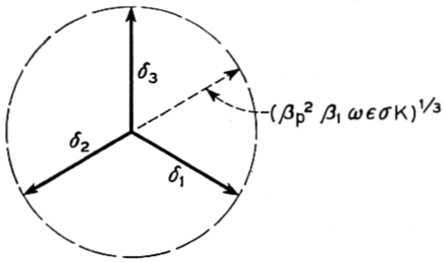

Fig. 2.3—Values of δ for the three forward waves of a traveling-wave tube when the electron velocity is equal to the velocity of the undisturbed wave.

This yields the usual three forward waves of the traveling-wave tube.

$$\delta = (\beta^2\beta_1^2\omega\epsilon\sigma K)^{1/3}e^{j(-\pi/2+2n\pi)/3}$$

$$\delta_1 = (\beta^2\beta_1^2\omega\epsilon\sigma K)^{1/3}(\sqrt{3}/2 - j/2)$$

$$\delta_2 = (\beta^2\beta_1^2\omega\epsilon\sigma K)^{1/3}(-\sqrt{3}/2 - j/2)$$

$$\delta_3 = (\beta^2\beta_1^2\omega\epsilon\sigma K)^{1/3}(j)$$

We see that δ_1 represents an increasing wave slower than the natural phase velocity of the circuit, δ_2 represents a decreasing wave slower than the natural phase velocity of the circuit, and δ_3 represents an unattenuated wave faster than the natural phase velocity of the circuit. The 3 δ's are illustrated in Fig. 2.3.

If $\beta_0 \neq \beta_1$, and if β_1 is complex (a lossy circuit) the equation for δ is more complicated, but δ can be obtained numerically.

In addition to the three forward waves, that is, waves in the direction of electron motion, there is a backward wave. This is very much out of synchronism with the electron stream, and the backward wave is essentially the same as the wave in the absence of electron flow.

III. Fitting Boundary Conditions; Gain

So far the discussion has been concerned with a differential equation and wave-type solutions of it. Let us now consider an overall problem. Suppose that we inject an unmodulated electron stream into a circuit of some finite length and apply a signal to the end of the circuit nearest the source of electrons. Suppose that we adjust the output termination so that there is no backward wave.* How will the field strength vary along the circuit? To answer this question, we must find out what combination in phase and amplitude of the three forward waves corresponds to these conditions. In terms of solving differential equations, we must fit the boundary conditions.

From Section I we have

$$\frac{\partial v}{\partial t} + u_0 \frac{\partial v}{\partial z} = -\frac{e}{m} E \tag{1.2}$$

or

$$v = \frac{j \dfrac{e}{m}}{u_0(\beta_0 - \beta)} E \tag{3.1}$$

with which we couple

$$i = \frac{-\beta_p^2}{(\beta_0 - \beta_p)^2} j\omega\epsilon E \tag{2.5}$$

In terms of

$$\beta = \beta_0 + j\delta$$

these relations become

$$\left(\frac{1}{u_0} \frac{e}{m}\right) v = \frac{1}{\delta} E \tag{3.2}$$

$$\left(\frac{-j}{\omega\epsilon\beta_p^2}\right) i = \frac{1}{\delta^2} E \tag{3.3}$$

These relations hold for each of the waves separately. Now, let us denote by E_1, E_2, E_3 the fields of the three waves, and by E the actual field on the circuit. Then at the beginning of the circuit, where E is E_0, the applied field, the amplitudes E_{10}, E_{20}, E_{30} of E_1, E_2 and E_3 must satisfy

$$E_{10} + E_{20} + E_{30} = E_0 \tag{3.4}$$

* This is a very special case, requiring a unique impedance terminating the +z end of the output circuit. See Section V.

Also, at the beginning of the helix, a-c. velocity and the a-c. convection current must be zero. This means that

$$\frac{E_{10}}{\delta_1} + \frac{E_{20}}{\delta_2} + \frac{E_{30}}{\delta_3} = 0 \tag{3.5}$$

$$\frac{E_{10}}{\delta_1^2} + \frac{E_{20}}{\delta_2^2} + \frac{E_{30}}{\delta_3^2} = 0 \tag{3.6}$$

For the case we have considered, $\beta_1 = \beta_0$, β_1 real,

$$\delta_2 = \delta_1 e^{-j(2\pi/3)}$$

$$\delta_3 = \delta_1 e^{+j(2\pi/3)}$$

and our equations become

$$E_{10} + E_{20} + E_{30} = E_0$$

$$E_{10} + E_{20} e^{j\frac{1}{3}(2\pi)} + E_{30}^{-j\frac{1}{3}(2\pi)} = 0$$

$$E_{10} + E_{20} e^{j\frac{2}{3}(2\pi)} + E_{30}^{-j\frac{2}{3}(2\pi)} = 0$$

We easily see that the solution is

$$E_{10} = E_{20} = E_{30} = \tfrac{1}{3} E_0 \tag{3.7}$$

If E is the field at a distance z along the helix

$$E = \tfrac{1}{3} E_0 e^{-j\beta_0 z}(e^{\delta_1 z} + e^{\delta_2 z} + e^{\delta_3 z}) \tag{3.8}$$

In Fig. 3.1,

$$20 \log_{10} \left| \frac{E}{E_0} \right|$$

is plotted vs CN, a factor proportional to distance.

We see that initially the amplitude does not change. This is necessarily so. The strength of the field can grow only through the electron stream giving energy to the circuit. The electron stream can give energy to the circuit only if it has an a-c. convection current. Initially the electron stream is unmodulated and hence it can give energy to the circuit only after it has traveled far enough to become modulated.

In the case we have considered, the amplitudes of the three wave components of the field are initially equal. Now, E_1 increases with distance, while E_2 decreases with distance and E_3 is unattenuated. Hence, if the tube is long enough, E_2 and E_3 will be negligible near the output of the tube; and the field at the output, a distance ℓ from the input, will be very nearly

$$E = \tfrac{1}{3} E_0 e^{-j\beta_0 \ell} e^{\delta_1 \ell}$$

Under these circumstances the gain G in db will be very nearly

$$G = 20 \log_{10} \left| \frac{E}{E_0} \right| = 20 \log_{10} \tfrac{1}{3} e^{\mathrm{Re}(\delta \ell)} \; db$$

$$G = -9.54 + \frac{20}{2.3} \frac{\sqrt{3}}{2} (\beta^2 \beta_1^2 \omega \epsilon \sigma K)^{1/3} \ell$$

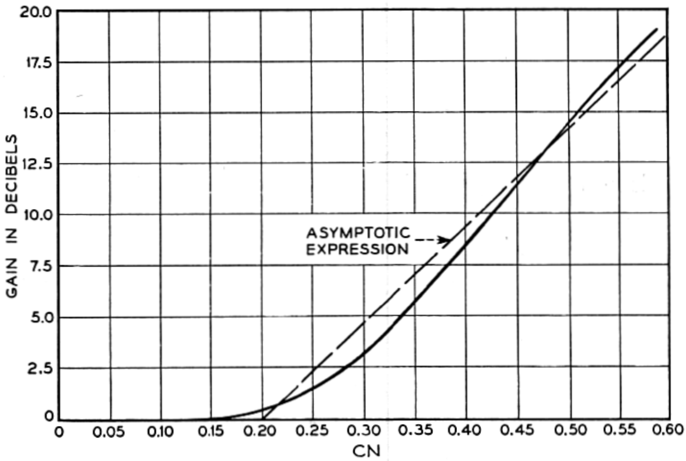

Fig. 3.1—Signal level along the helix of a traveling-wave tube.

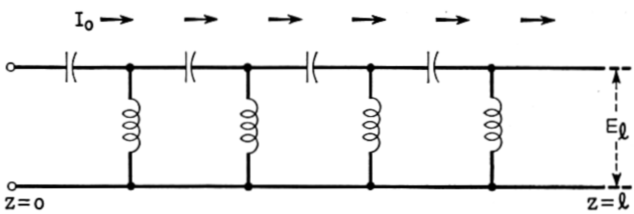

Fig. 4.1—A high-pass structure in which the phase velocity is in a direction opposite to that of power flow.

IV. Backward Waves and Other Peculiar Waves

It is important to notice that, for the usual traveling-wave tube, it is possible to express the overall gain in terms of the increasing wave alone only because of the relative amplitudes of the three waves which make up the solution of the particular problem considered. That this is by no means a trivial point can be demonstrated by considering a case in which the circuit is a high-pass filter, as shown in Fig. 4.1. For such a circuit, the phase constant β_1 is negative for a wave excited at the left end of the line which carries energy to the right. Such a wave will not interact with electrons moving to the right. A wave excited at the right with power flow to the left has a positive value of β and will interact with electrons traveling to the right. Let us consider such an interaction.

First, as to the δ's. We see that, for a wave which varies with distance as

$$e^{-j\beta z}$$

where β is a positive number and has power flow to the left, the sign of V/I must be the opposite of what it would be if the power flowed to the right. This can be taken into account by reversing the sign of K in (2.19), and making

$$\delta^3 = +j\beta_p^2\beta_1\omega\epsilon\sigma K \tag{4.1}$$

where K is now taken as a positive number. We can then take

$$\delta_1 = (\beta_p^2\beta_1^2\,\omega\epsilon\sigma K)^{1/3}\left(-\frac{\sqrt{3}}{2} + j/2\right)$$

$$\delta_2 = (\beta_p^2\beta_1^2\,\omega\epsilon\sigma K)^{1/3}\left(\frac{\sqrt{3}}{2} + j/2\right)$$

$$\delta_3 = (\beta_p^2\beta_1^2\,\omega\epsilon\sigma K)^{1/3}(-j)$$

Here δ_1 represents a wave whose amplitude increases to the left, that is, a wave which grows in the direction of energy flow. We might think that this would immediately imply a gain similar to that obtained for energy flow in the direction of electron motion, but this would be jumping at conclusions.

Suppose we taken $z = 0$ at the left-hand or output end of the circuit. There the electron stream enters unmodulated. There also we will assume the circuit to be terminated so as to prevent reflection of power. At the right-hand or input end of the circuit power will be fed in, giving an impressed field E_ℓ.

Suppose δ_1, δ_2 and δ_3 are the appropriate δ's for this case. We see that our boundary conditions are

$$e^{-j\beta_0\ell}(E_{10}e^{j\delta_1\ell} + E_{20}e^{j\delta_2\ell} + E_{30}e^{-j\delta_3\ell}) = E_\ell$$

$$\frac{E_{10}}{\delta_1} + \frac{E_{20}}{\delta_2} + \frac{E_{30}}{\delta_3} = 0$$

$$\frac{E_{10}}{\delta_1^2} + \frac{E_{20}}{\delta_2^2} + \frac{E_{30}}{\delta_3^2} = 0$$

We have a relation between the δ's

$$\delta_1 = \frac{\sqrt{3}}{2} + j\tfrac{1}{2}$$

$$\delta_2 = \delta_1 e^{j(2\pi/3)}$$

$$\delta_3 = \delta_1 e^{-j(2\pi/3)}$$

From this we easily we see that a solution of the last two equations is

$$E_{10} = E_{20} = E_{30}$$

Accordingly, the first equation becomes

$$E_{10} e^{-j\beta_0 \ell}(e^{j\delta_1 \ell} + e^{j\delta_2 \ell} + e^{j\delta_3 \ell}) = E_\ell$$

$$E_{10} = \frac{E_\ell e^{j\beta_0 \ell}}{(e^{j\delta_1 \ell} + e^{j\delta_2 \ell} + e^{j\delta_3 \ell})} \tag{4.2}$$

Let us now assume that the tube is very long. We easily see that in this case

$$| e^{j\delta_1 \ell} | \gg | e^{j\delta_2 \ell} |$$

$$| e^{j\delta_1 \ell} | \gg | e^{j\delta_3 \ell} |$$

So very nearly

$$E_{10} = E_{20} = E_{30} = \frac{E_\ell e^{j\beta_0 \ell}}{e^{j\delta_1 \ell}} \tag{4.2}$$

and the total field at the output end of the tube is

$$E = E_{10} + E_{20} + E_{30} = 3E_\ell e^{j\beta_0 \ell} e^{-j\delta_1 \ell} \tag{4.4}$$

This, however, is much smaller than the field E_ℓ at the input end of the tube.

What is the physical picture? The electrons are injected into the circuit as an unmodulated stream. In order to fit the boundary conditions at this point, the three waves must have comparable magnitudes at the point of injection. If this is the output, then any wave which "grows" from input toward output must be relatively very small at the input.

If boundary conditions are fitted for other cases, as, for an electron speed not equal to the circuit phase velocity ($\beta_0 \neq \beta_1$), it may be found that the output may be a little greater than the input under some circumstances; this represents a small gain achieved through a spatial interference of the three wave components.

A sure way of distinguishing conditions which will allow amplification from conditions which will not is through a solution of the differential equations together with a fitting of the boundary conditions. In the case of backward waves there are, however, considerations concerning the source of energy and its transfer to the circuit (or field) which are useful.

Suppose that an unmodulated electron stream enters a microwave amplifier, travels for some distance through it, and emerges. If the electromagnetic output power of the amplifier is greater than the input power, the additional power must have come from the kinetic energy of the electron stream. The average electron must leave the amplifier with less energy of motion than it had on entering it.

We can say a little more. Let us call the total velocity of electrons, a-c. and d-c., u. Then we have, corresponding to (1.1)

$$\begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial z} &= -\frac{e}{m} E \\ \frac{\partial u}{\partial t} + \frac{\partial}{\partial z}(u^2) &= -\frac{e}{m} E \end{aligned} \tag{4.5}$$

We will consider an amplifier in which the u and E at any z-position are truly periodic. Let us integrate over the period of a cycle, τ, and divide by τ

$$\frac{1}{\tau} \Big|_{t}^{t+\tau} u + \frac{1}{\tau} \frac{\partial}{\partial z} \int_{t}^{t+\tau} u^2 \, dt = \frac{1}{\tau} \int_{t}^{t+\tau} E \tag{4.6}$$

As u will be the same at t and $t + \tau$, the first term on the left is zero, and we have

$$\frac{\partial}{\partial z} \overline{u^2} = \overline{E} \tag{4.7}$$

Here $\overline{u^2}$ and $\overline{E}$ are time averages.

The field E is produced in a linear circuit by (1) the application of an a-c. signal, (2) by the presence of the electron stream. Certainly, the applied signal can produce no average field in a linear circuit. Further, unless electrons are turned back, the average electron convection current is independent of r-f level. In a linear circuit the average field must be proportional to the average impressed current, so the average field $\overline{E}$ must be zero or independent of r-f level. Thus, the *time* average of $\overline{u^2}$ at a given point must be independent of r-f level.*

This means that the electron stream cannot be slowed down bodily by

* L. A. MacColl pointed this out to the writer.

the r-f field. Energy is extracted from the stream only by a bunching process in which in the emerging beam the charge density is higher when the velocity is below average than it is when the velocity is above average. In other words, the kinetic energy averaged over electrons is reduced, even though the *time* average of u^2 is not changed. This means that the emerging beam must be strongly bunched if much power is to be abstracted.

In the conventional traveling-wave tube all is well. At the input the r-f field is small and the beam is unbunched. At the output the r-f field is high, and the beam is strongly bunched, having lost energy to the circuit.

Imagine a tube using a backward wave, however. The electrons are injected unbunched at the output, where the signal level is high. They emerge at the input where the signal level is low. If the tube is to give high power, the stream must emerge strongly bunched. The disturbance in the electron stream cannot gradually increase as the field amplitude increases.

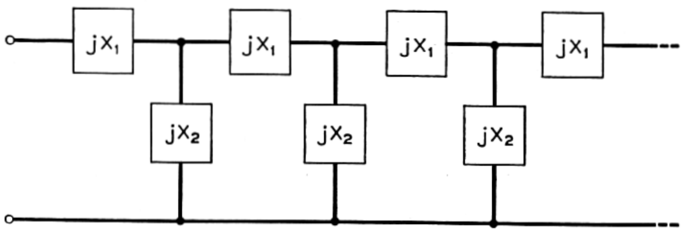

Fig. 4.2—A ladder network.

We have seen that one cannot draw conclusions about gain just by looking at the propagation constants of the waves. Waves are merely solutions of a differential equation connected with a physical system. To find the properties of the system one must examine, not various solutions of the differential equation, but the particular solution (which may be a combination of simple solutions) which applies to the system in question.

As a further example, we will examine another system whose differential equations yield "growing" solutions which turn out to be backward waves. Consider the ladder network of Fig. 4.2. This propagates an unattenuated wave if X_1 and X_2 have opposite signs, (X_1 inductive and X_2 capacitive, for instance). If, however, X_1 and X_2 are both capacitive or both inductive, then a wave excited in the circuit decays exponentially with distance. If we speak in terms of β_1, then

$$-j\beta_1 = -\alpha_1$$

$$\beta_1 = -j\alpha_1$$

where α_1 is a real number.

The characteristic impedance K is reactive, inductive if both impedances are inductive and capacitive if both impedances are capacitive. Let

$$K = jX_0$$

where X_0 is a real number. Then a positive value of X_0 means inductive elements, and a negative value capacitive elements.

The space-charge waves are given by (2.17)

$$\beta = \beta_0 \pm \beta_p \sqrt{\frac{-\omega\epsilon\sigma X_0 \alpha_1 \beta_0^2}{\alpha_1^2 + \beta_0^2}} \tag{4.8}$$

We see that the waves are unattenuated for negative, capacitive values of X_0, and are increasing and decreasing for positive, inductive values of X_0. It can be shown that the increasing space-charge waves can be used to obtain gain.

The forward circuit waves are given by using $K = jX_0$, $\beta_1 = -j\alpha_1$ in (2.16), $\beta = -j\alpha_1$ on the left and in the numerator on the right and $\beta = -j\alpha$ in the denominator on the right.

$$\alpha = \alpha_1 \left(1 + \frac{\omega\epsilon\sigma X_0 \alpha_1 \beta_p^2}{(j\alpha_1 + \beta_0)^2}\right)^{1/2} \tag{4.9}$$

As $\alpha = j\beta$, the variation with distance is as

$$e^{-\alpha z}$$

The backward wave is given by using $\beta = +j\alpha_1$ on the left of (2.16) and in the numerator on the right

$$\alpha = -\alpha_1 \left(1 - \frac{\omega\epsilon\sigma X_0 \alpha_1 \beta_p^2}{(j\alpha_1 - \beta_0)^2}\right)^{1/2} \tag{4.9}$$

If α differs little from $\pm\alpha_1$, we can expand the square root in (4.6) and (4.7), separate real and imaginary parts, and write:

Forward wave:

$$\alpha = \alpha_1 \left(1 + \frac{\omega\epsilon\sigma\beta_p^2 X_0 \alpha_1(\beta_0^2 - \alpha_1^2)}{2(\beta_0^2 + \alpha_1^2)} - \frac{j\omega\epsilon\sigma\beta_p^2 \beta_0 X_0 \alpha_1^2}{(\beta_0^2 + \alpha_1^2)^2}\right) \tag{4.10}$$

Backward wave:

$$\alpha = -\alpha_1 \left(1 + \frac{\omega\epsilon\sigma\beta_p^2 X_0 \alpha_1(\beta_0^2 - \alpha_1^2)}{2(\beta_0^2 + \alpha_1^2)} + \frac{j\omega\epsilon\sigma\beta_p^2 X_0 \alpha_1^2}{(\beta_0^2 + \alpha_1^2)}\right) \tag{4.11}$$

The circuit "waves" which were rapidly attenuated in the absence of electrons ($\beta_p = 0$) are a little more or less rapidly attenuated in the presence of electrons (more or less depending on whether X_0 is positive or negative, and on the relative magnitudes of β_0 and α_1), and they now have a phase constant, that is, an imaginary component of the propagation constant.

The phase velocity may be either positive or negative, depending on the sign of X_0. This added feature gives the solution a more "wavelike" quality, but physically we have merely a slight perturbation of the disturbance natural to the non-propagating ladder network.

In the absence of electrons, there is no real power flow in the modes of propagation of a purely reactive ladder network in which the shunt and series reactances have the same sign. Such a network can of course transmit power to a resistive load, but it transmits no power when terminated in its (reactive) characteristic impedance.

In the presence of electrons, there is a small power flow in the circuit. We can easily evaluate this. If, in (1.11), we assume a variation of the quantities with time and distance as

$$e^{-\alpha z}e^{j\omega t}$$

we obtain

$$I = \frac{\alpha}{j\omega L} V$$

Here ωL stands for the series reactance, which we may call X_1

$$\omega L = X_1$$

A positive value of X_1 means series inductance. For non-propagating ladders, X_1 and the characteristic reactance X_0 have the same sign.

We then have

$$I = -j \frac{\alpha}{X_1} V$$

The quantity $-j\alpha/X_1$ as evaluated in the presence of electrons will be the "hot" characteristic admittance.

The complex power flow P is

$$P = VI^*$$

So, in this case

$$P = \frac{j\alpha^*}{X_1} VV^*$$

Now, the "backward" wave, for which α is given by (4.12), "increases" in the direction of electron flow. For it, the real part of the power Re P is given by

$$\text{Re}\, P/VV^* = -\frac{\omega\epsilon\sigma\beta_p^2\alpha_1^3 X_0}{(\beta_0^2 + \alpha_1^2)X_1} \tag{4.12}$$

Note that X_0 and X_1 must have the same sign. Thus, the power flow for the wave which "increases" in the direction of electron flow is always in the direction opposite to the electron flow. The circuit power does not flow in the direction of increasing amplitude for the wave which "grows" in the direction of electron flow. We might have deduced this from the fact that the phase velocity for the wave is greater than the electron velocity (see (2.9)).

While the wave which increases in the direction contrary to electron flow has its power flow in the direction of increasing amplitude, it is a backward wave and hence not suitable for producing gain.

The disturbance on the non-propagating ladder is closely related to a passive or cut-off mode of a waveguide excited at a frequency less than the cutoff frequency for the mode in question. In this case, the analogue of the circuit power VI^* is the integral of the Poynting vector over the guide cross section. When electrons flow through a waveguide these cut-off modes are perturbed much as indicated by (4.19) and (4.12). Because the perturbed modes have a "wavelike" character in that the propagation constant is no longer purely real, and because the amplitude may increase in the direction of electromagnetic power flow, some workers have proposed to obtain gain from these "growing waves."[2]

V. Further Considerations Concerning Boundary Conditions

How necessary is it to fit boundary conditions in order to deduce what will happen? The suspect waves we have examined so far might be rejected as increasing in a direction contrary to the direction of electron flow,* or as having electromagnetic power flow in a direction opposite to the direction of growth. Can we find some method for separating waves useful in producing gain from waves which are not, without consideration, explicit or implicit, of boundary conditions?

Let us consider the problem of fitting boundary conditions for a circuit plus an electron stream. Imagine that the end of the circuit near to the electron source ("near" end) is connected to a load impedance Z and that the end away from the electron source ("far" end) is driven by a voltage V. Let the wave which increases most rapidly in the direction of electron flow vary in amplitude as exp (αz). Suppose that the length of the circuit is great, so that αL is a large number and exp (αL) is a very large number.

At the near end the various wave components must be so related that i and v are zero and that the circuit voltage is ZI. At the near end, all four waves must be used to fit the boundary conditions at a given voltage level. Disallowing very special values of Z, we would expect that at the near end the four waves will have comparable amplitudes (the amplitudes are related by linear simultaneous equations). Thus, at the far end of the circuit, the wave which increases most rapidly with distance should strongly predominate. It seems that the most rapidly increasing wave is naturally connected with excitation of the circuit at the far end.

* Though rejected only through considering boundary conditions.

On the other hand, assume that the far end of the circuit is terminated in some impedance Z. Consider the case in which the electron stream is velocity modulated at the source end and no exciting voltage is applied at the far end. We would expect that the required boundary conditions at the source end could be satisfied by using the waves excepting the one which increases most rapidly with distance. At the far end, to make $V = IZ$ it is necessary to add a component of the wave which increases most rapidly with distance, a component of magnitude comparable to the sum of other components present *at the far end*. However, this added component is so small at the near end that there it can be disregarded. Thus, the manifestation of large forward gain comes not from the mere presence of a wave which increases in the forward direction, but from special properties of the waves and/or the terminating impedances which can be determined with certainty only by fitting boundary conditions.

Are not these arguments at variance with the usual analyses of operation of the traveling-wave tube? Suppose, for instance, that the helix is terminated in an arbitrary impedance at the input (near) end and that a voltage V is applied at the output (far) end. What wave will predominate? For a lossless helix, the true answer is that the increasing (forward) wave, not the unattenuated backward wave, will predominate. This can be avoided only by (1) choosing a particular (matched) value of source impedance or (2) making the helix lossy enough so that the backward wave "increases" more rapidly in the $+z$ direction than any forward wave does. In tubes with a uniform loss along the helix, expedient (2) is adopted; when a center lossy section is used, both (1) and (2) are invoked, (1) in the output section and (2) in the center lossy section.

It is dangerous to consider the solutions of the linear differential equations of a physical system singly rather than in the combination which satisfies the boundary conditions. This sort of reasoning might lead one to believe that the problem of obtaining high voltages can be solved by finding a solution of Laplace's equation (say $V = 1/r$) for which the potential goes to infinity at some point.

Cautions against neglecting the problem of boundary conditions apply equally well to problems of instability (increase of disturbances with time) as to problems of amplification. Thus, electron flow may be unstable when none of the waves grows with time for real values of β[6]. On the other hand, in criticizing the work of Bohm and Gross,[1] R. Q. Twiss has shown[9] that electron flow is not necessarily unstable merely because some of the waves grow exponentially with time for real values of β.

[9] R. Q. Twiss, "On the Theory of Plasma Oscillations" Services Electronics Research Laboratory, Extracts from Quarterly Report No. 20, Oct. 1950, pp. 14–28.