ON THE OPTIMIZATION OF THE DRIFT LENGTHS OF STAGGER-TUNED
MULTI-CAVITY KLYSTRON AMPLIFIERS FOR SMALL SIGNALS

by

FRED GÜNTHER SCHRACK
B.A.Sc., University of British Columbia, 1958

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

In the Department of
Electrical Engineering

We accept this thesis as conforming to the
standards required from candidates for the
degree of Master of Applied Science

Members of the Department
of Electrical Engineering

The University of British Columbia
JUNE 1960
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical Engineering

The University of British Columbia,
Vancouver 8, Canada.

Date June 21, 1960
The drift length of a two-cavity klystron amplifier for highest power gain is one-quarter of the reduced space-charge wavelength $\lambda_q$. The presently accepted values for these optimum drift lengths for a multi-cavity klystron amplifier are also one-quarter of $\lambda_q$. These lengths are obtained from an extrapolation of the two-cavity klystron by regarding the multi-cavity klystron as a cascade amplifier consisting of several two-cavity stages.

In this thesis, the multi-cavity klystron amplifier is considered as an entity, and not as a cascade amplifier. It is shown that if the power-gain function of a flat-staggered design is optimized with respect to the drift lengths, a set of unequal lengths is obtained. The optimization of the drift lengths leads to an appreciable increase in gain with a negligible increase in overall tube length. Hence it is worthwhile not to regard the multi-cavity klystron amplifier as a cascade amplifier.

Only the small-signal theory is considered in this thesis.
# TABLE OF CONTENTS

List of Illustrations .......................................................... v
List of Tables ........................................................................ vi
Acknowledgement .................................................................... vii

1. Introduction ....................................................................... 1

2. The Two-Cavity Klystron Amplifier ..................................... 4
   2.1 First-Order Theory of Gap Action .................................. 7
   2.2 First-Order Bunching Theory (Ballistic Theory) .......... 9
      2.2 a) Bunched Beam Current ........................................ 11
      2.2 b) Optimum Drift Length ...................................... 14
   2.3 Second-Order Bunching Theory (Space-Charge Debunching) .... 14
      2.3 a) Webster's Debunching Theory ............................ 15
      2.3 b) Optimum Drift Length ...................................... 16
      2.3 c) The Hahn and Ramo Theory .............................. 17
      2.3 d) Optimum Drift Length ...................................... 19
   2.4 The Transadmittance .................................................... 19

3. The Multi-Cavity Klystron Amplifier .................................. 24
   3.1 The Signal-Flow Diagram ........................................... 24
   3.2 The Admittance of a Resonator ................................... 26
   3.3 The Transadmittance .................................................. 27
   3.4 The Voltage-Amplification Function ............................ 27
   3.5 The Power-Gain Function .......................................... 29
   3.6 Stagger-Tuning ......................................................... 32

4. The Maximization of the Power Gain .................................. 38
   4.1 Results Reported in the Literature .............................. 38
4.2 The Objective of this Thesis and the Proposed Solution. 40
4.3 The Power Gain as a Function of the Drift Angles. 41
4.4 The Method of Partial Derivatives. 42
4.5 The Gradient Method. 45

5. Numerical Computations and Results 49
5.1 The Choice of the Design Parameters 49
5.2 Calculations 50
5.3 Results 53

6. Conclusion 66

Appendix A: Harmonic Analysis of the Bunched Beam Current 68
Appendix B: Webster's Debunching Theory 70
Appendix C: The Hahn Theory 74
Appendix D: A Necessary Transformation for the Approximation Condition 85
Appendix E: The Gradient Method 88

References 95
## LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Applegate Diagram</td>
<td>6</td>
</tr>
<tr>
<td>2.</td>
<td>Notation for a Resonator Gap</td>
<td>7</td>
</tr>
<tr>
<td>3.</td>
<td>Bunched Beam Current at the Catcher of a Two-Cavity Klystron</td>
<td>13</td>
</tr>
<tr>
<td>4.</td>
<td>Equivalent Circuit of a Two-Cavity Klystron</td>
<td>20</td>
</tr>
<tr>
<td>5.</td>
<td>Signal-Flow Diagrams for Two-, Three-, and Four- Cavity Klystrons</td>
<td>24</td>
</tr>
<tr>
<td>6.</td>
<td>Equivalent Circuits of Input and Output Cavities</td>
<td>30</td>
</tr>
<tr>
<td>7.</td>
<td>Functions Pertinent to the Approximation Condition</td>
<td>36</td>
</tr>
<tr>
<td>8.</td>
<td>Increases in Gain due to Optimization</td>
<td>56</td>
</tr>
<tr>
<td>9.</td>
<td>Variation of Optimum Drift Angles with an Increase of Bandwidth, Tuning Pattern 2143</td>
<td>57</td>
</tr>
<tr>
<td>10.</td>
<td>Variation of Optimum Drift Angles with an Increase of Bandwidth, Tuning Pattern 1423</td>
<td>58</td>
</tr>
<tr>
<td>11.</td>
<td>Gain-Frequency Characteristics for the 2% Bandwidth Designs</td>
<td>59</td>
</tr>
<tr>
<td>12.</td>
<td>Gain-Frequency Characteristics for the 4% Bandwidth Designs</td>
<td>60</td>
</tr>
<tr>
<td>13.</td>
<td>Gain-Frequency Characteristics for the 6% Bandwidth Designs</td>
<td>61</td>
</tr>
<tr>
<td>14.</td>
<td>Gain-Frequency Characteristics for the 8% Bandwidth Designs</td>
<td>62</td>
</tr>
<tr>
<td>15.</td>
<td>Gain-Frequency Characteristics for the 10% Bandwidth Designs</td>
<td>63</td>
</tr>
<tr>
<td>16.</td>
<td>Gain-Frequency Characteristics for the Optimized Designs, Tuning Pattern 2143</td>
<td>64</td>
</tr>
<tr>
<td>C.1.</td>
<td>Cross-Section of Drift Tube</td>
<td>79</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Summary of Results of the Computer Study of the Modified EMI VX 5063 Tube</td>
<td>52</td>
</tr>
<tr>
<td>II. $Q'$s and Normalized Frequencies for the Stagger-Tuning of the Optimized Designs</td>
<td>65</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

The author would like to thank the supervisor of this project, Dr. A.D. Moore, for his help during the course of this research.

He would also like to acknowledge gratefully the assistance given by Mr. J.H.R. Dempster and other staff members of the Computing Centre of this University.

The cooperation of the author’s colleague, Mr. John Tsong Yuan, was most helpful.

This research was carried out under a Defence Research Board Grant (DRB 5503-04) granted to Dr. A.D. Moore with additional support from a National Research Council Grant (BT-68) granted to Dr. F. Noakes.
ON THE OPTIMIZATION OF THE DRIFT LENGTHS OF STAGGER-TUNED
MULTI-CAVITY KLYSTRON AMPLIFIERS FOR SMALL SIGNALS

1. INTRODUCTION

Multi-cavity klystron amplifiers are of particular interest in the microwave field because of the combination of two characteristic features which are not present in some other tubes. These are the klystron amplifiers' high output power capability and their adaptability for stagger-tuning, so that these amplifiers can be used as high-gain broad-band devices.

Theories of the klystron were advanced soon after their invention and two facts were recognized early 1,2,3:

a) in order to avoid non-linear equations, which are difficult to solve analytically, it is necessary to assume that only small signal amplitudes are present in the tube;

b) there exists a value of the drift length for which certain characteristics, such as the efficiency, the gain, or the gain-bandwidth product, will be optimized.

This thesis is concerned with the problem of maximizing the power gain at the band centre of a stagger-tuned multi-cavity klystron amplifier by a variation of its drift lengths using the small-signal theory.

The interest in small-signal theories is justified by the following facts:
a) A knowledge of the small-signal theory is necessary to give an understanding of the operation of the klystron;

b) The driving stages of many large-signal tubes, as well as the driving stages of some hybrid tubes, often work under small-signal conditions;

c) The results obtained from the small-signal theory can be used to predict large-signal behaviour.

With regard to the drift length, the generally accepted value for an optimum drift length for a two-cavity klystron is one-quarter of a wavelength at a certain characteristic frequency, the reduced plasma frequency, and this value is extrapolated to the multi-cavity klystron. In doing so, the multi-cavity klystron amplifier is regarded as a cascade amplifier consisting of several stages made up of two-cavity amplifiers. The situation, however, is not quite so simple, since there exists a forward coupling of signals between non-adjacent cavities in addition to the straight cascading of the adjacent ones. The multi-cavity klystron amplifier therefore cannot be compared with a conventional cascade amplifier. In fact, in this thesis it will be shown that it is necessary to regard the multi-cavity klystron amplifier as an entity. It is not adequate to extrapolate from the two-cavity klystron to the multi-cavity klystron to obtain optimum drift lengths.

In this thesis, the theories regarding the optimization of the drift length of a two-cavity klystron amplifier will be reviewed first, after which the power-gain
function of a multi-cavity amplifier will be derived. Taking this gain function as the quantity to be maximized, formulas will be developed suitable for a numerical evaluation yielding optimized drift lengths for a multi-cavity klystron.

Numerical results will be presented for a specific four-cavity klystron amplifier.

This thesis is concerned with the theoretical development of the equations leading to the optimized drift lengths and with a numerical solution of these equations for a variety of design data. No attempt has been made to verify these results experimentally in view of the fact that it has been shown previously that a theoretical analysis is capable of predicting the actual behaviour of the amplifier quite accurately.
2. THE TWO-CAVITY KLYSTRON AMPLIFIER

In a klystron amplifier, an electron beam emerging from an electron gun in a vacuum is shot down a metallic tube, called the drift tube. On its way, the beam encounters gaps in the drift tube, the gaps being the openings of re-entrant resonator cavities. The input signal is fed into the first cavity, the buncher, and sets up an alternating electric field at the gap. This field changes the velocity of the electrons, either by retarding or accelerating them, a process named velocity modulation. Beyond the gap, in the drift tube, if one can neglect fringing fields and space-charge effects, there are no fields to interact with the electrons. In this region, the electrons behave ballistically, with the faster electrons eventually overtaking the slower ones. In this way, bunches of electrons are formed. The gap of the next cavity is situated a certain distance, the drift length, from the buncher. The bunches of electrons passing it will excite this cavity and ac power can be extracted by coupling a load to the cavity.

In a multi-cavity klystron amplifier, rather than delivering power to a load, the second cavity will velocity-modulate the beam just as the first cavity did. Thus, the bunching is enhanced in the second drift length and on encountering more gaps this process is continued. As before, in the final cavity, the catcher, the ac power is
extracted by coupling the cavity to a load. Since the tube is converting dc power into ac power at the frequency of the input signal, the klystron acts as an amplifier.

A convenient way of representing the bunching process in the two-cavity klystron is a distance-time diagram called the Applegate Diagram (see Figure 1). In this diagram, the path of those electrons which have the same velocity is represented by a line. For zero buncher voltage, i.e. at $V_1 = 0$, the slope of the line corresponds to the velocity of the electrons emerging from the gun (called the dc velocity). For $V_1 > 0$, the electrons are accelerated, their velocity is greater than the dc velocity, and hence the slope is greater; similarly, for $V_1 < 0$, the electrons are retarded, their velocity is lower than the dc velocity, and hence the slope is smaller. After a certain time, unmodulated electrons will have overtaken slow electrons which passed the buncher during the previous quarter cycle; at the same instant, fast electrons which passed the buncher during the following quarter cycle will have overtaken the unmodulated electrons. This is the bunching action; it can be seen clearly in the Applegate Diagram. At the point of maximum bunching, the velocity modulation has been converted completely to density modulation. Beyond crossover, the density modulation is gradually reconverted to velocity modulation.
Figure 1. The Applegate Diagram
2.1 First-Order Theory of Gap Action

The axial field due to a gridded gap such as that shown in Figure 2 can be expressed mathematically as follows:

\[ E_z = \frac{V_1 \sin \omega t}{a} \quad \text{for } a << r, \]

where

- \( E_z \)\(^*\) the axial electric field intensity,
- \( V_1 \sin \omega t \) the input signal voltage at the buncher,
- \( a \) the gap spacing,
- and \( \omega \triangleq 2\pi f \)

\( \triangleq \) the signal frequency in radians.

* "\( \triangleq \)" denotes "equals by definition"
The equation of motion \( F_z = m \ddot{z} \) for an electron of mass \( m \) and charge \(-e\) is

\[
m \ddot{z} = -eE_z \tag{2-1}
\]

or

\[
\ddot{z} = \left(-\frac{\eta V_1}{a}\right) \sin\omega t / a
\]

where \( \eta = e/m \)

is the charge-to-mass ratio of an electron.

Integration gives

\[
\dot{z} = \left(-\frac{\eta V_1}{a}\right) \cos\omega t / a + C
\]

The dc velocity of an electron is determined by the accelerating voltage of the electron gun, \( V_o \), and hence is known:

\[
u_o = \left(2\eta V_o\right)^{\frac{1}{2}}
\]

where \( u_o \) is the velocity of an electron which has not been velocity-modulated. Such an electron will be called a dc electron.

Let \( t_0 \triangleq \) the time at which the electron passes the first grid of the gap;

then the integration constant is given by

\[
C = u_o + \left(\frac{\eta V_1 \cos\omega t_o}{a}\right) / a
\]

and so

\[
\dot{z} = u_o + \left(\frac{\eta V_1}{a}\right) \left(\cos\omega t_o - \cos\omega t\right) \tag{2-2}
\]

The ratio of the ac input amplitude to the beam accelerating voltage \( V_1/V_o \) is called the depth of modulation. For the reasons already cited, it will be assumed that only small signal amplitudes are present. The depth of modulation, then, is much smaller than unity, i.e.,

\[
V_1 << V_o
\]

and hence

\[
u_1 \approx u_o
\]
so that \( t_1 \approx t_0 + a/u_0 \)

where \( t_1 \) is the time at which the electron passes the second grid of the gap.

Let \( u_1 \) be the velocity at the second grid;

then Equation (2-2) becomes

\[
\frac{\eta V_1}{u_0 \phi_g} \cos \omega t_0 = \cos(\omega t_0 + \phi_g)
\]

where \( \phi_g \approx \omega a/u_0 \)

is the transit angle of the gap.

The transit angle is a convenient expression since it represents a normalized length. Using the trigonometric identity for the difference of two cosines,

\[
u_1 = u_0 + \frac{\eta V_1}{u_0 \phi_g} \sin \frac{1}{2} \phi_g \sin(\omega t_0 + \frac{1}{2} \phi_g)
\]

Therefore, \( u_1 = u_0 \left[1 + \frac{MV}{V_o} \sin (\omega t_0 + \frac{1}{2} \phi_g)\right]\)

The ratio \( M \triangleq \frac{\sin \frac{1}{2} \phi_g}{\frac{1}{2} \phi_g} \) is called the gap modulation coefficient or the beam coupling coefficient, it is dependent on the geometry of the gap. For non-gridded gaps, this coefficient becomes a ratio of Bessel functions of the gap transit angle and the normalized beam diameter. The coefficient takes into account the fact that the transit time in the gap is an appreciable portion of the cycle of gap voltage.

2.2 First Order Bunching Theory (Ballistic Theory)

The analysis of the electron motion in a drift length begins with the equation of motion in a way similar to that of the theory of the gap action. It is assumed that the gap length is negligible compared to the drift length,
but the finite gap transit time is accounted for by the beam coupling coefficient.

The energy acquired by an electron in passing the buncher gap is

\[ \frac{1}{2} m u_1^2 = e V_{\text{total}} \]

\[ = e (V_0 + M V_1 \sin \omega t_1') \]

where \( t_1' \triangleq \) the time at which the electron passed the centre of the buncher gap.

Therefore, \( u_1 = u_o \left( 1 + \frac{M V_1}{V_0} \sin \omega t_1' \right)^{\frac{1}{2}} \) \( ... (2-3) \)

This can be expanded binomially to give

\[ u_1 = u_0 \left[ 1 + \frac{1}{2} \frac{M V_1}{V_0} \sin \omega t_1' - \frac{1}{8} \left( \frac{M V_1}{V_o} \right)^2 \sin^2 \omega t_1' + ... \right] \]

Under the condition that \( \frac{M V_1}{2 V_0} \ll 1 \), this is approximated as

\[ u_1 \approx u_0 \left( 1 + \frac{1}{2} \frac{M V_1}{V_0} \sin \omega t_1' \right) \] \( ... (2-4) \)

The time of travel in the drift tube is given by \( d/u_1 \), i.e.,

\[ t_2 = t_1' + \frac{d}{u_1} \] \( ... (2-5) \)

where \( t_2 \triangleq \) the time at which the electron passed the centre of the catcher gap,

\( d \triangleq \) the drift length between the centres of the two gaps.

Substitution of \( u_1 \) from Equation (2-3) into (2-5) gives:

\[ t_2 = t_1' + \frac{d}{u_0} \left( 1 + \frac{M V_1}{V_0} \sin \omega t_1' \right)^{-\frac{1}{2}} \]

Let \( t_0 \triangleq \frac{d}{u_0} \)

\( \triangleq \) the time the dc electron takes to travel along the drift tube from the centre of the buncher to the centre of the catcher.
Binomial expansion of $t_2$ now yields

$$t_2 = t_1 + t_0 \left[ 1 - \frac{1}{2} \frac{MV_1}{V_0} \sin \omega t_1 + \frac{3}{8} \left( \frac{MV_1}{V_0} \right)^2 \sin 2\omega t_1 + \ldots \right]$$

Again assuming the small-signal condition, i.e., that $\frac{MV_1}{2V_0} << 1$, this is approximately

$$t_2 \approx t_1 + t_0 \left( 1 - \frac{1}{2} \frac{MV_1}{V_0} \sin \omega t_1 \right) \ldots (2-6)$$

Define the following drift angles:

$$\phi_0 \triangleq \omega t_0$$

$$\phi_1 \triangleq \omega t_1 \approx \omega t_1 \text{ if the gap transit time is neglected,}$$

and $$\phi_2 \triangleq \omega t_2$$

Multiplication of every term of (2-6) by $\omega$ yields

$$\phi_2 = \phi_1 + \phi_0 - X \sin \phi_1 \ldots (2-7)$$

where $$X \triangleq \frac{1}{2} \frac{MV_1}{V_0} \phi_0$$

The parameter $X$ occurs so often in klystron theory that it has acquired a specific designation, and is named the bunching parameter.

2.2 a) Bunched Beam Current

The average beam current $I_0$ represents the average number of electrons per unit time. By the principle of the conservation of charge, a given group of electrons
passing the buncher gap during a time interval $dt_1$ will take a time interval $dt_2$ to pass the catcher gap, i.e.,

$$i_2 dt_2 = i_1 dt_1$$

But at the input gap, $i_1 = I_0$, so

$$i_2 = I_0 \left(\frac{dt_1}{dt_2}\right) \quad \ldots(2-8)$$

where $i_2$ is the total beam current at the second gap.

In order to get $dt_1/dt_2$, differentiate Equation (2-6) with respect to $t$ and form the differential relationship

$$dt_2 = dt_1 - X \cos \omega t_1 dt_1$$

Substitution into Equation (2-8) yields

$$i_2 = I_0 / (1 - X \cos \phi) \quad \ldots(2-9)$$

A simultaneous solution of (2-7) and (2-9) will give the catcher current $i_2$ in terms of the arrival time of the catcher $t_2$. A graph of the resulting function (see Figure 3) shows the bunched character of $i_2$ and it also shows that $i_2$ is rich in harmonics.

A harmonic analysis of $i_2$ (see Appendix A, Equation (A-3)) yields an expression for the $n$-th harmonic component of $i_2$:

$$n^2 I_2 = 2 I_0 J_n(nX)$$

where $J_n(nX) = \text{the Bessel function of the first kind of order } n$.

Thus the fundamental component of $i_2$ is

$$I_2 = 2 I_0 J_1(X) \quad \ldots(2-10)$$

which is maximum for $X = 1.84$, when $J_1(1.84) = 1.16$. 
Figure 3. Bunched Beam Current at the Catcher of a Two-Cavity Klystron
This analysis is valid if the origin of time is chosen to be the instant at which the bunch centre passes the centre of the catcher gap. With this understanding, \( n I_2 \) may be treated as a phasor.

2.2 b) Optimum Drift Length

As can be seen from Figure 3, crossover of the paths of electrons occurs when the bunching parameter \( X \) equals unity, yet a higher current in the second cavity is obtained somewhat beyond this value when \( X = 1.84 \).

Since

\[
X = \frac{1}{2} MV_1 \phi_0 / V_0
\]

and

\[
\phi_0 = \omega d / u_0
\]

the optimum drift length is

\[
d_{\text{opt}} = 1.84 \left(2V_0 / MV_1 \right) \frac{u_0}{\omega}
\]

Here, then, the question of an optimum drift length arises for the first time. But, as will be seen later, there are other approaches also yielding values for the "optimum" drift length.

2.3 Second-Order Bunching Theory (Space-Charge Debunching)

Experimental evidence shows that the ballistic theory is not adequate. This is to be expected since a plasma, i.e., an electron cloud, is being considered, and so far no account has been taken of the effects of the mutual repulsion of the electrons. This phenomenon is known as space-charge debunching. There are two basic kinds of space-charge debunching:
i) **Transverse Debunching.** This is the radial dispersion of the electrons; if excessive it causes an appreciable loss of electrons to the walls. The metallic tunnel surrounding the beam is, of course, enhancing this effect since positive image charges are formed on it. In order to avoid this loss of beam current, most practical tubes are immersed in a strong dc longitudinal magnetic field, thus focussing the beam. The assumption that this is the case will be made and the effect of transverse debunching will therefore be ignored.

ii) **Longitudinal Debunching.** This is the dispersion of electrons parallel to the axis of the beam. The longitudinal debunching in effect prevents the formation of sharply defined electron bunches. Since this effect acts along the direction of motion, it will have to be taken into account.

### 2.3 a) Webster's Debunching Theory

Shortly after the invention of the klystron by the Heil and Varian Brothers, Webster gave a theory taking into account the longitudinal debunching. Hahn and Ram extended this theory soon afterwards.

Webster makes the following assumptions:

i) The bunch considered is a plane parallel sheet of electrons extending to infinity. Practically, this means that the drift-tube wall is distant from the beam, i.e., the tube diameter is much larger than that of the bunches.
ii) The theory is restricted to small signals.

iii) The transit time of any gap can be considered to be negligible.

Webster's theory modifies Equation (2-7) as follows (see Appendix B, Equation (B-5)):

$$\phi_2 = \phi_1 + \phi_o - X(\frac{\sin \phi_d}{\phi_d}) \sin \phi_1 \quad \ldots(2-11)$$

where $h$ is a parameter called the debunching wave number, given by $h \triangleq \left(I_o \eta / \epsilon_o A u_o^3\right)^{\frac{1}{2}}$

where $\epsilon_o$ is the permittivity of free space, and $A$ is the cross-sectional area of the beam.

The form of (2-11) suggests the definition of a modified bunching parameter $X'$:

$$X' \triangleq (\frac{\sin \phi_d}{\phi_d}) X$$

Then the fundamental component of the current in the catcher becomes simply

$$I_2 = 2 I_o J_1(X') \quad \ldots(2-12)$$

Thus, the space-charge effects are taken into account by the factor $(\sin \phi_d)/(\phi_d)$.

The Webster theory has two main drawbacks. Firstly, it takes into account neither fringing nor wall effects, and secondly, it is valid only up to, but not beyond, the crossover $\cdots$.

2.3 b) Optimum Drift Length

Previously, an optimum value for the drift length was found by letting the bunching parameter $X$ be equal to that value which makes $J_1(X)$ in Equation (2-10) a maximum ($X = \ldots$)
1.84). Since \( X \propto d \), this determined a value \( d_{\text{opt}} \).

Considering now the modified bunching parameter \( X' \), it is found that

\[
X' \propto \sin \theta d
\]

because

\[
X' = \left( \frac{MV_1 \omega}{2V_0 u_o h} \right) \sin \theta d
\]

For small signals it may happen that \((MV_1 \omega/2V_0 u_o h)\) is less than 1.84, so that it will be impossible to reach the first maximum of \( J_1(X') \). Therefore, for small signals the optimum value for \( X' \) is obtained by setting

\[
\sin \theta d = 1
\]

or

\[
\theta d = \frac{1}{2} \pi
\]

and so

\[
d_{\text{opt}} = \frac{\pi}{2h}
\]

Experimental evidence shows that this \( d_{\text{opt}} \) gives an overestimate of the space-charge debunching effect.6

2.3 c) The Hahn and Ramo Theory2,3

Hahn's theory starts out with Maxwell's equations and therefore takes into account automatically both fringing and wall effects for a beam and drift tube of finite diameter. The theory (see Appendix C) shows that there can exist an infinite number of modes of propagation of space-charge waves along the beam. This complicates matters considerably, although it is not impossible to carry out the mode summation, as has been shown by Feenberg.11

Ramo suggested considering the first mode. The accuracy
of this procedure has been experimentally verified recently\textsuperscript{12}.

The Hahn and Ramo theory shows that the debunching wave number is closely associated with the natural frequency of oscillation of the plasma. In fact,

$$h = \frac{\omega_p}{u_0}$$  \hspace{1cm} (2-14)

where

$$\omega_p \triangleq \left( \frac{\gamma I_0}{\varepsilon_0 u_0 A} \right)^{1/2}$$

and is called the \textit{plasma frequency}.

This theory furthermore shows that the plasma frequency is reduced by the effects of fringing and by the effects of the image charges at the walls. Thus, all expressions derived by the ballistic theory are still valid, provided that wherever \(\omega_p\) appears in an expression, it is to be replaced by the effective or \textit{reduced plasma frequency} \(\omega_q\).

The relationship between the two frequencies is given by:

$$F_n \omega_p = (\omega_q)_n$$

where

\(F_n \triangleq \text{the space-charge reduction factor for the } n\text{-th mode.}\)

\(F_n\) is a function of the beam and the tube geometry only and has been presented in convenient graphical form by Branch and Mihran\textsuperscript{13}. For the fundamental mode, \(F_n \triangleq F\), and \((\omega_q)_n \triangleq \omega_q^*\).

The \textit{finite debunching wave number} for the fundamental mode now becomes

$$h_f \triangleq \frac{\omega_q}{u_0}$$

and the \textit{finite bunching parameter} is

$$x_f^* \triangleq \left( \frac{\sin h_f d}{h_f d} \right) x = \frac{\omega MV_1}{\omega q^* v_0} \sin h_f d$$  \hspace{1cm} (2-15)
2.3 d) Optimum Drift Length

The optimum drift length found by the Webster theory from Equations (2-13) and (2-14) is

\[ h_d = \frac{\omega_p}{u_0} = \frac{1}{2} \pi \]

Replacement of \( \omega_p \) by \( \omega_q \) according to the Hahn-Ramo theory gives \( h_d = \frac{\omega_q}{u_0} \).

The right-hand side of this equation looks so similar to the expression for a drift angle that it is convenient to define it as the drift angle in terms of the reduced plasma frequency:

\[ \Theta_o = \omega_q / u_0 \]

The optimum drift angle now becomes

\[ \Theta_{opt} = \frac{1}{2} \pi \]

or \( d_{opt} = \frac{\pi u_0}{2 \omega_q} \)

Let \( \lambda_q \triangleq \frac{u_0}{f_q} \)

\( \triangleq \) the reduced plasma wavelength,

where \( f_q \triangleq \frac{\omega_q}{2\pi} \)

\( \triangleq \) the reduced plasma frequency, in cycles per second.

Then \( d_{opt} = \frac{1}{4} \lambda_q \)

i.e., the optimum drift angle for a two-cavity klystron is one-quarter of the reduced plasma wavelength. This is the currently accepted value of \( d_{opt} \).

2.4 The Transadmittance

It is convenient for the analysis of the multi-cavity klystron amplifier to define a transadmittance (transfer-admittance) relating the voltage at one gap to the resonator current at another gap.
For a two-cavity klystron, the transadmittance $Y_{21}$ is defined as the ratio of the fundamental component of complex resonator current of the catcher cavity to the complex signal voltage at the buncher gap.

A further assumption to simplify the discussion will be made by regarding all gap geometries to be alike, so that no distinction is made between the gap coupling coefficients of cavity 1 and cavity 2.

The magnitude of the transadmittance is

$$|Y_{21}| = \left|\frac{M}{V_1}\right| I_2$$

Note the difference between the fundamental component of beam current, $I_2$, flowing from the anode to the cathode, and the fundamental component of the resonator current, $-M I_2$, as shown in Figure 4.

Figure 4. Equivalent Circuit of a Two-Cavity Klystron
As may be seen from the Applegate diagram (Figure 1), $I_2$ lags $V_1$ by the drift angle of a dc electron minus $\frac{1}{2} \pi$.

Thus,

$$Y_{21} \triangleq \left| Y_{21} \right| (-e^{-j(\varphi_{21} - \pi/2)})$$

$$= \left| \frac{M_1 I_2}{V_1} \right| (-e^{-j(\varphi_{21} - \pi/2)}) \quad \ldots (2-16)$$

where

$$\varphi_{21} \triangleq \omega d_{21}/u_o = \phi_{21} \omega/\omega_q$$

and

$$\phi_{21} \triangleq \omega_q d_{21}/u_o = h_d d_{21}$$

$\triangleq$ the dc drift angle* between cavities 1 and 2, in terms of the reduced plasma frequency.

and $d_{21} \triangleq$ the drift length between cavities 1 and 2.

Neglecting for the time being the phase factor,

Equations (2-12) and (2-15) yield

$$\left| I_2 \right| = 2 I_o \left| J_1 (X_f^i) \right|$$

Thus

$$\left| Y_{21} \right| = \left| \frac{2 M I_o}{V_1} J_1 (X_f^i) \right|$$

For small signals, $X_f^i$ is small and then the ratio

$$\frac{J_1 (X_f^i)}{X_f} \rightarrow \frac{1}{2} \quad \text{as} \quad X_f^i \rightarrow 0$$

From Equation (2-15),

$$\left| Y_{21} \right| = \left| M I_o X_f^i/V_1 \right|$$

*The designation of the drift lengths between any two cavities of a multi-cavity klystron amplifier will be made according to the notation used by Kreuchen, Auld and Dixon5 rather than that used by Mason's signal-flow graph theory14. Thus, the drift angle $\varphi_{mh}$ is the drift length between cavities $h$ and $m$, where $h < m$. 
\[ |Y_{21}| = \left| \frac{m^2}{2} \frac{I_0}{V_o} \sin \theta_{21} \right| \]

Hence, \[ Y_{21} = -m^2 \left[ \frac{I_0}{V_o} \omega \right] \sin \theta_{21} \sin \left( e^{-j(\varphi_{21} - \pi/2)} \right) \]

\[ = -(j\omega) m^2 \left( \frac{I_0}{V_o} \right) \sin \theta_{21} e^{-j\varphi_{21}} \ldots(2-17) \]

\[ = j m^2 Y_o \sin \theta_{21} e^{-j\varphi_{21}} \]

where \( Y_o \triangleq -\frac{I_0}{V_o} \omega \) and is named the characteristic beam admittance \(^{15}\).

It can be seen from Equation (2-17) that, under the assumption made, the transadmittance varies sinusoidally along the beam. This is a result which has been stated previously by several authors \(^5, 16, 17, 18, 19\).

In order to extend the discussion to the more general case of the complex frequency variable \( s = \sigma + j\omega \), replace \( j\omega \) in the expression for \( Y_{21} \), Equation (2-17), by \( s \):

\[ Y_{21}(j\omega) = -j\omega m^2 \left( \frac{I_0}{V_o} \right) \sin \theta_{21} e^{-j\omega\tau_{21}} \]

and

\[ Y_{21}(s) = -s m^2 \left( \frac{I_0}{V_o} \right) \sin \theta_{21} e^{-s\tau_{21}} \]

where \( \tau_{21} \triangleq \frac{d_{21}}{u_o} \).

For convenience define a parameter \( C_{21} \) by:

\[ sc_{21} e^{-s\tau_{21}} \triangleq Y_{21} \ldots(2-18) \]

and

\[ C_{21} = C_{\text{max}} \sin \theta_{21} \ldots(2-19) \]

where \( C_{\text{max}} \triangleq M^2 Y_o / \omega \)
Since by Equation (2-18) the admittance is proportional to s, dimensionally the coefficient of s is a capacitance and hence could be named a transfer-capacitance or transcapacitance. Note that \( C_{\text{max}} \) is a tube parameter, independent of the drift angle and signal frequency, defining the amplitude of \( C_{21} \).

To summarize the assumptions made so far:

i) The small-signal theory holds, i.e., \( \frac{M V}{2 V_0} \ll 1 \).

ii) Only the first mode of propagation is considered.

iii) The gap lengths are negligible compared to the drift lengths.

iv) All gap geometries are alike.

v) A strong dc magnetic focusing field exists to prevent transverse debunching.

If it is desired to relax the constraint of equal bunching parameters, the transcapacitance can be defined as

\[
(C_{\text{max}})_{21} \triangleq -M_1 M_2 \frac{I_0}{2 V_0 \omega_q}
\]

where \( M_1 \) and \( M_2 \) are the beam coupling coefficients corresponding to gaps 1 and 2 respectively.
3. THE MULTI-CAVITY KLYSTRON AMPLIFIER

3.1 The Signal-Flow Diagram

So far, the effects of velocity modulation in only one drift length have been considered, i.e., the case of a two-cavity klystron. In a multi-cavity klystron, however, the electron beam encounters intermediate cavities before reaching the catcher cavity. These intermediate cavities are acted upon by the beam and act on the beam. The total effect is more complicated than in a cascade amplifier because all the possible forward paths of signals have to be taken into account. The situation is depicted best by a signal-flow diagram, Figure 5.

Figure 5. Signal-Flow Diagrams for Two-, Three-, and Four-Cavity Klystrons
The total amplification function of the multi-cavity klystron amplifier consists of the sum of the amplifications of all possible signal paths in the forward direction from the buncher to the catcher. Each path must include the effects of the various gaps (cavities) and the drift lengths.

To express mathematically the results expressed by the signal-flow graph, the following two variables are defined:

\[ a_{mh} \triangleq \text{the voltage amplification between gaps } h \text{ and } m \text{ with all intermediate gaps short-circuited,} \]

\[ A_{mh} \triangleq \text{the voltage amplification between gaps } h \text{ and } m \text{ with none of the intermediate gaps short-circuited.} \]

From the flow graphs of Figure 5, the following equations are obtained:

for \( n = 2 \), \( A_{21} = a_{21} \),

for \( n = 3 \), \( A_{31} = a_{31} + a_{32}a_{21} \),

for \( n = 4 \), \( A_{41} = a_{41} + a_{43}a_{31} + a_{42}a_{21} + a_{43}a_{32}a_{21} \)

...(3-1)

For reasons of symmetry, define

\[ A_{11} \triangleq 1 \]

Then

\[ A_{41} = a_{41}A_{11} + a_{42}A_{21} + a_{43}A_{31} \]

or, for any \( n \),

\[ A_{nh} = \sum_{h=1}^{n-1} a_{nh} A_{h1} \]

This is the desired result.
3.2 The Admittance of a Resonator

The admittance of a cavity resonator at the equivalent input terminals may be represented by a circuit consisting of a series connection of an infinite number of anti-resonant LRC circuits. The anti-resonances correspond to the modes of the resonator.

If it is assumed that the cavity is operating in the dominant mode, and that all other resonant frequencies are sufficiently far removed, the cavity may be represented by a single parallel LRC circuit.

The resistive part of the equivalent circuit has two components, one representing the external loading and the other representing the internal loading, consisting of the copper losses and the beam loading. Considering the m-th resonator in a multi-cavity tube, the admittance is given by

\[ Y_m(s) = sC_m + \frac{1}{sL_m} + \frac{1}{R_m} \]

where \( \frac{1}{R_m} = \frac{1}{R_{em}} + \frac{1}{R_{im}} \) \( \ldots (3-2) \)

and \( R_m \triangleq \) the total shunt resistance of the m-th resonator,

\( R_{em} \triangleq \) the external load of the m-th resonator,

\( R_{im} \triangleq \) the equivalent resistance of the m-th resonator representing the copper losses and the beam loading.

In terms of the anti-resonant frequency \( \omega_m \) and the effective \( Q_m \), the admittance becomes

\[ Y_m(s) = \left( \frac{s}{\omega_m} + \frac{\omega_m}{s} + \frac{1}{Q_m} \right) \left( \frac{Q_m}{R_m} \right) \]
It will be assumed that the ratio \((R_m/Q_m)\) is the same for all cavities. For simplicity, let
\[
(R_m/Q_m) \equiv (R/Q)
\]
Then \(Y_m(s) \triangleq 1/Z_m(s)\)
\[
= \frac{1}{(R/Q) \omega_m s} (s^2 + \frac{\omega_m}{Q_m} s + \omega_m^2)
\]...(3-3)

3.3 The Transadmittance

The transadmittance \(Y_{mh}(s)\) between any two resonators \(h\) and \(m\) is obtained in exactly the same way as it was obtained for the two-cavity amplifier. If the transadmittance between two non-adjacent cavities is of interest, the intermediate cavities are considered to be short-circuited. Thus, analogous to Equations (2-18) and (2-19),
\[
sC_{mh} e^{-s\tau_{mh}} = Y_{mh}(s)
\]
and
\[
C_{mh} = C_{max} \sin \theta_{mh}
\]
where
\[
C_{mh} \triangleq \text{the transcapacitance between gaps } h \text{ and } m,
\]
and
\[
\tau_{mh} \triangleq d_{mh}/v_0
\]
As for the two-cavity klystron, it is assumed that the gap geometries are alike, so that the beam coupling coefficients are all equal.

With these definitions an expression for the voltage amplification can be obtained without much difficulty.

3.4 The Voltage-Amplification Function

Consider the signal-flow path between resonators \(h\) and \(m\). The voltage amplification between cavities \(h\) and \(m\), with all intermediate cavities short-circuited, is simply
\[ a_{mh}(s) = \frac{V_m}{V_h} = Y_{mh} Z_m = s C_{mh} Z_m e^{-s \tau_{mh}} \]

Using the 4-cavity klystron as an example, Equation (3-1) becomes
\[
A_{41}(s) = (s^3 c_{43} c_{32} c_{21} z_4 z_3 z_2 + s^2 c_{43} c_{31} z_4 z_3 \\
+ s^2 c_{42} c_{21} z_4 z_2 + s c_{41} z_4) e^{-s \tau_{41}}
\]
\[
= (s^3 c_{43} c_{32} c_{21} z_4 z_3 z_2 \\
+ s^2 c_{43} c_{31} z_4 z_3 z_2 y_2 \\
+ s^2 c_{42} c_{21} z_4 z_3 y_3 z_2 + s c_{41} z_4 z_3 y_3 z_2 y_2) e^{-s \tau_{41}}
\]

Let \[ E(s) = s^4 c_{43} c_{32} c_{21} + s^3 c_{43} c_{31} y_2 \\
+ s^3 c_{42} c_{21} y_3 + s^2 c_{41} y_3 y_2 \] ...(3-4)

The function \( E(s) \) is a polynomial of fourth degree. This can be seen by inserting the values of \( Y_m \) (Equation (3-3)) and expanding.

By Equation (3-3), the product of the resonator impedances becomes
\[
Z_4 z_3 z_2 = s^3 \frac{(R/Q)^3 \omega_4 \omega_3 \omega_2}{(s^2 + \omega_4^2)(s^2 + \omega_3^2)(s^2 + \omega_2^2)}
\]

Thus \[ [s^{-3} z_4 z_3 z_2]^{-1} \] is a polynomial of sixth degree.

Then the voltage-amplification function becomes
\[
A_{41}(s) = [s^2 E(s)] [s^{-3} z_4 z_3 z_2] e^{-s \tau_{41}}
\]
The function \([s^2 E(s)]\) is the numerator of \(A_41(s)\) and therefore defines all zeros in the finite \(s\)-plane, while \([s^{-3} Z_4 Z_3 Z_2]\) is the denominator of \(A_41(s)\) and therefore defines all poles in the finite \(s\)-plane.

Generalizing for the \(n\)-cavity klystron, the voltage amplification is given by

\[
A_{nl}(s) = [s^2 E(s)] [s^{1-n} \prod_{h=2}^{n} Z_h] e^{-s\tau_{nl}}
\]

where \(E(s)\) is defined by an equation analogous to Equation (3-4), consisting of powers of \(s\), \(C_{mh}\), and \(Y_k\) over all possible signal-flow path combinations.

### 3.5 The Power-Gain Function

The power-gain function is defined as the ratio of the output power to the available power from the source. The input power may not be the actual power fed into the klystron amplifier.

In the equivalent circuit of the input resonator (Figure 6), \(R_{el}\) is the equivalent source resistance referred to the resonator, and hence includes the effect of impedance transformation from the input transmission line to the cavity. The power available from the source is

\[
P_{in} \triangleq \frac{1}{2} \left| \frac{I_1}{2} \right|^2 R_{el}
\]
The output power is defined as the power delivered to the load. The voltage $V_n$ appearing at the gap of the output resonator is

$$V_n = A_{nl}(s) V_1$$

Thus the output power for sinusoidal signals ($s=j\omega$) is

$$P_{out} = \frac{1}{2} \frac{|V_n|^2}{R_{en}} = \frac{|A_{nl}(s) V_1|^2}{2 R_{en}}$$

But

$$V_1 = I_1 Z_1$$

Thus

$$P_{out} = \frac{|A_{nl}(s) Z_1|^2 |I_1|^2}{2 R_{en}}$$
The power gain therefore becomes

\[ |G_{nl}(s)| = \frac{P_{out}}{P_{in}} = \frac{4 |A_{nl}(s) Z_1|^2}{R_{el} R_{en}} \]  \quad \ldots (3-6) 

The resistances \( R_{el} \) and \( R_{en} \) have yet to be expressed in quantities measurable in microwave circuits.

The \( Q \) is related to the susceptance of the resonator through the equation

\[ Q_m = \omega_m C_m R_m \]

Let \( Q_{im} \doteq \omega_m C_m R_{im} \)

and \( Q_{em} \doteq \omega_m C_m R_{em} \).

Then with the aid of Equation (3-2), the following relations hold:

\( Q_{im} = R_{im} (Q_m/R_m) = R_{im} (Q/R) \)

\( Q_{em} = R_{em} (Q_m/R_m) = R_{em} (Q/R) \) \quad \ldots (3-7)

\[ 1/Q_m = 1/Q_{em} + 1/Q_{im} \]

Using Equation (3-7), the power gain from Equation (3-6) now becomes

\[ |G_{nl}(s)| = \frac{4 |A_{nl}(s) Z_1|^2}{Q_{el} Q_{en} (R/Q)^2} \]

\[ = K |A_{nl}(s) Z_1|^2 \]

where

\[ K = \frac{4}{Q_{el} Q_{en} (R/Q)^2} \]

Now, for any complex function \( f(s) \), when \( s = j\omega \), it
is true that
\[ f(-s) = f^*(s) \]
where \( f^*(s) \) is the complex conjugate of \( f(s) \),
and so
\[ |f(s)|^2 = f(s) f(-s), \text{ for } s = j\omega. \]

The power-gain function, rewritten as the ratio of two polynomials, finally becomes

\[
|G_{nl}(s)| = K \left[ (s)^{1-n} (-s)^{1-n} \prod_{h=1}^{n} Z_h(s) Z_h(-s) \right] [s^4 E(s) E(-s)]
\]

...(3-8)

The first square bracket defines the denominator polynomial and the second the numerator polynomial.

\section{3.6 Stagger-Tuning}

Since the klystron amplifier in this thesis is regarded as a broad-band device, the amplifier will have to be stagger-tuned in order to get a maximally-flat gain-frequency response. Flat-staggering is achieved by a proper tuning and loading of all the resonators. The tuning can be accomplished with the aid of frequency vanes in the resonators, and the loading of the intermediate resonators is achieved by coupling dummy loads to them. The procedure for calculating the values for the loading (the Q's) and the tuning (the \( \omega \)'s) will now be described.

As was mentioned before, the power gain may be expressed as the ratio of two polynomials,

\[
G_{nl}(s) = \frac{N(s) N(-s)}{D(s) D(-s)}
\]
where $N(s) N(-s) \triangleq s^4 E(s) E(-s)$

\[ \triangleq \text{the numerator polynomial, representing all the zeros,} \]

and

\[
D(s) D(-s) \triangleq \text{the denominator polynomial, representing all the poles.}
\]

If all zeros were at infinity, i.e., if the gain function were an all-pole function, then

\[
N(s) N(-s) \equiv 1
\]

and the problem of flat-staggering would be the well-known one of the design of a filter, for which Butterworth showed that all poles must be evenly spaced on a semi-circle of radius defined by the half-power bandwidth in the left half of the complex-frequency plane, centred at the frequency $s = j\omega_0$.

In any case, the coefficients of the denominator polynomial are constrained by the desired response. This constraint is called the approximation condition.

For klystron amplifiers, however, $N(s) N(-s)$ is not identically equal to unity. In fact, as may be seen from Equation (3-4), there exist finite zeros and these zeros happen to depend on the poles. This dependence is called the realizability condition.

For situations of this kind, the iterative method devised by Moore is applicable. The method has been used already for a variety of applications in network theory, including the stagger-tuning of multi-cavity klystron amplifiers. Although its convergence properties are good,
divergence or oscillatory behaviour may occur; however, it has never been observed by this author. Since finite zeros are present, the poles are in different locations from those prescribed by Butterworth. The situation can best be visualized by the potential analogue theory.

As a first approximation the poles are taken to be located on the semi-circle, and from them the zeros are obtained. In the presence of these zeros, the poles must now be adjusted so that the gain function again satisfies the approximation condition (becomes maximally flat over the specified bandwidth). Once the new set of poles is found from the approximation condition, a new set of zeros may be obtained from the realizability condition. Thence, a third set of poles is calculated, and so on, until convergence is achieved.

i) The realizability condition. The zeros are easily obtained from the poles by solving the polynomial \( N(s) \), Equation (3-4). This polynomial depends on the impedance functions of the resonators and on the transcapacitances. Therefore the zeros are determined once the poles of the impedances are formed.

ii) The approximation condition. To obtain the poles from the new set of zeros, a function \( H(s) \) has to be determined which must have the following properties:

a) \( H(s) \) must be maximally flat within the passband,

b) \( H(s) \) must have the same magnitude at the two half-power points \( j\omega_1 \) and \( j\omega_2 \),
c) \( H(s) \) must have as its denominator the function containing all the zeros from the realizability condition, i.e., \( N(s) N(-s) \).

One such function is

\[
H(s) = (s^2 - s^2_m)^{2n}/N(s) N(-s) \quad \ldots (3-9)
\]

where \( s_m = j\omega_m \) is determined by condition b):

\[
H(j\omega_1) = H(j\omega_2) \quad \ldots (3-10)
\]

Thus from Equations (3-9) and (3-10)

\[
\left( \frac{\omega_1^2 - \omega_2^2}{\omega_m^2 - \omega_m^2} \right)^{2n} = \left| \frac{N(j\omega_1)}{N(j\omega_2)} \right|^2 \Delta k^{2n}
\]

Hence \( \omega_m = (\omega_1^2 - k \omega_2^2)/(1-k) \)

A plot of \( H(j\omega) \) versus \( \omega \) is given in Figure 7. As may be seen from this plot, the function is not yet of the desired form. So, the function \( J(s) \) is formed,

\[
J(s) \triangleq H(s) + H(j\omega_1) \quad \ldots (3-11)
\]

which is then inverted to give

\[
M(s) \triangleq 1/J(s)
\]

\[
= \frac{N(s) N(-s)}{(s^2 - s_m^2)^{2n} + N(s) N(-s) H(j\omega_1)} \quad \ldots (3-12)
\]

A plot of \( M(j\omega) \) versus \( \omega \) is also given in Figure 7. It can be seen that \( M(s) \) is the desired function. The function \( M(s) \) is maximally flat at \( \omega_m \), has a magnitude at \( \omega_1 \) and \( \omega_2 \) which is one-half of that at \( \omega_m \), and its zeros are
Figure 7. Functions Pertinent to the Approximation Condition
those of $N(s) N(-s)$. The denominator of $M(s)$ is the polynomial whose zeros are the desired poles of transmission.

For numerical reasons it is necessary that a transformation be made from the $s$-plane to a plane in which the half-power points $j\omega_1$ and $j\omega_2$ are normalized to $+1$ and $-1$ respectively. Due to this transformation (see Appendix D), it is desirable to choose as a "centre" frequency the root-mean-square of the half-power points, i.e.,

$$\omega_0 \triangleq \left[ \frac{\omega_1^2 + \omega_2^2}{2} \right]^{\frac{1}{2}} \text{ ...(3-13)}$$

The choice of this centre frequency avoids round-off errors.

Another simplification results for the calculations in the $s$-plane if frequency is normalized with respect to an arbitrary frequency. The choice of $\omega_0$ as that frequency will normalize the band centre to unity. In the chapters to follow, $s$ will be used as a normalized frequency variable in this sense.

An impedance transformation does not seem to be necessary.

It should be mentioned here that the calculations involved in these iterations could hardly be carried out without the use of a high-speed digital computer due to the necessity for solving for the zeros of polynomials of high degree.
4. THE MAXIMIZATION OF THE POWER GAIN

In order to maximize the power gain, a variety of design parameters of the klystron amplifier could be considered as variables, e.g., the gap lengths, the gap geometries or the drift lengths. A variation of the gap lengths and gap geometries affects the beam coupling coefficients $M$ and the capacitances $C_{\text{max}}$ and hence the transcapacitances $C_{h,h-1}$; a variation of the drift lengths affects only the factor $\sin \theta_{h,h-1}$ in the expression for the transcapacitances. Theoretically and practically it is much easier to vary the drift lengths than to vary the gap configurations. Therefore, it was decided to investigate the former case only.

4.1 Results Reported in the Literature

Attempts at finding a set of optimum drift lengths for multi-cavity klystron amplifiers which regard the amplifier as an entity have not been reported in the literature, although there are a few papers which extend large signal results obtained for a two-cavity klystron to the multi-cavity klystron.

In May 1958, Curnow\textsuperscript{27} stated "...it would appear that the best arrangement for a multi-cavity klystron is to have the cavities separated by $\frac{1}{4}\lambda_q$...". This conclusion he draws from the following considerations. Since there exists a coupling between non-adjacent cavities in the
multi-cavity klystron, the gain function is not an all-pole function; instead, there are finite zeros present. The effect of these zeros is to lower the gain-frequency characteristic in their vicinity. Curnow assumes that the total available drift length is a fixed proportion of the reduced plasma wavelength, say $\frac{3}{4}\lambda_q$. This may be used with four cavities, separated by $\frac{1}{4}\lambda_q$, for five separated by $3\lambda_q/16$, etc. The contribution of the zeros to the gain-frequency characteristic are then shown in a graph which reveals that, as the number of the cavities is increased and their separation correspondingly decreased, the effect of the zeros becomes more pronounced. Thus, he reasons that the cavities should be separated by $\frac{1}{4}\lambda_q$, and the beam made as long as is practicable. By reasoning in this way, the effect of a reduction of the drift lengths is compared with a successive increase of the number of cavities, whereas it would seem that a more realistic approach would be to compare a variation of the drift lengths for a fixed number of cavities. Curnow also neglects "the small effect of the detuning of the cavities", i.e., he does not consider at all the influence of the poles on the zeros. Consequently his arguments are not very convincing.

A month before Curnow's work appeared, Webber published some results of a theoretical computer study for the large-signal analysis of a two-cavity klystron. This study indicates that the optimum drift length should be reduced,
i.e., should be less than $\frac{1}{4}\lambda_q$. This result was later verified experimentally by Mihran\textsuperscript{10} and Stephenson\textsuperscript{29}.

Mihran's experiments, conducted to investigate the effects of space-charge on a two-cavity klystron, led to some conclusions regarding the drift lengths of a synchronously tuned multi-cavity klystron. It was recommended that a length of one-quarter of an effective space-charge wavelength should be used for all but the last two drift lengths "because they operate at small-signal levels". The second last and the last drift angles should be, respectively, 65 to 45, and 30 degrees of $\lambda_q$, because of the large-signal conditions existing there.

4.2 The Objective of this Thesis and the Proposed Solution

The specific objective of this thesis is to regard the multi-cavity klystron amplifier as an entity and not as a cascade amplifier consisting of several stages made up of two-cavity klystrons, and thence to find a set of drift lengths which will give the highest possible power gain for small signals. Stated mathematically, the objective is to maximize the power-gain function $G(j\omega)$ by a variation of the drift angles $\theta_{h,h-1}$. For obvious reasons, it is desired to maximize a power-gain function which is simultaneously flat-staggered.

The maximization process should raise the entire level of the power gain in the pass band. This means that the power gain should be considered over a wide frequency band. The gain, however, is to be regarded in this section as a
function of the drift angles, and not as a function of the frequency. Thus it is necessary to make the frequency constant. For convenience, the frequency is taken to be the mid-band frequency, \( s = j\omega_0 = j1 \).

Making the frequency constant, i.e., fixing it at one particular point, means in turn that the power gain is to be maximized at that one particular point only, and not over the entire pass band.

The following procedure will be adopted. For a flat-staggered design, the power gain will be maximized at the band centre by a variation of the drift angles. With the new set of drift angles, the power gain will no longer be flat in the pass band, so that the gain must be flat-staggered anew with the new set of drift angles. The new power gain, which should be higher than the previous one, must be maximized in turn at the band centre. Continuing in this fashion, an iterative procedure emerges. The power gain must be flat-staggered and maximized alternately until convergence is achieved. The final design gives an optimized, flat-staggered multi-cavity klystron amplifier.

4.3 The Power Gain as a Function of the Drift Angles

On inspection of the expression for the power gain \( G(s) \), Equation (3-8), it becomes apparent that it will be sufficient to maximize the function \( N(s) N(-s) \), since only this expression is an explicit function of the drift angles.
Although the denominator polynomial and the constant $K$ of $G(s)$ are influenced by the stagger-tuning, they are constant with respect to the $\theta$'s. This simplifies matters somewhat.

Two optimization methods have been applied, but only one of them has given reasonable results. Both methods are numerical ones. The complexity of the equations involved is so high that it appears to be impossible to find analytical solutions. The case of a four-cavity klystron has been taken for the numerical calculations, although the procedure may be extended without difficulty to any number of cavities.

For $s = \imath l$, $N(s) = E(s)$; therefore, from Equation (3-4), the function $N(s)$ for the four-cavity klystron can be written in terms of the drift angles as

$$N(s) = D_1(s) \sin \theta_{43} \sin \theta_{32} \sin \theta_{21}$$
$$+ D_2(s) \sin \theta_{43} \sin \theta_{31} + D_3(s) \sin \theta_{42} \sin \theta_{21}$$
$$+ D_4(s) \sin \theta_{41}$$

...(4-1)

where

$$D_1(s) = s^4 C_{\text{max}}^3$$
$$D_2(s) = s^3 Y_2(s) C_{\text{max}}^2$$
$$D_3(s) = s^3 Y_3(s) C_{\text{max}}^2$$
$$D_4(s) = s^2 Y_2(s) Y_3(s) C_{\text{max}}$$

where the $Y_m$'s are given by Equation (3-3), $C_{\text{max}}$ by (2-20), and the $D$'s are to be evaluated at $s = \imath l$.

4.4 The Method of Partial Derivatives

The first, obvious, method to maximize $N(s) N(-s)$
is by forming the first partial derivatives with respect to the drift angles and by equating them to zero, and then solving the resulting set of simultaneous equations for the unknown optimum \( \theta \)'s. This method proved to be useless, but, nevertheless, will be described for completeness' sake.

For simplicity, let

\[ \theta_{h,h-1} = \theta \quad (h = 2, 3, 4, \ldots, n) \]

then

\[ \frac{\partial}{\partial \theta} [N(s)N(-s)] = 0 \]

\[ N(s)\frac{\partial}{\partial \theta} [N(-s)] + N(-s)\frac{\partial}{\partial \theta} [N(s)] = 0 \]

But, for \( s = j\omega \),

\[ N(-s) = N^*(s) \]

so

\[ N(s)\frac{\partial}{\partial \theta} [N(s)]^* + N^*(s)\frac{\partial}{\partial \theta} [N(s)] = 0 \]

\[ \ldots(4-2) \]

Let

\[ N(s) = \text{Ev}[N(s)] + \text{Od}[N(s)] \]

where \( \text{Ev}[N(s)] \) and \( \text{Od}[N(s)] \) denote respectively those parts of \( N(s) \) containing even and odd powers of \( s \).

Thus, for \( s = j\omega \),

\[ \text{Ev}[N(s)] = \text{Ev}[N(j\omega)] = \text{Re}[N(j\omega)] \]

\[ \text{Od}[N(s)] = \text{Od}[N(j\omega)] = j\text{Im}[N(j\omega)] \]

Equation (4-2) becomes now

\[ \left( \text{Ev}[N(s)] + \text{Od}[N(s)] \right) \frac{\partial}{\partial \theta} \left( \text{Ev}[N(s)] + \text{Od}[N(s)] \right) = 0 \]

\[ \left( \text{Ev}[N(s)] + \text{Od}[N(s)] \right) \frac{\partial}{\partial \theta} \left( \text{Ev}[N(s)] + \text{Od}[N(s)] \right) = 0 \]

Multiplying out and simplifying

\[ \text{Ev}[N(s)] \frac{\partial}{\partial \theta} \text{Ev}[N(s)] - \text{Od}[N(s)] \frac{\partial}{\partial \theta} \text{Od}[N(s)] = 0 \]

\[ \ldots(4-3) \]
Both $E_v[N(s)]$ and $O_d[N(s)]$ consist of sums of products of sines and are quite complicated expressions. The products $E_v[N(s)] \frac{d}{ds} E_v[N(s)]$ and $O_d[N(s)] \frac{d}{ds} O_d[N(s)]$ are, of course, even more complicated yet. In fact, it is necessary to resort to an iterative method. To do this, Equation (4-3) is rewritten as

$$E_v[N_1(s)] \frac{d}{ds} E_v[N_2(s)] + O_d[N_1(s)] \frac{d}{ds} O_d[N_2(s)] = 0$$

...(4-4)

where $N_1(s)$ is the function $N(s)$ evaluated with the $\theta$'s of a previous iteration and hence regarded constant. By doing this, Equation (4-3) simplifies considerably and the set can be solved for the $\theta$'s with much less difficulty. After having obtained a new set of $\theta$'s, $E_v[N_1(s)]$ and $O_d[N_1(s)]$ are recomputed with the newly obtained $\theta$'s, and Equation (4-4) solved once more, with the hope that an improved set of $\theta$'s is obtained, i.e., that the whole process converges.

In the course of this analysis it is convenient to change the variables $\theta_h, h-1$ to a new set of variables. Then a system of non-linear transcendental equations is obtained which can be solved by the usual Newton-Raphson method.

This process, then, which started out to be a direct one, had to be modified to be an iterative one. As soon as iterative methods are dealt with, questions arise concerning the convergence and stability properties which, on the whole, are difficult to answer. Indeed, by experimentation with the Alwac III-E digital computer, for
which a program was written to solve the set of Equations (4-4), it was found that even if a good estimate of the $\theta$'s is given at the outset, either instability or convergence toward non-optimum sets of drift angles occurs.

4.5 The Gradient Method

After a few designs with the previous method, it became obvious that it was imperative to gain a better understanding of the behaviour of the function $N(s)N(-s)$ evaluated at $s = jl$:

$$\left[ N(s)N(-s) \right]_{(s=jl)} = \left| N(jl) \right|^2$$

To do this for the four-cavity amplifier, consider two of the three drift angles to be fixed. For example, consider $\left| N(jl) \right|^2$ as a function of $\theta_{21}$ only. It can be shown (see Appendix E) that if $\left| N(jl) \right|^2$ is defined as $F(\theta_{21})$ when $\theta_{32}$ and $\theta_{43}$ are constant, then

$$F(\theta_{21}) = T_{21} \sin (2\theta_{21} + \delta_{21}) + L_{21} \ldots (E-3)$$

where $T_{21} \triangleq$ the amplitude of the sinusoidal component of $F(\theta_{21})$,

$L_{21} \triangleq$ the mean value of $F(\theta_{21})$,

and $\delta_{21} \triangleq$ the phase angle of $F(\theta_{21})$.

The constants $T_{21}$, $L_{21}$, and $\delta_{21}$ are functions of the tuning and loading of cavities 2 and 3, of $(R/Q)$, $C_{\text{max}}$, and of the two remaining drift angles $\theta_{32}$ and $\theta_{43}$. Their explicit expressions are given in Appendix E.

Similarly it can be shown that...
\[ F(\theta_{32}) = T_{32} \sin(2\theta_{32} + \delta_{32}) + L_{32} \]

and \[ F(\theta_{43}) = T_{43} \sin(2\theta_{43} + \delta_{43}) + L_{43} \]

Thus, if only one of the drift angles is permitted to vary at a time, the function \( |N(jl)|^2 \) is periodic with a period of \( \pi \) radians of drift angle, an amplitude of \( T \), a phase shift of \( \delta \), and a mean value of \( L \), which are generally all different for the three cases. This fact suggests immediately a gradient method for obtaining the optimum drift angles.

Let \( k = h^{-1} \) and call the value \( \theta_{hk} \) for which \( F(\theta_{hk}) \) is a maximum \( (\theta_{hk})_1 \), i.e., \( (\theta_{hk})_1 \) is the solution of the equation

\[ \frac{\partial}{\partial \theta_{hk}} F(\theta_{hk}) = 2T_{hk} \cos(2\theta_{hk} + \delta_{hk}) = 0 \]

or, taking the principal value,

\[ 2\theta_{hk} + \delta_{hk} = \frac{1}{2} \pi \]

or \( (\theta_{hk})_1 = \frac{1}{4}(\pi - 2 \delta_{hk}) \) \( \ldots (4-5) \)

The following steps are to be taken:

i) Pick arbitrary values for two of the three variables, say for \( \theta_{32} \) and \( \theta_{43} \).

ii) Find the value \( (\theta_{21})_1 \) as given by Equation (4-5).

iii) Take this \( (\theta_{21})_1 \), together with the arbitrary \( \theta_{43} \) and substitute into \( F(\theta_{32}) \) and find the value \( (\theta_{32})_1 \) for which \( F(\theta_{32}) \) is a maximum. This new value of \( (\theta_{32})_1 \) will, of course, disturb \( F(\theta_{21}) \), i.e., now \( F(\theta_{21})_1 \) is no longer a maximum. For
the time being, however, this fact is ignored.

iv) Repeat the process for $F(\theta_{43})$ and $(\theta_{32})_1$ and $(\theta_{21})_1$ as the fixed variables.

v) Return to the first step to get a new value of $\theta_{21}$, say $(\theta_{21})_2$.

This process is continued until all three $\theta$'s have converged to the desired degree of accuracy, i.e., when

$$(\theta_{hk})_i = (\theta_{hk})_{i-1} + \varepsilon$$

where $|\varepsilon| \leq$ the tolerance.

The gradient method was also tested on the Alwac III-E digital computer. The program that was written calculated the function $|N(j_1)|^2$ as a function of any of the three drift angles, i.e., $F(\theta)$.

According to the location in the memory at which the calculations were to start, either $F(\theta_{21}), F(\theta_{32})$ or $F(\theta_{43})$ was evaluated. The machine would start at any specific $\theta$ and then proceed to find values of $F(\theta)$, each time increasing the drift angle $\theta$ by a certain specified increment. In this way, the output of the computer was in a form suitable for plotting.

In carrying out a design, the steps i) to v) described above were followed, but the values of $(\theta_{hk})_i$ were obtained by an interpolative method using Equation (E-1) because of the complexity of $\delta_{hk}$ in Equation (4-5).

This procedure showed that the convergence of the gradient method was a definite one and also that it was very
rapid. Once convergence was achieved, the sums of the
amplitudes and the levels of the three function $F(\theta)$ had
the same value, i.e.,

$$T_{21} + L_{21} = T_{32} + L_{32} = T_{43} + L_{43}$$
5. NUMERICAL COMPUTATIONS AND RESULTS

The calculations involved necessitate the use of a digital computer, as was pointed out before.

Programs have been written for the Alwac III–E in the Computing Centre at this University. The Alwac III–E is a medium-sized, electronic, general-purpose digital machine with an 8192-word magnetic-drum memory. The programs are stored internally. Floating-point arithmetic with a 26 binary-digit accuracy for the mantissa and a 6 binary-digit accuracy for the exponent was used for all calculations except for the solving of the zeros of polynomials, where an 8-24 floating-point arithmetic system was used. Since this gives an accuracy of seven decimal digits, it is adequate. A program for finding the zeros of polynomials which uses Bairstow's method was made available by the computer library. At times use was made of a compiler program which interprets a certain simple algebraic language in absolute (machine) language.

The reliability of the machine was very high and when mistakes did occur they were detected immediately.

5.1 The Choice of the Design Parameters

The choice of the numerical design parameters was dictated by the one consideration that they be quantities which were feasible in practice.

Kreuchen, Auld, and Dixon have reported details of a tube (VX 5063) which has been built by them at the Electric
and Musical Industries Research Laboratories Limited, Middlesex, England. Because they have given a complete set of the design parameters for this tube, they were the ones chosen.

This amplifier operated at the frequency of 2940 Mc with a peak output power of approximately 20 Kw over a bandwidth of 300 Mc. The maximum transconductance is given as \( (6.8) \times 10^{-4} \) mhos, the unloaded \( Q \)'s as 350, and the ratio \( (R/Q) \) as 120. The drift lengths were fixed at \( \lambda/\lambda_{\text{opt}} = 0.37 \), but for the purpose of this thesis it was, of course, assumed that they be variable.

5.2 Calculations

When designing a stagger-tuned, multi-cavity klystron amplifier, the choice of the tuning pattern to be used influences the midband gain considerably, as is pointed out by Yuan. The term tuning pattern refers to the allocation of the poles to the various cavities. For drift lengths of one-quarter of \( \lambda_q \), the arrangement giving the highest gain at the band centre is 2143, i.e., cavity 2 is tuned lowest, cavity 3 highest, and cavities 1 and 4 are to be tuned between the other two. Another arrangement which gives fairly high gain is 1423, where cavity 1 is tuned lowest and cavity 3 highest. This tuning pattern, incidentally, gives the highest gain for the actual drift lengths of the EMI tube.

Since it is desired to show the difference in gain obtained between the presently accepted value of \( \frac{1}{2} \pi \) for
the optimum drift angles and the values obtained by the gradient method, the calculations were carried out in all cases first for \( \theta_{21} = \theta_{32} = \theta_{43} = \frac{1}{2} \pi \) and then these designs were optimized with the gradient method.

Designs for the following set of bandwidths were made for both of the above-mentioned tuning patterns: \( \frac{1}{2}, 2, 4, 6, 8, \) and 10 percent. Percent bandwidth was defined as 
\[
(100 \times \text{bandwidth})/(\text{frequency of maximum gain}).
\]

After carrying out a flat-staggered design with all drift angles equal to \( \frac{1}{2} \pi \), the first set of optimized drift angles was obtained. With these drift angles, another flat-staggered design was done, the midband gain of which is shown in Table I in the column headed by "1st opt. set". This design was then optimized again, yielding the second set of optimized drift angles.

For the tuning pattern 2143, it was found that the second set of drift angles differed only very little, if at all, from the first set, and no improvement could be obtained for the midband gain.

For the tuning pattern 1423, however, the second set of optimum drift angles differed enough from the first to warrant another flat-staggered design. The midband gain of this design is shown in Table I in the column headed "2nd opt. set". A third attempt at optimizing this design produced no improvement.

As may be seen from Table I, the midband gain of the designs for the 8- and 10-percent bandwidths actually were lower by 0.1 db as compared to the gain of the first
<table>
<thead>
<tr>
<th>Percent Bandwidth</th>
<th>Power Gain (decibels)</th>
<th>Diff. in Gain</th>
<th>Optimum Drift Angles (radians)</th>
<th>Overall Drift Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[\theta's = \frac{\pi}{2}]</td>
<td>[\theta_{43}]</td>
<td>[\theta_{32}]</td>
<td>[\theta_{21}]</td>
</tr>
<tr>
<td>Tuning Pattern 2143</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[\frac{1}{2}]</td>
<td>69.8</td>
<td>-</td>
<td>0.0</td>
<td>1.52</td>
</tr>
<tr>
<td>2</td>
<td>40.7</td>
<td>-</td>
<td>0.4</td>
<td>1.36</td>
</tr>
<tr>
<td>4</td>
<td>24.0</td>
<td>-</td>
<td>1.4</td>
<td>1.20</td>
</tr>
<tr>
<td>6</td>
<td>14.8</td>
<td>-</td>
<td>2.2</td>
<td>1.10</td>
</tr>
<tr>
<td>8</td>
<td>8.8</td>
<td>-</td>
<td>2.8</td>
<td>1.03</td>
</tr>
<tr>
<td>10</td>
<td>4.5</td>
<td>-</td>
<td>3.2</td>
<td>1.00</td>
</tr>
<tr>
<td>Tuning Pattern 1423</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[\frac{1}{2}]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.52</td>
</tr>
<tr>
<td>2</td>
<td>34.5</td>
<td>35.9</td>
<td>36.2</td>
<td>1.7</td>
</tr>
<tr>
<td>4</td>
<td>17.3</td>
<td>21.9</td>
<td>22.2</td>
<td>4.9</td>
</tr>
<tr>
<td>6</td>
<td>10.9</td>
<td>15.4</td>
<td>15.5</td>
<td>4.6</td>
</tr>
<tr>
<td>8</td>
<td>7.5</td>
<td>11.3</td>
<td>11.2</td>
<td>3.8</td>
</tr>
<tr>
<td>10</td>
<td>4.8</td>
<td>8.2</td>
<td>8.1</td>
<td>3.4</td>
</tr>
</tbody>
</table>

Table I. Summary of Results of the Computer Study of the Modified EMI VX 5063 Tube
optimized designs. This seeming discrepancy is explained as follows: Actually, the power gain is not being maximized, but only its numerator polynomial. As has been shown, there exists a dependence of the numerator polynomial on the denominator polynomial. The reverse can also be said to be true; the denominator depends on the numerator, due to the approximation condition in the flat-staggering process. This implicit interdependence is a weak one but sufficient at high bandwidths to depress the gain by a small fraction of a decibel. The error so produced is considered to be unimportant.

The optimum drift angles shown in Table I are the ones corresponding to the highest gains.

5.3 Results

As may be seen from Table I, the sets of optimum drift angles, obtained by the gradient method, gave higher gains than the sets of $\Theta$'s $= \frac{1}{2}\pi$ in all designs for a bandwidth larger than $\frac{1}{2}$ percent. The difference in gain has been plotted versus the percent bandwidth in Figure 8. Since a 3-db increase in gain constitutes a doubling of the power, an appreciable increase in power gain can indeed be obtained by a proper adjustment of the drift lengths.

In order to show the variation of the optimum drift angles for the various bandwidths, the $\Theta$'s were plotted versus the percent bandwidth in Figures 9 and 10.
The results from the optimization process depend heavily on the chosen tuning pattern. This is clearly seen in Figure 8, as well as by comparing Figures 9 and 10.

For the tuning pattern 2143, the difference in gain rises monotonically with an increase in percent bandwidth, reaching a little over 3 db at 10%. The drift angle closest to the gun ($\theta_{21}$) increases with an increase of Bandwidth, while the centre drift angle ($\theta_{32}$) stays very nearly at a value of $\frac{1}{2}\pi$ and the final drift angle ($\theta_{43}$) decreases at the same time. The overall drift angle ($\theta_{41}$), corresponding to the overall length of the tube, remains remarkably constant at approximately $3\pi/2$ for all designs.

For the tuning pattern 1423, the situation is quite different. Here the improvement in gain rises sharply between the two to four percent bandwidth designs, then falls somewhat to reach nearly the same value at 10% as was obtained for the other tuning pattern. The drift angles all decrease with a decrease of bandwidth. The first and last drift angles are about equal to each other, whereas the centre drift angle is smaller than the other two. The overall length of the tube decreases considerably as the Bandwidth is increased.

Comparing the designs for a 10% bandwidth for the two tuning patterns, it may be noticed that the tuning arrangement 1423 actually gave a midband gain which is higher by one-half of a decibel, in contrast to its behaviour at the other bandwidths. Since at the same time the overall
tube length is considerably shorter, this fact might be of some advantage in practice.

The gain–frequency plots obtained for the various designs are shown in Figures 11 to 15, each figure showing four designs for a specific bandwidth.

As a summary, all optimized designs for the 2143 pattern were plotted on Figure 16. In Table II are displayed the Q's and the normalized frequencies for each of the four cavities for the optimized flat–staggered designs.
Figure 8. Increases in Gain due to Optimization.
Figure 9. Variation of Optimum Drift Angles with an Increase of Bandwidth, Tuning Pattern 2143
Figure 10. Variation of Optimum Drift Angles with an Increase of Bandwidth, Tuning Pattern 1423
Figure 11. Gain–Frequency Characteristics for the 2% Bandwidth Designs.
Figure 12. Gain-Frequency Characteristics for the 4% Bandwidth Designs
Figure 13. Gain-Frequency Characteristics for the 6\% Bandwidth Designs
Figure 14. Gain-Frequency Characteristics for the 8%-Bandwidth Designs
Figure 15. Gain-Frequency Characteristics for the 10% Bandwidth Designs
Figure 16. Gain-Frequency Characteristics for the Optimized Designs, Tuning Pattern 2143
<table>
<thead>
<tr>
<th>Percent Bandwidth</th>
<th>Cavity 1</th>
<th>Cavity 2</th>
<th>Cavity 3</th>
<th>Cavity 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Q_1$</td>
<td>$\omega_1$</td>
<td>$Q_2$</td>
<td>$\omega_2$</td>
</tr>
<tr>
<td>1/2</td>
<td>216.20</td>
<td>0.9990</td>
<td>523.34</td>
<td>0.9977</td>
</tr>
<tr>
<td>2</td>
<td>53.07</td>
<td>0.9959</td>
<td>133.38</td>
<td>0.9906</td>
</tr>
<tr>
<td>4</td>
<td>25.39</td>
<td>0.9908</td>
<td>69.91</td>
<td>0.9808</td>
</tr>
<tr>
<td>6</td>
<td>16.13</td>
<td>0.9845</td>
<td>49.04</td>
<td>0.9706</td>
</tr>
<tr>
<td>8</td>
<td>11.56</td>
<td>0.9775</td>
<td>38.39</td>
<td>0.9603</td>
</tr>
<tr>
<td>10</td>
<td>8.89</td>
<td>0.9698</td>
<td>31.81</td>
<td>0.9498</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/2</td>
<td>520.85</td>
<td>0.9977</td>
<td>216.62</td>
<td>1.0010</td>
</tr>
<tr>
<td>2</td>
<td>124.51</td>
<td>0.9907</td>
<td>54.76</td>
<td>1.0045</td>
</tr>
<tr>
<td>4</td>
<td>52.96</td>
<td>0.9819</td>
<td>30.04</td>
<td>1.0123</td>
</tr>
<tr>
<td>6</td>
<td>30.64</td>
<td>0.9739</td>
<td>23.03</td>
<td>1.0212</td>
</tr>
<tr>
<td>8</td>
<td>20.86</td>
<td>0.9663</td>
<td>19.71</td>
<td>1.0299</td>
</tr>
<tr>
<td>10</td>
<td>15.54</td>
<td>0.9590</td>
<td>17.44</td>
<td>1.0386</td>
</tr>
</tbody>
</table>

Table II: $Q$'s and Normalized Frequencies for the Stagger-Tuning of the Optimized Designs.
6. CONCLUSION

The computations carried out in this study show that there exists a set of optimum drift lengths for a stagger-tuned multi-cavity klystron amplifier which gives a higher small-signal power gain than the presently accepted value of one-quarter of the reduced space-charge wavelength.

The increase in power gain was as large as 4.9 db, depending on the bandwidth and the tuning pattern. Since a 3-db increase in gain represents a doubling of the output power, this is quite a substantial improvement.

An initial attempt at using a partial-derivative method for the optimization failed. Subsequently a gradient method was devised which has good convergence properties. Because the equations involved are too complex to permit an analytical solution, only numerical methods seem to be feasible.

The gradient method has been described in detail for the four-cavity klystron amplifier. It can be extended readily to any number of cavities, although the complexity of the equations rises quickly as the number of cavities increases.

The behaviour of the optimization process was investigated for only two specific tuning patterns from among a number of possible ones. It might be advisable to study this behaviour for other tuning patterns as well.
The theory developed is valid only for the small-signal, fundamental-mode case. Also, the power gain has been maximized by a variation of the drift lengths only. Further work could be done to investigate the influence of the large-signal, higher-order mode theory. The optimization of the power gain by a variation of other tube parameters such as the gap lengths, as well as the maximization of other characteristics such as phase linearity, could be explored. The gradient method might very well lend itself to these studies.
A Fourier analysis of the bunched beam current will yield the frequency components of $i_2$. Since the curves for $i_2$ are symmetrical about the bunch centre (see Figure 3), the resultant terms in its frequency composition will be cosine terms if the origin of the angle variable is centred arbitrarily at the bunch centre by neglecting $\phi_0$, the dc transit angle. This amounts to neglecting the phase factor, which can be put in later by inspection.

Designate the fundamental component of $i_2$ by $I_{21}$, and the $n$-th harmonic by $I_{2n}$. Then

$$nI_2 = \frac{1}{\pi} \int_{-\pi}^{\pi} i_2(\phi_2) \cos(n\phi_2) \, d\phi_2 \quad \text{...(A-1)}$$

The expression for $i_2$, Equation (2-9), gives $i_2$ as a function of $\phi_1$, whereas $i_2$ is wanted as a function of $\phi_2$. Therefore change variables, using Equation (2-7):

$$\phi_2 = \phi_1 + \phi_0 - X \sin \phi_1$$

Differentiating,

$$d\phi_2/d\phi_1 = 1 - X \cos \phi_1$$

Substituting Equation (2-9) and the above equation into (A-1) gives

$$nI_2 = \frac{1}{\pi} \int_{-\pi}^{\pi} I_o \cos n(\phi_1 + \phi_0 - X\sin\phi_1) \, d\phi_1$$
\[ nI_2 = \frac{I_0}{\pi} \int_{-\pi}^{\pi} \cos n(\phi_1 - x \sin \phi_1) \, d\phi_1 \]  

...(A-2)

The integrand can be expressed in the form

\[ \cos n(\phi_1 - x \sin \phi_1) = \cos n\phi_1 \cos(n x \sin \phi_1) + \sin n\phi_1 \sin(n x \sin \phi_1) \]

But, from the theory of Bessel functions

\[ \cos(n x \sin \phi_1) = J_0(nX) + 2 \sum_{m=1}^{\infty} J_{2m}(nX) \cos(2m\phi_1) \]

\[ \sin(nX \sin \phi_1) = 2 \sum_{m=1}^{\infty} J_{2m-1}(nX) \sin(2m-1)\phi_1 \]

Substitution into (A-2) gives:

\[ nI_2 = \frac{I_0}{\pi} J_0(nX) \int_{-\pi}^{\pi} \cos n\phi_1 \, d\phi_1 \]

\[ + \frac{I_0}{\pi} 2 \sum_{m=1}^{\infty} J_{2m}(nX) \int_{-\pi}^{\pi} \cos n\phi_1 \cos(2m\phi_1) \, d\phi_1 \]

\[ + \frac{I_0}{\pi} 2 \sum_{m=1}^{\infty} J_{2m-1}(nX) \int_{-\pi}^{\pi} \sin n\phi_1 \sin(2m-1)\phi_1 \, d\phi_1 \]

Because of the orthogonality property of trigonometric functions, all integrals above but one are zero for any integer value of \( n \) and, in each case, for \( n \) even or odd,

\[ nI_2 = 2 I_0 J_n(nX) \]  

...(A-3)
The following derivation of Webster's results is originally due to Savelyev\(^3\), as translated by Ware\(^4\).

To take into account the effect of electron repulsion, Equation (2-5) is modified as follows:

\[
t_i = t_1 + \int_0^{z_i} \left( \frac{1}{u_i} \right) \, dz
\]

where \( t_i \) is the time of arrival of an electron at \( z = z_i \), where the buncher lies at \( z = 0 \).

Differentiate this equation with respect to \( t_1 \):

\[
\frac{dt_i}{dt_1} = 1 - \int_0^{z_i} \frac{u_i}{u_i^2 \frac{dt_i}{dt_1}} \, dz
\]

\( \frac{du_i}{dt_1} \) is the acceleration of an electron at \( z_i \) in the space-charge field. The concentration of charge causes a field whose magnitude can be calculated. At \( z_i \), the current is \( i_i \), so the excess space-charge density is given by

\[
\rho = -(i_i - I_o)/u_i A
\]

where \( A \) is the cross-sectional area of the beam.

From Maxwell's equations,

\[
dE_z/dz = \rho/e_0 = -(i_i - I_o)/e_0 u_i A
\]

Integrating,

\[
E_z = -\frac{I_o}{e_0 A} \int_{z_i}^{k} \frac{i_i - I_o}{u_i I_o} \, dz
\]

where \( (E_z)_{z=k} = 0 \).

From Equation (2-8),

\[
i_i/I_o = dt_1/dt_i
\]
and therefore

$$E_z = \frac{I_o}{\varepsilon_0 A} \int \frac{z_i}{u} \left( \frac{dt_i}{dt_1} - 1 \right) \, dz$$

Equation (2-1) states

$$\dot{z} = \frac{du_i}{dt} = -\eta E_z$$

Substituting the last two equations into (B-1) yields

$$\frac{dt_i}{dt_1} - 1 = \frac{\eta I_o}{\varepsilon_0 A} \int \frac{1}{u^2} \left[ \int \frac{1}{u} \left( \frac{dt_i}{dt_1} - 1 \right) \, dz \right] \, dt_i \, dz$$

This equation is rigorous, but now it is necessary to approximate with the small-signal assumption: $V_1/V_o \ll 1$, hence $u_i \approx u_o$. Also, if the resonator is near the point of formation of the first bunch, $dt_i/dt_1$ can be taken as constant over the integration and moved into the square bracket; so

$$\frac{dt_i}{dt_1} - 1 = \frac{\eta I_o}{\varepsilon_0 A u_0^3} \int \int \left( 1 - \frac{dt_i}{dt_1} \right) \, dz \, dt_i$$

Let

$$h^2 \triangleq \frac{\eta I_o}{\varepsilon_0 A u_o^3} = \frac{\eta I_o}{\varepsilon_0 A (2\eta V_o)^{3/2}} \propto \frac{I_o}{V_o^{3/2}} \ldots (B-2)$$

$\triangleq$ the debunching wave number.

Since the perveance of a gun is defined as the ratio $I_o/V_o^{3/2}$ it can be seen that this parameter is closely associated with the perveance.
Let \( Y(z) \triangleq \frac{dt_i}{dt_1} - 1; \)
then Equation (B-2) becomes

\[
Y(z) = -h^2 \int_0^z \int_k Y(z) \, dz_j \, dz
\]

Now differentiate \( Y(z) \) twice with respect to \( z \):

\[
d^2 Y(z)/dz^2 = -h^2 Y(z)
\]
or \( d^2 Y(z)/dz^2 + h^2 Y(z) = 0 \)

The solution of this second-order differential equation is

\[
Y(z) = A \sin hz + B \cos hz ....(B-3)
\]

In this equation, \( B = 0 \), because \( Y(0) = 0 \).

Also, from (B-1),

\[
dY(z)/dz = \frac{d}{dz} \left[ \frac{1}{u_i^2} \left( \frac{du_i}{dt_1} \right) \right]
\]

From Equation (2-4),

\[
\frac{du_i}{dt_1} = X \frac{u_0 \omega}{\Phi_0} \cos \omega t_1
\]

Thus

\[
\left[ \frac{dY(z)/dz}{dz} \right]_{z=0} = -\left( \frac{u_0 \omega}{u_1^2 \Phi_0} \right) X \cos \omega t_1
\]

\[
= -(X/d) \cos \omega t_1
\]

\[
= h \bar{A}, \text{ from Equation (B-3)}.
\]

Hence \( \bar{A} = -(X/hd) \cos \omega t_1 \)

The solution of \( Y(z) \) is then
\[ T(z) = -X \left( \sin \frac{hz}{hd} \right) \cos \omega t_1 \]
\[ = \frac{dt_1}{dt_1} - 1 \]

Rearranging,
\[ dt_1 = \left( 1 - X \frac{\sin \frac{hz}{hd}}{\omega} \cos \omega t_1 \right) dt_1 \]

Integration gives
\[ t_i = t_1 \left( 1 - \frac{X}{\omega} \frac{\sin \frac{hz}{hd}}{\omega} \sin \omega t_1 \right) + C \quad \ldots (B-4) \]

An electron passing the buncher at \( t_1 = 0 \) must be in the centre of a bunch and unaffected by the electric field in the drift tube.

Hence, for \( z_i = d \),
\[ t_i = t_1 + \frac{d}{u_o} \]
\[ C = \frac{d}{u_o}. \]

Multiply (B-4) by \( \omega \), and the required result is obtained:
\[ \theta_2 = \theta_0 + \theta_1 - X \left( \sin \frac{hd}{hd} \right) \sin \theta_1 \quad \ldots (B-5) \]
The starting point of the Hahn theory is Maxwell's equations:

\[
\begin{align*}
\text{curl } \hat{E} &= -\frac{\partial \hat{B}}{\partial t} \\
\text{curl } \hat{H} &= \frac{\partial \hat{D}}{\partial t} + \hat{J} \\
\text{div } \hat{D} &= \rho \\
\text{div } \hat{B} &= 0
\end{align*}
\]

where

- \( \hat{E} \) = the electric field intensity vector,
- \( \hat{B} \) = the magnetic flux density vector,
- \( \hat{H} \) = the magnetic field intensity vector,
- \( \hat{D} \) = the electric flux density vector,
- \( \rho \) = the charge density,
- \( \hat{J} \) = the current density vector,
- \( c \) = the dielectric constant or permittivity,
- \( \mu \) = the permeability.

Note: In the following paragraphs, subscript "1" denotes an ac quantity, subscript "o" a dc quantity.

Since this theory is to take into account fringing and wall effects, boundary conditions must be applied. The case of a cylindrical beam coaxial with a conducting cylindrical wall will be investigated, because this is the configuration of most practical tubes.

Both TM and TE fields are able to satisfy the boundary conditions. In the case of a TE field, however, there is no electric field in the direction of motion of the electrons.
and therefore it is of no interest as far as interaction is concerned. Thus, only a TM field is considered in the following analysis.

The quantities \( \rho, u, \) and \( J \) are time periodic, i.e., one is dealing with oscillating fields which contain the factor \( \exp j(\omega t - \beta z) \), where \( \beta = \) the propagation constant. For simplicity, this factor will be omitted, but its presence is always implicit. This means that the partial derivatives can be replaced as follows:

\[
\frac{\partial}{\partial t} = j\omega, \quad \frac{\partial}{\partial z} = -j\beta.
\]

The Lorentz equation

\[
m(\frac{du}{dt}) = -e[\vec{E} + (\vec{u} \times \vec{B})]
\]

becomes, if magnetic forces are neglected, in one dimension,

\[
m(\frac{du}{dt}) = -eE_z
\]

or

\[
\frac{du}{dt} = -\gamma E_z
\]

\[
= \frac{\partial u}{\partial t} + \frac{\partial u}{\partial z} \frac{dz}{dt}
\]

\[
= \frac{\partial u}{\partial t} + u_0 \frac{\partial u}{\partial z}
\]

The ac part, then, is

\[
\frac{\partial u}{\partial t} + u_0 \frac{\partial u}{\partial z} = -\gamma E_z
\]

which becomes

\[
j(\omega - \beta u_0) u_1 = -\gamma E_z l
\]

...(C-2)

The continuity equation (the law of conservation of charge) is given by

\[
\text{div} \; \vec{J} = -\frac{\partial \rho}{\partial t}
\]
which becomes
\[ \frac{\partial J_z}{\partial z} + \frac{\partial \rho}{\partial t} = 0 \]
or \[ J_{z1} = \omega \rho_1 / \beta \] \( \cdots \) (C-3)
By definition,
\[ J_z \triangleq u \rho \]
and therefore the total quantities are
\[ (J_o + J_1)_z = (\rho_o + \rho_1)(u_0 + u_1) \]
\[ J_{z1} = \rho_o u_1 + \rho_1 u_0 + \rho_1 u_1 \]
For the small-signal theory, the product of the ac quantities can be neglected. Thus \( \rho_1 u_1 \approx 0 \) and
\[ J_{z1} \approx \rho_o u_1 + \rho_1 u_0 \] \( \cdots \) (C-4)
From the vector identity
\[ \text{curl curl} \ \vec{A} = \text{grad div} \ \vec{A} - \nabla^2 \vec{A} \]
and from Maxwell's Equations (C-1) is obtained
\[
\text{grad} \ (\rho_1 / \epsilon_o) - \nabla^2 E_1 = -j\omega (\nabla \times \vec{B}_1)
\]
\[ = -j\omega (\mu_o \vec{J}_1 + j\omega \mu_o \epsilon_o \vec{E}_1) \]
So,
\[ \nabla \frac{\rho_1}{\epsilon_o} - \nabla^2 \vec{E}_1 = -j\omega \mu_o \vec{J}_1 + k^2 \vec{E}_1 \]
where \( k^2 = \omega^2 \mu_o \epsilon_o = (\omega/c)^2 \)
and \( c \) = the velocity of light in free space.
Rearranging,
\[ \nabla^2 \vec{E}_1 + k^2 \vec{E}_1 = \nabla \frac{\rho_1}{\epsilon_o} + j\omega \mu_o \vec{J}_1 \] \( \cdots \) (C-5)
\[ \nabla^2 \vec{E}_1 = \frac{1}{r^2} \frac{\partial}{\partial r} \left[ r \frac{\partial \vec{E}_1}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 \vec{E}_1}{\partial \phi^2} + \frac{\partial^2 \vec{E}_1}{\partial z^2} \]
For rotationally symmetric TM waves, \( \frac{\partial^2 E_1}{\partial \phi^2} = 0 \); hence

\[
\nabla^2 E_1 = \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial E_1}{\partial r} \right] - \beta^2 E_1
\]

and

\[
\nabla^2 E_{z1} = \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial E_{z1}}{\partial r} \right] - \beta^2 E_{z1}
\]

Further, considering the z-component only,

\[
\frac{\partial \rho_1}{\partial z} = -j \beta \rho_1
\]

Substituting into (C-5),

\[
\frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial E_{z1}}{\partial r} \right] - (\beta^2 - k^2) E_{z1} = -j (\lambda/\varepsilon_0) \rho_1 + j \omega \mu_0 J_{z1}
\]

From (C-2),

\[
u_1 = \frac{j (\lambda E_{z1})}{(\omega - \beta u_0)}
\]

Substituting (C-3) into (C-4),

\[
J_{z1} = \rho_o u_1 + u_0 (\beta/\omega) J_{z1}
\]

or

\[
J_{z1} = (\omega \rho_o u_1)/(\omega - u_0 \beta)
\]

Combining this with Equation (C-6) yields

\[
J_{z1} = j (\omega \rho_o \eta)/(\omega - u_0 \beta)^2 E_{z1}
\]

But

\[
\rho_o = -I_o/\omega u_0 A_s
\]

hence

\[
J_{z1} = -j \frac{\omega \eta I_o}{A u_0 (\omega - u_0 \beta)^2} E_{z1}
\]

Now define a quantity called the plasma frequency, \( \omega_p \), which occurs often in connection with electron beams,
as follows:

\[ \omega_p = \left( \frac{\eta I_o}{\varepsilon_o u_o A} \right)^{\frac{1}{2}} = \left( -\frac{\eta p_o}{\varepsilon_o} \right)^{\frac{1}{2}} \]

\[ = 1.83 \left( 10^8 \right) \frac{10^{12}}{\sqrt{\varepsilon_o}} \text{ radians} \]

Then

\[ J_{z1} = -j \frac{\omega \varepsilon_o \omega_p}{(\omega - u_o \beta)^2} E_{z1} \quad \ldots \text{(C-8)} \]

Combining Equations (C-3), (C-6), and (C-8),

\[ \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial E_{z1}}{\partial r} \right] + \left( \beta^2 - k^2 \right) \left( \frac{\omega_p^2}{(\omega - u_o \beta)^2} - 1 \right) E_{z1} = 0 \]

Let \( \Gamma_r^2 \triangleq (\beta^2 - k^2) \left( \frac{\omega_p^2}{(\omega - u_o \beta)^2} - 1 \right) \quad \ldots \text{(C-9)} \]

where \( \Gamma_r \triangleq \text{the radial propagation constant} \).

Then

\[ \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial E_{z1}}{\partial r} \right] + \Gamma_r^2 E_{z1} = 0 \quad \ldots \text{(C-10)} \]

or

\[ \frac{\partial^2 E_{z1}}{r^2} + \frac{1}{r} \frac{\partial E_{z1}}{\partial r} + \Gamma_r^2 E_{z1} = 0 \]

because \( E_{z1} = E_z \) within the drift tube.

This is the required wave equation, which is a Bessel equation of zero order.

The solution of the wave equation depends on the boundary conditions. These are (refer to Figure 17)

(i) at \( r = 0 \), \( E_z \) is finite,

(ii) at \( r = a \), \( E_z \) is zero,
There are two regions which must be considered separately:

Region I, $0 \leq r \leq b$, which is the electron beam proper and hence contains space charge;

Region II, $b < r < a$, which is the space between the beam and the wall, and hence contains no space charge.

The solution of Equation (C-10) for Region I, taking into account boundary condition (i), is,

$$ E_{zI} = A J_0(\Gamma_r r) \quad \ldots \ldots \text{(C-11)} $$

The solution of Equation (C-10) for Region II is,

$$ E_{zII} = B I_0(\Gamma_o r) + C K_0(\Gamma_o r) \quad \ldots \ldots \text{(C-12)} $$

where $J_0$ is the Bessel function of the first kind of order zero, $I_0$ and $K_0$ are modified Bessel functions of the first order.
and second kind, respectively, of order zero.

The solution for Region II is justified as follows:

Since in this region there exists no space charge, $\omega_p = 0$, and the radial propagation constant, $\Gamma_r$, becomes

$$r^2_r = -(\beta^2 - k^2) \triangleq -\Gamma_0^2$$

i.e., define

$$\Gamma_0^2 \triangleq (\beta^2 - k^2),$$

Equation (C-10) now becomes a modified Bessel equation

$$\frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial E_z}{\partial r} \right] - \Gamma_0^2 E_z = 0$$

whose solution is given by (C-12).

For a beam in space, not surrounded by walls, the radial propagation constant $\Gamma_r$ approaches zero as the radius of the beam, $b$, becomes large,

or $\Gamma_r^2 \to 0$ as $b \to \infty$.

For $\Gamma_r^2 = 0$

two possibilities exist:

(a) $\beta^2 = k^2$

or $\beta_{1,2} = \pm k = \pm \omega/c$

and (b) $\omega_p^2 = (\omega - u_o \beta)^2$

or $\beta_{3,4} = (\omega^+ - \omega_p)/u_o$

...(C-13)

Case (a) represents a pair of waves propagating with the velocity of light and since negligible interaction occurs with the slow beam, these waves are of no interest. Case (b)
represents a pair of slow waves, one slightly faster and one slightly slower than the beam, because for beams in all normal klystrons $\omega_p/\omega < 1$.

If the assumption of an infinite beam is removed, i.e., if the beam radius is finite and metallic walls are present, the same phenomena occur; there will be pairs of slow waves interacting with the beam, and fast waves having negligible interaction.

Let $\beta_e \equiv \omega/u_o$

$\triangleq$ the phase constant of electrons;

then Equation (C-13) becomes

$$\beta_{3,4} = \beta_e \pm \beta_p$$

where $\beta_p \equiv \omega_p/u_o$

Since $\omega_p << \omega$, the above equation becomes

$$\beta \approx \beta_e$$

Also, since $k << \beta$ unless relativistic beam velocities exist,

$$\Gamma_o \approx \beta \approx \beta_e$$

...(C-14)

The integration constants $A$, $B$, and $C$ of the Equations (C-11) and (C-12) can now be determined by applying the boundary conditions:

(ii) at $r = a$, $E_{zII} = 0$

so $C/B = -I_0(\Gamma_o a)/K_0(\Gamma_o a)$

and $E_{zII} = B\left[I_o(\Gamma_o r) - \frac{I_0(\Gamma_o a)}{K_0(\Gamma_o a)} K_0(\Gamma_o r)\right]$

(iii) at $r = b$, the tangential and radial components of the electric fields are continuous. Hence
\[ E_{zI} = E_{zII} \]

from which
\[ A/B = \frac{1}{\mathbf{J}_0(\Gamma_0 a)} \left[ I_0(\Gamma_0 b) - \frac{I_0(\Gamma_0 a)}{K_0(\Gamma_0 a)} K_0(\Gamma_0 b) \right] \]

Furthermore, since for TM waves
\[ E_r \propto \frac{\partial E_z}{\partial r} \]

one obtains, by condition (iii),
\[ \frac{\partial E_{zI}}{\partial r} = \frac{\partial E_{zII}}{\partial r} \]

For Bessel functions, the following relations hold:
\[ \frac{dJ_0(x)}{dx} = -J_1(x) \]
\[ \frac{dI_0(x)}{dx} = I_1(x) \]
\[ \frac{dK_0(x)}{dx} = -K_1(x) \]

Hence
\[ \frac{\partial E_{zI}}{\partial r} = -\Gamma_r A J_1(\Gamma_r r) \]

\[ \frac{\partial E_{zII}}{\partial r} = \Gamma_r B \left[ I_1(\Gamma_r r) + \frac{I_0(\Gamma_0 a)}{K_0(\Gamma_0 a)} K_1(\Gamma_0 r) \right] \]

and therefore
\[ A/B = -\frac{(\Gamma_0 b)}{(\Gamma_r b)J_1(\Gamma_r b)} \left[ I_1(\Gamma_0 b) + \frac{I_0(\Gamma_0 a)}{K_0(\Gamma_0 a)} K_1(\Gamma_0 b) \right] \]

Equating Equations (C-14) and (C-15),
To solve this equation, a graphical method is employed.

Since, from Equation (C-14),

\[ \Gamma_o = \beta_e \]

\[ f_2(\Gamma_o) \approx f_2(\beta_e) = f_2\left(\frac{\omega}{u_o}\right) \quad \ldots\text{(C-18)} \]

and thus \( f_2 \) can be computed. Then \( f_1(\Gamma_r) \) is plotted as a function of \( \Gamma_r \), yielding an infinite number of intersections which correspond to an infinite number of modes.

The solution of (C-17) is best given in terms of an effective or reduced plasma frequency, \( \omega_q \), which is smaller than the plasma frequency, \( \omega_p \).

In anticipation of this reduced plasma frequency, redefine the propagation constant as follows (compare with Equation (C-13)),

\[ \beta = \frac{\omega \pm \omega_q}{u_o} \quad \ldots\text{(C-19)} \]

From (C-9), neglecting \( k^2 \),

\[ \Gamma_r^2 \approx \beta^2 \left( \frac{\omega_p^2}{(\omega - u_o \beta)^2} - 1 \right) \quad \ldots\text{(C-20)} \]

Now define a constant, called the space-charge reduction factor, \( F_n \), for the n-th mode, as follows:
Replacing $\omega_q$ in (C-19) and substituting $\beta$ into the bracket of (C-20) yields, for the fundamental mode,

$$\Gamma_r^2 = \beta^2 \left( \frac{1}{F^2} - 1 \right)$$

or

$$F = \left[ \frac{1}{(\Gamma_r/\beta)^2 + 1} \right]^{\frac{1}{2}}$$

Also, since $\beta \equiv \beta_e$ and $\omega_q \ll \omega$,

$$F \approx \left[ \frac{1}{(\Gamma_r/\beta_e)^2 + 1} \right]^{\frac{1}{2}}$$

But from (C-18), values of $\Gamma_r/\beta_e$ are obtained graphically, and thus $F_1$ and $\omega_q$ can be determined. The space-charge reduction factor $F$ can be shown to be dependent on the tube and beam geometry only, i.e., on the ratio $a/b$.

For $1.0 \leq a/b \leq \infty$ it can be shown that $0 \leq F_1 \leq 1.0$.

Factors for a wide variety of practical tube geometries have been published in the literature.
APPENDIX D: A NECESSARY TRANSFORMATION FOR THE APPROXIMATION CONDITION

The variable $s$ is assumed to represent normalized frequency, where the frequency of normalization is given by (3-13).

From the realizability condition, $N(s)N(-s)$ is known. This function is a polynomial of degree $4(n-1)$ for an $n$-cavity klystron amplifier. In terms of its zeros, the polynomial becomes

$$N(s)N(-s) = \alpha \prod_{i=1}^{n-1} (s^2 - z_i^2)(s^2 - z_i^*) \quad \ldots(D-1)$$

where $z_i \triangleq$ the $i$-th zero of $N(s)$, and $\alpha \triangleq$ a constant.

Now define a new variable given by

$$\Gamma \triangleq K_n(s^2 + K_s)$$

or

$$s = \left( \frac{\Gamma}{K_n} - K_s \right)^{\frac{1}{2}} \quad \ldots(D-2)$$

where $K_n \triangleq 2/(\omega_2^2 - \omega_1^2)$

and $K_s \triangleq \frac{1}{2}(\omega_1^2 + \omega_2^2) = 1.0$

With this change of variables,

$N(s)N(-s)$ transforms into $n(\Gamma)$,
$H(s)$ into $h(\Gamma)$,
$s = j \omega_m$ into $\Gamma_m$,
$s = j \omega_1$ into $\Gamma_1 = 1.0$,
$s = j \omega_2$ into $\Gamma_2 = -1.0$,
\[ s = z_i \text{ into } \Gamma_{z\text{i}}. \]

Hence, from Equation (3-9),

\[
h(\Gamma) = K_n^{2n}(\Gamma - \Gamma_m)^{2n} / n(\Gamma) \quad \cdots (D-3)
\]

\( \Gamma_m \) is so chosen that from Equation (3-10)

\[
h(1) = h(-1) \quad \cdots (D-4)
\]

From (D-1),

\[
n(\Gamma) = a K_n^2(1-n) \prod_{i=1}^{n-1} (\Gamma - \Gamma_{z_i})(\Gamma - \Gamma_{z_i}^*)
\]

From Equations (D-3) and (D-4),

\[
\frac{(1 - \Gamma_m)^{2n}}{n(1)} = \frac{(-1 - \Gamma_m)^{2n}}{n(-1)} \quad \cdots (D-5)
\]

Define the following constant:

\[ p^{2n} \triangleq \frac{n(1)}{n(-1)} \]

Then (D-5) becomes

\[
\frac{1 - \Gamma_m}{1 + \Gamma_m} = p
\]

The constant \( p \) is a positive, real number because \( |\Gamma_m| < 1 \), and \( \Gamma_m \) is real.

The transformation of \( J(s) \) from Equation (3-11) gives

\[
j(\Gamma) \triangleq h(\Gamma) + h(1) \quad \cdots (D-6)
\]

which, when inverted, gives

\[
m(\Gamma) \triangleq 1/j(\Gamma)
\]

The function of interest is the denominator of \( m(\Gamma) \), which will give the desired poles in the \( \Gamma \)-plane.
If the denominator polynomial of \( m(\Gamma) \) is called \( f(\Gamma) \), then a substitution of (D-3) into (D-6) gives

\[
 f(\Gamma) = (\Gamma - \Gamma_m)^{2n} + \frac{(1 - \Gamma_m)^{2n}}{n(1)} n(\Gamma) 
\]

...(D-7)

In this expression, \((1 - \Gamma_m)^{2n}/n(1)\) is a constant; in fact, it is equal to \( h(1)k_n^{2n} = h(-1)k_n^{2n} \).

Expanding \((\Gamma - \Gamma_m)^{2n}\) as well as \(n(\Gamma)\) and performing the operations indicated in Equation (D-7) gives \( f(\Gamma) \) as a polynomial of degree \(2n\):

\[
 f(\Gamma) = \sum_{i=1}^{2n} c_i \Gamma^i 
\]

Now if the zeros of the polynomial \( f(\Gamma) \) are found and transformed back into the s-plane with the aid of Equation (D-2), then they are the poles required.

The power-gain function of Equation (3-8) involves the factor \( s^4 E(s)E(-s) \) which equals by definition \( N(s)N(-s) \). For computational reasons it is desirable to make the assumption that the relative bandwidth is small so that \( s^4 \) is approximately equal to \( 1 \) over the passband if \( s \) is taken to be the normalized frequency variable. With this assumption \( E(s)E(-s) \cong N(s)N(-s) \) near \( s = j1 \). It is not absolutely necessary to do this; in fact, it might be desirable for higher accuracy or large relative bandwidth to take the zeros at the origin into account.
APPENDIX E: THE GRADIENT METHOD

The behaviour of the function \( N(s)N(-s) \) at \( s = jl \) is to be investigated.

Let

\[
\begin{align*}
    x & \triangleq \sin \theta_{43} \sin \theta_{32} \sin \theta_{21} \\
    y & \triangleq \sin \theta_{43} \sin \theta_{31} \\
    z & \triangleq \sin \theta_{42} \sin \theta_{21} \\
    w & \triangleq \sin \theta_{41}
\end{align*}
\]

Then \( N(s) \), from Equation (4-1), becomes:

\[
N(s) = D_1(s) x + D_2(s) y + D_3(s) z + D_4(s) w
\]

Performing the multiplication \( N(s)N(-s) \) for \( s = jl \),

\[
|N(jl)|^2 = B_1 x^2 + B_2 y^2 + B_3 z^2 + B_4 w^2 + B_5 xy + B_6 xz \\
+ B_7 xw + B_8 yz + B_9 yw + B_{10} zw \quad \ldots (E-1)
\]

where the coefficients are

\[
\begin{align*}
    B_1 & \triangleq |D_1(jl)D_1(-jl)| = |D_1|^2 \\
    B_2 & \triangleq |D_2(jl)D_2(-jl)| = |D_2|^2 \\
    B_3 & \triangleq |D_3|^2 \\
    B_4 & \triangleq |D_4|^2 \\
    B_5 & \triangleq D_1(jl)D_2(-jl) + D_1(-jl)D_2(jl) \\
    & \triangleq 2 \text{Re}[D_1]\text{Re}[D_2] \\
    B_6 & \triangleq D_1(jl)D_3(-jl) + D_1(-jl)D_3(jl) \\
    & \triangleq 2 \text{Re}[D_1]\text{Re}[D_3]
\end{align*}
\]
\[ B_7 = 2 \text{Re} \left[ D_1 \right] \text{Re} \left[ D_4 \right] \]
\[ B_8 = 2 \text{Re} \left[ D_2 \right] \text{Re} \left[ D_3 \right] + \text{Im} \left[ D_2 \right] \text{Im} \left[ D_3 \right] \]
\[ B_9 = 2 \text{Re} \left[ D_2 \right] \text{Re} \left[ D_4 \right] + \text{Im} \left[ D_2 \right] \text{Im} \left[ D_4 \right] \]
\[ B_{10} = 2 \text{Re} \left[ D_3 \right] \text{Re} \left[ D_4 \right] + \text{Im} \left[ D_3 \right] \text{Im} \left[ D_4 \right] \]

For Example,
\[ B_5 = 2 \left( s_4^4 c_{\text{max}}^3 \right) \text{Re} \left[ (-s)^3 I_2 (-s) c_{\text{max}}^2 \right] s = j1 \]
\[ = 2 c_{\text{max}}^5 \frac{(1 - \omega_2^2)}{(R/Q) \omega_2} \]

All variables will have to be expressed in terms of single drift angles. For simplicity, let
\[ \sin \theta_{h,h-1} = s_{h-1} \]
\[ \cos \theta_{h,h-1} = c_{h-1} \]

Using the relation for the sine and cosine of the sum of two angles, the above variables becomes
\[ x = s_3 s_2 s_1 \]
\[ y = s_3 (s_2 c_1 + c_2 s_1) \]
\[ z = s_1 (s_3 c_2 + c_3 s_2) \]
\[ w = c_1 (s_3 c_2 + c_3 s_2) + s_1 (c_3 c_2 - s_3 s_2) \]

Substituting the above equations for \( x, y, z, \) and \( w \) into (E-1) yields the following equation:
\[
|N(j1)|^2 = (B_1 + B_4 - B_7) s_3^2 s_2^2 s_1 + B_2 s_3 s_2 c_1^2 \\
+ (B_2 + B_3 + B_8) s_3^2 c_2^2 s_1 + B_3 c_3^2 s_2^2 s_1 \\
+ B_4 c_3^2 s_2^2 c_1 + B_4 s_3 c_2^2 c_1 + B_4 c_3 c_2^2 s_1 \\
+ (B_5 + B_6 + B_9 - B_{10}) s_3^2 s_2 c_2^2 s_1 \\
+ (B_5 - B_9) s_3^2 s_2 s_1 c_1 + (B_6 - B_{10}) s_3 c_3^2 s_2^2 s_1 \\
+ B_9 s_3^2 s_2 c_2^2 s_1 + B_9 s_3 c_3^2 s_2^2 c_1 \\
+ (B_9 + B_{10}) s_3^2 c_2^2 s_1 c_1 + (B_9 + B_{10}) s_3 c_3 c_2^2 s_1 \\
+ B_{10} c_3^2 s_2 s_1 c_1 + B_{10} c_3 s_2 c_2^2 s_1 \\
+ (2B_2 - 2B_4 + B_7 + B_8) s_3^2 s_2 c_2 s_1 c_1 \\
+ (2B_3 - 2B_4 + B_7 + B_8) s_3 c_3 s_2 c_2 s_1^2 \\
+ 2B_4 s_3 c_3 c_2 s_1 c_1 + 2B_4 s_3 c_2 c_2 c_1^2 \\
+ (-2B_4 + B_7 + B_8) s_3 c_3 s_2 s_1 c_1 \\
+ 2B_4 c_3^2 s_2 c_2 s_1 c_1 + (2B_9 + 2B_{10}) s_3 c_3 s_2 c_2 s_1 c_1 
\]

Consider now \( \theta_{32} \) and \( \theta_{43} \) to be fixed, i.e., \( s_2, c_2, s_3 \), and \( c_3 \) are constant. A regrouping of the above equation gives

\[
|N(j1)|^2 \equiv F(\theta_{21}) \\
= f_{21} s_1^2 + g_{21} s_1 c_1 + h_{21} c_1^2 \\
= f_{21} \sin^2 \theta_{21} + g_{21} \sin \theta_{21} \cos \theta_{21} + h_{21} \cos^2 \theta_{21} \\
\ldots (E-2)
\]
Let \( k \triangleq h^{-1} \), then the functions \( f_{21}, g_{21}, \) and \( h_{21} \) are given by the following matrix equation:

\[
\mathcal{F}_{hk} = \mathcal{B}_{hk} \mathcal{F}_{hk}
\]

where

\[
\mathcal{F}_{hk} \triangleq \begin{bmatrix} f_{hk} \\ g_{hk} \\ h_{hk} \end{bmatrix}
\]

\[
\mathcal{B}_{21} \triangleq \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} & b_{17} & b_{18} & b_{19} \\ b_{21} & b_{18} & b_{16} & 0 & b_{19} & b_{26} & b_{27} & b_{26} & b_{29} \\ b_{31} & b_{14} & b_{14} & 0 & b_{35} & 0 & b_{35} & 0 & b_{26} \end{bmatrix}
\]

\[
\mathcal{B}_{32} \triangleq \begin{bmatrix} b_{11} & b_{31} & b_{13} & b_{14} & b_{21} & b_{16} & b_{17} & b_{35} & b_{27} \\ b_{15} & b_{35} & b_{16} & 0 & b_{19} & b_{26} & b_{19} & b_{26} & b_{29} \\ b_{12} & b_{14} & b_{14} & 0 & b_{18} & 0 & b_{18} & 0 & b_{26} \end{bmatrix}
\]

\[
\mathcal{B}_{43} \triangleq \begin{bmatrix} b_{11} & b_{31} & b_{12} & b_{14} & b_{21} & b_{18} & b_{15} & b_{35} & b_{19} \\ b_{17} & b_{35} & b_{18} & 0 & b_{27} & b_{26} & b_{19} & b_{26} & b_{29} \\ b_{13} & b_{14} & b_{14} & 0 & b_{16} & 0 & b_{16} & 0 & b_{26} \end{bmatrix}
\]
\[
\mathcal{F}_{hk} \triangleq \begin{bmatrix}
(s_i s_j^2) \\
(s_i^2 c_j) \\
(s_i^2 s_j) \\
(c_i^2 s_j) \\
(c_i^2 c_j) \\
(s_i^2 s_j c_j) \\
(c_i^2 s_j c_j) \\
(s_i c_i s_j^2) \\
(s_i c_i c_j^2) \\
(s_i c_i s_j c_j)
\end{bmatrix}
\]

For \( \mathcal{F}_{21} \), \( i = 3, j = 2 \),
for \( \mathcal{F}_{32} \), \( i = 3, j = 1 \),
for \( \mathcal{F}_{43} \), \( i = 2, j = 1 \).

The coefficients of the \( \mathcal{B} \)-matrices are given by

\[
\begin{align*}
b_{11} & \triangleq (B_1 + B_4 - B_7) \\
b_{12} & \triangleq (B_2 + B_3 + B_8) \\
b_{13} & \triangleq B_3 \\
b_{14} & \triangleq B_4 \\
b_{15} & \triangleq (B_5 + B_6 + B_9 - B_{10}) \\
b_{16} & \triangleq B_{10} \\
b_{17} & \triangleq (B_6 - B_{10}) \\
b_{18} & \triangleq (B_9 + B_{10}) \\
b_{19} & \triangleq (2B_3 - 2B_4 + B_7 + B_8) \\
b_{21} & \triangleq (B_5 - B_9)
\end{align*}
\]
To remove the products of sines and cosines in Equation (E-2), rewrite that equation in terms of sines and cosines of double angles:

\[
F(\theta_{21}) = \frac{1}{2}f_{21}(1 - \cos 2\theta_{21}) + \frac{1}{2}g_{21} \sin 2\theta_{21}
\]

\[
+ \frac{1}{2}h_{21}(1 + \cos 2\theta_{21})
\]

\[
= \frac{1}{2}g_{21} \sin 2\theta_{21} + \frac{1}{2}(h_{21} - f_{21}) \cos 2\theta_{21}
\]

\[
+ \frac{1}{2}(f_{21} + h_{21})
\]

\[
= \frac{1}{2}[g_{21}^2 + (h_{21} - f_{21})^2]^{\frac{1}{2}} \sin(2\theta_{21} + \delta_{21})
\]

\[
+ \frac{1}{2}(f_{21} + h_{21})
\]

Let \( T_{21} = \frac{1}{2}[g_{21}^2 + (h_{21} - f_{21})^2]^{\frac{1}{2}} \)

\( \delta_{21} = \tan^{-1}\left[ \frac{(h_{21} - f_{21})}{g_{21}} \right] \)

and \( L_{21} = \frac{1}{2}(f_{21} + h_{21}) \)

Then, finally,

\[
F(\theta_{21}) = T_{21} \sin(2\theta_{21} + \delta_{21}) + L_{21} \quad \ldots (E-3)
\]

which is the relation that was wanted.

For the five-cavity klystron amplifier, the procedure for the gradient method is exactly analogous to the present one. First, the multiplication \( N(s)N(-s) \) is performed;
then all variables are expressed in terms of simple drift angles. Next, a grouping of the fixed variables yields a 3 by 27 matrix for the $b_{ij}$ coefficients; then the functions $f_{hk}$, $g_{hk}$, and $h_{hk}$ for $h = 2, 3, 4, 5$ are determined, from which the optimum drift angles are found.


18. Bers, A., "High-Power Microwave Tubes", Massachusetts Institute of Technology Quarterly Progress Reports, (October 1956)


20. Schelkunoff, S.A., "Representation of Impedance Functions in terms of Resonant Frequencies", Proc. IRE, 32:33, (February 1944)


32. Watson, G. N., Theory of Bessel Functions, Cambridge, University Press, 1922
