

6762

NATIONAL LIBRARY
OTTAWA



BIBLIOTHÈQUE NATIONALE
OTTAWA

NAME OF AUTHOR... JAN FREDERIK VANELDİK.....

TITLE OF THESIS... AUTOMATIC MICROWAVE.....
LOAD MATCHING.....

.....

UNIVERSITY... UNIVERSITY OF ALBERTA.....

DEGREE FOR WHICH THESIS WAS PRESENTED... Ph.D......

YEAR THIS DEGREE GRANTED... 1970.....

Permission is hereby granted to THE NATIONAL LIBRARY
 OF CANADA to microfilm this thesis and to lend or sell copies
 of the film.

The author reserves other publication rights, and
 neither the thesis nor extensive extracts from it may be
 printed or otherwise reproduced without the author's
 written permission.

(Signed)... J. Vaneldik.....

PERMANENT ADDRESS:

40 GRENFELL AVENUE
ST. ALBERT.....
ALBERTA.....

DATED... Sept. 30.....1970

THE UNIVERSITY OF ALBERTA
AUTOMATIC MICROWAVE LOAD MATCHING

by



JAN FREDERIK VANELDIK

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

FALL, 1970

UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read,
and recommend to the Faculty of Graduate Studies for
acceptance, a thesis entitled "Automatic Microwave Load
Matching", submitted by Jan Frederik Vaneldik in partial
fulfilment of the requirements for the degree of Doctor
of Philosophy.

..... *V. J. Kingma*
Supervisor

..... *Peter R. Smy*

..... *L. B. Walker*

..... *Cooper*

..... *J. S. Reger*

..... *P. H. Schubert*
External Examiner

Date... *28-IX-'70*

.....

ABSTRACT

This thesis presents the complete analysis and synthesis of a device, which will automatically, rapidly, and continuously match time-varying microwave loads to a feeder waveguide.

The thesis describes, in detail, the selection of an advantageous matching scheme, the development of a novel loading condition sensing method, and the construction of electromechanically actuated matching components of unusual design.

A closed-loop implementation of the automatic matching device, which becomes a non-linear multi-variable system, is analyzed in terms of stability and performance. Both theoretical and computer simulation results are presented. Some consideration is also given to optimization and adaptive control techniques to improve the system performance.

Finally, a detailed description of the construction and actual performance of an X-band prototype device, together with recommendations for use of the device in medium and high power microwave systems, completes the thesis.

ACKNOWLEDGEMENTS

The author wishes to thank the staff members of the Department of Electrical Engineering for the many valuable discussions and their continued support. Particularly, the author wishes to thank Mr. H. A. Woods for his generous loan of books and information on microwave theory.

The author especially acknowledges his great indebtedness to Professor Y. J. Kingma, the supervisor of this project, for his advice, assistance, and encouragement.

Furthermore, the author is appreciative of the financial assistance received from the National Research Council, Atomic Energy of Canada Ltd., and the Department of Electrical Engineering without which this work would not have been possible.

Finally, the author wishes to express his thanks to Mr. G. R. Kampjes and Sherritt Gordon Mines Ltd. for their aid in the actual production of this thesis.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
CHAPTER 1: GENERAL SYSTEM CONSIDERATIONS	5
1-1 System Concept	5
1-2 System Concept Details	7
1-3 Matching Methods	9
1-4 Double Stub and E-H Tuners	10
1-5 The Single Stub Tuner	14
1-6 Measurement Methods	18
CHAPTER 2: MATCHING DEVICE COMPONENT DESIGN	21
2-1 Phaseshifter Microwave Design Formulae	22
2-2 Dielectric Slab and Slab Support Design	28
2-3 Phaseshifter Servomotor Design	31
2-4 Rate Coil Design	33
2-5 The Complete Phaseshifter	33
2-6 Phaseshifter Performance	37
CHAPTER 3: MEASURING DEVICE DEVELOPMENT	39
3-1 Single Detector E-Field Sampling	39
3-2 Single Detector H-Field Sampling	43
3-3 Three Detector E-Field Sampling	49
3-4 Three Detector H-Field Sampling	52
3-5 H-Field Probe Coupling and Attenuation	55
3-6 The Complete Three Probe H-Field Measuring Section	58
CHAPTER 4: THE CLOSED LOOP SYSTEM	63
4-1 The Closed Loop System Configuration	64
4-2 The System Differential Equations	67

4-3	The System Non-Linearities	71
4-4	Some Properties of the Non-Linearities	76
4-5	Equilibrium and Uniqueness	80
CHAPTER 5:	SYSTEM STABILITY AND RESPONSE ANALYSIS	88
5-1	Stability Analysis Preliminaries	88
5-2	The System Transfer Function Matrix	93
5-3	The Positive Real Associated Matrices	97
5-4	The Non-Linearities and Stability	103
5-5	Stability of the Stub Phaseshifter Loop	113
5-6	Stability of the Complete System	118
5-7	System Response Speed	128
CHAPTER 6:	SYSTEM SIMULATION AND PRACTICAL CONSIDERATIONS	133
6-1	Computer Simulation of the System	134
6-2	System Simulation Results	137
6-3	System Improvement and Optimization	148
6-4	Non-Ideal Properties of the System and its Components	160
6-5	Stability of the Actual System	165
CHAPTER 7:	THE ACTUAL SYSTEM AND ITS PERFORMANCE	169
7-1	The Microwave System Arrangement	169
7-2	Special Phaseshifter Features	171
7-3	The Electronic Control Unit	174
7-4	System Adjustments	180
7-5	Actual System Performance	184
7-6	System Operation at High Microwave Power Levels	188
	CONCLUSIONS AND RECOMMENDATIONS	191
	BIBLIOGRAPHY	194

INDEX OF TABLES

	Page
TABLE 1: EQUATIONS FOR THE SYSTEM NON-LINEARITIES	75

INDEX OF FIGURES

	Page
FIGURE 1: CONVENTIONAL CLOSED-LOOP CONTROL SYSTEM	5
FIGURE 2: PROPOSED MATCHING CONTROL SYSTEM	6
FIGURE 3: MORE COMPLETE SYSTEM PROPOSAL	8
FIGURE 4: THE DOUBLE STUB TUNER	10
FIGURE 5: SMITH CHART FOR THE DOUBLE STUB AND E-H TUNERS	11
FIGURE 6: IMPROVED DOUBLE STUB TUNER	13
FIGURE 7: THE SINGLE STUB TUNER	15
FIGURE 8: SMITH CHART FOR THE SINGLE STUB TUNER	16
FIGURE 9: DIELECTRIC SLAB PERTURBED WAVEGUIDE	24
FIGURE 10: THE DIELECTRIC PHASESHIFTER SLAB	29
FIGURE 11: THE LINEAR MOTION DC SERVOMOTOR	31
FIGURE 12: CROSS-SECTIONAL VIEW OF THE COMPLETE PHASESHIFTER	34
FIGURE 13: PARTIAL CROSS-SECTION OF SUPPORT ROD ARRANGEMENT	35
FIGURE 14: AN E-FIELD PROBE AND DETECTOR	40
FIGURE 15: THE REFLECTION COEFFICIENT VECTOR DIAGRAM FOR VOLTAGE	42
FIGURE 16: A RECTANGULAR WAVEGUIDE SECTION	45
FIGURE 17: THE REFLECTION COEFFICIENT VECTOR DIAGRAM FOR CURRENT	46
FIGURE 18: E-FIELD PROBE PLACEMENT	49
FIGURE 19: H-FIELD PROBE PLACEMENT	52
FIGURE 20: THE H-FIELD PROBE AND ATTENUATOR	55
FIGURE 21: THE H-PLANE TEE AND MEASURING SECTION	59
FIGURE 22: H-FIELD SAMPLING PROBE ON A CUT-OFF MODE WAVEGUIDE	61
FIGURE 23: THE CLOSED LOOP MATCHING SYSTEM	66
FIGURE 24: THE PHASESHIFTER ELECTROMECHANICAL SYSTEM	67
FIGURE 25: COUPLED NON-LINEAR FEEDBACK LOOPS	69

FIGURE 26:	TYPICAL SMITH CHART TRAJECTORIES	77
FIGURE 27:	ERROR SIGNALS AND POLARITIES ON THE SMITH CHART	83
FIGURE 28:	THE GENERAL MULTI-DIMENSIONAL NON-LINEAR SYSTEM	90
FIGURE 29:	THE TWO DIMENSIONAL NON-LINEAR SYSTEM	94
FIGURE 30:	SYSTEM BLOCK DIAGRAM SHOWING THE SEPARATE STUB PHASESHIFTER CONTROL LOOP	96
FIGURE 31:	PLOTS OF $K_{RH}(x,y,Y_L)$ FOR VARIOUS VALUES OF y AND Y_L	106
FIGURE 32:	PLOTS OF $K_{RH}(x,y,Y_L)$ VERSUS x WITH y ADJUSTED FOR $K_{RV}=0$	107
FIGURE 33:	PLOTS OF $K_{RV}(0,y,Y_R)$ VERSUS y FOR VARIOUS VALUES OF Y_R	108
FIGURE 34:	ILLUSTRATION OF PERIOD ENDPOINT LOCATION WITH RESPECT TO THE EQUILIBRIUM VALUES OF x	110
FIGURE 35:	THE STUB CONTROL LOOP WITH RESTRICTED PHASESHIFTER SLAB MOTION	115
FIGURE 36:	DIELECTRIC SLAB POSITIONS SATISFYING ANDERSON'S CONDITIONS ON THE NON-LINEARITIES	120
FIGURE 37:	THE SYSTEM BLOCK DIAGRAM WITH RESTRICTED STUB PHASESHIFTER SLAB MOTION	121
FIGURE 38:	POSITIVE DEFINITE NATURE OF THE SYSTEM LYAPUNOV FUNCTION	125
FIGURE 39:	SIGNAL FLOW GRAPH FOR THE LINEARIZED SYSTEM	130
FIGURE 40:	SYSTEM SIMULATION ON A HYBRID COMPUTER	136
FIGURE 41:	COMPUTER SIMULATION: MATCHING TRAJECTORIES FOR AN "OVERDAMPED" SYSTEM	139
FIGURE 42:	COMPUTER SIMULATION: MATCHING TRAJECTORIES FOR A "CRITICALLY DAMPED" SYSTEM	140
FIGURE 43:	COMPUTER SIMULATION: MATCHING TRAJECTORIES FOR AN "UNDERDAMPED" SYSTEM	141
FIGURE 44:	MATCHING TRAJECTORIES FOR A FASTER STUB LOOP (UPPER) OR A FASTER SERIES LOOP (LOWER)	144

FIGURE 45:	CHANGES IN THE REFLECTION COEFFICIENT MAGNITUDE AS MATCHING PROCEEDS	146
FIGURE 46:	TOTAL MICROWAVE ENERGY REFLECTED FROM START OF MATCHING PROCEDURE, NORMALIZED TO UNITY REFLECTION COEFFICIENT	147
FIGURE 47:	SMITH CHART GEOMETRY FOR LINEARIZED ERROR SIGNALS	153
FIGURE 48:	MATCHING TRAJECTORIES FOR A PARTIALLY LINEARIZED SYSTEM	155
FIGURE 49:	MATCHING TRAJECTORIES AND REFLECTION COEFFICIENT CHANGES FOR AN ADAPTIVE SYSTEM	159
FIGURE 50:	SCHEMATIC DIAGRAM OF PHASESHIFTER TEST ARRANGEMENT	162
FIGURE 51:	PHASESHIFTER TEST CIRCUIT, SLAB POSITION SQUARE WAVE RESPONSE	163
FIGURE 52:	THE PROTOTYPE MICROWAVE COMPONENT ARRANGEMENT	170
FIGURE 54:	SERIES PHASESHIFTER PHOTOELECTRIC LIMIT STOP	172
FIGURE 55:	THE ELECTRONIC CONTROL UNIT CIRCUIT DIAGRAM	176
FIGURE 56:	THE SERIES LOOP POLARITY SWITCHING CIRCUIT	178
FIGURE 57:	SYSTEM RESPONSE TO STUB LOOP SQUARE WAVE FORCING	185
FIGURE 58:	SYSTEM RESPONSE TO SERIES LOOP SQUARE WAVE FORCING	185
FIGURE 59:	ACTUAL SYSTEM MATCHING TRAJECTORIES FOR TYPICAL CONDITIONS	186

INTRODUCTION

In the electrical engineering and physics fields, the measurement of admittance or impedance has always been of considerable importance. Of particular interest were measurement means involving comparison of unknown impedances against known standards. These rapidly evolved into a variety of bridge methods of impedance or admittance measurement.

With the growth in importance of automatic control engineering a new dimension was added to the field of measurement. It became possible to make impedance or admittance bridges fully automatic. This allowed direct read-out of the impedance or admittance value of an unknown, while retaining all the advantages of the basic bridge type of measurement. Considerable work in this area was carried out in the Department of Electrical Engineering at the University of Alberta. A specific example is the thesis entitled "Self Balancing Capacitance Bridge" (1).

In that area of electrical engineering concerned with the generation of large amounts of radio frequency power, for use in both communication and industrial fields, the problem of matching a dissipative load to its power source is of some importance. There are several major reasons for desiring a matched condition between a radio frequency power source and its load.

Firstly, mismatched conditions between the power source and its load entail losses in efficiency, since not all of the available power is transferred to the load.

Secondly, in the communications field, mismatch conditions mean that unwanted reflections and "echos" exist on the transmission line

connecting the power source and load, which can cause distortion of a transmitted, modulated signal.

Thirdly, the power reflected from a mismatched load, if no further precautions are taken, is transmitted back to the power source to be dissipated there. Some specialized radio frequency power sources are very intolerant of high degrees of load mismatch that persist for any length of time.

Many methods have been developed and documented for matching of certain static radio frequency loads to their power sources. These need no further consideration here. Problems arise when the radio frequency load changes rapidly and unpredictably. Under these conditions, manual adjustment of matching components could become at least tedious, and often impossible. It might be feasible, only, to match to certain average loading conditions. However, completely satisfactory results would not readily be obtained.

Since radio frequency load matching involves, in a sense, a specialized form of bridge type measurement of impedance or admittance, in that the load properties are compared to the feeder line characteristics, it seems only natural that the work done, in this Department of Electrical Engineering, on automatic impedance bridges, should evolve into consideration of an automatic load matching device. Automatic load matching devices could, conceivably, maintain a matched condition between a radio frequency power source and its load, even if the load were to change with time. Several workers have indeed devoted their attention to the development of automatic matching devices suitable for the frequency range from 2 to 30 MHz (2,3).

In recent years, microwave power has increasingly come to be used

for industrial purposes. It is now utilized in many processes such as food preparation, batch and conveyor belt drying of materials, drying of films, hardening of glues, and so forth. It is characteristic of this usage of microwave power that the power levels are high and that the loads are not constant, but vary with time. To maintain efficiency at all times and to protect the power sources from damage, continual attention must, thus, be devoted to the matching of the loads and power sources.

Even though work has been done on automatic load matching devices for the lower radio frequency ranges, the microwave frequency range has so far been neglected in this respect. The work reported on in this thesis attempts to fill the gap.

Impetus for the development of a microwave automatic load matching device was further provided by some problems arising out of the ING Project as proposed by Atomic Energy of Canada Limited. The ING Project involved the design and construction of an Intense Neutron Generator. It was planned that the intense neutron beam be produced by impinging a high intensity, high energy proton beam upon a splallation target. The continuous proton beam would have been accelerated in a linear accelerator consisting of a long chain of alternate microwave acceleration cavities and drift regions. The electric field intensity in each cavity as well as the phasing of the microwave power signal from cavity to cavity, would have had to be very carefully controlled in order to produce acceleration. Both temperature variations of the cavity walls and intentional or unintentional beam intensity changes would have caused varying load impedances or admittances to be presented to the power sources. The large number of cavities and power sources involved, to-

gether with the partially unpredictable nature of the load changes would have made automatic matching very desirable.

The above considerations should make it obvious that there is a definite need for an automatic microwave matching device capable of matching time-varying waveguide loads to a feeder waveguide rapidly and accurately. The matching device should be relatively simple and economical compared to the complexity and cost of the main microwave power system. This thesis is devoted entirely to the analysis and synthesis of such a device.

CHAPTER ONE

GENERAL SYSTEM CONSIDERATIONS

In conceiving of a control system which will automatically match a time-varying load to a feeder waveguide, one is faced with a choice of several matching and measurement methods and components. This chapter is devoted to a discussion of the system arrangement, and the selection of specific control system components and matching methods.

1-1 System Concept

In developing an automatic matching device that, is relatively simple in construction, has predictable accuracy, continuously monitors and corrects for mismatch conditions without involving memory, and which is insensitive to component parameter changes, only closed-loop control systems may be considered.

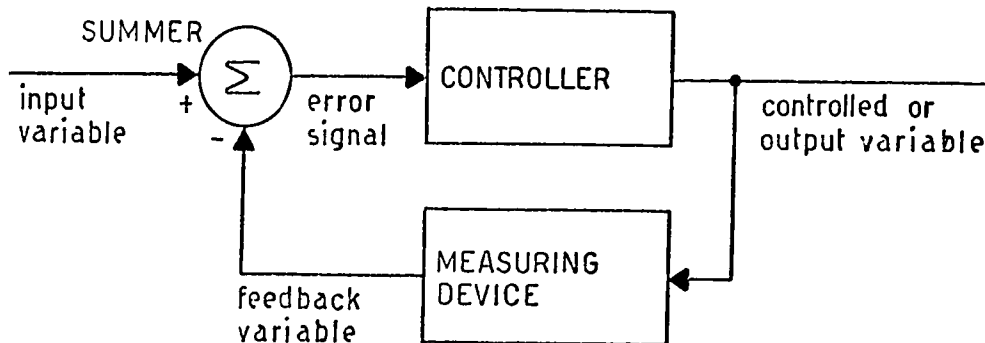


FIGURE 1. CONVENTIONAL CLOSED-LOOP CONTROL SYSTEM

Figure 1 schematically illustrates a conventional closed-loop control system. The conventional system contains a controller which amplifies the error signal, and which uses the amplified error signal to make adjustments to the output or controlled variable. A measuring device monitors the output variable and converts output variable information into a feedback variable. A summer, as shown, takes the difference between the input variable and the feedback variable to produce the error signal. The closed-loop control system shown operates in such a manner as to minimize the error signal at all times, thus causing the output variable to follow the input variable with predictable accuracy.

A similar system arrangement may be devised to match a microwave load automatically. A schematic diagram of a feasible system arrangement is shown in figure 2. This diagram illustrates some of the important similarities as well as differences between a conventional control system and an automatic microwave load matching device. The measuring device produces an error signal which is mathematically related to the degree of mismatch, and which is processed by the computational and

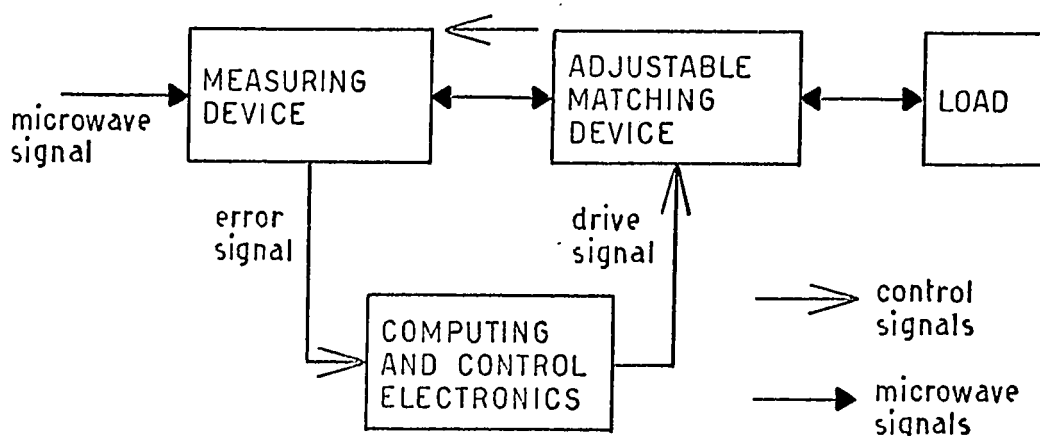


FIGURE 2. PROPOSED MATCHING CONTROL SYSTEM

and control electronics. If the combined load and matching device present a matched load condition to the feeder waveguide, the measuring device senses this and adjustment ceases.

The important similarities and differences between the conventional control system of figure 1 and the system shown in figure 2 may be tabulated as follows:

1. The input variable in figure 2 is a constant and corresponds to a matched load condition.
2. The load is a disturbing signal in the control loop, the effect of which the control system tries to minimize.
3. The feedback variable in figure 2 is represented by the microwave signal reflected from the load.
4. A summer which takes the difference between the input variable and the feedback variable is difficult to define as such. The mixing of the input and feedback variables takes place along the waveguide in the microwave system itself.
5. Most of the operations in the control system of figure 2 are highly non-linear. The matching operation itself is non-linear, the mixing of the input and feedback variables is non-linear, and the measuring operation is also non-linear.

1-2 System Concept Details

The system scheme represented by figure 2 is a deceptive simplification of the actually required matching system arrangement.

Sinusoidal electric or magnetic field variations, and equally so load reflection coefficients or impedance and admittance quantities,

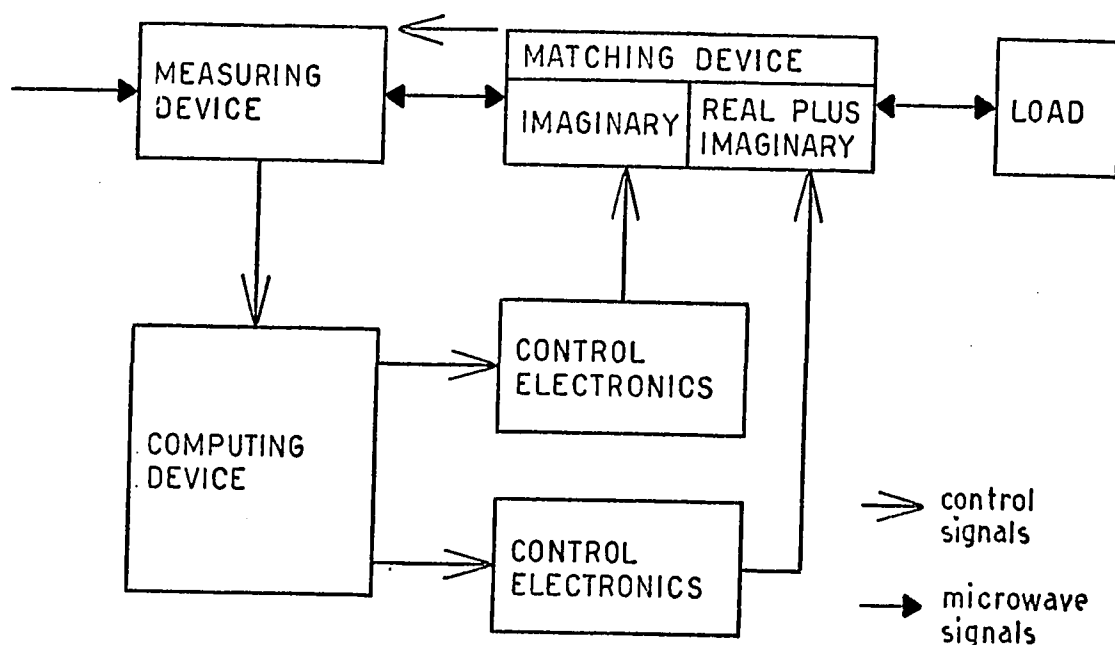


FIGURE 3. MORE COMPLETE SYSTEM PROPOSAL

can be adequately characterized only by complex variables consisting of either real and imaginary parts, or of magnitude and angular parts. Since matching of a microwave load means, in effect, reducing the reflection coefficient of a combined load and matching unit to zero, two separate adjustments are always required. The real and imaginary parts of the reflection coefficient individually have to be reduced to zero. Similar reasoning applies if impedance and admittance or angle and magnitude are considered.

From the previous discussion it is obvious that a single control loop will not suffice. A minimum of two control loops is always required in any automatic load matching device. Unfortunately, the nature of microwave systems is such that two control loops cannot be made

mathematically independent.

The interdependence of the two control loops and the non-linear nature of the system add complexities in system analysis and design that need further consideration in later chapters.

Figure 3 represents a more complete control system which takes into account that two control loops are necessary and which also shows that the control loops are not independent.

1-3 Matching Methods

Many methods are available that allow the matching of a waveguide load to its feeder waveguide. These have been well documented (4,5,6). Matching methods can roughly be divided into two classes. First, there are those that add reflections of predetermined magnitude at selected positions along a waveguide. The slide-screw tuner and the three-screw tuner belong to this class. Secondly, there are those tuners which, for their operation, depend upon the addition of series and/or parallel waveguide admittances and impedances. The latter class includes E-H tuners, double stub tuners, and single stub tuners. Several variations of the tuners of class two are possible also.

The reflection type tuners of class one are totally unsuited to use in an automatic tuning device. The main reasons are:

1. It is difficult to match loads that represent a high degree of mismatch.
2. Protrusions of various forms parallel to the E-field into waveguides that are used at medium or high power levels often cause breakdown difficulties.

3. The devices do not lend themselves readily to electro-mechanical adjustment procedures carried out at reasonably high speeds.

4. Three, instead of two control loops may be required, thus adding unnecessary complications.

This class of devices, though eminently useful in certain applications, will be considered no further.

1-4 Double Stub and E-H Tuners

The double stub and E-H tuners are most easily discussed with reference to a Smith chart ⁽⁷⁾, even though a characteristic waveguide impedance may be difficult to define. Arbitrarily, normalized admittance rather than impedance values will be used.

Since admittance is used in the discussion, the double stub tuner shown in figure 4 will be assumed to contain two shorted stubs connected

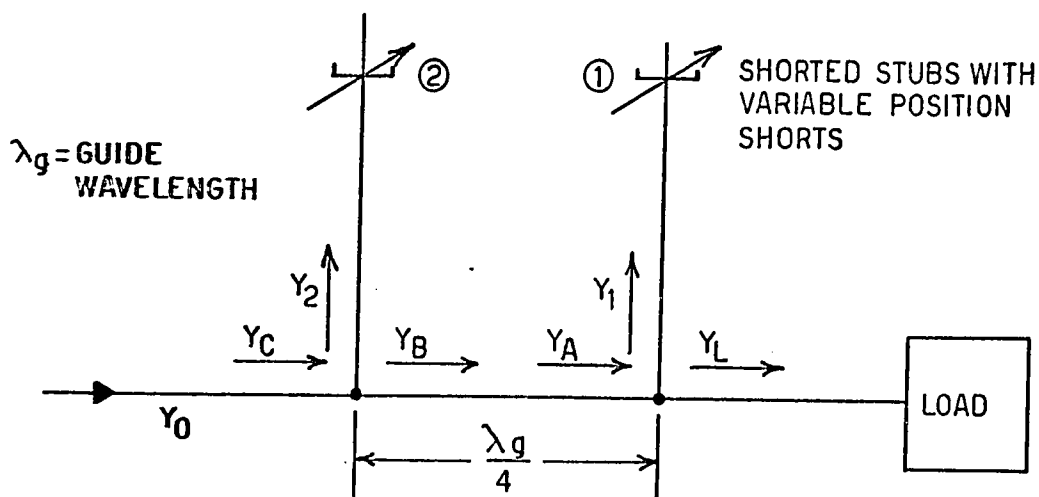


FIGURE 4. THE DOUBLE STUB TUNER

ted to the main waveguide by means of H-plane tees. The spacing between the stubs is one-quarter electrical guide wavelength.

Figure 5 shows a Smith chart applicable to the discussion of double stub tuners. Referring to figures 4 and 5, the operation of the double stub tuner may be explained as follows:

An ideal stub, shorted at one end by a movable short, and connected to a main waveguide by means of an H-plane tee, is capable of adding to the main waveguide admittance any value of susceptance between

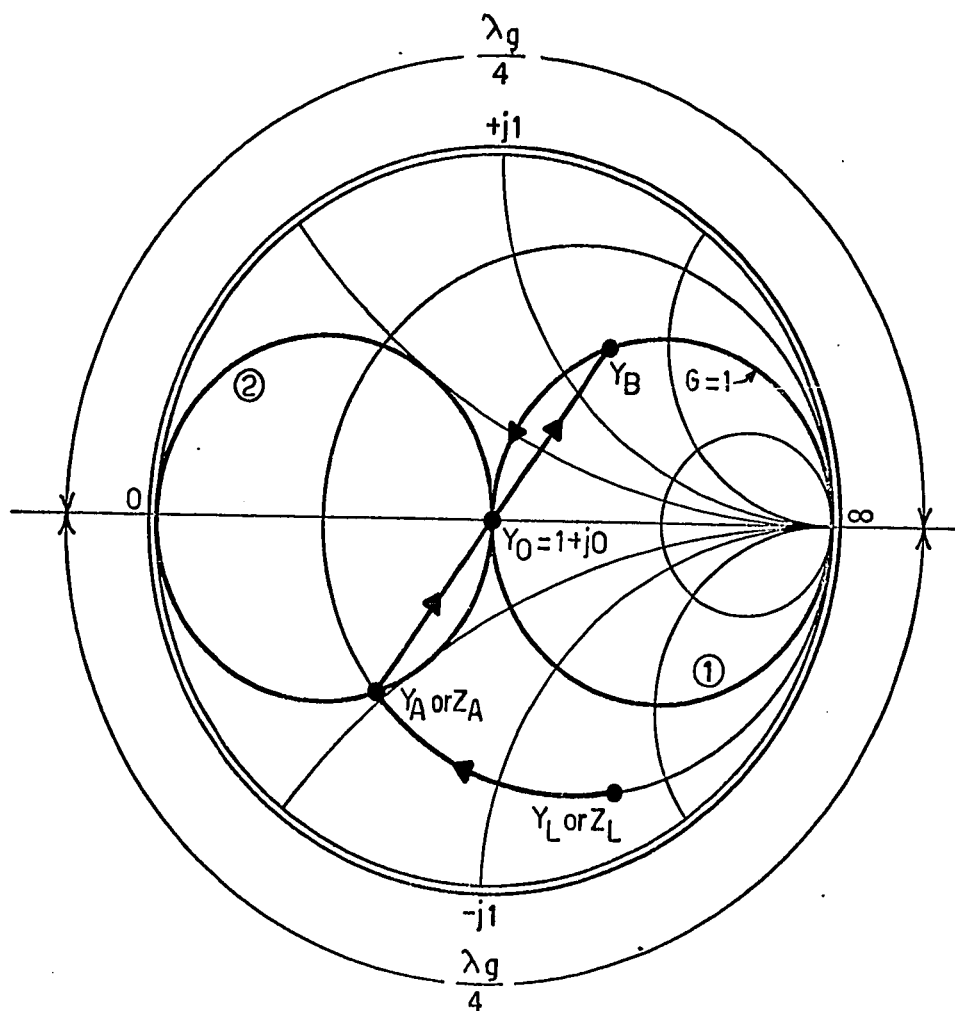


FIGURE 5. SMITH CHART FOR THE DOUBLE STUB AND E-H TUNERS

$+j\omega$ and $-j\omega$. The imaginary part of a load admittance, as it appears at some point along a feeder waveguide, may thus be changed at will by adjustment of a parallel connected shorted stub.

A load admittance, Y_L , if susceptance is added by means of stub 1, may be altered along a constant conductance circle on the Smith chart until it becomes admittance Y_A , which lies on circle 2 of figure 5. Circle 2 is the locus of those admittances which, if transformed or rotated by means of a quarter wavelength section of waveguide, will become the admittances of circle 1. The quarter wavelength waveguide section between stubs 1 and 2 will thus transform Y_A into Y_B . Admittance Y_B lies on circle 1, which is the locus of all admittances for which the real part is equal to one. By means of stub 1 susceptance may again be added, such that the imaginary part of Y_B is cancelled. This changes Y_B into Y_C , which equals Y_0 at the origin of the Smith chart. The load Y_L has now been matched.

The operation of the E-H tuner is similar to that of the double stub tuner. The transformation of a certain impedance into its equivalent admittance on a Smith chart is entirely analogous in every respect to transforming Y_A into Y_B by means of a quarter wavelength guide section. Referring again to figure 5, a load impedance, Z_L , may be changed along a constant resistance line into Z_A by the addition of series reactance. Impedance Z_A is equivalent to admittance Y_B , which may be changed to Y_0 by the addition of further susceptance. The H-plane stub 1 and the quarter wavelength guide section of figure 4 may, thus, be replaced by an E-plane shorted stub at the location of stub 2. The explanation of the operation of the E-H tuner, as presented above, has however, ignored some of the fundamental difficulties in analysis that

arise if two extended discontinuities, such as an E- and H-tee, are present at one position along a waveguide.

Matching devices of the type just described, even though they can be adapted to automatic operation, have one serious drawback. Unmatched loads, Y_L , that lie inside the unity conductance circle 1 on the Smith chart of figure 5, cannot be matched by a double stub or E-H tuner. For loads, Y_L , inside circle 1 the simple addition of pure susceptance cannot place Y_A on circle 2 since the constant conductance circles inside circle 1 do not intersect circle 2.

The above difficulties may be remedied by several methods, one of which is shown in figure 6. Here, a phaseshifter has been added between the actual load and stub 1. Proper adjustment of the phaseshifter of figure 6 will rotate or transform Y_L into Y_L' which does lie in the region outside circle 1 of the Smith chart, and can thus be matched.

Although the double stub tuner or E-H tuner with a phaseshifter added could possibly be used in an automatic matching device, these tu-

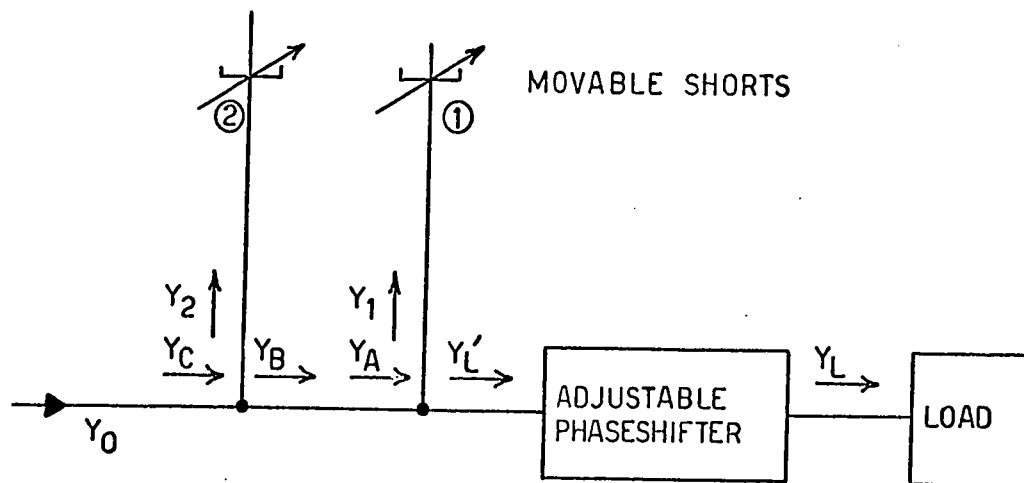


FIGURE 6. IMPROVED DOUBLE STUB TUNER

ners were not considered to be sufficiently suitable. The extra phase-shifter necessitates the addition of a third control loop. From an automatic control point of view the third control loop introduces some redundancy which may cause the automatic matching device to search for a matched condition without reaching it. Because of the third adjustment, the matching operation is no longer unique. Uniqueness of adjustment is an essential requirement for successful automation of the matching operation. Undesirable complexities are also added by the fact that the measuring device, still to be discussed, would have to detect if the load to be matched lies inside the unity conductance circle of figure 5.

A more advantageous matching method is available and will be discussed in the next section.

1-5 The Single Stub Tuner

The most suitable matching method for automatic control purposes is a single stub tuner, or some variation of the single stub tuner. This type of tuner is capable of matching any load, other than complete open or short circuits or purely reactive loads, with just two adjustable components. This means, that the theoretical minimum of two control loops can be used to perform the matching operation. Figure 7 shows a schematic diagram of a single stub tuner arrangement.

The operation of the single stub tuner is again most easily explained by means of a Smith chart as shown in figure 8, and bears some resemblance to the double stub tuner.

Adjustment of a series phaseshifter as shown in figure 8, can

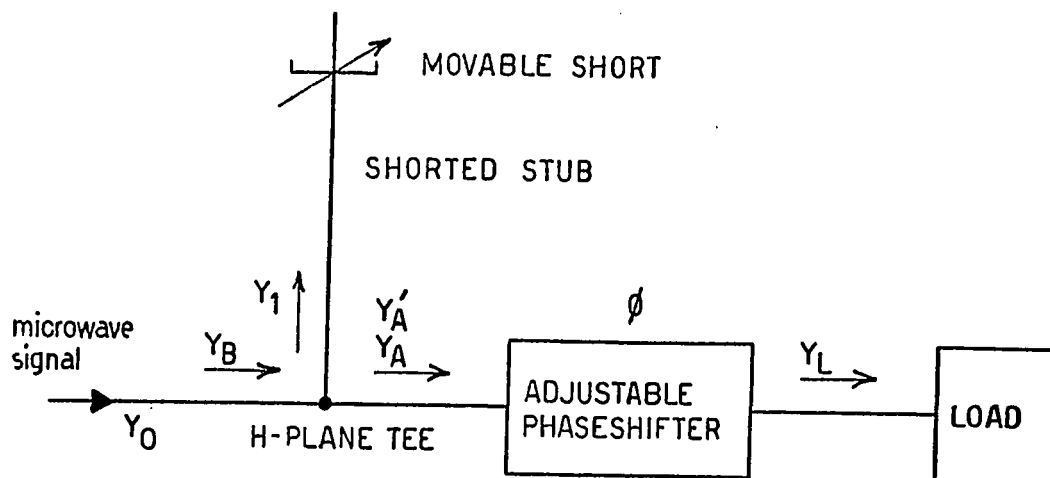


FIGURE 7 THE SINGLE STUB TUNER

transform any load admittance, Y_L , into any other admittance which leaves the magnitude of the reflection coefficient unaltered. Expressed differently, a series phaseshifter can rotate a load admittance along a circle which passes through the original load admittance and which has the origin of the Smith chart as center. The family of concentric circles, that have the Smith chart origin as centers, all intersect the unity conductance circle 2 in figure 8. By proper adjustment of the series phaseshifter any load, Y_L , can, therefore, be transformed into admittances Y_A or Y'_A , both of which lie on the unity conductance circle 2.

Once Y_A or Y'_A lies on circle 2, the imaginary part of admittance Y_A or Y'_A can be cancelled by adding susceptance of appropriate sign by means of the movable shorted stub shown in figure 7. This makes the composite load admittance equal to Y_0 , and the load is matched.

In anticipation of the discussions in later chapters, it should be

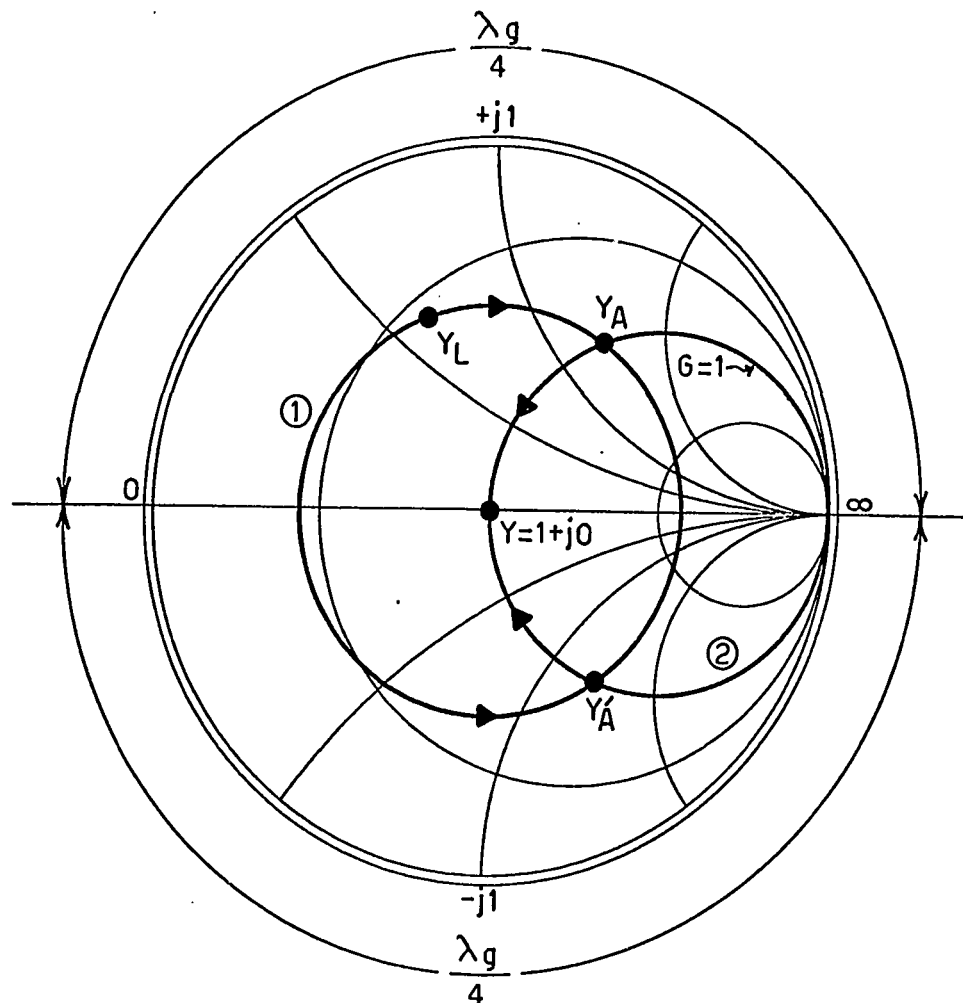


FIGURE 8. SMITH CHART FOR THE SINGLE STUB TUNER

mentioned that the selection of the sign of the loop gain associated with the series phasestifter allows one to choose either the upper or the lower half of the unity conductance circle 2 of figure 8 as the stable locus of admittance values Y_A to which the phasestifter will always adjust.

Again anticipating later chapters, certain load changes, that effectively represent repeated encirclements of the Smith chart origin,

would cause the series phaseshifter to seek equilibrium adjustment positions corresponding to ever increasing phase angles. However, alternate selection of the upper and lower unity conductance semi-circles by appropriate methods can limit the actually required phaseshifts to a range of plus and minus 90° . Careful attention to the points just mentioned will make the matching adjustments entirely unique.

The immediate advantages of using the single stub matching method in an automatic matching device are as follows:

1. Any load, other than loads on the perimeter of the Smith chart, can be matched.
2. The number of adjustable components, and therefore, the number of control loops required, is equal to the theoretical minimum of two.
3. Some limited degree of separation between the two adjustments exists in that the susceptance added by the stub does not affect the real part of the load admittance, while phaseshifter adjustment corresponds to a purely angular change in the reflection coefficient.
4. The adjustment of the matching device for any particular load is unique once the phaseshifter range has been limited by automatic selection of the proper unity conductance semi-circle.

The single stub tuner was chosen as most suitable for an automatic matching device.

1-6 Measurement Methods

Automation of microwave matching procedures requires the development of a device which will detect if a matched condition exists and which, in the absence of a matched condition, will produce an error signal that is mathematically related to the degree of mismatch. More specifically, the error signal should be an electrical signal which is zero if the load is matched, but which also contains polarity information related to the required direction of adjustment to obtain a matched condition. These requirements appear simple, but are, in fact, rather stringent.

The two normal measurement methods often employed in microwave engineering are reflectometry and standing wave ratio measurement.

Reflectometry is based on the fact that a load is completely characterized, in terms of the feed line properties, by its reflection coefficient (8). This fact can be deduced from the following equation:

$$k_R = \frac{1 - Y_L}{1 + Y_L} \quad 1.1$$

where k_R is the reflection coefficient, and Y_L is the load admittance. Under matched conditions k_R is zero.

The magnitude squared, and therefore the magnitude, of the reflection coefficient may be measured by sampling the power reflected from the load (by means of a directional coupler). The ratio of reflected power to incident power is equal to the square of the reflection coefficient magnitude. Often however, the actual ratio is not determined, but

1-6 Measurement Methods

Automation of microwave matching procedures requires the development of a device which will detect if a matched condition exists and which, in the absence of a matched condition, will produce an error signal that is mathematically related to the degree of mismatch. More specifically, the error signal should be an electrical signal which is zero if the load is matched, but which also contains polarity information related to the required direction of adjustment to obtain a matched condition. These requirements appear simple, but are, in fact, rather stringent.

The two normal measurement methods often employed in microwave engineering are reflectometry and standing wave ratio measurement.

Reflectometry is based on the fact that a load is completely characterized, in terms of the feed line properties, by its reflection coefficient ⁽⁸⁾. This fact can be deduced from the following equation:

$$k_R = \frac{1 - Y_L}{1 + Y_L} \quad 1.1$$

where k_R is the reflection coefficient, and Y_L is the load admittance. Under matched conditions k_R is zero.

The magnitude squared, and therefore the magnitude, of the reflection coefficient may be measured by sampling the power reflected from the load (by means of a directional coupler). The ratio of reflected power to incident power is equal to the square of the reflection coefficient magnitude. Often however, the actual ratio is not determined, but

only the reflected power is measured.

The angular part of the reflection coefficient is not easily determined. The angular part of the reflection coefficient is equal to the phase angle difference between the E-field of the incident wave and the E-field of the reflected wave. To measure this phase difference, some rather sophisticated equipment, such as a phase discriminator, has to be designed. Again, in many applications the angular part of the reflection coefficient is ignored, even though it contains vital information.

The main disadvantages of using reflectometry to derive error signals useful for automation of the matching apparatus are:

1. The reflection coefficient magnitude signal does not contain polarity information to indicate the direction of required adjustment.
2. The equipment required to measure the angular part of the reflection coefficient is complex.
3. Angular and magnitude information about the reflection coefficient is not directly useful for automation. Trigonometric calculations would have to be performed to extract the real and imaginary parts of the reflection coefficient, but, trigonometric calculations are not readily carried out by analogue methods.

Instead, standing wave ratio measurement methods are better suited to use in automated matching equipment.

If two electromagnetic waves, of the same frequency, but travelling in opposite directions through a waveguide, interfere, they produce a characteristic space pattern of maxima and minima along the waveguide.

The incident wave and the wave reflected from a mismatched load interfere in exactly this manner. The ratio of the field intensity at a maximum to that at a minimum is known as the standing wave ratio; it is related to the magnitude of the reflection coefficient. The positions of the maxima and minima along the waveguide, in turn, are related to the angular part of the reflection coefficient. The nature of the load is, therefore, completely characterized, in terms of the feeder guide properties, by the standing wave pattern in the feeder guide, provided the feeder guide exhibits low loss properties and provided also that the power source is matched to the feeder waveguide.

A simple measuring device, incorporating three microwave crystal detector probes, may be designed which directly extracts two signals from the standing wave pattern, which respectively, are proportional to the real and imaginary components of the reflection coefficient. The signals so extracted do contain polarity information indicative of the desired direction of adjustment of the matching device.

Complete details of this measuring device, which samples the standing wave pattern, will be given in a subsequent chapter.

CHAPTER TWO

MATCHING DEVICE COMPONENT DESIGN

In chapter one, use of the single stub matching method for an automatic matching device was advocated. The single stub matching method employs one phaseshifter and one adjustable shorted stub to accomplish the matching of a load. For reasons of symmetry, and because phaseshifters are more easily adjusted electromechanically, the movable short in the shorted stub may be replaced by a fixed short and a phaseshifter. Then, only one basic type of phaseshifter need be developed for the entire matching unit.

From a variety of possible phaseshifter designs, the dielectric slab phaseshifter was specifically chosen for the following reasons:

1. The adjustment of the phaseshifter involves only linear motion, while the necessary range of motion is small.
2. The mechanical construction of the phaseshifter is simple, and the moving parts have relatively low mass.
3. The phaseshifter is readily adaptable to electromechanical actuation.
4. Careful design of the phaseshifter will allow operation at microwave power levels of 10 to 20 kilowatts.

Design of a medium power dielectric slab phaseshifter is possible if the following points are borne in mind:

1. To reduce losses and prevent overheating, only very low loss dielectric materials, such as teflon or polystyrene, should be used in the phaseshifter.

2. To prevent breakdown and arcing, no areas of excessive electric field strength concentration must be allowed.
3. To reduce the effects of unwanted reflections, the dielectric slab and its supports must be properly matched into the waveguide.

Further attention will be given to these considerations in subsequent sections of this chapter.

2-1 Phaseshifter Microwave Design Formulae

Expressions for the phaseshifts obtainable with a dielectric slab of given dimensions may be derived using perturbation formulae. Among others, Altman ⁽⁶⁾ develops a perturbation formula for rectangular waveguide with a longitudinally placed slab of material inside. His expression for waveguide perturbation is:

$$\gamma - \gamma_0 = \frac{\omega \int_{\Delta S} [(\epsilon - \epsilon_0) \bar{E}_0^* \bar{E}' + (\mu - \mu_0) \bar{H}_0^* \bar{H}'] da}{2 \int_S (\bar{E}_0^* \times \bar{H}') \cdot d\bar{a}} \quad 2.1$$

where:

\bar{E}_0 and \bar{H}_0 are the electric and magnetic fields, respectively, in an unperturbed waveguide.

\bar{E} and \bar{H} are the fields in the perturbed waveguide

ΔS is the cross-sectional area of the slab of perturbing material.

S is the cross-sectional area of the waveguide.

da is a differential element of area.

γ_0 and γ are the complex propagation constants of the unperturbed and perturbed waveguides respectively.

μ_0 and ϵ_0 are the permeability and dielectric constants of free space.

μ and ϵ are the permeability and dielectric constants of the perturbing material.

ω is the radian frequency of the microwave signal.

* indicates the complex conjugate of the quantity.

The fields \bar{E}'_0 , \bar{E}' , \bar{H}'_0 , and \bar{H}' are defined by the equations:

$$\begin{aligned}\bar{E}_0 &= \bar{E}'_0 e^{j(\omega t - \gamma_0 z)} \\ \bar{E} &= \bar{E}' e^{j(\omega t - \gamma z)} \\ \bar{H}_0 &= \bar{H}'_0 e^{j(\omega t - \gamma_0 z)} \\ \bar{H} &= \bar{H}' e^{j(\omega t - \gamma z)}\end{aligned}$$

2.2

where t is time, z is the direction of propagation, and j is the imaginary operator of complex variables.

Consider now a lossless, purely dielectric slab in a lossless waveguide as shown in figure 9. Thus, with specific reference to figure 9:

$$\gamma_0 = \beta_0 - j\alpha_0 \rightarrow \beta_0; \gamma = \beta - j\alpha \rightarrow \beta; \mu_0 = \mu \quad 2.3$$

If the dielectric material is homogeneous and isotropic, then, by use

of 2.3, equation 2.1 may be simplified to:

$$\beta - \beta_0 = \frac{\omega(\epsilon - \epsilon_0) \int_{\Delta S} \bar{E}'_0^* \cdot \bar{E}' da}{2 \int_S (\bar{E}'_0^* \times \bar{H}') \cdot d\bar{a}} \quad 2.4$$

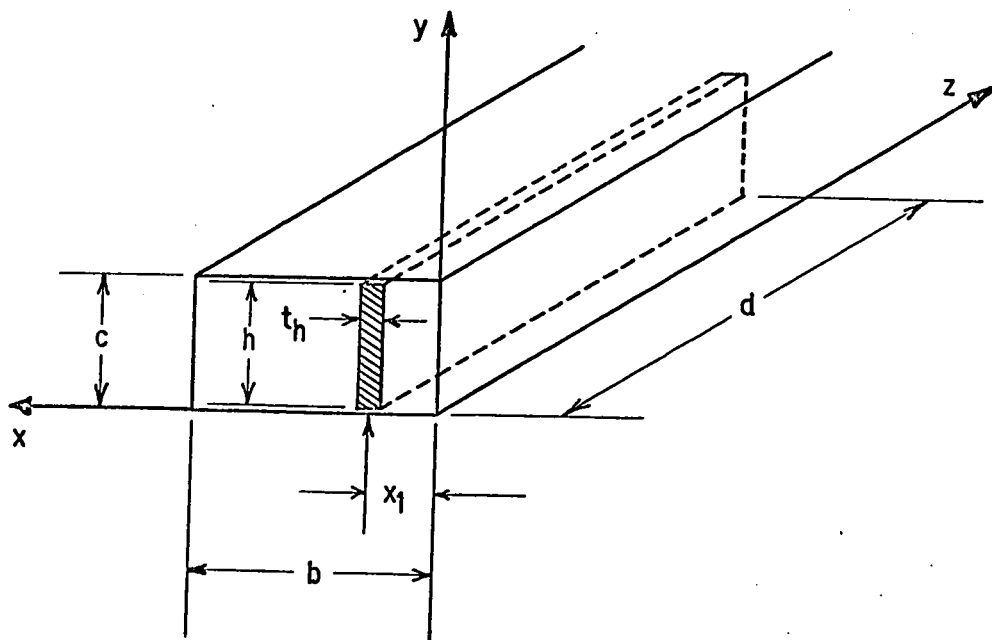


FIGURE 9. DIELECTRIC SLAB PERTURBED WAVEGUIDE

For the common TE_{10} mode of waveguide operation, one has the following expressions for the electromagnetic field components (see for instance Altman (6)):

$$E_x = E_z = H_y = 0 \quad 2.5$$

and:

$$E_y = -A \left(\frac{\mu_0}{\epsilon_0} \right)^{1/2} \frac{\lambda_{g_0}}{\lambda} \frac{\pi}{b} \left[\sin \left(\frac{\pi x}{b} \right) \right] e^{j(\omega t - \beta_0 z)} \quad 2.6$$

$$H_x = A \frac{\pi}{b} \left[\sin \left(\frac{\pi x}{b} \right) \right] e^{j(\omega t - \beta_0 z)} \quad 2.7$$

where:

A is an amplitude constant.

λ_{g_0} is the guide wavelength in the unperturbed guide.

λ is the free space wavelength.

The component H_z exists, but will contribute nothing to the integral $\int_S (\bar{E}_0^* \times \bar{H}') \cdot d\bar{a}$, since H_z is parallel to $d\bar{a}$. Thus, H_z will not be considered.

Assuming that $\Delta S \ll S$ and that, as a result, the fields outside the dielectric are practically equal to the fields in an unperturbed waveguide, the denominator integral of equation 2.4 may be evaluated as follows, taking into consideration that there is no change in E_y and H_x in the y -direction:

$$\int_S (\bar{E}_0^* \times \bar{H}') \cdot d\bar{a} = \int_S -(\bar{k} E_y H_x) \cdot d\bar{a} = -c \int_0^b E_y H_x dx \quad 2.8$$

Substituting equations 2.6 and 2.7 into equation 2.8, and knowing that $bc=S$, one has:

$$2 \int_S (\bar{E}_0^* \times \bar{H}') \cdot d\bar{a} =$$

$$2cA^2 \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} \frac{\lambda_{g_0}}{\lambda} \frac{\pi^2}{b^2} \int_0^b \sin^2\left(\frac{\pi x}{b}\right) dx =$$

$$SA^2 \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} \frac{\lambda_{g_0}}{\lambda} \frac{\pi^2}{b^2}$$

2.9

To evaluate the numerator integral of equation 2.4, several assumptions must again be made. The fields in the perturbed and the unperturbed waveguides are taken as being equal. Also, considering that $\Delta S \ll S$, and that the tangential electric field component across an air-dielectric boundary must be continuous, the electric field is assumed to be constant across the width, t_h , of the dielectric slab and is taken as being equal to its value in an empty guide at the midpoint of the slab. Therefore:

$$E_y \Big|_{at x_1} = -A \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} \frac{\lambda_{g_0}}{\lambda} \frac{\pi}{b} \left[\sin\left(\frac{\pi x_1}{b}\right)\right] e^{j(\omega t - \beta_0 z)} \quad 2.10$$

Thus:

$$\int_{\Delta S} \bar{E}_0^* \cdot \bar{E}' da =$$

$$A^2 \frac{\mu_0}{\epsilon_0} \frac{\lambda_{g_0}^2}{\lambda^2} \frac{\pi^2}{b^2} \sin^2\left(\frac{\pi x_1}{b}\right) \int_{\Delta S} da =$$

$$A^2 \frac{\mu_0}{\epsilon_0} \frac{\lambda_{g_0}^2}{\lambda^2} \frac{\pi^2}{b} (\Delta S) \sin^2\left(\frac{\pi x_1}{b}\right)$$

2.11

Substitution of the results of equations 2.9 and 2.11 into equation 2.4 gives, after simplification:

$$\beta - \beta_0 = \frac{(\epsilon - \epsilon_0) \omega (\mu_0 \epsilon_0)^{1/2} \lambda_{g_0} \Delta S \sin^2\left(\frac{\pi x_1}{b}\right)}{\epsilon_0 \lambda S} \quad 2.12$$

Realizing that $\omega \sqrt{\mu_0 \epsilon_0} = \frac{2\pi}{\lambda}$, and setting $\frac{\epsilon}{\epsilon_0} = \epsilon'$, equation 2.12 may be written into its final form:

$$\frac{\beta}{2\pi} - \frac{\beta_0}{2\pi} = \frac{1}{\lambda_g} - \frac{1}{\lambda_{g_0}} = (\epsilon' - 1) \frac{\lambda_{g_0}}{\lambda^2} \frac{\Delta S}{S} \sin^2\left(\frac{\pi x_1}{b}\right) \quad 2.13$$

where λ_g is the wavelength in the perturbed waveguide, and ϵ' is the relative dielectric constant of the dielectric slab material.

The differential phaseshift, DPS, that is, the difference in phaseshifts between an empty guide and the same guide with the dielectric slab inserted is:

$$DPS = \left(\frac{1}{\lambda_g} - \frac{1}{\lambda_{g_0}} \right) 2\pi d = 2\pi d (\epsilon' - 1) \frac{\lambda_{g_0}}{\lambda^2} \frac{\Delta S}{S} \sin^2\left(\frac{\pi x_1}{b}\right) \quad 2.14$$

Similarly, the incremental phaseshift, IPS, i.e. the change in phaseshift in moving the dielectric slab from position x_1 to position x_2 is:

$$IPS = 2\pi d (\epsilon' - 1) \frac{\lambda_{g_0}}{\lambda^2} \frac{\Delta S}{S} \left[\sin^2\left(\frac{\pi x_1}{b}\right) - \sin^2\left(\frac{\pi x_2}{b}\right) \right] \quad 2.15$$

Equations 2.13, 2.14, and 2.15 are the basic dielectric slab phaseshifter design equations. These three equations are not exact, especially if the dielectric slab is not very thin. However, they will be found to offer sufficient accuracy for engineering purposes.

2-2 Dielectric Slab and Slab Support Design

For a given dielectric slab thickness, t_h , usually taken as less than 10% of the waveguide width, and for a certain value of maximum slab displacement for the range of phaseshifts desired, the required slab length may be determined by use of equation 2.15.

If the dielectric slab ends were abruptly terminated, large reflections from them would occur. The best method of drastically reducing these reflections is to taper the ends. The effectiveness of tapers in reducing reflections from the slab ends is almost independent of perturbed guide wavelength, and therefore, of slab position in the guide. Tapers were thus chosen over some other method of matching the slab into the guide. Experimentally it was found that tapers, equal in length to one empty guide wavelength, are adequate since the voltage standing wave ratio caused by the slab alone, under these conditions, will be less than 1.03.

In calculating the required length, d , of the slab, the effective electrical lengths of the tapers may be considered to be one half of their actual lengths. This is so because, according to equation 2.15, the differential phaseshift is proportional to ΔS , the cross-sectional area of the slab, and ΔS varies linearly along the length of each taper if the slab thickness is constant.

The main features of the dielectric slab are illustrated in figure 10.

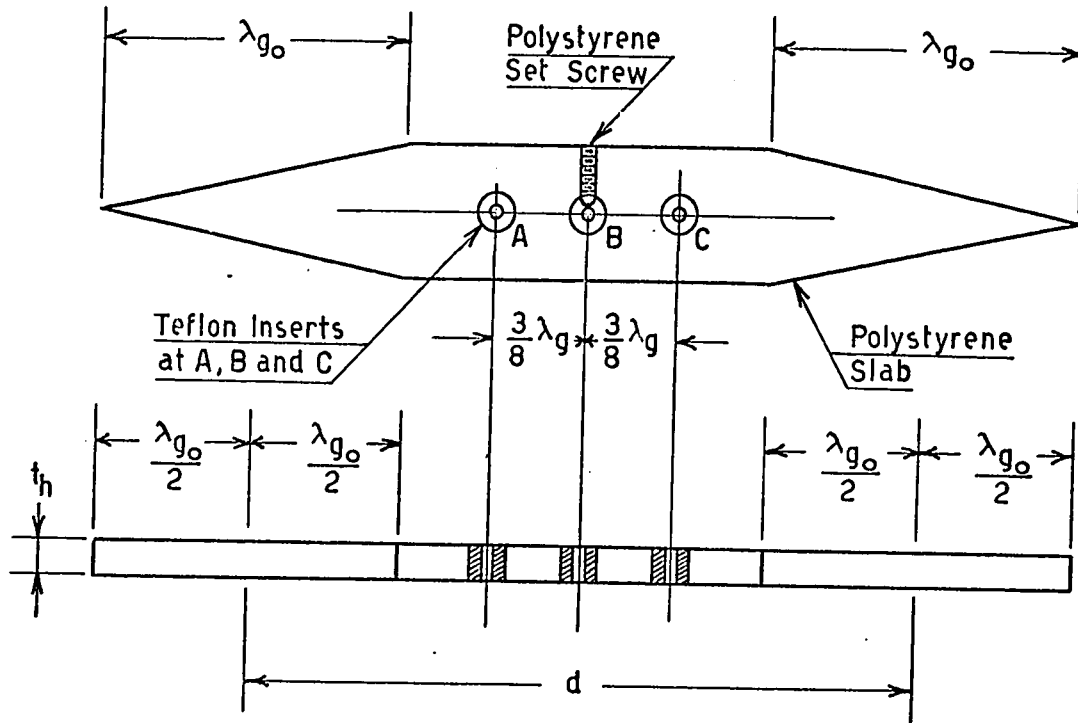


FIGURE 10. THE DIELECTRIC PHASESHIFTER SLAB

The low loss polystyrene slab, shown in figure 10, should be suspended in the waveguide by means of three entirely equally constructed supports. The reflections from any two equal supports spaced at a distance of $\lambda_g/4$ would cancel perfectly at one particular value of guide propagation constant β , but, the guide propagation constant in the section containing the dielectric slab changes with varying slab positions in the guide. Broadbanding is thus required. This is accomplished by using three supports spaced at a distance of $3\lambda_g/8$. The value of λ_g used is calculated from equation 2.13 for an average value of slab displacement x_1 .

Very low friction sliding supports are required for the dielectric slab. Polished rhodium sliding on teflon exhibits the necessary low coefficient of friction. In addition, teflon is another low loss material with a dielectric constant that is not very different from that of polystyrene. Teflon inserts, at A and C in figure 10, may thus be used without causing excessive disturbances of the electric fields. The dielectric slab slides, at A and C, on highly polished, rhodium plated, hollow, brass rods which pass through the waveguide sidewalls by means of holes ten one-thousandths of an inch larger than the rod diameter. The support rods at A and C, which should have a diameter of less than 10% of the waveguide height, are firmly attached to the outside of the waveguide through use of a teflon plate with brass cover plate.

The central teflon insert, at B in figure 10, is used for symmetry reasons only. The rod at B is the pusher rod, which extends well beyond the waveguide sidewalls through sidewall holes, also ten one-thousandths of an inch larger than the rod diameter. The slab is fixed to rod B with a polystyrene setscrew, as shown. By means of the pusher rod, motion may thus be imparted to the slab from outside the waveguide. Rod B is supported in a bearing hole drilled into the same teflon plate, on the outside of the waveguide, which also fixes the rods at A and C. Further constructional details will be discussed in a later section.

If the phaseshifter is constructed according to the principles set forth, its insertion loss will be low, while the voltage standing wave ratio, caused by the phaseshifter, over a phaseshift range of 180° will be no greater than 1.1. Since the entire matching device is capable, to

some degree, of compensating for its own errors, this is adequate.

2-3 Phaseshifter Servomotor Design

To use the phaseshifter in an automated matching device, it must be possible to adjust the position of the dielectric slab in the waveguide by means of an externally applied electrical signal.

The preferred method of directly producing linear motion is to move a long coil in a short magnetic gap. Some details of the proposed method are shown in figure 11. As indicated, a standard heavy loud-speaker magnet is used, but with the polepieces modified such as to allow a much longer than normal coil to be inserted into the magnetic

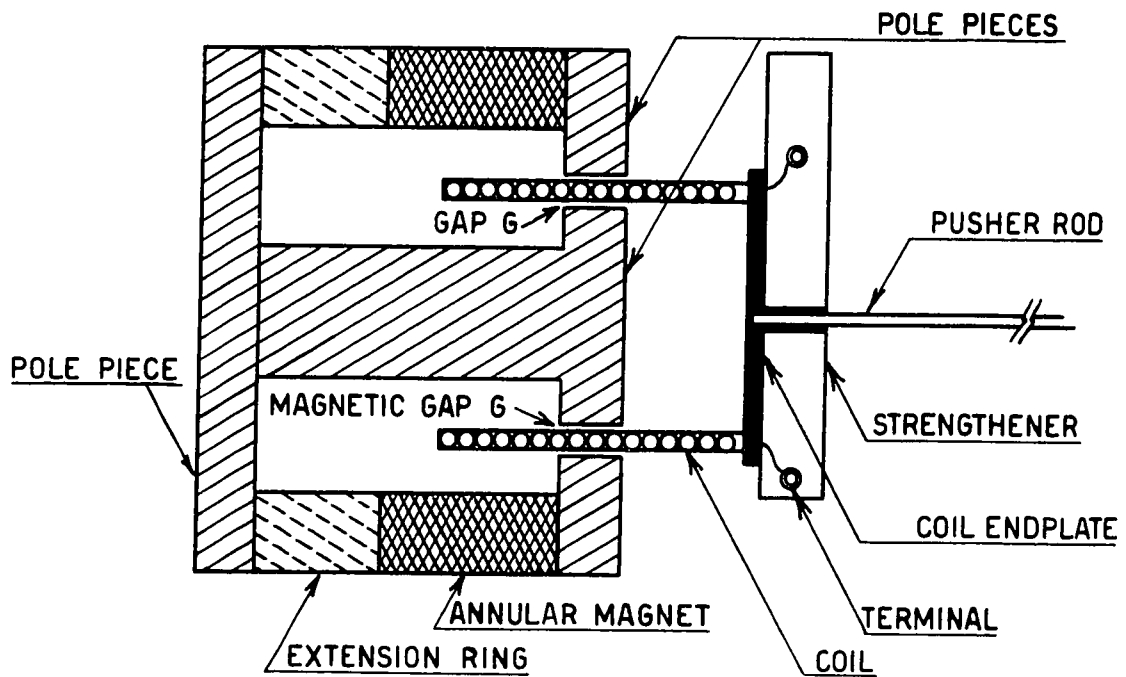


FIGURE 11. THE LINEAR MOTION DC SERVOMOTOR

gap. The circular gap shown was also enlarged in diameter by modification of the polepieces. With this type of gap construction, practically all of the magnetic flux generated by the magnet is concentrated in a fairly uniform field in the short gap region G in figure 11. The coil itself is made long enough so that, over the desired range of motion, the ends of the coil never come close to entering the gap region G. Under these circumstances, for any coil position within its range of motion, the same number of turns is contained within the effective magnetic gap. Therefore, for a given constant value of coil current, the force produced by the coil is independent of coil position.

The drive coil consists of two or four layers of enameled copper wire wound into wet epoxy cement such as to produce a self-supporting coil, without a coil form to use up valuable gap space. The coil wire size and the number of layers should be chosen appropriately to give a coil DC resistance of about 6 or 7 Ohms. The coil may then be driven from a conventional transistorized DC power amplifier. The coil wire size does, of course, depend on the transient current magnitudes occurring during adjustment, which must be withstood without overheating.

The following is the design formula for the drive coil:

$$F = \pi D N \frac{l}{s} B i \quad 2.16$$

where:

F is the force in newtons produced by the coil.

D is the average coil diameter in meters.

N is the number of coil layers.

s is the wire diameter in millimeters.

l is the gap length in millimeters.

B is the average field strength in the gap in teslas.

i is the coil current in amperes.

The coil resistance may be determined from the total wire length in the coil and its resistance per meter.

2-4 Rate Coil Design

As will become obvious in later chapters, it is necessary to obtain an electrical signal proportional to the rate of motion, or velocity, of the dielectric slab, which is used for damping of the control loops. The rate signal is obtained by means of a coil and magnet structure that is exactly analogous in construction to the arrangement of the servomotor drive coil and magnet.

The magnet used for the rate of motion structure is somewhat smaller than that of the servomotor. A smaller coil diameter is therefore used, but the coil is now wound with 6 or 8 layers of number 40 enameled copper wire formed into a self-supporting coil by means of epoxy cement.

The magnetic gap and coil arrangement is again such that, for a constant rate of motion, the voltage generated by the coil is virtually independent of the position of the coil in the gap.

2-5 The Complete Phaseshifter

A complete and detailed cross-sectional view of the entire phase-

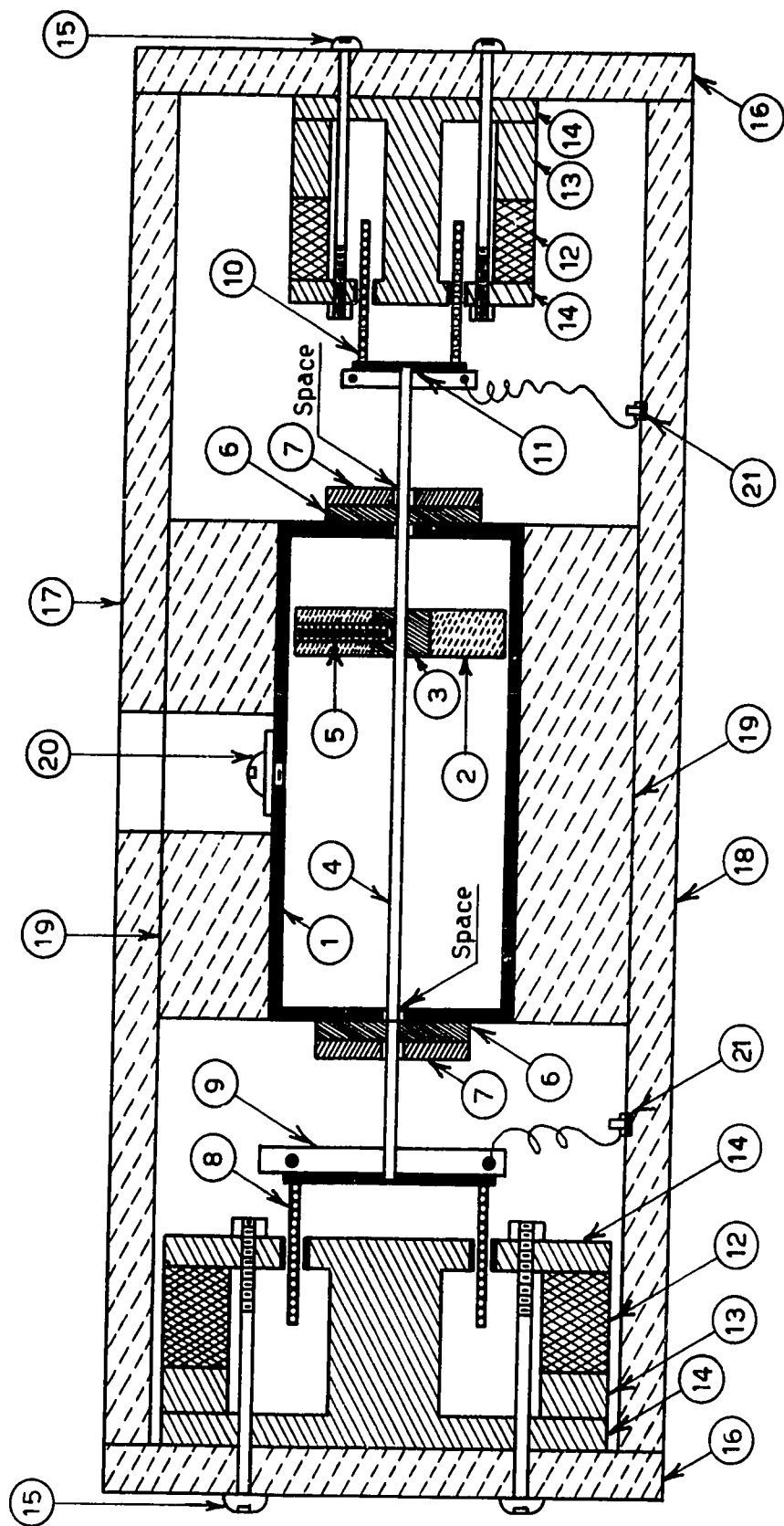


FIGURE 12. CROSS-SECTIONAL VIEW OF THE COMPLETE PHASESHIFTER

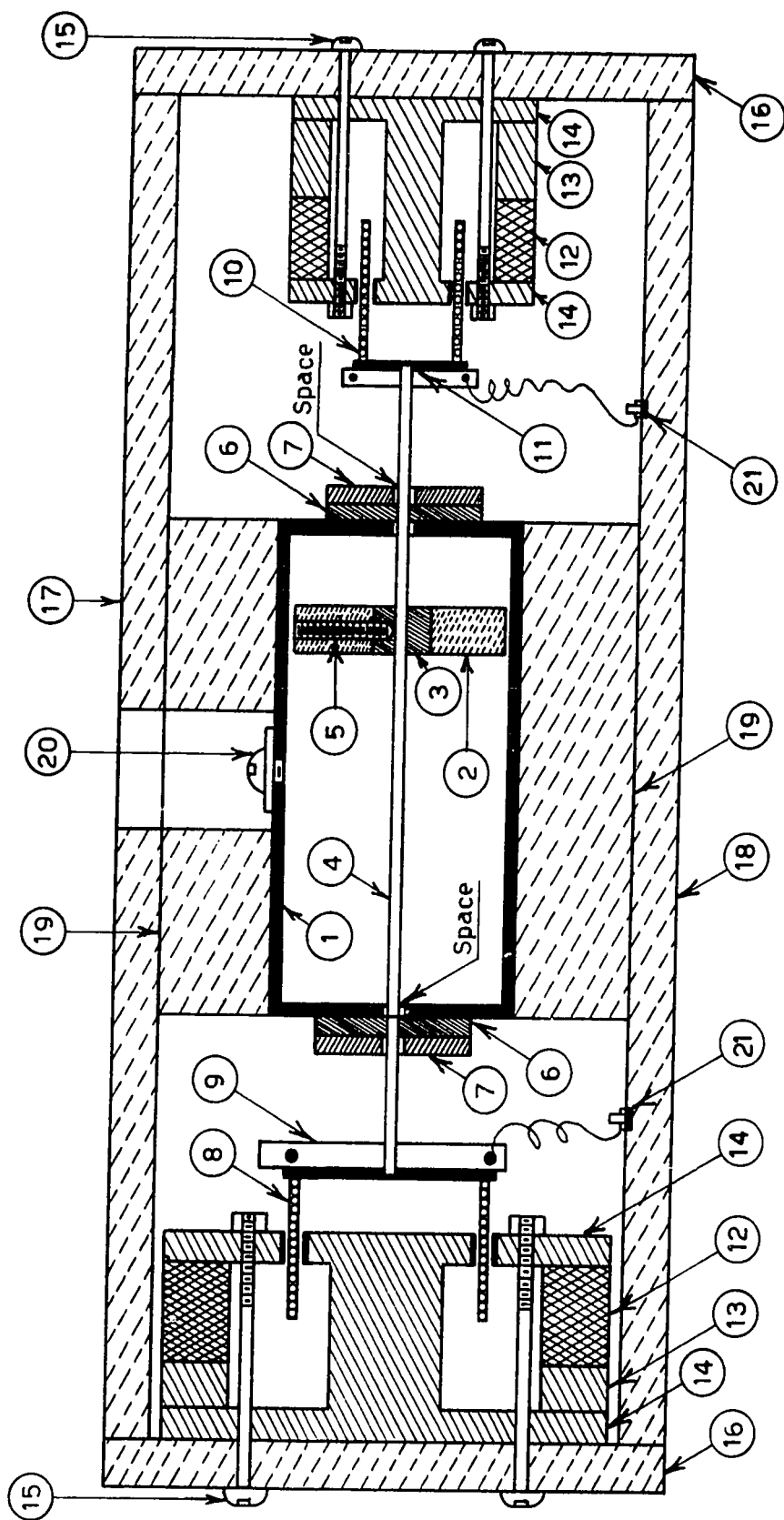


FIGURE 12. CROSS-SECTIONAL VIEW OF THE COMPLETE PHASESHIFTER

shifter, taken at a central plane through the pusher rod, is shown in figure 12. Figure 13 is a cross-sectional view, taken at a vertical plane through a support rod, to explain the method of attaching the support rods to the waveguide walls. Lack of space on the drawings makes it necessary to use numbers to refer to the various parts.

Referring to figure 12, waveguide section (1) contains dielectric slab (2) with teflon insert (3), fastened to pusher rod (4) by means of setscrew (5). The pusher rod protrudes from each sidewall of the waveguide and rests in bearing holes in teflon plates (6) which are held in place by brass plates (7) screwed to the waveguide sidewalls. One end of the pusher rod carries drive coil (8) attached to the

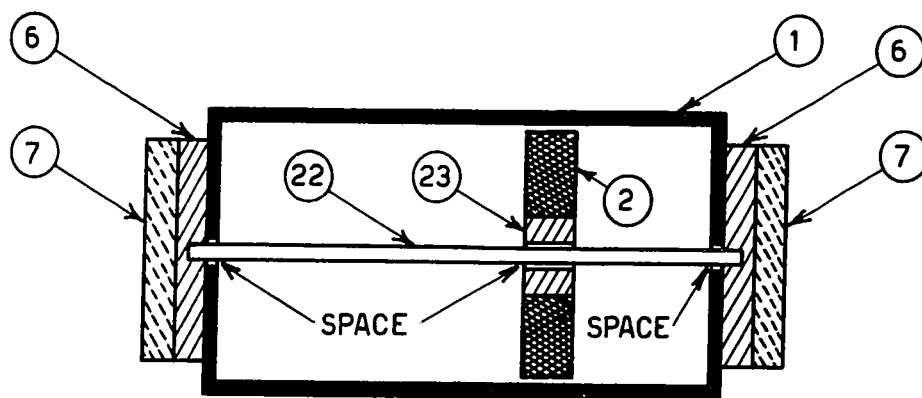


FIGURE 13. PARTIAL CROSS-SECTION OF SUPPORT ROD ARRANGEMENT

pusher rod by means of an end plate and stiffeners (9), while the other pusher rod end carries rate coil (10) with end plate (11). Each magnet structure, in which the coils move, consists of a magnet (12), a magnet extension ring (13), and polepieces (14). The magnet structures are

fastened to a support structure by means of bolts (15) passing through oversize holes in the support sideplates (16) to allow position adjustment of the magnets. The magnet support structure, consisting of support sideplates (16), a top plate (17), a bottom plate (18), and two bearing plates (19), is clamped to the waveguide section with four long bolts through the top and bottom plates next to the waveguide sidewalls. For clarity of the diagram, these bolts are not shown. Set-screw (5) may be reached through a hole in the top wall of the waveguide, which is later closed by means of a screw (20) which is flush with the inside surface of the waveguide top wall. The connections to the coils are brought out to terminals (21) mounted on the bottom support plate. The leads from the coils to the terminals consist of very fine copper wire spirals.

Figure 13 explains how the dielectric slab (2) in waveguide section (1) slides on support rod (22) by means of a teflon insert (23) to obtain low friction. The pusher rod of figure 12 had to bear on teflon only, and therefore, was made to pass through the waveguide sidewall in such a manner as to leave five one-thousandths of an inch space around the pusher rod. To maintain complete symmetry at $3\lambda_g/8$, the support rod (22) also passes through the sidewall of the waveguide with 0.005 inches of space around the rod. Support rod (22) in figure 13 is clamped to the waveguide sidewall as shown using the same teflon plates (6) which serve as bearings for the pusher rod. Brass plates (7) hold the teflon bearing plates in place. Finally, the screws that secure the brass and teflon plates to the waveguide (screws not shown) are ground flush with the inner surface of the waveguide sidewall.

2-6 Phaseshifter Performance

It has already been mentioned that the phaseshifter insertion loss is low, and that the voltage standing wave ratio for the phaseshifter is better than 1.1. Moreover, the reflections caused by the phaseshifter tend to be almost purely reactive. The shorted stub in the matching device cancels the imaginary part of the load admittance by adding the appropriate amount of reactive admittance of opposite sign. The stub can, therefore, also cancel the small reactive part contributed by the phaseshifters.

The currents generated in the pusher rods and support rods have no axial components along the rods. At the same time, the rod support mechanism consisting of the waveguide sidewall holes, the teflon plates and the brass plates acts, together with the rods, like a very low impedance mismatched coaxial line section. As a result there is practically no leakage of microwave signal from the phaseshifters.

Great care has been taken to make the entire phaseshifter as linear as possible, but, equations 2.14 and 2.15 show that the phaseshift obtained from the phaseshifter varies as a squared sine function of the slab displacement. Nevertheless, measurement has shown that the finite thickness of the dielectric slab has an averaging effect over the small displacement range necessary to obtain phaseshifts of plus and minus 90° . In addition, the 180° phaseshift range is only a part of the total phaseshift available from the device. Consequently, the phaseshift appears to be an almost linear function of slab displacement. In the system analysis work of later chapters, it will be assumed that the phaseshift is indeed a linear function of slab displacement. This will

clearly not introduce objectionable inaccuracies.

CHAPTER THREE

MEASURING DEVICE DEVELOPMENT

The measuring device, which develops the control loop error voltages as functions of the loading conditions, is a vital part of the automatic matching device. As previously discussed, the measuring device will use the standing wave pattern to extract signals proportional to the real and imaginary parts of the reflection coefficient.

Relationships, expressing the output voltage from a single microwave crystal detector in terms of the reflection coefficient and the detector position along the waveguide, will first be developed. Using these results, the proper waveguide positions of three probes for obtaining reflection coefficient information will be selected. Finally, the measuring device will be designed.

Two separate cases will be considered, E-field sampling, and H-field sampling. The H-field sampling method will be discussed in greater detail, since it is more suitable for high power systems.

3-1 Single Detector E-Field Sampling

The probe section analysis is most easily carried out by establishing an analogy between waveguide transmission on the one hand, and wire line transmission on the other. Such an analogy has, in effect, already been used in discussing waveguide matching problems with the aid of a Smith chart.

Without loss of generality, it may be assumed that the voltage ap-

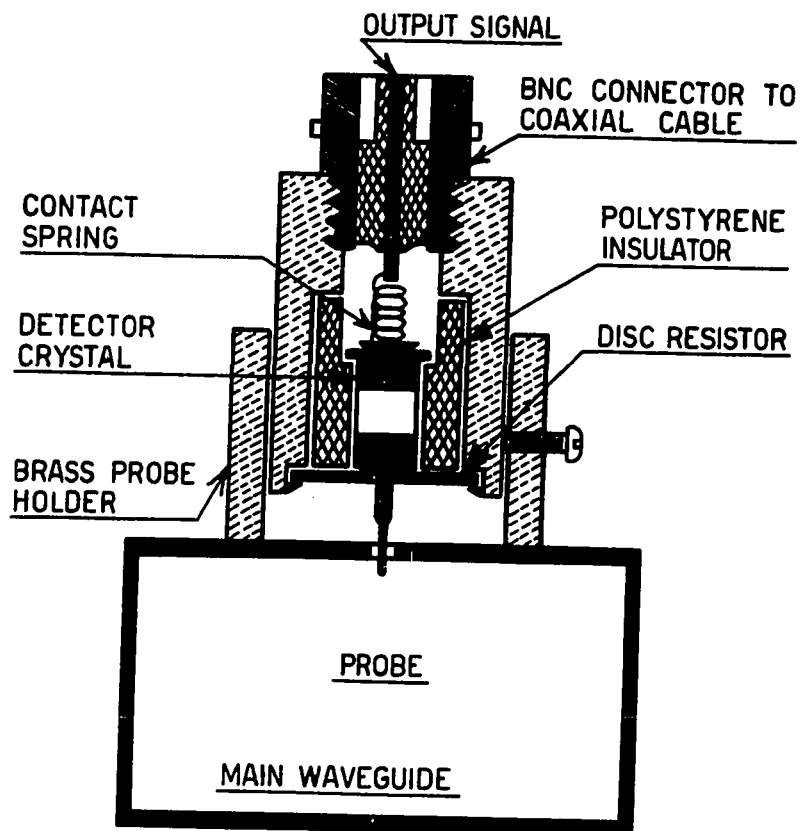


FIGURE 14. AN E-FIELD PROBE AND DETECTOR

pearing on a vertical E-plane probe, in the center of the top wall of a waveguide operating in the TE_{10} mode, is analogous to the voltage on a wire transmission line. If the probe penetration is properly adjusted to sample only a small part of the total E-field, the voltage on such a probe may be rectified and detected by a microwave crystal detector operating in its square law region, to give a crystal output voltage that is proportional to the square of the E-field magnitude in the center of the waveguide.

A typical E-field sampling probe and detector assembly is shown in

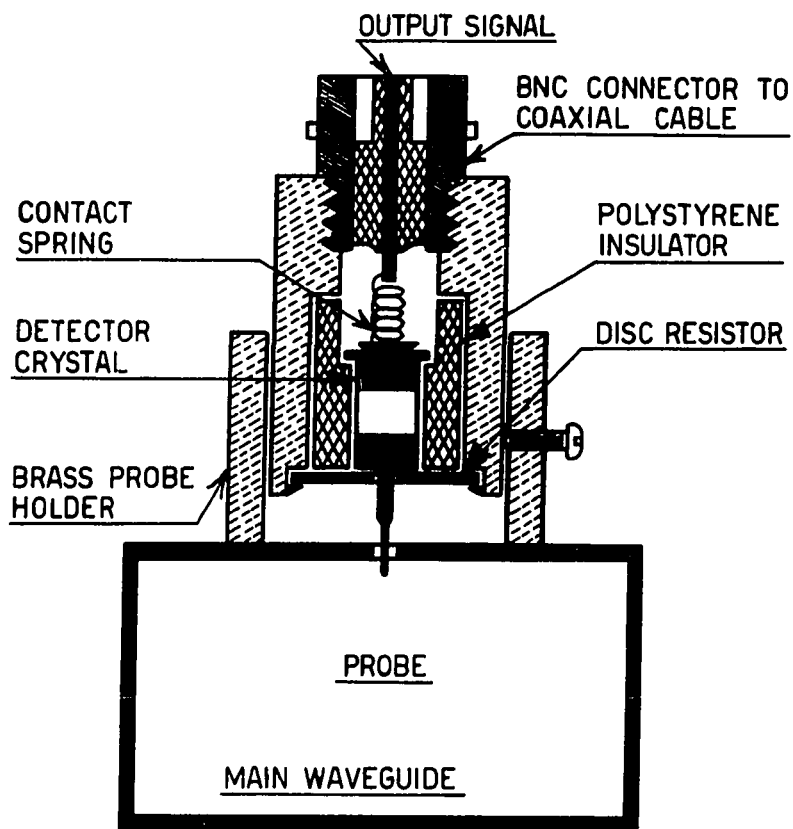


FIGURE 14. AN E-FIELD PROBE AND DETECTOR

pearing on a vertical E-plane probe, in the center of the top wall of a waveguide operating in the TE_{10} mode, is analogous to the voltage on a wire transmission line. If the probe penetration is properly adjusted to sample only a small part of the total E-field, the voltage on such a probe may be rectified and detected by a microwave crystal detector operating in its square law region, to give a crystal output voltage that is proportional to the square of the E-field magnitude in the center of the waveguide.

A typical E-field sampling probe and detector assembly is shown in

cross-section in figure 14, and is self-explanatory.

Returning now to the wire transmission line analogy, the output voltage from the crystal detector may be expressed in terms of the incident voltage, the load reflection coefficient, and the distance between the load and the probe. The relationships between these quantities may be expressed first by means of a vector diagram as shown in figure 15.

In figure 15, let the incident voltage at the load be E_R^+ and let the reflected voltage be E_R^- . Then:

$$E_R^- = k_R E_R^+ = K_R E_R^+ e^{j\theta_R} \quad 3.1$$

where K_R is the magnitude of the reflection coefficient, k_R , and θ_R is the angular part of the reflection coefficient.

Now, let the incident voltage at a distance d from the load be E^+ , and let the reflected voltage at a distance d from the load be E^- . A phase difference will occur between the voltages at the load and those measured at a distance d from the load, which may be expressed by the following relations:

$$E^+ = E_R^+ e^{j\beta d} \quad 3.2$$

$$E^- = E_R^- e^{-j\beta d} \quad 3.3$$

where β is the propagation constant for a lossless line. Substituting

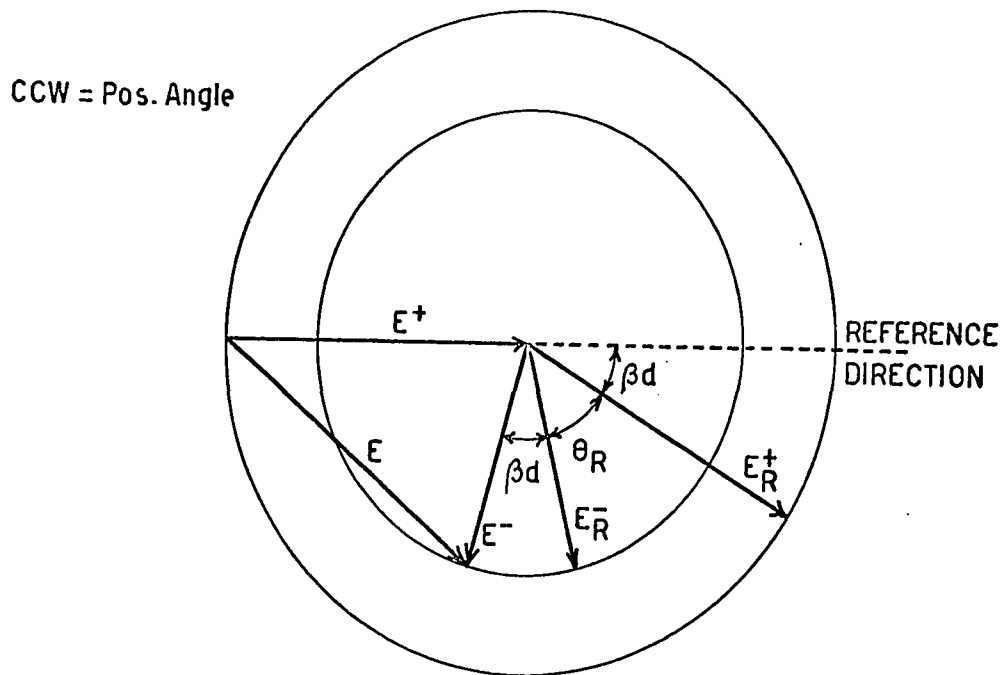


FIGURE 15. THE REFLECTION COEFFICIENT VECTOR DIAGRAM FOR VOLTAGE

equation 3.1 into equation 3.3 and then combining equations 3.2 and 3.3 gives:

$$E^- = K_R E^+ e^{j(\theta_R - 2\beta d)} \quad 3.4$$

The total voltage at the probe, denoted by E , is the vector sum of E^+ and E^- . Application of the cosine rule to the vector diagram of figure 15 allows E to be expressed as follows:

$$\begin{aligned} (E)^2 &= (E^+)^2 + (E^-)^2 - 2(E^+)(E^-) \cos[180^\circ - (\theta_R - 2\beta d)] = \\ &= (E^+)^2 + (E^-)^2 + 2(E^+)(E^-) \cos(\theta_R - 2\beta d) \end{aligned} \quad 3.5$$

Substituting of equation 3.4 into 3.5, and considering voltage magnitudes only gives:

$$|E|^2 = |E^+|^2 + K_R^2 |E^+|^2 + 2K_R |E^+|^2 \cos(\theta_R - 2\beta d) \quad 3.6$$

or:

$$\frac{|E|^2}{|E^+|^2} = 1 + K_R^2 + 2K_R \cos(\theta_R - 2\beta d) \quad 3.7$$

For a fixed source power level, the diode sensitivity and the probe attenuation may be combined into a proportionality factor η_1 . Since the detector DC output voltage, E_d , is proportional to the square of the voltage applied to it by the probe, one finally has:

$$E_d = \eta_1 [1 + K_R^2 + 2K_R \cos(\theta_R - 2\beta d)] \quad 3.8$$

Equation 3.8 expresses the microwave crystal detector output voltage as a function of the reflection coefficient magnitude K_R , the reflection coefficient angle θ_R , the position of the probe along the transmission line, d , and the probe and diode sensitivity η_1 . This is the desired result.

3-2 Single Detector H-Field Sampling

In view of the fact that, for a TE_{10} mode of rectangular waveguide operation, the magnetic field has a transverse and a longitudinal component, two methods of H-field sampling may be employed. The trans-

verse H-field component may be sampled by means of a pick-up loop directed along a longitudinal vertical plane, while the longitudinal H-field component may be sampled with a vertical loop in a transverse plane.

By application of Poynting's vector theorem ⁽⁸⁾ to a waveguide section, it may be shown that only the transverse magnetic field component is involved in power transfer along the waveguide. However, direct sampling of the transverse magnetic field component implies that a discontinuity must be introduced into the waveguide through the top wall, resulting in areas of electric field concentration. This is undesirable at high microwave power levels.

Let one examine the magnetic field components in a rectangular waveguide operating in the TE₁₀ mode. Referring to figure 16, the magnetic field components are:

$$H_x = A \frac{\pi}{b} \sin\left(\frac{\pi x}{b}\right) e^{j(\omega t - \beta_0 z)} \quad 3.9$$

$$H_z = -jA \frac{k_c^2}{\beta_0} \cos\left(\frac{\pi x}{b}\right) e^{j(\omega t - \beta_0 z)} \quad 3.10$$

where:

H is the magnetic field intensity.

A is an amplitude constant.

β_0 is the guide propagation constant at a certain frequency.

k_c is the waveguide cut-off propagation constant.

If the z-component of the magnetic field is sampled with a proper-

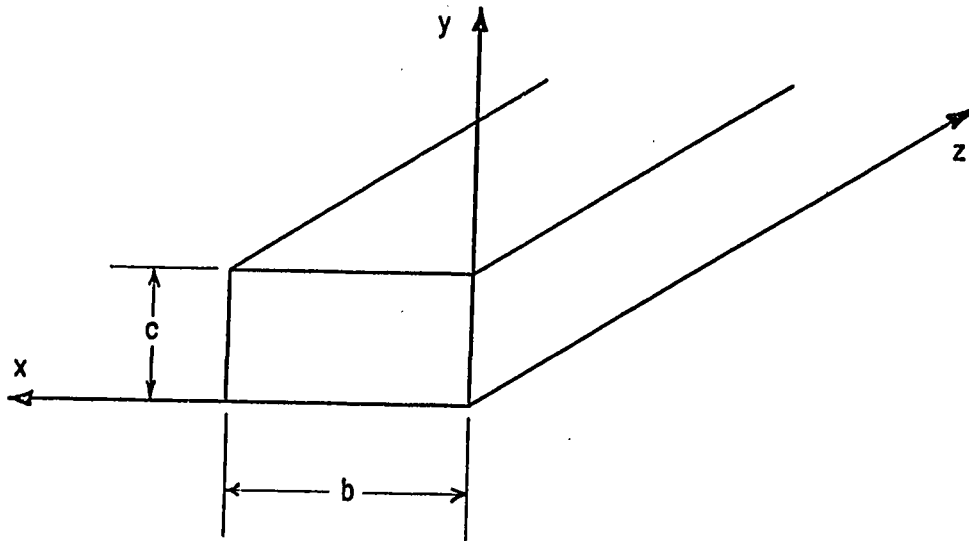


FIGURE 16. A RECTANGULAR WAVEGUIDE SECTION

ly oriented pick-up loop, the voltage generated in the loop will be proportional to the time rate of change of H_z . Thus, from 3.10:

$$\begin{aligned} \frac{\partial H_z}{\partial t} &= \frac{\partial}{\partial t} \left[-jA \frac{k_c^2}{\beta_0} \cos\left(\frac{\pi x}{b}\right) e^{j(\omega t - \beta_0 z)} \right] \\ &= A \frac{k_c^2 \omega}{\beta_0} \cos\left(\frac{\pi x}{b}\right) e^{j(\omega t - \beta_0 z)} \end{aligned} \quad 3.11$$

From equation 3.9, the transverse H-field component involved in the transfer of power, at the center of the waveguide, is:

$$H_x \Big|_{x=\frac{b}{2}} = A \frac{\pi}{b} e^{j(\omega t - \beta_0 z)} \quad 3.12$$

Similarly, from expression 3.11, the time rate of change of H_z , at a sidewall of the waveguide, is proportional to the voltage generated in a properly oriented pick-up loop at the waveguide sidewall. This time rate of change is:

$$\left. \frac{\partial H_z}{\partial t} \right|_{x=0} = A \frac{k_c^2 \omega}{\beta_0} e^{j(\omega t - \beta_0 z)} \quad 3.13$$

Comparison of expressions 3.12 and 3.13 shows that, provided the power source frequency and the waveguide dimensions are fixed, the time rate of change of H_z at a waveguide sidewall is simply related by a constant to H_x in the center of the waveguide. The sidewall pick-up loop voltage is thus proportional to the transverse component of the magnetic field.

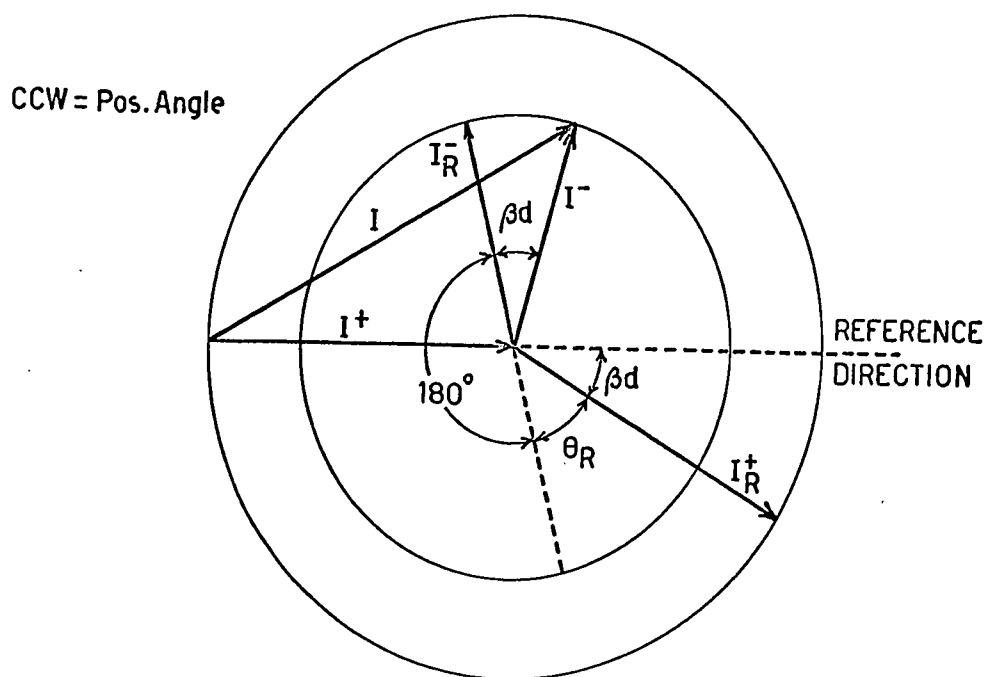


FIGURE 17. THE REFLECTION COEFFICIENT VECTOR DIAGRAM FOR CURRENT

The actual method of sampling the longitudinal H-field component at a sidewall, where the electric field is zero and will cause no arcing problems, will be discussed in detail later.

The transverse magnetic field component in a rectangular waveguide operating in the TE_{10} mode may, again, be considered to be entirely analogous to the current in a wire transmission line. And again, by use of transmission line theory, an expression may be derived for the DC voltage obtained from a single H-field sampling detector probe (loop) and crystal, in terms of the load reflection coefficient magnitude and angle, the distance between the load and probe, and the combined sensitivities of the probe and detector diode.

Starting with a vector diagram as shown in figure 17, the currents on the line may be expressed as follows. Let the incident current at the load be I_R^+ and let the reflected current at the load be I_R^- . Then:

$$I_R^- = -K_R I_R^+ e^{j\theta_R} = K_R I_R^+ e^{j(180^\circ + \theta_R)} \quad 3.14$$

where K_R is the magnitude of the reflection coefficient, and θ_R is its angle

At a distance d from the load, and for a propagation constant β , the incident current, I^+ , and the reflected current, I^- , may be written as follows:

$$I^+ = I_R^+ e^{j\beta d} \quad 3.15$$

$$I^- = I_R^- e^{-j\beta d} \quad 3.16$$

Combining the results of expressions 3.14, 3.15, and 3.16 gives finally:

$$I^- = K_R I^+ e^{j(180^\circ + \theta_R - 2\beta d)} \quad 3.17$$

Then, by application of the cosine rule to the vector diagram of figure 17:

$$(I)^2 = (I^+)^2 + (I^-)^2 - 2(I^+)(I^-) \cos(\theta_R - 2\beta d) \quad 3.18$$

Substituting of equation 3.17 into 3.18, and considering current magnitudes only, results in:

$$|I|^2 = |I^+|^2 + K_R^2 |I^+|^2 - 2 K_R |I^+|^2 \cos(\theta_R - 2\beta d) \quad 3.19$$

Equation 3.19 may be normalized to the square of the incident current.

Doing this gives:

$$\frac{|I|^2}{|I^+|^2} = 1 + K_R^2 - 2 K_R \cos(\theta_R - 2\beta d) \quad 3.20$$

Now returning to waveguide transmission, the sensitivities of the detector crystal operating in its square law region, the attenuation of the probe assembly, and the proportionality factor for conversion from transverse to longitudinal H-field may be combined into a factor η_2 , to yield:

$$E_d = \eta^2 \left[1 + \kappa_R^2 - 2 \kappa_R \cos(\theta_R - 2\beta d) \right] \quad 3.21$$

as the final expression for the DC voltage, E_d , obtained from the detector diode.

3-3 Three Detector E-Field Sampling

Three E-field sampling detector probes may be so placed along a waveguide that the DC output voltages can be combined to yield the real and imaginary components of the reflection coefficient. The derivation of this result follows.

Let three separate E-field probes be positioned along a waveguide

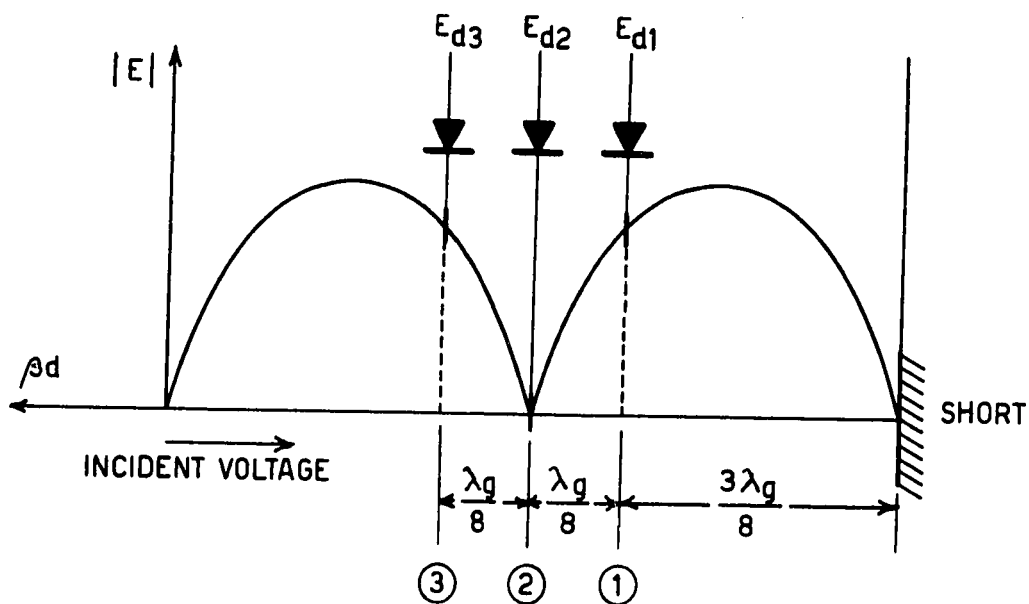


FIGURE 18. E-FIELD PROBE PLACEMENT

as shown in figure 18. The reference is taken to be a short-circuit at the plane of the load, although an arbitrary number of half wavelengths may be added between the probes and the load. The short is used only to define the position of the minima. The distances from the load to the probes, shown in guide wavelengths in figure 18, may be converted to electrical degrees. These distances in degrees are:

$$\begin{aligned}
 \text{Probe 1:} & \quad 135^\circ \text{ or } 2/\beta d_1 = 270^\circ \\
 \text{Probe 2:} & \quad 180^\circ \text{ or } 2/\beta d_2 = 360^\circ \\
 \text{Probe 3:} & \quad 225^\circ \text{ or } 2/\beta d_3 = 450^\circ
 \end{aligned}
 \tag{3.22}$$

By substituting the results of 3.22 into equation 3.8, the expressions for the DC output voltages from the detector diodes become:

$$\begin{aligned}
 E_{d1} &= \eta_1 [1 + K_R + 2 K_R \cos(\theta_R - 270^\circ)] \\
 &= \eta_1 [1 + K_R - 2 K_R \sin \theta_R]
 \end{aligned}
 \tag{3.23}$$

$$\begin{aligned}
 E_{d2} &= \eta_1 [1 + K_R + 2 K_R \cos(\theta_R - 360^\circ)] \\
 &= \eta_1 [1 + K_R + 2 K_R \cos \theta_R]
 \end{aligned}
 \tag{3.24}$$

$$\begin{aligned}
 E_{d3} &= \eta_1 [1 + K_R + 2 K_R \cos(\theta_R - 450^\circ)] \\
 &= \eta_1 [1 + K_R + 2 K_R \sin \theta_R]
 \end{aligned}
 \tag{3.25}$$

Forming [(3.25)-(3.23)] gives:

$$\begin{aligned}
 (E_{d3} - E_{d1}) &= \eta_1 [2 K_R \sin \theta_R + 2 K_R \sin \theta_R] \\
 &= 4 \eta_1 K_R \sin \theta_R
 \end{aligned}
 \tag{3.26}$$

Since $(K_R \sin \theta_R)$ is equal to K_{RV} , the imaginary component of the reflection coefficient, one has:

$$K_{RV} = \frac{E_{d3} - E_{d1}}{4\eta_1} \quad 3.27$$

Equation 3.27 clearly shows that the difference in DC output voltages from detector crystals one and three is proportional to the imaginary component of the reflection coefficient.

Similarly, forming $[-(3.23)-(3.25)+2(3.24)]$ produces:

$$\begin{aligned} -E_{d1} - E_{d3} + 2E_{d2} &= \\ &= \eta_1 [2K_R \sin \theta_R - 2K_R \sin \theta_R + 4K_R \cos \theta_R] \\ &= 4\eta_1 K_R \cos \theta_R \end{aligned} \quad 3.28$$

The quantity $(K_R \cos \theta_R)$ is equal to K_{RH} , the real component of the reflection coefficient. Thus:

$$K_{RH} = \frac{2E_{d2} - E_{d1} - E_{d3}}{4\eta_1} \quad 3.29$$

It is evident from equation 3.29 that twice the output voltage from probe two minus the sum of the output voltages from detectors one and three is directly proportional to the real component of the reflection coefficient.

The above illustrates how the reflection coefficient may be meas-

ured simply and conveniently, using three E-field sampling detector probes.

3-4 Three Detector H-Field Sampling

Three H-field sampling probes may equally readily be positioned along a waveguide such that the reflection coefficient can be measured. The derivation is analogous to that of the three probe E-field sampling arrangement.

Figure 19 shows the special placement of three H-field detector probes along a waveguide, where a reference short at the plane of the load is again used to define the positions of the minima. The spacing between the probes is $\lambda_g/8$ as before, but since the current at a short is maximum, the distance from the load to the first minimum is now

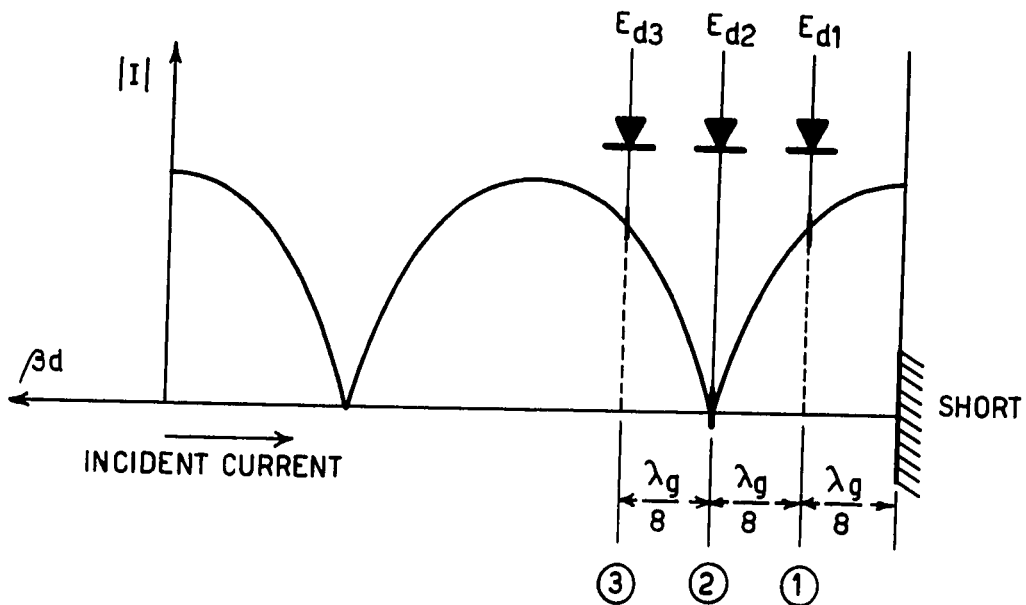


FIGURE 19. H-FIELD PROBE PLACEMENT

$[(\lambda_g/4)+n(\lambda_g/2)]$, where n is an arbitrary positive integer.

The distances from the load to the probes, if translated into electrical degrees are:

$$\begin{aligned}
 \text{Probe 1 :} & \quad 45^\circ \text{ or } 2/\beta d_1 = 90^\circ \\
 \text{Probe 2 :} & \quad 90^\circ \text{ or } 2/\beta d_2 = 180^\circ \\
 \text{Probe 3 :} & \quad 135^\circ \text{ or } 2/\beta d_3 = 270^\circ
 \end{aligned} \tag{3.30}$$

Substitution of the values for $(2/\beta d)$ from 3.30 into equation 3.21 gives the detector output voltages as:

$$\begin{aligned}
 E_{d1} &= \eta_2 [1 + K_R^2 - 2K_R \cos(\theta_R - 90^\circ)] \\
 &= \eta_2 [1 + K_R^2 - 2K_R \sin \theta_R]
 \end{aligned} \tag{3.31}$$

$$\begin{aligned}
 E_{d2} &= \eta_2 [1 + K_R^2 - 2K_R \cos(\theta_R - 180^\circ)] \\
 &= \eta_2 [1 + K_R^2 + 2K_R \cos \theta_R]
 \end{aligned} \tag{3.32}$$

$$\begin{aligned}
 E_{d3} &= \eta_2 [1 + K_R^2 - 2K_R \cos(\theta_R - 270^\circ)] \\
 &= \eta_2 [1 + K_R^2 + 2K_R \sin \theta_R]
 \end{aligned} \tag{3.33}$$

By forming [(3.33)-(3.31)] one obtains:

$$\begin{aligned}
 E_{d3} - E_{d1} &= \eta_2 [2K_R \sin \theta_R + 2K_R \sin \theta_R] \\
 &= 4\eta_2 K_R \sin \theta_R
 \end{aligned} \tag{3.34}$$

But, K_{RV} , the imaginary component of the reflection coefficient, equals

$(K_R \sin \theta_R)$. Equation 3.34 can thus be expressed as:

$$K_{RV} = \frac{E_{d3} - E_{d1}}{4\eta_2} \quad 3.35$$

Also, taking $[-(3.31)-(3.33)+2(3.32)]$ one obtains:

$$\begin{aligned} -E_{d1} - E_{d3} + 2E_{d2} &= \\ &= \eta_2 [-2K_R \sin \theta_R + 2K_R \sin \theta_R + 4K_R \cos \theta_R] \\ &= 4\eta_2 K_R \cos \theta_R \end{aligned} \quad 3.36$$

Again, $(K_R \cos \theta_R)$ equals the real part of the reflection coefficient, and therefore, equation 3.36 becomes:

$$K_{RH} = \frac{2E_{d2} - E_{d1} - E_{d3}}{4\eta_2} \quad 3.37$$

These results are the same as those derived for the E-field probes.

It should be recognized that the results derived above are correct only if a set of three perfectly matched detector crystals, operating in their square law regions, is used. However, use of a closed loop system for automatic matching implies that deviations from square law operation of the crystals, which might occur during matching of highly mismatched loads, are of little consequence. The main feature of the

measuring device must be that, under matched conditions over a range of power levels, the error voltages generated are zero. Differences in sensitivities and tracking errors of the crystals may be compensated for, over a limited range of microwave power levels, by static adjustment of the individual probe attenuators.

3-5 H-Field Probe Coupling and Attenuation

The desirability of sampling the longitudinal magnetic field component in a rectangular waveguide operated in a TE_{10} mode was discussed in a previous section.

The microwave power level at which the matching device is expected to be used is 1 kW or more. However, normal microwave detector crystals must be operated at power levels of about 1 mW in order to insure square law response from these crystals. An attenuation of at least 60 dB is, therefore, required between the main waveguide and the detector probe pick-up loop. Because the microwave power level is not

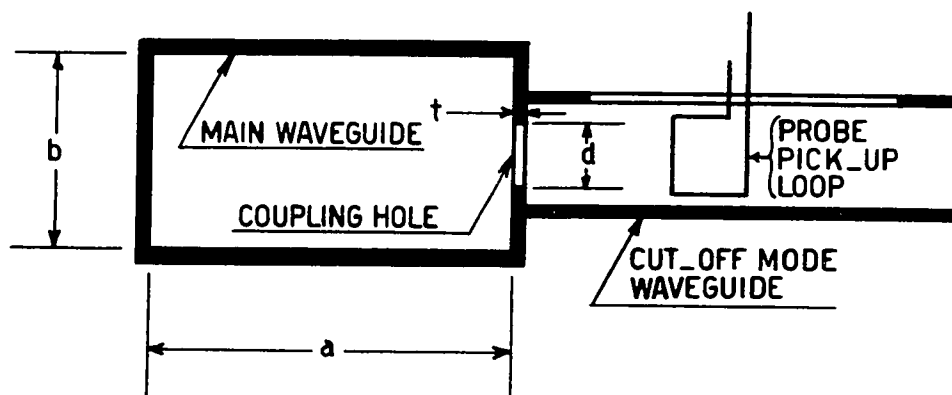


FIGURE 20. THE H-FIELD PROBE AND ATTENUATOR

definite, and because the detector crystal sensitivities and tracking must be compensated for, it is furthermore required that the attenuation between the main guide and the probe be adjustable.

The stated requirements are most easily met by designing an adjustable position probe which samples the magnetic field in a cut-off waveguide, while coupling main guide energy into the cut-off guide by means of a main waveguide sidewall coupling hole. Figure 20 schematically illustrates the probe arrangement.

Consider, first, the sidewall coupling hole. Bethe has developed design formulae for the attenuation through coupling holes in waveguide walls. His work is presented in several reference texts such as, for instance, Montgomery⁽⁹⁾. Bethe's expression for the attenuation through a single sidewall coupling hole (see figure 20) is:

$$C = 20 \log_{10} \left[\frac{\pi d^3 \lambda_g}{12 a^3 b} - 16 \left(1 - \frac{1.71 d}{\lambda} \right)^2 \right]^{1/2} \cdot \frac{t}{d} \quad 3.38$$

where:

C is the coupling factor for the coupling hole, given in dB.

d is the coupling hole diameter in cm.

t is the main guide sidewall thickness in cm.

a is the main guide width in cm.

b is the main guide height in cm.

λ_g is the guide wavelength in cm.

λ is the free space wavelength in cm.

Expression 3.38 allows one to find the hole diameter, d, necessary to give a fixed attenuation of about 60 dB for a specific size of

main waveguide and a fixed frequency. Only the longitudinal magnetic field component will couple through the hole.

Further adjustable attenuation may be obtained from the cut-off mode rectangular waveguide shown in figure 20, into which coupling takes place. The coupling hole will attempt to excite a TE_{10} mode in the cut-off guide, but no electromagnetic wave can be propagated by this guide. Instead, the fields will decay exponentially, away from the excitation point, thus producing a constant logarithmic attenuation per unit length of cut-off guide.

Montgomery⁽⁹⁾ gives the design formulae for cut-off mode waveguide. They are:

$$\alpha = 8.686 \frac{2\pi}{\lambda_c} \left[1 - \left(\frac{\lambda_c}{\lambda} \right)^2 \right]^{1/2} \quad 3.39$$

and:

$$\lambda_c = \frac{2(\epsilon')^{1/2}}{\left[\left(\frac{m}{a} \right)^2 + \left(\frac{n}{b} \right)^2 \right]^{1/2}} \quad 3.40$$

where:

α is the attenuation constant in dB/cm.

λ_c is the cut-off wavelength in cm.

λ is the free space wavelength in cm.

ϵ' is the relative dielectric constant of the material filling the guide.

a is the width of the waveguide in cm.

b is the height of the guide in cm.

m and n are the mode numbers of the wave in the guide.

The cut-off wavelength, λ_c , may be obtained from equation 3.40, for a particular guide size. This value may then be substituted into expression 3.39 to obtain α . The attenuation may readily be made adjustable over a 40 dB range, so that practically any power level can be accommodated.

3-6 The Complete Three Probe H-Field Measuring Section

As discussed previously, resolution of the voltages measured by the three probe device, into signals proportional to the real and imaginary components of the load reflection coefficient, requires special positioning of the probes with respect to the load. In fact, of the three probes at $\lambda_g/8$ spacing, the central one should be located at a standing wave minimum as defined by a short substituted for the load at the load reference plane. Moreover, the definition of a load reference plane position directly affects the angular part of the reflection coefficient as measured.

In a closed-loop system for automatic matching, the load comprises the actual load and, also, the matching device (a single stub tuner). In defining a reference plane for the combined load and matching device, consideration should be given to the fact that the signal, which is proportional to the imaginary part of the reflection coefficient, will be used as the error voltage for the control loop that adjusts the stub. The stub affects only the imaginary part of the total

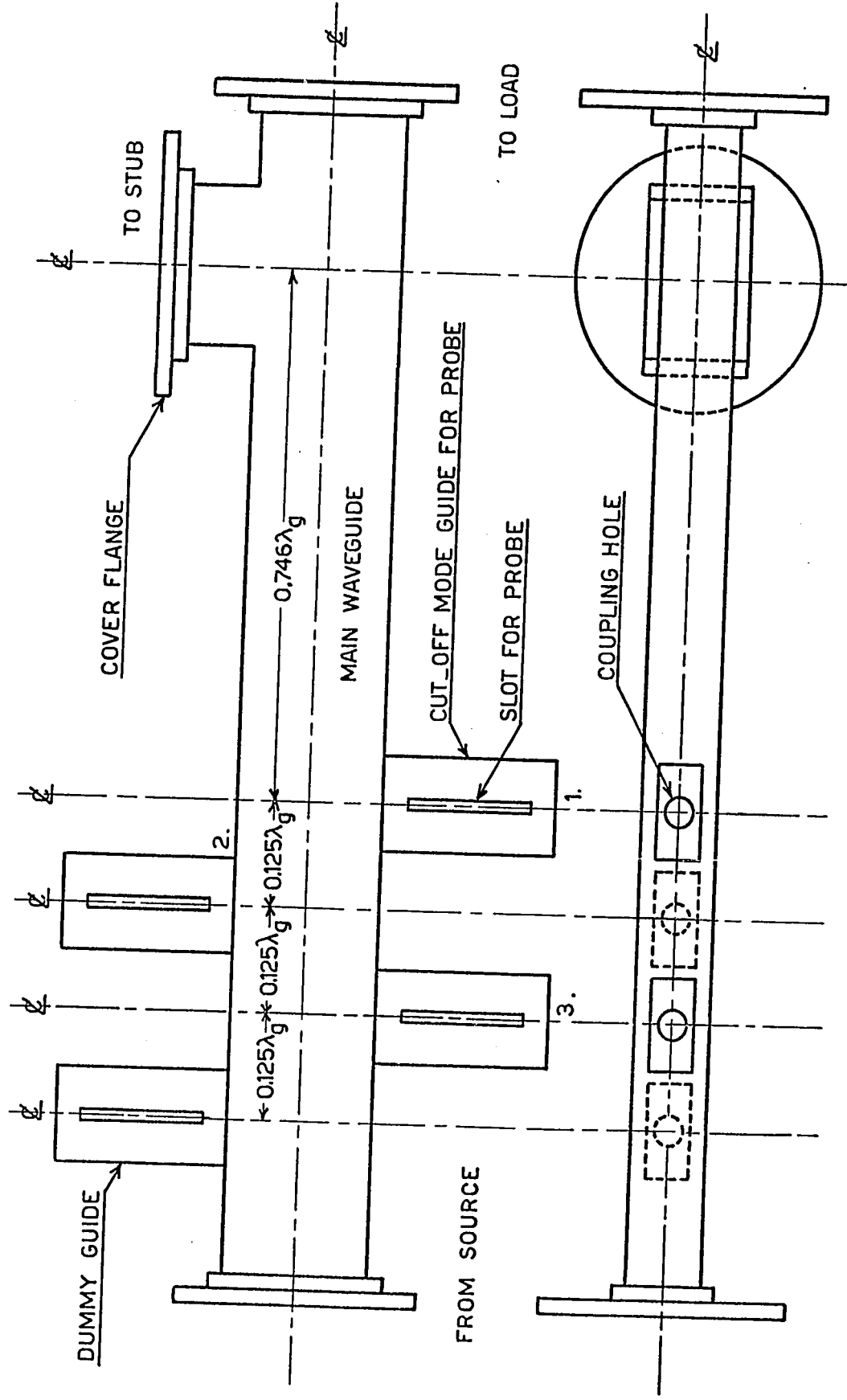


FIGURE 21. THE H-PLANE TEE AND MEASURING SECTION

load admittance. The combined load reference plane should, therefore be chosen at the null plane of the stub H-plane tee.

Altman⁽⁸⁾, and also Montgomery⁽⁹⁾, show that the actual position of the null plane for a stub and H-plane tee does not coincide with the vertical plane of symmetry through the stub, because the tee forms an extended discontinuity in the main waveguide. If the stub is so adjusted that it appears to place a short across the main waveguide, the first minimum in the H-field standing wave pattern will not occur at a distance of $\lambda_g/4$ away from the stub plane of symmetry. Rather, that distance will be somewhat less than $\lambda_g/4$.

Measurement has shown that for the H-plane tee and stub used, the distance from the plane of symmetry to the first H-field null is $0.246\lambda_g$. Later analysis will demonstrate that small errors in this distance are not detrimental.

Since the distances from the probes to the H-plane tee are fixed for a fixed microwave power source frequency, it seems reasonable to combine the measuring section and H-plane tee into one unit.

Figure 21 shows the complete measuring section and tee arrangement as discussed. To conserve space, the cut-off mode guides are alternately mounted on opposite sidewalls of the main waveguide. Furthermore, a fourth coupling hole and cut-off mode guide (a dummy) are included at a distance of $\lambda_g/8$, to help cancel any reflections caused by the measuring section.

A detailed cross-sectional view of one probe and detector crystal, showing the adjustable mounting method on the cut-off guide, is included as figure 22. The probe pick-up loop is inserted into the cut-

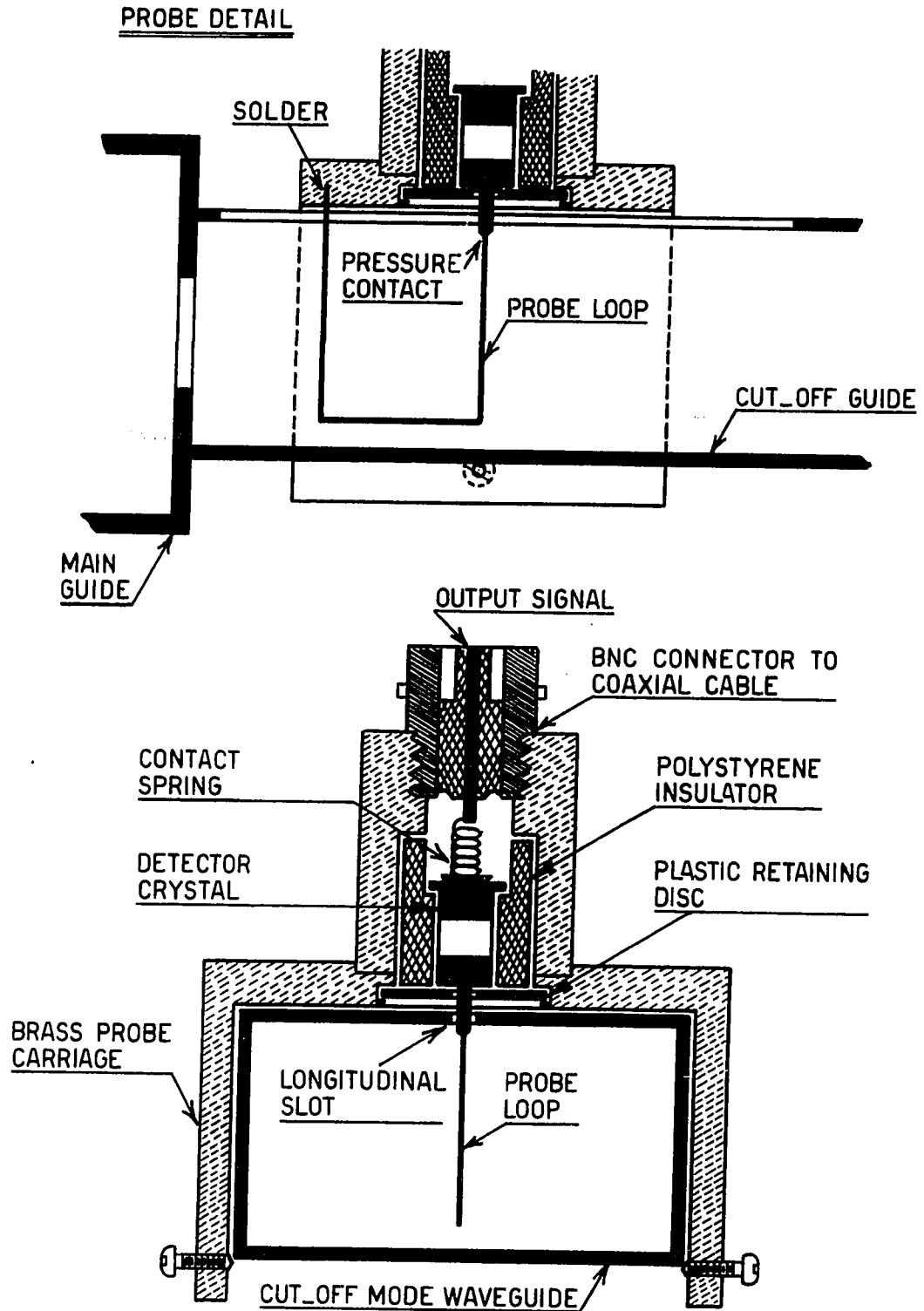


FIGURE 22. H-FIELD SAMPLING PROBE ON A CUT-OFF MODE WAVEGUIDE

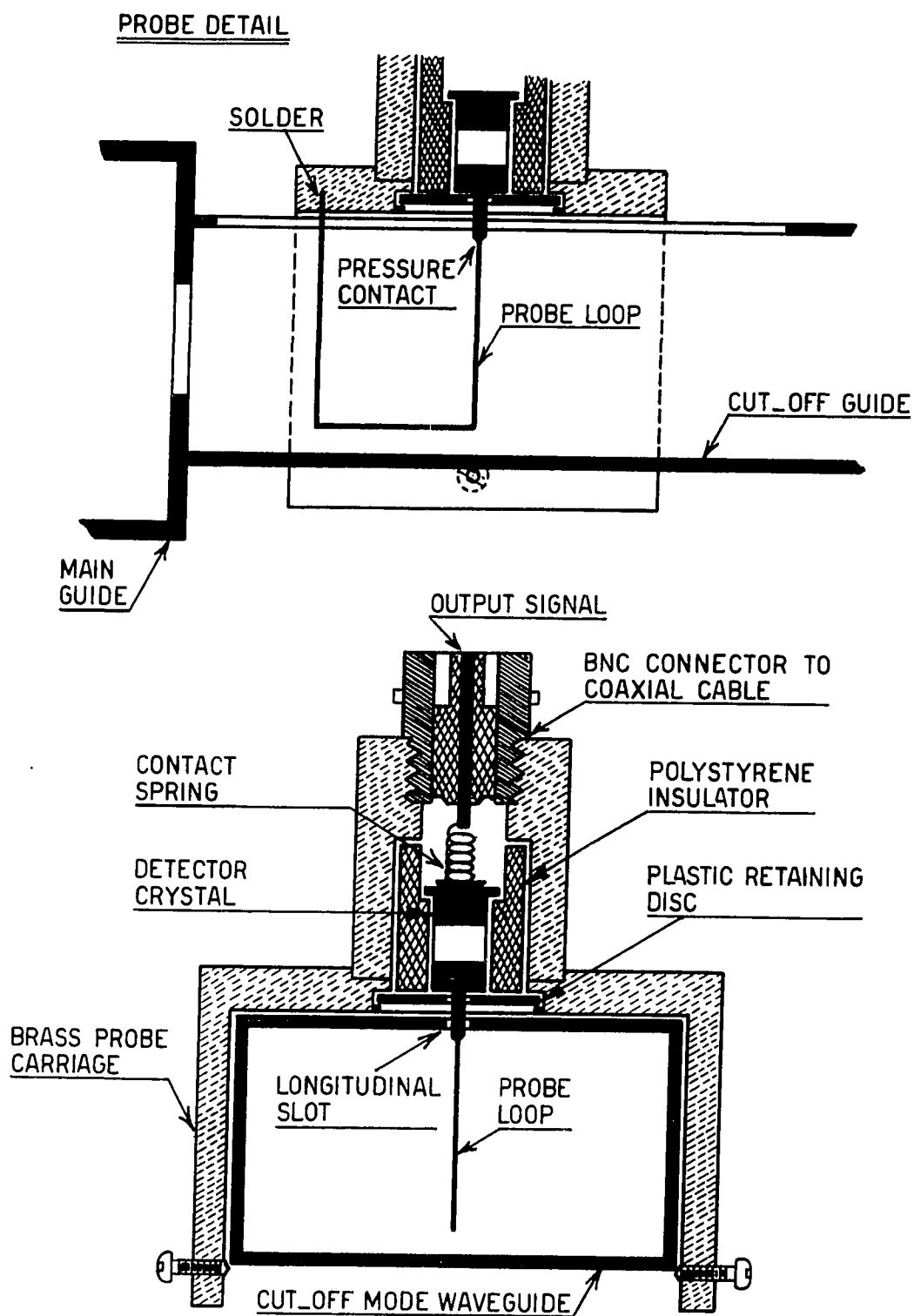


FIGURE 22. H-FIELD SAMPLING PROBE ON A CUT-OFF MODE WAVEGUIDE

off mode waveguide through a central longitudinal slot in the top wall of the cut-off guide.

Further constructional details of the measuring section and H-plane tee should be obvious from the diagrams of figures 21 and 22.

This completes the discussion of the automatic matching device microwave components. Subsequent chapters will deal with the analysis of the closed-loop system, and will describe a specific implementation of the automatic matching device. Some consideration will also be given to the electronic part of the system.

CHAPTER FOUR

THE CLOSED LOOP SYSTEM

The components described in previous chapters may be combined with some control electronics, for amplifying and conditioning of the error signals, to form an automatic microwave load matching device.

However, the automatic matching apparatus will be useful only if:

1. The control loops are stable under all microwave load conditions; i.e. a stable equilibrium can be reached for all microwave loads, and the control loop output variables remain bounded during load variations.
2. The equilibrium values of the output variables corresponding to each microwave load value are unique, and always represent a matched load condition.
3. The equilibrium conditions are reached rapidly, but without excessive damped oscillations about the equilibrium point.

The existence of the three conditions listed above is not guaranteed, and careful analysis is required to show that they can be attained.

This chapter will be concerned mainly with the system configuration, the input-output relationships of the microwave portion of the system, the system differential equations, and uniqueness of the solutions. Stability will be treated in another chapter.

4-1 The Closed-Loop System Configuration

Before proceeding with the system analysis, it is advantageous to discuss the system configuration. The system can be considered as consisting of the following:

1. A load admittance Y_L , which acts as a disturbing signal to the control loops.
2. A series phaseshifter, which "rotates" the load Y_L , allowing Y_L to become Y_R such that only the angular part of the reflection coefficient is affected.
3. A parallel shorted stub, made adjustable by means of a phaseshifter, which affects only the imaginary part of admittance Y_R , and which changes Y_R into Y .
4. A three detector-probe measuring device with associated analogue operational amplifiers to combine the detector outputs to yield two signals, one proportional to the real part of the reflection coefficient of Y , and another proportional to the imaginary part of the reflection coefficient of Y .
5. Amplifiers in each control loop to add gain and make the signals (that are proportional to the real and imaginary parts of the reflection coefficient) powerful enough to drive the phaseshifters.
6. Minor feedback loops within each of the two main control loops to provide electrical damping of the main control loops.

If damping terms are not considered, the electromechanical portions of the two phaseshifters may each be represented, in an analogue fashion, by two cascaded integrators. This is so because, according to

equation 2.16, the drive coil current produces a force that is proportional to coil current. That force accelerates the moving mass of the phaseshifter at a rate which is proportional to coil current. But, dielectric slab acceleration, if integrated twice, gives the position of the dielectric slab. Slab position corresponds to the output variable of the control loop.

It should immediately be recognized that a closed control loop, containing just two integrators and linear or non-linear gain, would tend to oscillate if an input or disturbing signal were to be applied to the loop. Therefore, it is always necessary to include some form of damping to obtain stable operation.

In the actual system, damping may be derived from the drive coil back e.m.f., from friction in the phaseshifter teflon bearings, from viscous air damping action on the dielectric slab, and from feedback of the rate coil signal. The air and frictional damping terms were found to be sufficiently small that they could be ignored. The main damping terms, derived from the drive coil back e.m.f. and from the rate coil signal, on the other hand, may be combined into a single damping feedback loop around one of the integrators discussed above.

The entire closed loop system is represented in block diagram form in figure 23. All admittance variables shown are normalized to the feeder line characteristic admittance and are considered as consisting of a real part, the conductance, and an imaginary part, the susceptance.

The diagram of figure 23 is basic and will serve for analysis purposes.

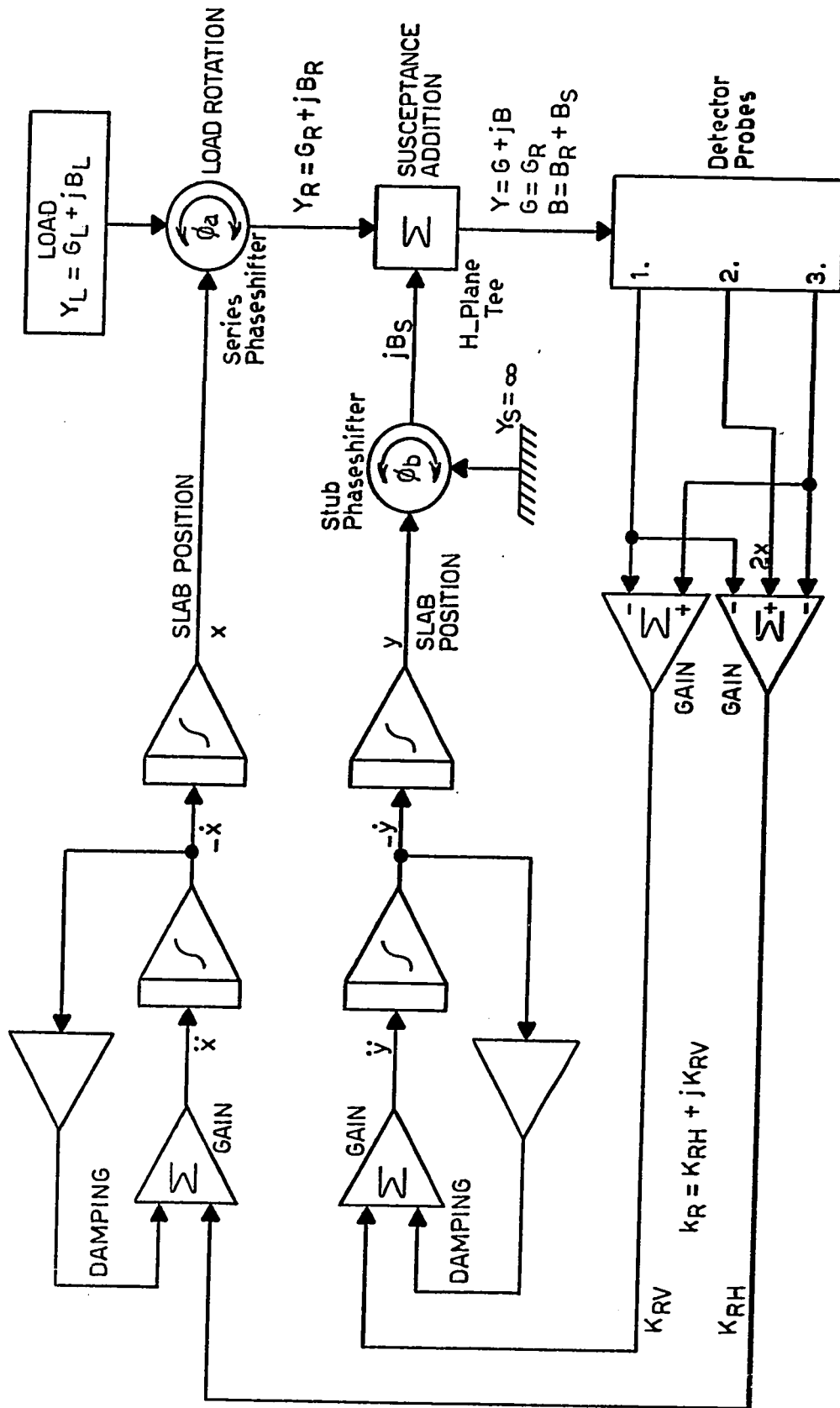


FIGURE 23. THE CLOSED LOOP MATCHING SYSTEM

4-2 The System Differential Equations

To derive the system differential equations, consider the phase-shifter moving mass and drive coil arrangement with rate feedback, as shown in figure 24. Let the drive coil be fed from an operational type power amplifier which has practically zero output impedance. For the time being, assume that the drive coil and rate feedback coil have zero inductance, so that the voltage current relationships are not frequency dependent.

Define the following quantities:

i_1 is the drive coil current.

e_i is the amplifier input voltage.

e_0 is the amplifier output voltage.

e_b is the drive coil back e.m.f.

e_r is the rate coil output voltage.

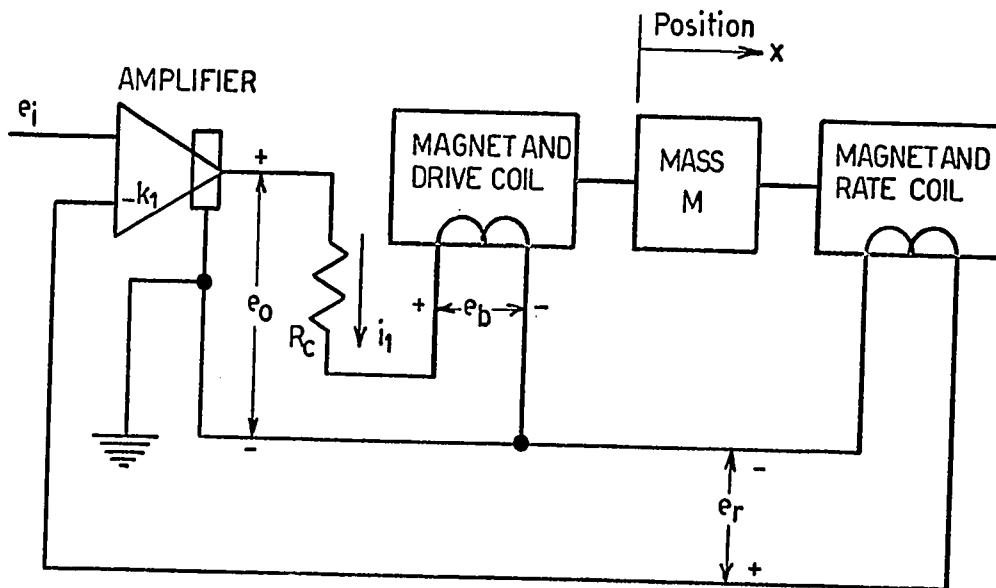


FIGURE 24. THE PHASESHIFTER ELECTROMECHANICAL SYSTEM

R_c is the drive coil resistance.

M is the phaseshifter moving mass.

x is the phaseshifter moving mass (slab) position.

t is time.

k_1 is the rate feedback amplifier input gain.

k_2 , k_3 , and k_4 are proportionality constants.

The quantities so defined, which refer to figure 24, are interrelated as expressed by the following equations:

$$i_1 = \frac{e_o - e_b}{R_c} \quad 4.1$$

$$e_r = k_3 \frac{dx}{dt} \quad 4.2$$

$$e_b = k_4 \frac{dx}{dt} \quad 4.3$$

$$e_o = -e_i - k_1 e_r \quad 4.4$$

Combining of equations 4.1, 4.2, 4.3, and 4.4 gives:

$$i_1 = -\frac{e_i}{R_c} - \left(\frac{k_1 k_3}{R_c} + \frac{k_4}{R_c} \right) \frac{dx}{dt} \quad 4.5$$

The force produced by the drive coil is proportional to the drive coil current. Thus:

$$M \frac{d^2x}{dt^2} = k_2 i_1 \quad 4.6$$

Equation 4.6 may be substituted into 4.5 which, after rearranging of

terms, yields:

$$\frac{d^2x}{dt^2} + \left(\frac{k_1 k_2 k_3}{R_c M} + \frac{k_2 k_4}{R_c M} \right) \frac{dx}{dt} + \frac{k_2}{R_c M} e_i = 0 \quad 4.7$$

The result of equation 4.7 justifies the analogue representation of the phaseshifter electromechanical parts given in figure 23.

The actual differential equations are readily obtained if the diagram of figure 23 is redrawn into the equivalent diagram of figure 25. From equation 4.7 and figure 25, the system differential equations may immediately be written down as equations 4.8 and 4.9. The symbols are defined as follows:

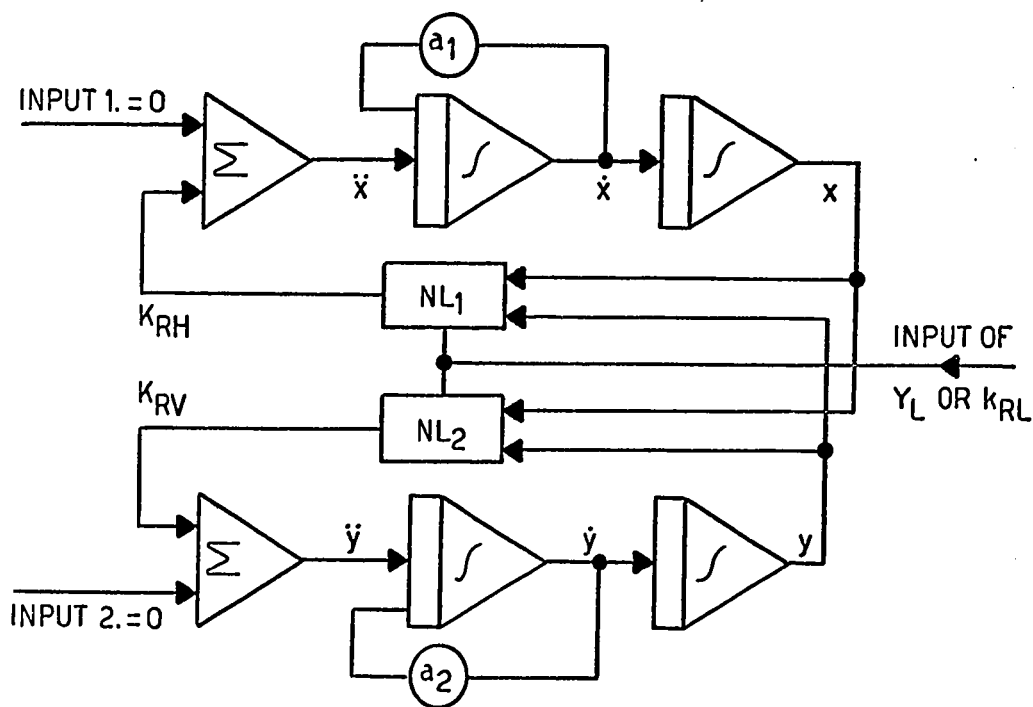


FIGURE 25. COUPLED NON-LINEAR FEEDBACK LOOPS

"." indicates the time derivative d/dt .

x is the series phaseshifter slab position.

y is the stub phaseshifter slab position.

a_1 and a_2 are the damping loop gains.

c_1 and c_2 are the main static loop gains.

$$\ddot{x} + a_1 \dot{x} + c_1 K_{RH}(x, y, Y_L) = 0 \quad 4.8$$

$$\ddot{y} + a_2 \dot{y} + c_2 K_{RV}(x, y, Y_L) = 0 \quad 4.9$$

Notice that the input signals (1 and 2) in figure 25, corresponding to the desired values of K_{RH} and K_{RV} , are zero. The load admittance Y_L acts merely as a disturbing signal applied to the non-linearities NL_1 and NL_2 . To each initially unmatched load there corresponds a new set of equilibrium values of x and y . A step change in load admittance Y_L makes the system behave as if new initial conditions on x and y had been applied. The system can thus be regarded as being autonomous in nature.

System equations 4.8 and 4.9 may also be expressed in state variable form as in equation set 4.10:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -a_1 x_2 - c_1 K_{RH}(x, y, Y_L) \\ \dot{y}_1 &= y_2 \\ \dot{y}_2 &= -a_2 y_2 - c_2 K_{RV}(x, y, Y_L) \end{aligned} \quad 4.10$$

The system equations are not complete until K_{RH} and K_{RV} have been

defined in terms of x , y , and Y_L . This will be done next.

4-3 The System Non-Linearities

Expressions for the non-linearities, $K_{RH}(x,y,Y_L)$ and $K_{RV}(x,y,Y_L)$, may be obtained using standard transmission line theory^(5,10).

Referring to figure 23, load admittance Y_L is transformed by the series phaseshifter, and any interconnecting waveguide, into an admittance $Y_R = G_R + jB_R$ on the load side of the neutral plane of the H-plane tee. The interconnecting waveguide contributes only a constant phase-shift angle to the reflection coefficient and is considered no further here.

The shorted stub adds, by means of the H-plane tee, a variable amount of susceptance to Y_R , thus transforming Y_R into $Y = G + jB$ on the generator side at the neutral plane of the H-plane tee.

The reflection coefficient, k_R , corresponding to Y at the neutral plane of the H-plane tee, is measured by the three-probe detector device. The measuring device, therefore, does not measure the actual load admittance Y_L , but measures rather the reflection coefficient of Y , which corresponds to Y_L transformed by the matching device.

Without loss of generality, assume that the incremental phase-shifts θ_a and θ_b (in radians) produced by the series and stub phase-shifters respectively, are numerically equal to their respective slab displacements in centimeters. This is legitimate since the necessary proportionality constants can be considered as part of the static loop gains c_1 and c_2 in equations 4.8 and 4.9.

Then, for a series incremental phaseshift θ_a , one has:

$$Y_R = \frac{Y_L + j \tan \phi_a}{1 + j Y_L \tan \phi_a} = \frac{Y_L + j \tan x}{1 + j Y_L \tan x} \quad 4.11$$

or, by setting $Y_L = G_L + jB_L$ and $Y_R = G_R + jB_R$, and expanding:

$$G_R = \frac{G_L(1 + \tan^2 x)}{(1 - B_L \tan x)^2 + G_L^2 \tan^2 x} \quad 4.12$$

$$B_R = \frac{\tan x + B_L(1 - \tan^2 x) - (G_L^2 + B_L^2) \tan x}{(1 - B_L \tan x)^2 + G_L^2 \tan^2 x} \quad 4.13$$

For a short-circuit admittance $Y_S = \infty$ and a stub incremental phase-shift ϕ_b , the susceptance added by the parallel stub at the H-plane tee is:

$$jB_S = j \tan \phi_b, \text{ or } B_S = \tan y \quad 4.14$$

Moreover, susceptance addition by the H-plane tee results in:

$$G = G_R = \frac{G_L(1 + \tan^2 x)}{(1 - B_L \tan x)^2 + G_L^2 \tan^2 x} \quad 4.15$$

$$\begin{aligned} B &= B_R + B_S \\ &= \frac{\tan x + B_L(1 - \tan^2 x) - (G_L^2 + B_L^2) \tan x}{(1 - B_L \tan x)^2 + G_L^2 \tan^2 x} + \tan y \end{aligned} \quad 4.16$$

The three-probe detector device measures k_R , the reflection coef-

ficient corresponding to $Y=G+jB$. From transmission line theory:

$$k_R = \frac{1-Y}{1+Y} = \frac{(1-G) - jB}{(1+G) + jB} \quad 4.17$$

Separating expression 4.17 into real and imaginary parts and setting

$k_R = K_{RH} + jK_{RV}$ gives:

$$K_{RH} = \frac{1-G^2 - B^2}{(1+G)^2 + B^2} \quad 4.18$$

$$K_{RV} = \frac{-2B}{(1+G)^2 + B^2} \quad 4.19$$

Substitution of equations 4.15 and 4.16 into the expressions for K_{RH} and K_{RV} , as given by 4.18 and 4.19, completely defines K_{RH} and K_{RV} , and thus NL_1 and NL_2 , in terms of x , y , and Y_L . The complexity of the expressions prevents this substitution from being carried out.

Even though the results just derived will be useful in subsequent analysis work, equally applicable and convenient results may be derived using polar coordinate representations of the non-linearities.

Let the load admittance, Y_L , be characterized by its reflection coefficient $k_{RL} = K_{RL} e^{j(\theta_{RL})}$, where e is the exponential operator. Then k_{RR} , the reflection coefficient corresponding to Y_R (see figure 23), is

$$k_{RR} = K_{RL} e^{j(\theta_{RL} + \phi_a)} = K_{RL} e^{j(\theta_{RL} + x)} = K_{RL} e^{j\psi} \quad 4.20$$

where θ_a is the phaseshift contributed by the series phaseshifter.

Now:

$$\begin{aligned}
 Y_R &= \frac{1 - k_{RR}}{1 + k_{RR}} = \frac{1 - K_{RL} e^{j\psi}}{1 + K_{RL} e^{j\psi}} = \\
 &= \frac{1 - K_{RL} \cos \psi - j K_{RL} \sin \psi}{1 + K_{RL} \cos \psi + j K_{RL} \sin \psi} \quad 4.21
 \end{aligned}$$

Assuming that the susceptance added by the shorted stub is

$B_S = \tan \theta_b = \tan y$, one has:

$$\begin{aligned}
 Y &= Y_R + j \tan y = \\
 &= \frac{1 - K_{RL} \cos \psi - K_{RL} \sin \psi \tan y}{1 + K_{RL} \cos \psi + j K_{RL} \sin \psi} + \\
 &\quad + j \frac{\tan y + K_{RL} \cos \psi \tan y - K_{RL} \sin \psi}{1 + K_{RL} \cos \psi + j K_{RL} \sin \psi} \quad 4.22
 \end{aligned}$$

Again, from transmission line theory, the reflection coefficient, k_R , corresponding to the admittance, Y , at the neutral plane of the H-plane tee, is given by:

$$\begin{aligned}
 k_R &= \frac{1 - Y}{1 + Y} = \quad 4.23 \\
 &= \frac{2 K_{RL} \cos \psi - K_{RL} \sin \psi \tan y}{(2 - K_{RL} \sin \psi \tan y) + j (\tan y + K_{RL} \cos \psi \tan y)} + \\
 &\quad - j \frac{\tan y + K_{RL} \cos \psi \tan y - 2 K_{RL} \sin \psi}{(2 - K_{RL} \sin \psi \tan y) + j (\tan y + K_{RL} \cos \psi \tan y)}
 \end{aligned}$$

$$\begin{aligned}
K_{RH}(x, y, y_L) &\equiv K_{RH}(x, y, k_{RL}) \equiv K_{RH}(x, y, k_{RL}, \theta_{RL}) = \\
&= \frac{4K_{RL}\cos(\theta_{RL}+x) - [4K_{RL}\sin(\theta_{RL}+x) + 4K_{RL}^2\sin(\theta_{RL}+x)]\cos(\theta_{RL}+x)}{4 - [4K_{RL}\sin(\theta_{RL}+x)]\tan y + [2K_{RL}\cos(\theta_{RL}+x) + K_{RL}^2]\tan^2 y} \tan y + [1 + 2K_{RL}\cos(\theta_{RL}+x) + K_{RL}^2]\tan^2 y
\end{aligned}$$

4.26

$$\begin{aligned}
K_{RV}(x, y, y_L) &\equiv K_{RV}(x, y, k_{RL}) \equiv K_{RV}(x, y, k_{RL}, \theta_{RL}) = \\
&= \frac{4K_{RL}\sin(\theta_{RL}+x) - [2 + 4K_{RL}\cos(\theta_{RL}+x) + 2K_{RL}^2]\tan y - [2K_{RL}\sin(\theta_{RL}+x) + 2K_{RL}\sin(\theta_{RL}+x)]\cos(\theta_{RL}+x)}{4 - [4K_{RL}\sin(\theta_{RL}+x)]\tan y + [2K_{RL}\cos(\theta_{RL}+x) + K_{RL}^2]\tan^2 y} \tan^2 y
\end{aligned}$$

4.27

TABLE 1. EQUATIONS FOR THE SYSTEM NON-LINEARITIES

The components of k_R , as actually measured by the three probe measuring device, are:

$$K_{RH} = \text{Re} [k_R] \quad 4.24$$

$$K_{RV} = \text{Im} [k_R] \quad 4.25$$

Separation of equation 4.23 into real and imaginary components according to 4.24 and 4.25, and simultaneous substitution of $\Psi = (\theta_{RL} + x)$, will yield defining equations for K_{RH} and K_{RV} , and hence for NL_1 and NL_2 , in terms of x , y , and Y_L or k_{RL} . Because of their length, the resulting expressions for $K_{RH}(x, y, Y_L)$ and $K_{RV}(x, y, Y_L)$ are given in table 1 as equations 4.26 and 4.27.

The results expressed by the equations of table 1 are entirely equivalent to those given by equations 4.15, 4.16, 4.18 and 4.19. Either set of equations completely characterizes the two system nonlinearities, NL_1 and NL_2 .

In subsequent parts of this report an analysis of the system operation will be attempted based on the results of this and previous sections.

4-4 Some Properties of the Non-Linearities

It is of some importance to examine the behaviour of the non-linear functions defined by equations 4.26 and 4.27. This may be done with reference to a Smith chart as in figure 26.

The relationship between admittances and reflection coefficients,

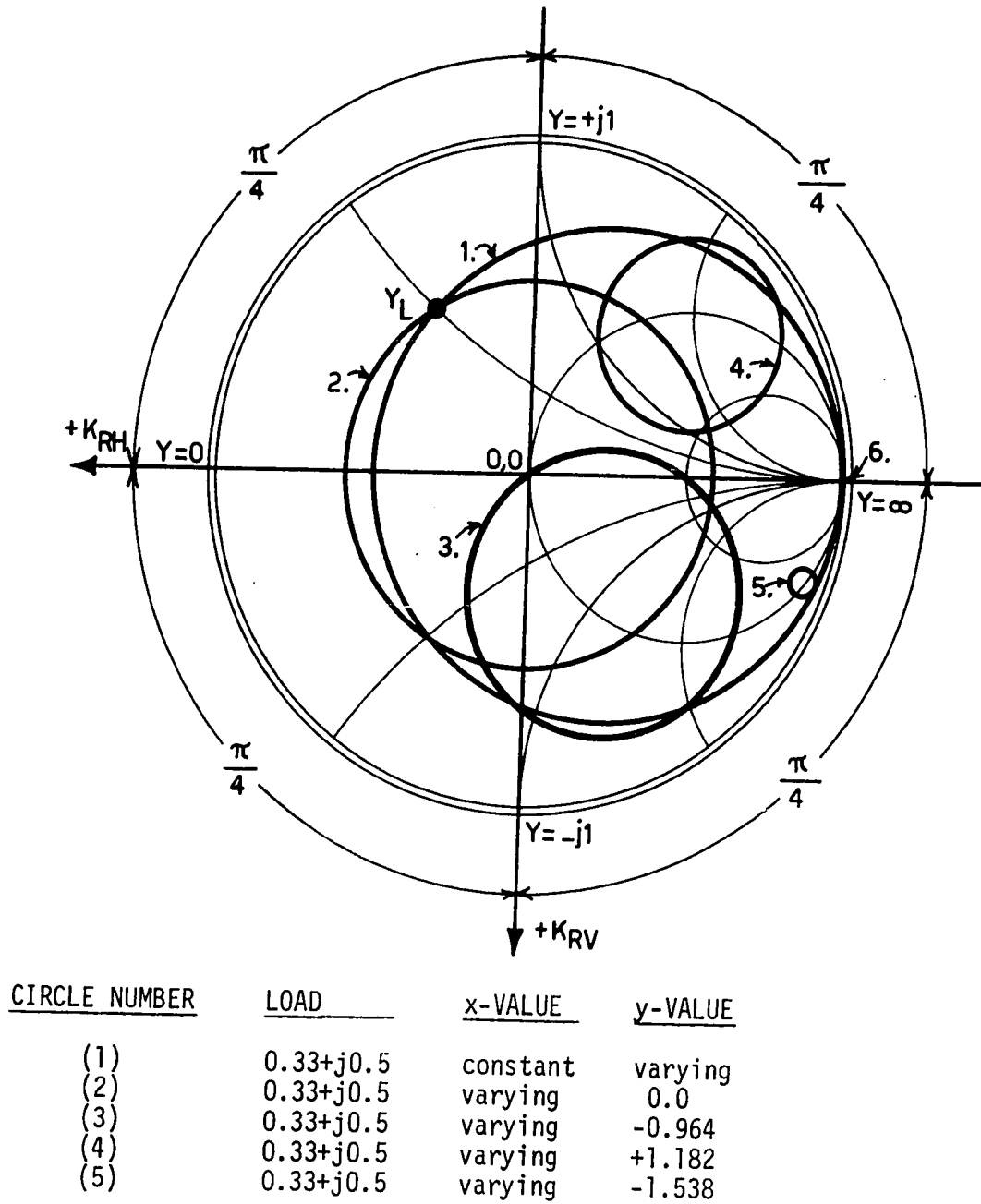


FIGURE 26. TYPICAL SMITH CHART TRAJECTORIES

as expressed by equation 4.17, is a bi-linear transformation from an admittance plane to a reflection coefficient plane (the Smith chart).

One of the basic properties of a bi-linear transformation is that straight lines and circles (in the admittance plane) are transformed into straight lines and circles (in the reflection coefficient plane). As a consequence of this property, the Smith chart trajectories that result from changes in x and y , for any given load admittance Y_L , are all circles. Unfortunately, the expressions for K_{RH} and K_{RV} are too complex for analytical derivation of information about the Smith chart trajectories. Where necessary, a digital computer was employed to calculate and plot K_{RH} and K_{RV} simultaneously for specific values of Y_L , as x and y were varied separately. Some typical trajectories are shown in figure 26.

The properties of the non-linearities may be summarized as follows:

1. For any initial value of Y_R (for definitions of the admittance variables used see fig.23) determined by the values of Y_L and x , the shorted stub adds pure susceptance B_S to Y_R to produce Y . Varying y will, therefore, leave the real part of Y unaffected. Thus, changing y inevitably causes motion only along constant conductance circles on the Smith chart. The radii of the constant conductance circles are functions of the real part of admittance Y , and all their centers lie on the K_{RH} -axis between 0 and -1. Changing either Y_L or x , or both, merely shifts Y_R to a new constant conductance circle along which motion will proceed if y is again varied. The changes in the trajectories resulting from changes in x are not profound. It is especially important to note that, for motion along any constant conductance circle, there is always

- a value of y which will cause K_{RV} to be equal to zero.
2. For any load admittance Y_L , and for $y=0$ such that the shorted stub adds zero susceptance, the series phasemitter affects only the angular part of the reflection coefficient corresponding to admittance Y . Varying x will, consequently, cause motion along circles whose centers are always the Smith chart origin. The radii of these circles are directly proportional to the magnitude of the reflection coefficient corresponding to Y .
 3. For constant values of y other than zero, varying x has less effect than for $y=0$ as discussed under 2 above, since parallel susceptance has now been added which tends to "shield" Y from changes in Y_R . The trajectories, as x is varied, are still circles, but the radii are smaller and the centers no longer lie between $K_{RH}=0$ and $K_{RH}=-1$ on the K_{RH} -axis. Rather, the centers now lie on circular arcs above and below the K_{RH} -axis that connect the points $K_{RH}=0$ and $K_{RH}=-1$ on the K_{RH} -axis. As $|y|$ increases, the centers move away from the Smith chart origin and the radii decrease. Also, as k_{RL} corresponding to Y_L increases the radii increase.
 4. The non-linearities NL_1 and NL_2 are periodic in x and y , with a period of π radians. All the circular trajectories discussed above are traversed once for a change of π radians in x or y , as the case may be.

The properties of the non-linearities, as given here, will be used in chapter five to prove stability.

4-5 Equilibrium and Uniqueness

Even though the results to be derived in this section will appear mathematically trivial, they are nevertheless of vital importance with regard to proper operation and stability of the control system.

Assume that the control system as described by equations 4.8 and 4.9 has come to rest; i.e. an equilibrium point has been reached. Under these conditions all time derivatives of the system variables must be zero, otherwise deviations from equilibrium would occur. Thus:

$$\ddot{x} = \dot{x} = \ddot{y} = \dot{y} = 0 \quad 4.28$$

Substitution of conditions 4.28 into equations 4.8 and 4.9 leads to:

$$K_{RH} = K_{RV} = 0 \quad 4.29$$

Equations 4.29 indicate that, in terms of K_{RH} and K_{RV} on the Smith chart, there is only one equilibrium point which simultaneously corresponds to a matched condition ($k_R=0$). Therefore, if the system is stable and exhibits no limit cycles, a matched condition will be reached as time goes to infinity.

For a given load admittance $Y_L = G_L + jB_L$, the values of x and y which produce a matched load condition are by no means unique. This can be shown by solving equations 4.15 and 4.16 for x and y . If, first, conditions 4.29 are substituted into equations 4.18 and 4.19, the expected equivalent requirements for a matched load condition become:

$$G = 1 ; B = 0 \quad 4.30$$

Setting G and B in equations 4.15 and 4.16 equal to the values of 4.30 and solving yields:

$$x = \tan^{-1} \frac{B_L \pm (G_L^3 - 2G_L^2 + G_L B_L^2 + G_L)^{1/2}}{G_L^2 + B_L^2 - G_L} \pm m\pi \quad 4.31$$

$$y = \tan^{-1} \frac{(G_L^2 + B_L^2) \tan x - \tan x - B_L(1 - \tan^2 x)}{(1 - B_L \tan x)^2 + G_L^2 \tan^2 x} \pm n\pi \quad 4.32$$

where:

x and y are the respective phaseshifter dielectric slab positions.

G_L and B_L are the normalized unmatched load conductance and susceptance respectively.

m and n are integers.

The values of x in equation 4.32 are equal to all possible values of x obtained from equation 4.31.

As demonstrated by the solutions to 4.31 and 4.32, to each load admittance Y_L there corresponds an infinite set of periodically recurring equilibrium values of x and y. Physical limitations on the ranges of slab motion prevent many of these pairs of equilibrium values from being reached. Furthermore, as n changes by unit increments, and for a closed loop system, the equilibria are alternately stable and unstable. More importantly, for a closed loop system, as the equilibrium values of y approach $[(\pi/2) \pm n\pi]$, the stable equilibria tend to

become saddle points. Therefore, if the system exhibits any overshoot, successive sets of stable equilibria, under these conditions, may be bypassed altogether, thus preventing the system from coming to rest. This would render the system useless.

The above difficulties may be avoided, and the equilibria may, at least partly, be made unique by the methods described hereafter.

Let the range of motion of the stub phaseshifter slab be mechanically limited such that the allowable phaseshift range is slightly less than $\pm\pi/2$ radians. This will restrict y in equation 4.32 to its principal values, while the parallel admittance B_S , as given by expression 4.14, will only assume values between $\pm\infty$, contained between the tangent function asymptotes at $y=\pm\pi/2$. Moreover, for any given set of values G_L , B_L , and x , the solution for y obtained from equation 4.32 will now be unique.

Note also, by reference to the Smith chart of figure 27 that, for motion along any constant conductance circle, the sign of K_{RV} is negative in the upper semi-circular region, positive in the lower semi-circular region, and zero on the real axis of the Smith chart. Proper arrangement of negative feedback, in a loop involving K_{RV} and the stub phaseshifter, will thus reduce K_{RV} to zero for any Y_L and x , provided the system is asymptotically stable. At equilibrium values of y close to $\pm\pi/2$, the mechanical slab motion stop will limit the overshoot so that, even under these conditions, equilibrium of the stub phaseshifter loop can be attained.

The situation with regard to the series phaseshifter is not as simple. Again, let the range of motion of the series phaseshifter slab be mechanically limited to $\pm\pi/2$ radians. In these circumstances it

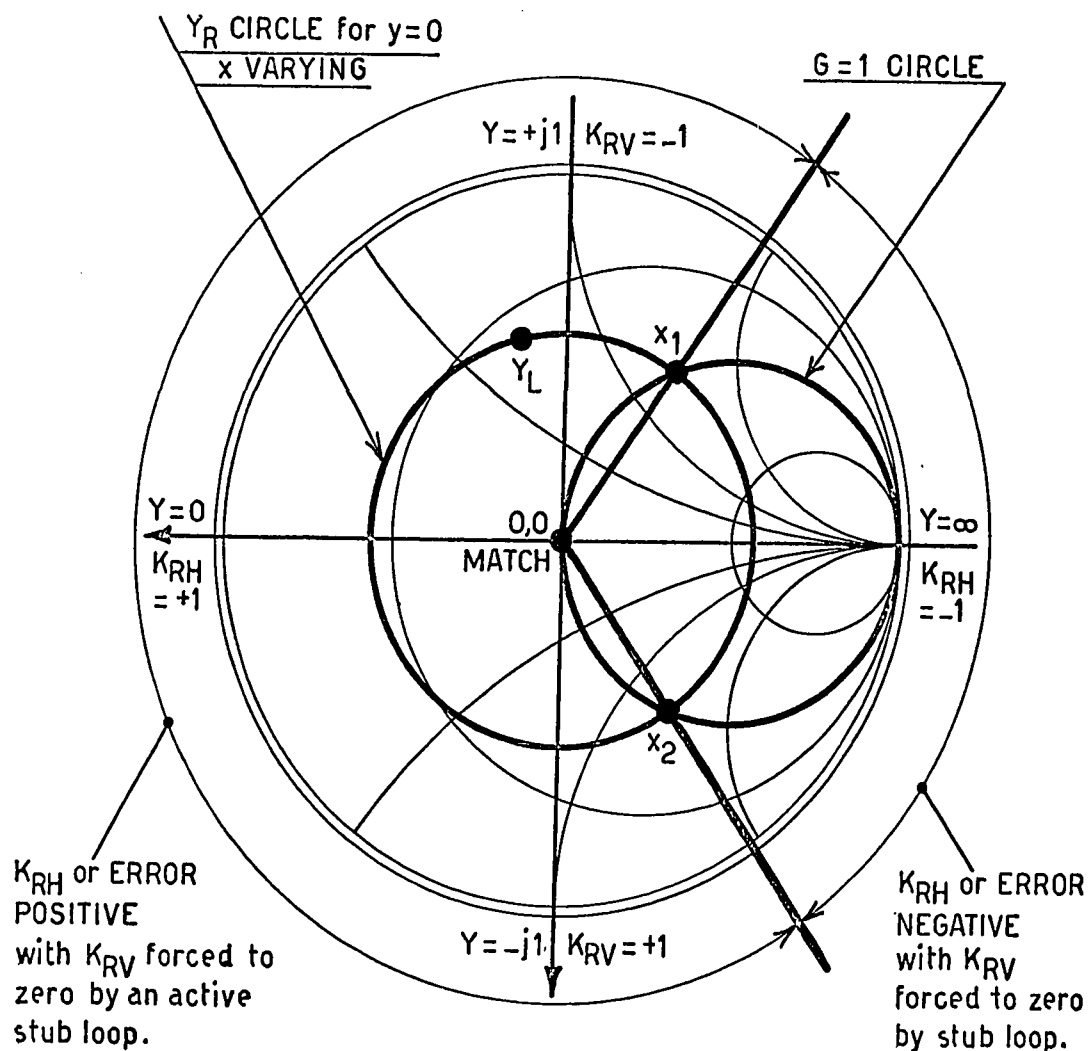


FIGURE 27. ERROR SIGNALS AND POLARITIES ON THE SMITH CHART

might be expected that, from expression 4.31, x is limited to its principal values, but this is, in fact, not necessarily so. The explanation follows.

The reflection coefficient components as measured by the three detector measuring device are those corresponding to the total admittance appearing at the neutral plane of the H-plane tee. Measurement allows determination of the contribution by the parallel shorted stub to the

admittance at the neutral plane of the H-plane tee. A proper stub phaseshifter reference point, which is independent of the load admittance or the series phaseshifter adjustment, may be chosen as that particular stub phaseshifter adjustment, which causes zero admittance to be added by the stub to the main waveguide admittance at the neutral plane of the H-plane tee.

Measurement of the admittance at the neutral plane of the H-plane tee in the main waveguide does not allow the contribution by the series phaseshifter to be separately determined. The angle of the reflection coefficient of Y , at the neutral plane of the H-plane tee, is the sum of the phase angle contributions of the load admittance, the series phaseshifter, and the interconnecting waveguide. Changes in phase angle resulting from adjustment of the series phaseshifter can be measured, but a zero phaseshift reference point cannot be defined; it is indeterminate. The series phaseshift reference point will be discussed again later.

The above difficulties should not imply, however, that the phaseshifter matching adjustments cannot be made unique. Reference to the Smith chart of figure 27 shows that, for $y=0$, varying x through its range of π radians causes the reflection coefficient vector to move in a circle about the Smith chart origin. On this circle there are two points at which K_{RH} is zero. In the left hand semi-circular region of the Smith chart K_{RH} is positive, while in the right hand semi-circular region K_{RH} is negative. More specifically, with K_{RV} always maintained at $K_{RV}=0$ while x is adjusted, there are two values of x (i.e. x_1 and x_2 in figure 27) that correspond to system equilibrium and matched load conditions. For the range of values of x in the left hand circular sec-

tor between x_1 and x_2 , K_{RH} or the error signal will be positive, while for values of x in the right hand sector between x_1 and x_2 , K_{RH} or the error signal will be negative. A feedback loop, involving the series phaseshifter activated by K_{RH} , will thus adjust to either one of the two equilibrium points x_1 or x_2 , depending on the sign of the loop gain. One of the two equilibrium points will be stable, while the other will be an unstable equilibrium point. On the Smith chart of figure 27, note particularly that the phase angle distances in radians from x_1 to x_2 are unequal in the two possible directions of motion, but that the sum of the two exactly equals π radians. Expression 4.31 indicates also that, within any series phaseshifter adjustment range π radians wide, there are only two values of x which could correspond to matched load conditions, one stable and one unstable (determined by the plus and minus signs in front of the square root sign).

Two difficulties arise in automating the matching adjustments. For a sudden change in load admittance Y_L , and depending on the previous adjustment of x , the necessary change in x in the direction specified by the sign of the loop gain, may exceed the remaining allowable range of phaseshifter adjustment in that direction. Furthermore, continual changes in Y_L , which cause the phase angle of the reflection coefficient to increase by more than π radians in one direction, would make the series phaseshifter loop try to adjust in that one direction only to follow the reflection coefficient phase change. Either of these conditions would force the series phaseshifter against its mechanical stop, upon which adjustment would cease without equilibrium having been reached.

The above difficulties may be overcome by a method of electrically reversing the polarity of the series phasehifter loop gain, each time

the series phaseshifter reaches one of its two mechanical stops. This guarantees that, while the phaseshifter was previously driven against one of its stops, now it will proceed away from that stop in the opposite direction. After reversal of the loop gain sign, there is always one stable equilibrium within the phaseshifter range of adjustment, as was demonstrated above.

The response speed of the series phaseshifter control loop, for those loads or load changes which cause the phaseshifter to reach one of its mechanical stops, may be improved by adding a short duration electrical pulse signal into the control loop, which will impart some initial velocity, in the opposite direction, to the phaseshifter slab each time the loop gain polarity is reversed at one of the stops. This initial velocity signal will also help the control loop to bypass an unstable equilibrium point, if necessary. Usually though, the first equilibrium point encountered, after sign reversal of the loop gain, will be a stable equilibrium.

Electrical circuit details of the sign reversing and initial velocity component blocks in the automatic matching device will be given later.

It has been shown that mechanical limit stops, incorporated into both phaseshifters, which limit the allowable phaseshift ranges to π radians, cause there to be two and only two distinct equilibrium adjustment positions for the phaseshifters, which correspond to a matched load condition. At any one time, only one of these represents a stable equilibrium. On the Smith chart the matched load condition is entirely unique.

However, the pair of equilibrium points is repeated over and over

again in consecutive periods of series phaseshifter adjustment. The loop gain sign switching, at the mechanical limit stops of the series phaseshifter, has the effect of folding the entire series phaseshift range at the limit stops, such that one single period covers the entire range of phaseshifts from $-\infty$ to $+\infty$. This is possible, because the different periods are indistinguishable one from another.

After a sudden change in load admittance, and provided the previous adjustment positions of the phaseshifters are known (equivalent to initial conditions), the equilibrium point towards which the system will move and the Smith chart trajectory, including a possible automatic loop gain sign reversal at the series phaseshifter limit stop, may be predicted. In this sense the equilibrium position and the adjustment procedure may be considered as being unique. Therefore, stability analysis may proceed on this basis.

CHAPTER FIVE

SYSTEM STABILITY AND RESPONSE ANALYSIS

In the previous chapter it was shown that, provided the control system is stable, the equilibrium positions of the phaseshifter slabs are unique. Furthermore, the non-linear differential equations were completely defined and the properties of the non-linearities were examined. It now remains to show stability. System stability will be demonstrated by three methods:

1. Theoretical analysis.
2. Computer simulation of the entire system.
3. Building of an actual prototype matching device.

This chapter will be devoted to theoretical analysis of the system.

5-1 Stability Analysis Preliminaries

Historically, a great deal of attention has been devoted to the study of non-linear differential equations, which characterize many control systems. Lyapunov⁽¹¹⁾ was the first worker to develop theory which allowed the stability of a system to be considered without the necessity of solving the differential equations. Many others have considered the stability of systems containing a single memory-less non-linearity, for example Lur'e⁽¹²⁾, Popov⁽¹³⁾, and Kalman⁽¹⁴⁾. Popov, especially, has contributed a great deal by presenting a frequency domain stability criterion applicable to non-linear closed loop control

systems containing a single non-linearity. Popov's results have since been extended to cases that place greater restrictions on the non-linearities, but fewer restrictions on the linear part of the system (see for instance Brockett and Willems⁽¹⁵⁾).

Also of great interest are those studies, which have attempted to extend Popov's result to systems containing multiple non-linearities. Ibrahim and Rekasius⁽¹⁶⁾, Jury and Lee⁽¹⁷⁾, Ku and Chieh⁽¹⁸⁾, and Anderson⁽¹⁹⁾ have all derived multi-dimensional frequency domain stability criteria for multiple non-linearity systems.

Unfortunately, because the non-linearities are restricted too much by these criteria, not all of the multi-dimensional stability criteria developed by the above workers are applicable to the system under consideration in this report. However, Anderson's, and also Ku and Chieh's work is applicable. Anderson's stability criterion, particularly, may be applied if the non-linearities of this system are given special consideration.

The theorem derived by Anderson^(19,20), and the conditions under which it applies, are stated next.

Let a multi-dimensional non-linear control system be represented by the block diagram of figure 28, where \underline{u} and \underline{z} are column vectors corresponding to the multiple inputs and outputs, respectively, to the linear part of the system. Let $\underline{W}(s)$ be the complex frequency domain transfer function matrix of the linear part of the system. Also, let $\underline{f}(\underline{z})$ be a non-linear vector function of its input \underline{z} . The output from the non-linearity is then identified with \underline{u} , the vector input to the linear part.

Furthermore, let $\underline{W}(s)$ be a matrix of stable, rational functions

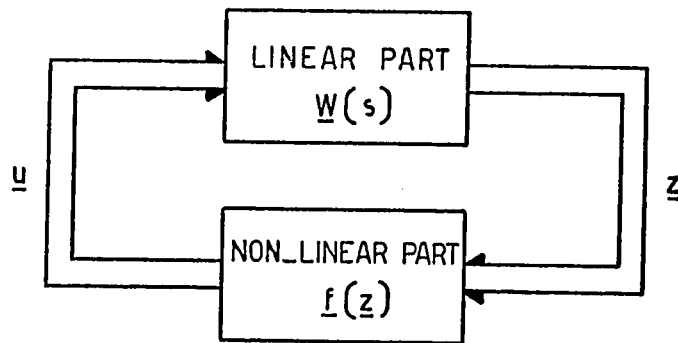


FIGURE 28. THE GENERAL MULTI-DIMENSIONAL NON-LINEAR SYSTEM

of s such that:

$$\underline{W}(s) \Big|_{s \rightarrow \infty} = \underline{0} \quad 5.1$$

which means that each non-zero element of the transfer function matrix must have more finite poles than it has finite zeros.

And, let the vector non-linearity be restricted to generalized n -dimensional sectors by the inequality:

$$\underline{f}^T(z) z \geq \underline{f}^T(z) \underline{K} \underline{f}(z) \quad 5.2$$

where the superscript T indicates the matrix transpose, and where \underline{K} is a non-negative definite matrix of constants.

For a system as shown in figure 28, and under the restrictions imposed by expressions 5.1 and 5.2, Anderson's theorem states that:

The system of figure 28 is stable (Lyapunov type stability) if there exist constants α and β such that:

$$\alpha \geq 0 \quad ; \quad \beta \geq 0 \quad ; \quad \alpha + \beta > 0 \quad 5.3$$

and such that:

$$\underline{M}(s) = \alpha \underline{K} + (\alpha + \beta s) \underline{W}(s) \text{ is positive real.} \quad 5.4$$

A matrix \underline{M} of functions of a complex variable s is called Positive Real if and only if:

1. $\underline{M}(s)$ has elements which are analytic for $\text{Re}[s] > 0$.
2. $\underline{M}^*(s) = \underline{M}(s^*)$ for $\text{Re}[s] > 0$, where $*$ indicates the complex conjugate.
3. The Hermitian matrix $\underline{N}(s) = \underline{M}(s) + \underline{M}^T(s^*)$ is non-negative definite for $\text{Re}[s] > 0$.

The non-negative definite character of the Hermitian matrix $\underline{N}(s)$ may be established with the aid of the method of Principal Minors⁽²¹⁾.

Anderson proves his stability theorem in two parts. In an initial paper⁽²⁰⁾ he shows that a transfer function matrix $\underline{Z}_1(s)$, with a minimum dimension realization:

$$\begin{aligned} \dot{\underline{x}} &= \underline{F}_1 \underline{x} + \underline{G}_1 \underline{u} \\ \underline{y} &= \underline{H}_1^T \underline{x} \end{aligned} \quad 5.5$$

and such that:

$$\underline{Z}_1(s) = \underline{H}_1^T (s \underline{I} - \underline{F}_1)^{-1} \underline{G}_1 \quad 5.6$$

is Positive Real if and only if there is a positive definite matrix \underline{P}_1 , and if there are matrices \underline{L}_1 and \underline{W}_{01} such that:

$$\begin{aligned}\underline{P}_1 \underline{F}_1 + \underline{F}_1^T \underline{P}_1 &= -\underline{L}_1 \underline{L}_1^T \\ \underline{P}_1 \underline{G}_1 &= \underline{H}_1 - \underline{L}_1 \underline{W}_{01} \\ \underline{W}_{01} \underline{W}_{01}^T &= \underline{Z}_1(\infty) + \underline{Z}_1^T(\infty)\end{aligned}\quad 5.7$$

Then, in a second paper⁽¹⁹⁾, he considers the stability of a system characterized by:

$$\begin{aligned}\dot{\underline{x}} &= \underline{F} \underline{x} - \underline{G} f(\underline{H}^T \underline{x}) \\ f^T(\underline{H}^T \underline{x}) (\underline{H}^T \underline{x}) &\geq f^T(\underline{H}^T \underline{x}) \underline{K} f(\underline{H}^T \underline{x})\end{aligned}\quad 5.8$$

with transfer function matrix $\underline{W}(s)$ for the linear part defined in a manner similar to that of 5.6. He relates a triple $[\underline{F}_2, \underline{G}_2, \underline{H}_2]$ associated with a minimum realization of $[\underline{\alpha} \underline{K} + (\underline{\alpha} + \beta s) \underline{W}(s)]$ to the triple $[\underline{F}, \underline{G}, \underline{H}]$ associated with $\underline{W}(s)$ for the actual system. This yields expressions for \underline{P} , \underline{L} and \underline{W}_0 , defined according to 5.7, in terms of $\underline{\alpha}$, β , and \underline{F} , \underline{G} and \underline{H} of the actual system.

Based on the above, he proceeds to show that a positive definite Lyapunov function for the system, defined and restricted by 5.8, is:

$$V(\underline{x}) = \underline{x}^T \underline{P} \underline{x} + 2\beta \int_0^{\underline{H}^T \underline{x}} f^T(\underline{\sigma}) d\underline{\sigma}\quad 5.9$$

and that, again under the restrictions on the non-linearity, the time derivative of $V(\underline{x})$ as given by:

$$\dot{V}(\underline{x}) = -[\underline{x}^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{x}) \underline{W}_0] [\underline{L} \underline{x} - \underline{W}_0 \underline{f}(\underline{H}^T \underline{x})] - 2\alpha [\underline{f}^T(\underline{H}^T \underline{x}) (\underline{H}^T \underline{x}) - \underline{f}^T(\underline{H}^T \underline{x}) \underline{K} \underline{f}(\underline{H}^T \underline{x})] \quad 5.10$$

is negative definite.

Stability is thus proven if $\underline{M}(s)$ is Positive Real and the non-linearity obeys condition 5.8.

Anderson's theorem will be used in the next sections.

5-2 The System Transfer Function Matrix

As a first step in the stability analysis of the automatic matching system, the transfer function matrix, $\underline{W}(s)$, of the linear part of the system needs to be determined. This is most easily accomplished by rearranging the system block diagram once more, and by simultaneously examining the system differential equations 4.8 and 4.9. The redrawn block diagram appears in figure 29.

Referring to the diagram of figure 29, let \underline{u} and \underline{z} be defined as:

$$\underline{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} K_{RH}(x, y, Y_L) \\ K_{RV}(x, y, Y_L) \end{bmatrix}$$

$$\underline{z} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Then, for the linear part of the system, and in terms of Laplace transforms:

$$\underline{z}(s) = \underline{W}(s) \underline{u}(s) \quad 5.11$$

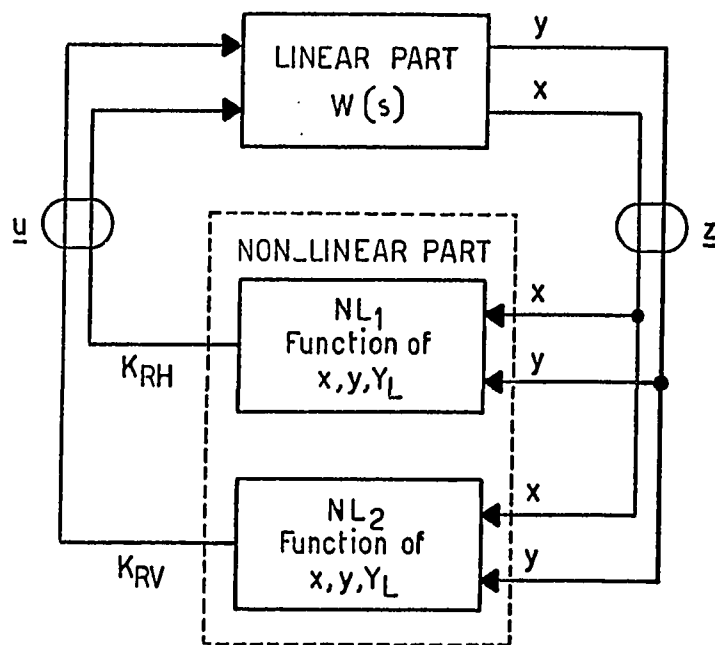


FIGURE 29. THE TWO DIMENSIONAL NON-LINEAR SYSTEM

System differential equations 4.8 and 4.9 may be rearranged, substituting the components of \underline{u} for K_{RH} and K_{RV} . After Laplace transformation and solving for $X(s)$ and $Y(s)$ in terms of $U_1(s)$ and $U_2(s)$ one has:

$$X(s) = \frac{-c_1}{s(s+a_1)} U_1(s) ; Y(s) = \frac{-c_2}{s(s+a_2)} U_2(s) \quad 5.12$$

Therefore, $\underline{W}(s)$ is as given by expression 5.13 where, as is common in control engineering, the minus signs appearing in equations 5.12 have been omitted and are considered as being part of a summing device external to $\underline{W}(s)$.

$$\underline{W}(s) = \begin{bmatrix} \frac{c_1}{s(s+a_1)} & 0 \\ 0 & \frac{c_1}{s(s+a_1)} \end{bmatrix} \quad 5.13$$

Since, as will be shown later, the non-linearities NL_1 and NL_2 have some peculiar properties which make them disobey Anderson's conditions on the non-linearities, it will also be necessary to consider the stability of the stub phasemifter control loop alone (this loop involves K_{RV}). To this end the block diagram of figure 29 may be redrawn as in figure 30.

Let the load admittance, Y_L , be transformed into Y_R by the series phasemifter (see figure 23). The non-linear dependence of Y_R upon x and Y_L may be represented by non-linearity NL_{1a} . Similarly, K_{RH} , one of the inputs to the linear part of the system, is a non-linear function of Y_R and y . The latter non-linear function may be represented by non-linearity NL_{1b} . In effect, NL_1 has been split into two cascaded non-linearities NL_{1a} and NL_{1b} . The nature of non-linearity NL_2 has not changed by this rearrangement. Changes in Y_L and x are entirely equivalent, in their effect upon NL_2 , to appropriate changes in Y_R alone.

That the above interpretation is correct may also be demonstrated by referring to some equations developed in chapter four. Equations 4.15 and 4.16 show that the effects of y (i.e. a quantity equal to the function $(j \tan y)$) are simply added to Y_R . Another separate non-linear function (equation 4.18), which is equivalent to NL_{1b} , converts the

result into K_{RH} . And, as shown by equations 4.12 and 4.13, the dependence of Y_R upon Y_L and x (non-linearity NL_{1a}) is independent of the addition of the effects of y .

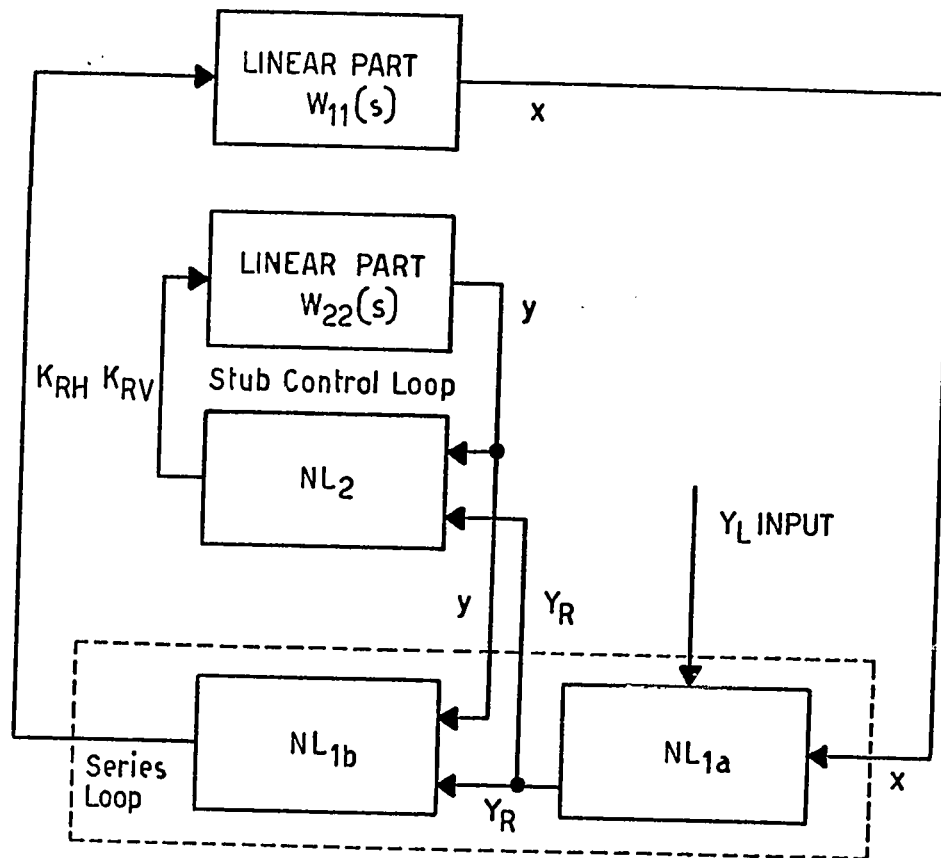


FIGURE 30. SYSTEM BLOCK DIAGRAM SHOWING THE SEPARATE STUB PHASESHIFTER CONTROL LOOP

Figure 30 shows that the stub phasemitter control loop is quite separate and independent within the overall system. The only "connection" between the stub phasemitter control loop and the rest of the system is through input Y_R into the stub control loop and output y from it.

It must be emphasized that the series phasemitter loop may, under

no circumstances, be similarly independently considered. The stub control loop is, effectively, a loop within a loop and is always part of the series phaseshifter control loop.

From the above discussion and from figure 30, it should be obvious that the stability of the stub phaseshifter control loop may be considered separately. To do this the one-dimensional case of Anderson's theorem, which is equivalent to Popov's stability criterion, will be used.

The applicable transfer function for the stub loop may be obtained from the transfer function matrix $\underline{W}(s)$ given in 5.13. Thus:

$$W_{22}(s) = \frac{c_2}{s(s+a_2)} \quad 5.14$$

Again, the minus sign will not be considered as being part of the transfer function.

5-3 The Positive Real Associated Matrices

Before stability can be proven it must be shown that the $\underline{M}(s)$ matrices, discussed in connection with Anderson's theorem, are Positive Real.

Consider first the stub phaseshifter control loop with linear part $W_{22}(s)$ and non-linearity NL_2 (see figure 30).

The expression for $W_{22}(s)$, from 5.14, may be substituted into the one-dimensional case of expression 5.4 to yield:

$$\begin{aligned}
 M_1(s) &= \alpha k + (\alpha + \beta s) W_{22}(s) \\
 &= \alpha k + (\alpha + \beta s) \frac{c_2}{s(s + a_2)}
 \end{aligned}
 \tag{5.15}$$

where $M_1(s)$ must be Positive Real.

Initially, assume only that α and β conform to the requirements of 5.3, and that a_2 and k are non-negative definite. Clearly, under these assumptions, $M_1(s)$ is analytic in the right half s -plane. Also, $M_1^*(s) = M_1(s^*)$ in the right half s -plane. Setting $s = \sigma + j\omega$, one has:

$$\begin{aligned}
 N_1(s) &= M_1(s) + M_1^T(s^*) = \\
 &= 2\alpha k + \frac{c_2[\alpha + \beta(\sigma + j\omega)]}{(\sigma + j\omega)[(\sigma + j\omega) + a_2]} + \frac{c_2[\alpha + \beta(\sigma - j\omega)]}{(\sigma - j\omega)[(\sigma - j\omega) + a_2]}
 \end{aligned}$$

After algebraic manipulation, this finally leads to:

$$N_1(s) = 2\alpha k + c_2 \frac{2\sigma[(\alpha + \beta\sigma)(\sigma + a_2) + \beta\omega^2] + 2\omega^2(\beta a_2 - \alpha)}{(\sigma^2 + \omega^2)[(\sigma + a_2)^2 + \omega^2]}
 \tag{5.16}$$

$N_1(s)$, as given by expression 5.16, must be non-negative definite for all values of ω , and for all values of σ such that $\sigma > 0$.

Assume that $k=0$; this corresponds to the least conservative restrictions on the non-linearity. Then, since σ may approach zero arbitrarily closely from the positive direction, and since the denominator of expression 5.16 is clearly positive definite, the non-negative def-

finite nature of $N_1(s)$ is assured if:

$$c_2 \geq 0 ; (\beta a_2 - \alpha) \geq 0 \quad 5.17$$

Conditions 5.17 may be divided into two separate cases, each of which conform with requirements 5.3 of Anderson's theorem, as follows:

$$c_2 \geq 0 ; \alpha = 0 ; \beta > 0 ; a_2 \geq 0 \quad 5.18a$$

$$c_2 \geq 0 ; \alpha > 0 ; \beta > 0 ; a_2 \geq \alpha/\beta \quad 5.18b$$

Thus, the stub phaseshifter loop is stable (system asymptotic stability will be proven later) provided conditions 5.18a or 5.18b are satisfied, and provided non-linearity NL_2 obeys the restriction:

$$K_{RV}(0, y, Y_R) y \geq 0 \quad \text{for all } y \text{ and } Y_R \quad 5.19$$

where the x and Y_L dependences of K_{RV} have been replaced with Y_R .

Consider now the entire two-dimensional control system.

The transfer function matrix is given by 5.13. According to expression 5.4 the associated matrix is:

$$\underline{M}(s) = \alpha \underline{K} + (\alpha + \beta s) \underline{W}(s)$$

Therefore, $\underline{M}(s)$ becomes:

$$\underline{M}(s) = \begin{bmatrix} \left[\alpha k_{11} + \frac{(\alpha + \beta s) c_1}{s(s + a_1)} \right] & 0 \\ 0 & \left[\alpha k_{22} + \frac{(\alpha + \beta s) c_2}{s(s + a_2)} \right] \end{bmatrix} \quad 5.20$$

Note that $\underline{W}(s)$ contains a free integrator in every non-zero element. This fact will be used later.

By similarity of the non-zero elements of $\underline{M}(s)$ to the single element of $M_1(s)$ just considered, it is clear that $\underline{M}(s)$ has elements which are analytic in the right half s -plane, and that $\underline{M}^*(s) = \underline{M}(s^*)$ for $\text{Re}[s] > 0$.

Algebraic procedures exactly analogous to those used for $M_1(s)$ yield an expression for the Hermitian matrix $\underline{N}(s)$ as follows:

$$\underline{N}(s) = \underline{M}(s) + \underline{M}^T(s^*) = \begin{bmatrix} n_{11}(s) & 0 \\ 0 & n_{22}(s) \end{bmatrix} \quad 5.21$$

$$n_{11}(s) = 2\alpha k_{11} + \frac{2\sigma [(\alpha + \beta\sigma)(\sigma + a_1) + \beta\omega^2] + 2\omega^2(\beta a_1 - \alpha)}{(\sigma^2 + \omega^2)[(\sigma + a_1)^2 + \omega^2]} \cdot c_1$$

$$n_{22}(s) = 2\alpha k_{22} + \frac{2\sigma [(\alpha + \beta\sigma)(\sigma + a_2) + \beta\omega^2] + 2\omega^2(\beta a_2 - \alpha)}{(\sigma^2 + \omega^2)[(\sigma + a_2)^2 + \omega^2]} \cdot c_2$$

The matrix $\underline{N}(s)$ must be non-negative definite for all values of ω , and for all values of σ such that $\sigma > 0$. For this simple two by two diagonal matrix, the method of principal minors for determining if $\underline{N}(s)$ is non-

negative definite reduces to the requirement that each of the diagonal elements must be non-negative definite. As a result of the similarity of the elements of $\underline{N}(s)$ to the single element of $N_1(s)$ just treated it is again apparent that such is the case if:

$$\begin{aligned} c_1 &\geq 0 ; (\beta a_1 - \alpha) \geq 0 \\ c_2 &\geq 0 ; (\beta a_2 - \alpha) \geq 0 \end{aligned} \quad 5.22$$

Two separate cases of conditions 5.22 may, once more, be recognized:

$$c_1 \geq 0 ; c_2 \geq 0 ; \alpha = 0 ; \beta > 0 ; a_1 \geq 0 ; a_2 \geq 0 \quad 5.23a$$

$$c_1 \geq 0 ; c_2 \geq 0 ; \alpha > 0 ; \beta > 0 ; a_1 \geq \alpha/\beta ; a_2 \geq \alpha/\beta \quad 5.23b$$

Conditions 5.23 conform with requirements 5.3 of Anderson's theorem.

The non-negative definite nature of $\underline{N}(s)$ is not affected by \underline{K} . Any non-negative definite matrix \underline{K} is, therefore, acceptable including the least conservative case:

$$\underline{K} = \underline{0} \quad 5.24$$

The entire system is now stable if conditions 5.23 are satisfied and if the non-linearities obey:

$$\underline{f}^T(\underline{z}) \underline{z} \geq \underline{f}^T(\underline{z}) \underline{K} \underline{f}(\underline{z}) \quad (5.2)$$

where:

$$\underline{f}(\underline{z}) = \begin{bmatrix} K_{RH}(x, y, Y_L) \\ K_{RV}(x, y, Y_L) \end{bmatrix}; \quad \underline{z} = \begin{bmatrix} x \\ y \end{bmatrix}$$

By condition 5.24 the above inequality reduces to:

$$\underline{f}^T(\underline{z}) \underline{z} \geq 0 \quad 5.25$$

Expansion of inequality 5.25 gives:

$$K_{RH}(x, y, Y_L) x + K_{RV}(x, y, Y_L) y \geq 0$$

for all x, y and Y_L 5.26

Inequality 5.26 is, here, given in a form which assumes that the origin is the equilibrium point. The non-linearities of the actual system do not comply with this requirement. However, it will be shown in the next section that the origin may be translated to the equilibrium point, but, as a consequence of this fact, expression 5.26 will have to be put into a slightly different form.

It should be noted that the conditions derived in this section imply that both the stub phasemifter control loop and the entire system are absolutely stable, since there are no gain restrictions on the non-linearities. The resulting stability requirements are, nevertheless, rather deceptive. Condition 5.26 does, in fact, represent a rather stringent restriction on the total system non-linearity, if the system is not a single variable system.

Moreover, the system non-linearities are periodic, which increases

the difficulties of proving stability. The problems of dealing with periodic non-linearities, and of complying with inequality 5.26 can be resolved only by considering the actual Lyapunov function associated with Anderson's theorem (equations 5.9 and 5.10). This will be done in a later section. In the next section the non-linearities will be tested for compliance with conditions 5.18 and 5.26.

5-4 The Non-Linearities and Stability

In chapter four it was established that, to each admittance value Y_L , there corresponds a pair of equilibrium values of x and y , which cause the load to become matched, and which are different for every new value of Y_L . Moreover, in view of the restrictions on the ranges of allowable motion of the phaseshifters, the pair of equilibrium values is unique for every load admittance value but one. The one exception is a load admittance $Y_L = 1 + j0$, which is already matched, but it will be shown later that a matched load Y_L presents no particular problem.

Let the equilibrium values of x and y , for a particular load admittance Y_L , be denoted by x_0 and y_0 . The values of x_0 and y_0 are almost never zero. Since all non-zero elements of the linear system part transfer function matrix, $\underline{W}(s)$, contain a free integrator, the system can come to rest with x_0 and y_0 not equal to zero, provided the error signals K_{RH} and K_{RV} are zero at x_0 and y_0 . But, x_0 and y_0 were defined to be precisely those values of x and y for which $K_{RH} = K_{RV} = 0$ for a particular Y_L .

Most stability criteria which, for their proof, depend on Lyapunov theorems, require that the origin be the equilibrium point (i.e. in

our case $x=y=0$ must be the equilibrium). If the equilibrium point and the origin do not coincide, the origin is usually translated to the equilibrium point. The inclusion of the free integrators in the transfer function matrix makes the origin translation quite legitimate.

Origin translation, if applied to the control system under discussion, changes conditions 5.18 and 5.26 to:

$$K_{RV}(0, y, Y_R)(y - y_{01}) \geq 0 \quad \text{for all } y \text{ and } Y_R \quad 5.27$$

$$K_{RH}(x, y, Y_L)(x - x_0) + K_{RV}(x, y, Y_L)(y - y_0) \geq 0$$

$$\text{for all } x, y \text{ and } Y_L \quad 5.28$$

where y_{01} is the equilibrium value of y which makes $K_{RV}=0$ for the stub loop alone, and where x_0 and y_0 are as defined before.

Inequality 5.28 includes the more conservative conditions:

$$K_{RH}(x, y, Y_L)(x - x_0) \geq 0 \quad \text{for all } x, y \text{ and } Y_L \quad 5.29$$

$$K_{RV}(x, y, Y_L)(y - y_0) \geq 0 \quad \text{for all } x, y \text{ and } Y_L \quad 5.30$$

It is important to note that expression 5.29 implies that y must appear in $K_{RH}(x, y, Y_L)$ merely as a single polarity gain parameter. Similarly, 5.30 implies that x must act only as a single sign gain parameter in $K_{RV}(x, y, Y_L)$. More specifically, if $K_{RV}(0, y, Y_R)$, $K_{RH}(x, y, Y_L)$, and $K_{RV}(x, y, Y_L)$ are plotted as functions of y , x , and y respectively with y_{01} , x_0 , and y_0 chosen as the origins, and with x , y , Y_R and Y_L varied as applicable, the resulting curves must lie entirely within the first and third quadrants and must pass through their respective origins.

The functions K_{RH} and K_{RV} are the projections upon the K_{RH} -axis

and K_{RV} -axis, respectively, of motion along the circular Smith chart trajectories discussed in section 4-4, particularly as x and y are varied separately. Closer examination of the non-linearity properties given in section 4-4 can show if conditions 5.27, 5.29 and 5.30 may be satisfied.

Clearly, since $K_{RV}(0, y, Y_R)$ is not a function of x , and since $K_{RV}=0$ at $y=y_{01}$ and y_{01} is the origin, and also since the constant conductance circles have their centers on the K_{RH} -axis of the Smith chart, it is indeed possible to satisfy condition 5.27 for all Y_R over the range and domain of definition of K_{RV} as determined by the dielectric slab motion stops.

On the other hand, conditions 5.29 and 5.30 cannot be satisfied, because the circular Smith chart trajectories of admittance Y , as x or y are varied (see section 4-4), change in radius as functions of y and Y_L , and have centers that are located in the first and fourth quadrants of the Smith chart on neither the K_{RH} -axis nor the K_{RV} -axis. Furthermore, none of these circles (except the ones for $y=y_0$) pass through the Smith chart origin. In fact, under certain conditions, these circular trajectories do not even intersect the K_{RH} -axis and K_{RV} -axis. Therefore, in plots of K_{RH} and K_{RV} versus x and y respectively, offsets from the origins x_0 and y_0 are certain to occur. The curves can, thus, not lie exclusively in the first and third quadrants, and 5.29 and 5.30 are not satisfied.

The points just discussed are brought out more clearly in sample plots of K_{RH} and K_{RV} versus x and y for various values of Y_L , as shown in figures 31, 32, and 33, with the applicable conditions marked directly on the figures. Also shown on these figures are the equilibrium

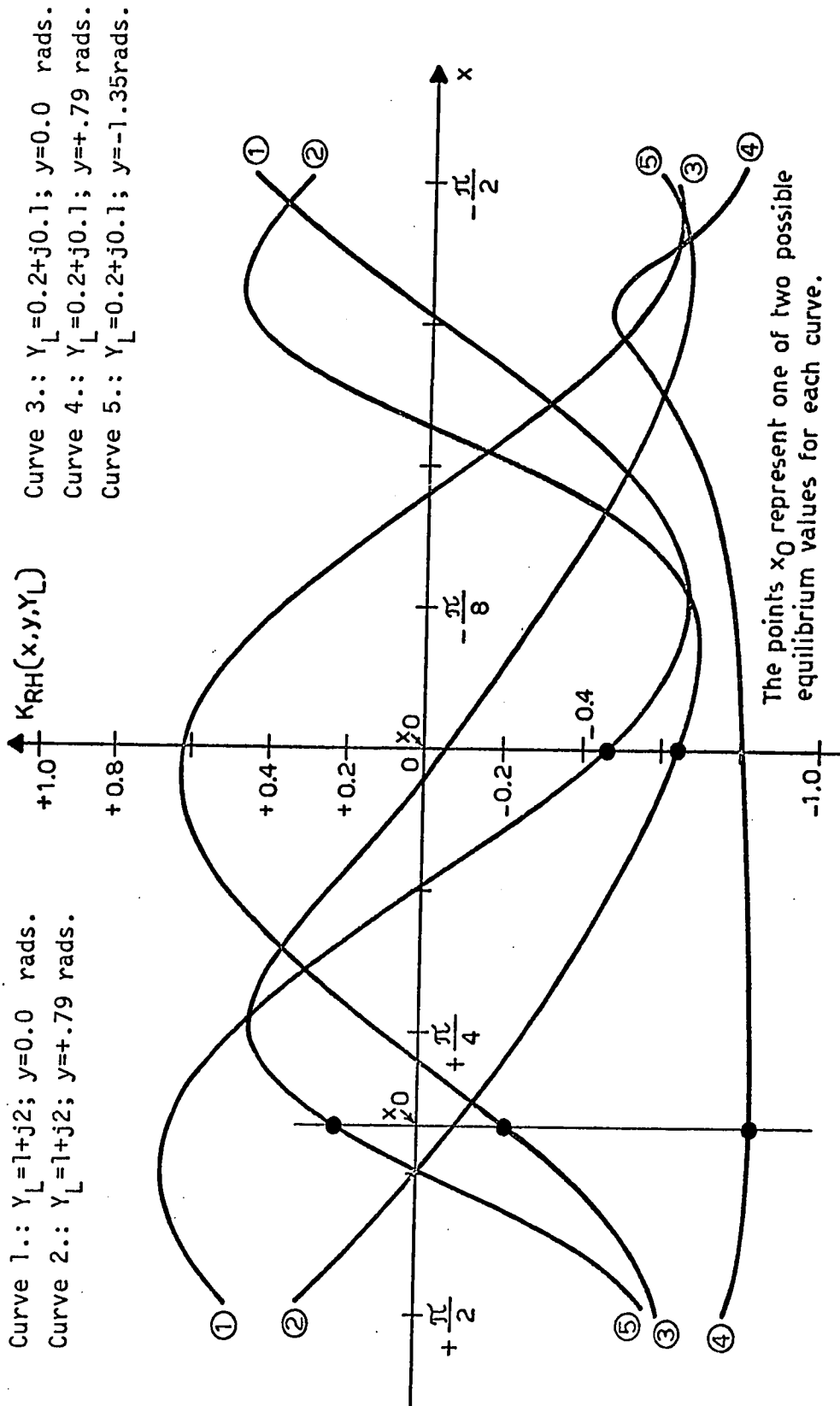


FIGURE 31. PLOTS OF $K_{RH}(x, y, Y_L)$ FOR VARIOUS VALUES OF y AND Y_L

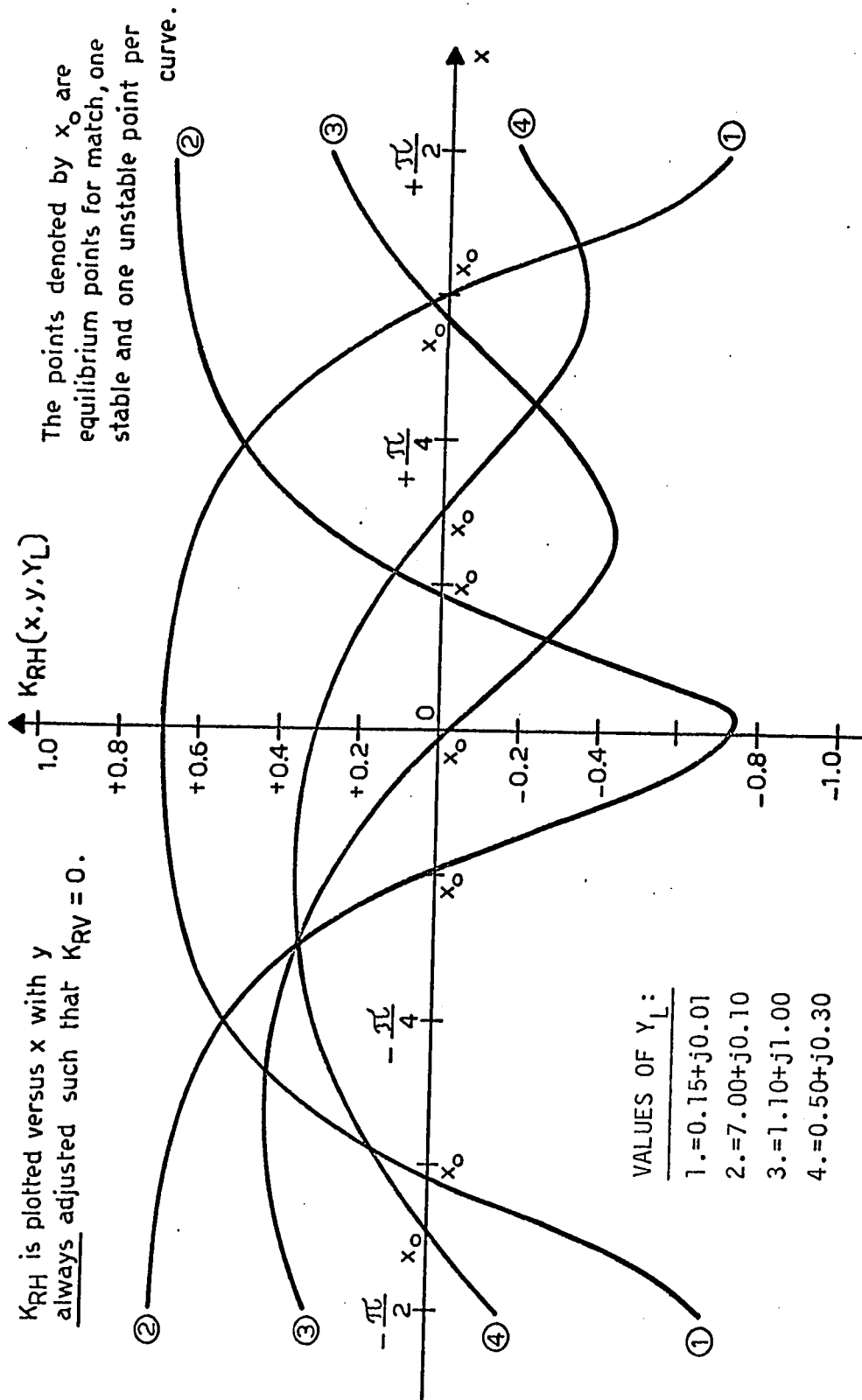


FIGURE 32. PLOTS OF $K_{RH}(x, y, Y_L)$ VERSUS x WITH y ADJUSTED FOR $K_{RV} = 0$

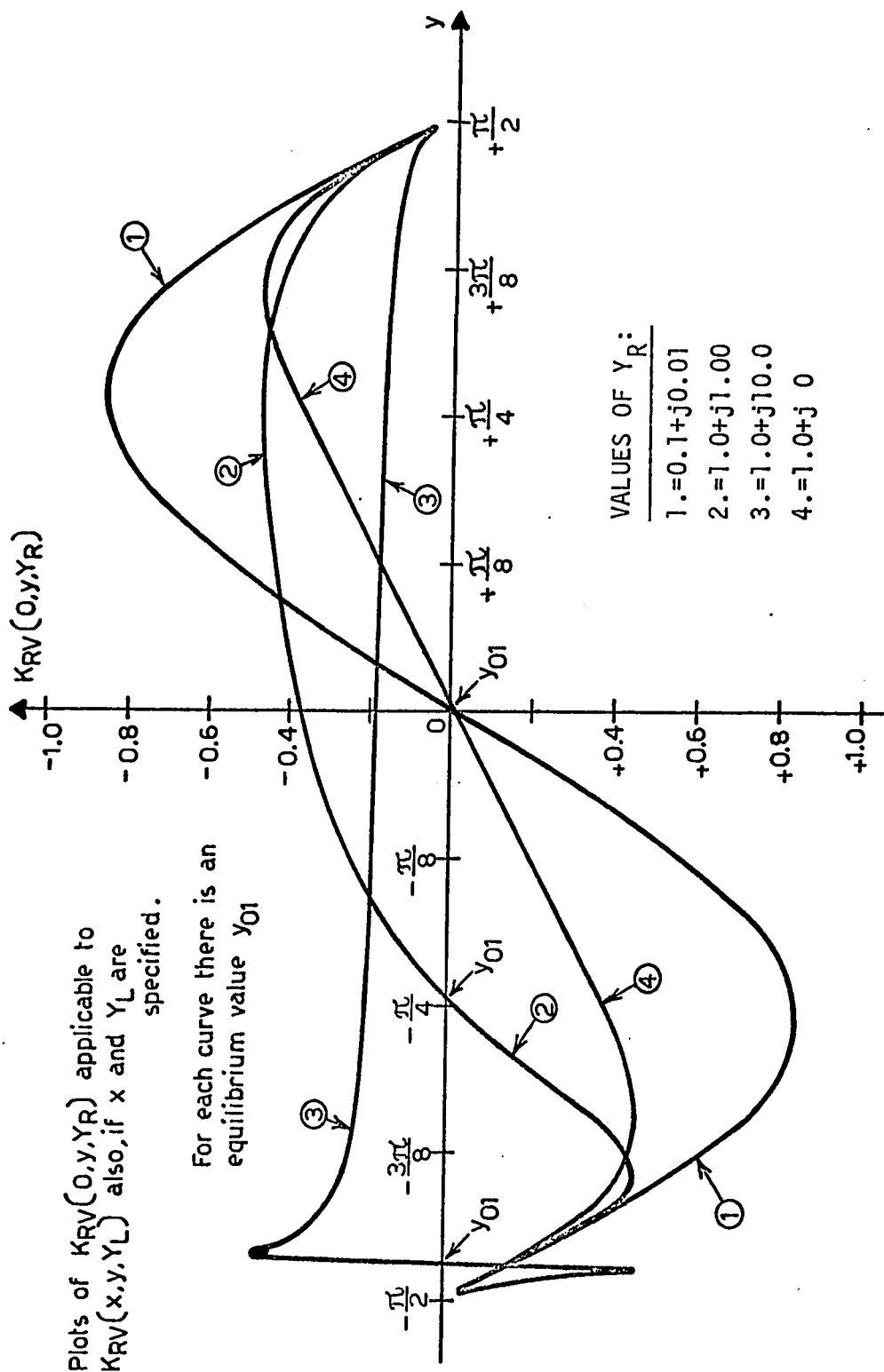


FIGURE 33. PLOTS OF $K_{RV}(0,y,Y_R)$ VERSUS y FOR VARIOUS VALUES OF Y_R

points corresponding to each curve. This makes the offsets just discussed rather obvious.

In figure 31, the K_{RH} curves, as x is varied, are plotted for one period of x only. This single period for each curve extends between specific phaseshift points, which will be discussed next.

Lyapunov stability theory, on which Anderson's theorem is dependent, requires that disturbances from equilibrium be considered as disturbances about the origin. For each specific load admittance value, Y_L , the corresponding equilibrium values of x and y (i.e. x_0 and y_0) may indeed be chosen as the origins. However, since the phaseshift functions are periodic, proper period endpoints must be chosen such that a single period is divided into two distinct portions, one with a positive error signal, and the other with a negative error signal.

For the stub phaseshifter, the period endpoints are readily established as $\pm\pi/2$ radians away from that value of y (slab position) for which the stub reactive admittance is zero. For $y=\pm\pi/2$, the stub admittance, and therefore Y , the admittance at the neutral plane of the H-plane tee, is $\pm\infty$. Since $Y=Y_R+Y_S$, where Y_S is the stub reactive admittance, and since the values $Y=\pm\infty$ and $Y=\pm j\infty$ all plot as a single point on the Smith chart (see figure 34), the period endpoints for the stub phaseshifter are not functions of either x or Y_L .

The choice of period endpoints for the series phaseshifter is governed by the sign of the error signal and the direction of the resultant motion for initial values of x close to the period endpoints. For specific values of Y_L and y , the circular Smith chart trajectory of Y may intersect the K_{RV} -axis twice. The period endpoints are chosen as those values of x farthest removed from the equilibrium value x_0

for which $K_{RH}=0$. this accomplishes the required division of the period into positive and negative error signal portions about x_0 , and at $y=y_0$. Even for values of y other than $y=y_0$, this choice of period endpoints is appropriate, since investigation has shown that these particular endpoints correspond to the largest initial values of x which do not cause matching device adjustment to an equilibrium point in an adjacent period. The above demonstrates that the period endpoints, and hence

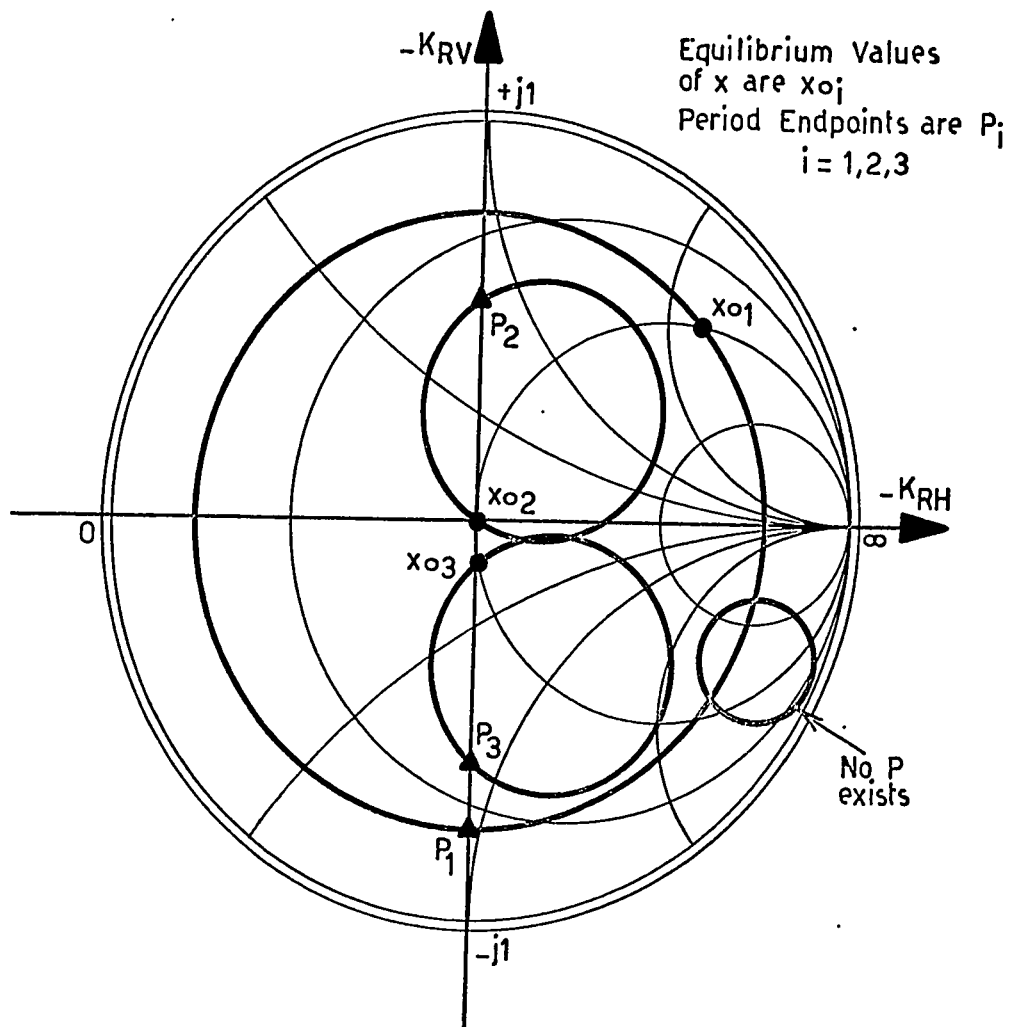


FIGURE 34. ILLUSTRATION OF PERIOD ENDPOINT LOCATION WITH RESPECT TO THE EQUILIBRIUM VALUES OF x

the definitions of the principal period, are variable and are functions of y and Y_L , in the case of the series phaseshifter. Some specific examples, where the period endpoints are denoted by P_i , are shown in figure 34.

Unfortunately, as explained in chapter four, for each specific load admittance Y_L , there are two critical values of y beyond which the circular trajectories of Y , as x is varied, do not intersect the K_{RV} -axis. As y is increased or decreased from y_0 , the intersection of the circles in question with the K_{RH} -axis usually disappears first, and as y is increased or decreased further, the intersection with the K_{RV} -axis disappears also. Whenever y and Y_L are such that no intersection with the K_{RV} -axis occurs, the period endpoints cannot be defined and are indeterminate. It is clear, however, that $(x-x_0)$ cannot be greater than π radians, the width of the period.

To return now to consideration of Anderson's conditions on the non-linearities.

It has already been established that conditions 5.29 and 5.30 cannot be satisfied, but it may still be possible to satisfy inequality 5.28. Unfortunately, the complexity of expressions 4.26 and 4.27 in table 1 is such that it is very difficult to prove compliance with inequality 5.28 by analytical means. Therefore, a series of numerical computations was performed, with the aid of a digital computer, to calculate the left hand side of inequality 5.28 for various conditions.

For each of a number of representative values of load admittance Y_L , the following computational steps were performed:

1. The equilibrium values x_0 and y_0 were computed from expressions 4.31 and 4.32.

2. A single period of y from $-\pi/2$ to $+\pi/2$ radians was divided into intervals 0.2 radians wide. Also, a single period of x from x_0 to $(x_0 - \pi)$ was divided into intervals 0.2 radians wide. An x - y grid was thus created.
3. For each grid value of y , K_{RV} and K_{RH} were computed for all grid values of x (equations 4.26 and 4.27).
4. The proper period endpoints of x , for each value of y , were detected from the first sign change in K_{RH} , and the period of x was shifted appropriately.
5. The left hand side of inequality 5.28 was computed for each pair of grid values of x and y .

The results obtained from the above computations may be summarized as follows:

1. For every load admittance Y_L , except a matched load or a load on the Smith chart perimeter, there is a range of values of y containing the origin (x_0, y_0) for which the Y circles resulting from varying x do intersect the K_{RV} -axis. For those values of y for which an intersection with the K_{RV} -axis exists, the left hand side of inequality 5.28 is zero at the origin and positive for all other values of x in the single period as defined.
2. For those values of y for every load, for which no K_{RV} -axis intersection exists, a proper period endpoint for x cannot be defined, and the sign of the left hand side of inequality 5.28 is indeterminate.
3. For those values of y and Y_L for which no K_{RV} -axis intersection exists, there are several possible choices of period

endpoints for x , which could cause inequality 5.28 to be satisfied.

The above shows that Anderson's conditions on the non-linearities are certainly not satisfied everywhere, although, for all values of x and for a limited range of values of y about the origin (x_0, y_0) , the left hand side of inequality 5.28 is positive definite.

It should be noted that, in the special case where Y_L is already a matched load, Y_R equals Y_L independent of the adjustment of the series phaseshifter. Therefore, no Y circle results from adjustment of x , and Anderson's conditions are satisfied at the origin only. However, this need not be a concern because, under these conditions, the effective gain of the non-linearity as a function of x , and hence the series phaseshifter loop gain, is zero. The series phaseshifter slab adjustment is thus indeterminate, but no instability can result.

Since Anderson's conditions cannot be met, stability of the system cannot be proven by direct application of Anderson's theorem. Nevertheless, for a system with a linear part transfer function matrix as given by expression 5.13, stability may still be established. This will be done in the next sections.

5-5 Stability of the Stub Phaseshifter Loop

Stability of the stub phaseshifter loop alone must first be established in order to prove asymptotic stability of the entire system.

The stub phaseshifter control loop stability conditions and non-linearity properties, derived in sections 5-3 and 5-4 respectively, may be summarized as follows:

1. The "matrix" $M_1(s)$ is Positive Real if either of the following sets of conditions holds:

$$k \geq 0; c_2 \geq 0; \alpha > 0; \beta > 0; a_2 \geq \alpha/\beta \quad (5.18b)$$

$$k \geq 0; c_2 \geq 0; \alpha = 0; \beta > 0; a_2 \geq 0 \quad (5.18a)$$

where c_2 is the gain of the linear part, a_2 is the damping loop gain, and $M_1(s)$ is as defined before.

2. In accordance with the restrictions on the motion of the actual phaseshifter slab, let y be restricted to $|y| < (\frac{\pi}{2} - \delta)$, where δ is an arbitrarily small positive constant. Then, for any value of Y_R for which an equilibrium value of y exists within the restricted range of y , the non-linear function $K_{RV}(0, y, Y_R)$ satisfies condition 5.27.

It should be emphasized at this point that, theoretically, δ may approach zero. Uniqueness of the stub phaseshifter adjustment is assured by restricting the slab motion as indicated (see chapter four). In a practical system, however, δ must be approximately 0.05 radians in order to make absolutely certain that, despite system inaccuracies, the K_{RV} error signal will change polarity only at the "origin", and not at the phaseshifter adjustment extremes also. For $\delta = 0.05$ radians, loads on or close to the Smith chart perimeter cannot be matched by the automatic device. But, this will be no drawback, because very low dissipation loads, such as these, cannot be matched in any case.

The mechanical restriction of the phaseshifter slab motion may be incorporated into the stub control loop block diagram by combining a

saturating non-linearity, NL_3 , with the non-linearity described by $K_{RV}(0, y, Y_R)$. This is shown in figure 35. The output from NL_3 normally equals the input, but, whenever y is greater than $(\frac{\pi}{2} - \delta)$, or less than $-(\frac{\pi}{2} - \delta)$, the output from NL_3 remains at $(\frac{\pi}{2} - \delta)$, or $-(\frac{\pi}{2} - \delta)$ respectively. Consideration of the Smith chart indicates that, in those circumstances where $|y| \geq (\frac{\pi}{2} - \delta)$, the output from the composite non-linearity is such that $|K_{RV}| > \epsilon$, where ϵ is a small positive constant. Clearly therefore, y may now be regarded as being unrestricted, while the composite non-linearity shown in figure 35 will satisfy condition 5.27 for all values of y .

Consider now the Lyapunov function and its time derivative, associated with the proof of Anderson's theorem, as given by expressions 5.9 and 5.10. These expressions may be rewritten in terms of a state vector:

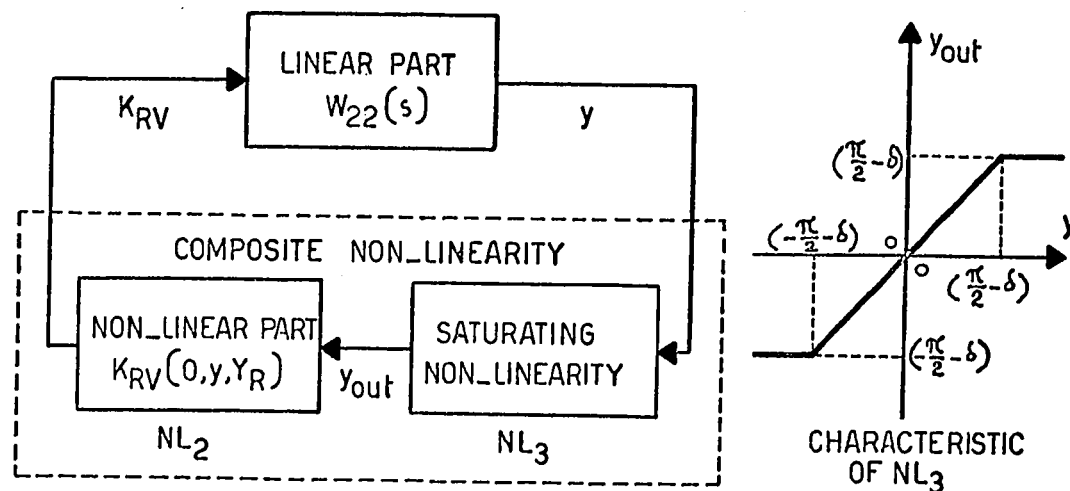


FIGURE 35. THE STUB CONTROL LOOP WITH RESTRICTED PHASESHIFTER SLAB MOTION

$$\underline{z}_y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

with $y_1=y$ and $y_2=dy_1/dt$ (see equation 4.10). The expressions for the Lyapunov function and its time derivative become:

$$V(\underline{z}_y) = \underline{z}_y^T \underline{P} \underline{z}_y + 2\beta \int_0^{\underline{H}^T \underline{z}_y} \underline{f}^T(\underline{q}) d\underline{q} \quad 5.31$$

$$\begin{aligned} \dot{V}(\underline{z}_y) = & - [\underline{z}_y^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}_y) \underline{W}_0^T] [\underline{L} \underline{z}_y - \underline{W}_0 \underline{f}(\underline{H}^T \underline{z}_y)] \\ & - 2\alpha [\underline{f}^T(\underline{H}^T \underline{z}_y) \underline{H}^T \underline{z}_y - \underline{f}^T(\underline{H}^T \underline{z}_y) \underline{K} \underline{f}(\underline{H}^T \underline{z}_y)] \end{aligned} \quad 5.32$$

Since the "matrix" $M_1(s)$ is Positive Real (see point 1 above), the matrix \underline{P} is positive definite and the matrix \underline{L} exists by Anderson's theorem. It was also shown that the composite non-linearity, consisting of the saturating non-linearity and $K_{RV}(0,y,Y_R)$, obeys condition 5.27 and lies entirely in the first and third quadrants for any value of y , and passes through the origin corresponding to each particular value of Y_R .

Therefore:

$$\begin{aligned} V(\underline{z}_y) & \text{ is Positive Definite} \\ V(\underline{z}_y) & \rightarrow \infty \text{ for } \|\underline{z}_y\| \rightarrow \infty \end{aligned} \quad 5.33$$

The row matrix $[\underline{z}_y^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}_y) \underline{W}_0^T]$, occurring in expression 5.32, generally has elements which are zero, positive, or negative. At least one of its elements may change polarity at points other than the origin.

Since:

$$[\underline{z}_y^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}_y) \underline{W}_0^T] = [\underline{L}^T \underline{z}_y - \underline{W}_0 \underline{f}(\underline{H}^T \underline{z}_y)]^T$$

the product

$$[\underline{z}_y^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}_y) \underline{W}_0^T] [\underline{L}^T \underline{z}_y - \underline{W}_0 \underline{f}(\underline{H}^T \underline{z}_y)]$$

represents a scalar, which may be zero at points other than the origin, but which is otherwise positive.

Suppose that conditions 5.18a are chosen as the pertinent stability conditions. This allows the damping loop gain a_2 to be zero, but, since α is zero, $\dot{V}(\underline{z}_y)$ is at best negative semi-definite. Thus, mere stability with the possibility of a stable limit cycle has been demonstrated.

Instead, select conditions 5.18b as appropriate stability conditions. These conditions indicate that the damping loop gain may assume any positive value, but may not be zero. By condition 5.18b, α is positive only. Also, since the non-linearity satisfies condition 5.27, the scalar

$$-2\alpha [\underline{f}^T(\underline{H}^T \underline{z}_y) \underline{H}^T \underline{z}_y - \underline{f}^T(\underline{H}^T \underline{z}_y) \underline{K} \underline{f}(\underline{H}^T \underline{z}_y)]$$

is zero at the origin only, and negative elsewhere.

Therefore:

$$\dot{V}(\underline{z}_y) \text{ is Negative Definite}$$

According to Lyapunov's second method stability theorem, conditions 5.33 and 5.34 assure asymptotic stability in the large of the stub control loop, provided the equilibrium value y_0 corresponding to a particular value of Y_R lies within the restricted range of phase-shifter slab motion, and provided also that $c_2 \geq 0$ (negative feedback), and $a_2 > 0$ (some loop damping is incorporated).

The stub phaseshifter control loop, in fact, exhibits absolute stability, since any non-linearity, which satisfies condition 5.27, leaves the stub control loop stable. Moreover, the stub phaseshifter loop will stably reach an equilibrium from any arbitrary set of initial conditions.

In an actual system, external access to the control loop is obtained only through admittance Y_R . Since the stub control loop non-linearity is a function of phaseshifter slab position y and admittance Y_R (see chapter four), a step change in the value of Y_R affects the nature of the non-linearity, and also suddenly causes a translation of the equilibrium value of y to a new value y_0 . Changes in the non-linearity do not affect the loop, because the loop is absolutely stable. The step translation of the equilibrium point y_0 (origin) merely corresponds to a set of new initial conditions, which the loop can readily deal with in a stable manner.

5-6 Stability of the Complete System

For the most part, the proof of stability of the entire system closely parallels the stability proof for the stub control loop given in the previous section.

A summary of pertinent system stability conditions and non-linearity properties, as derived in sections 5-3 and 5-4, may be presented as follows:

1. The matrix $\underline{M}(s)$ is Positive Real if either one of the following conditions holds:

$$c_1 \geq 0 \quad c_2 \geq 0 \quad \alpha = 0 \quad \beta > 0 \quad a_1 \geq 0 \quad a_2 \geq 0 \quad (5.23a)$$

$$c_1 \geq 0 \quad c_2 \geq 0 \quad \alpha > 0 \quad \beta \geq 0 \quad a_1 \geq \frac{\alpha}{\beta} \quad a_2 \geq \frac{\alpha}{\beta} \quad (5.23b)$$

where c_1 and c_2 are the gains of the linear parts, and a_1 and a_2 are the damping loop gains. Matrix $\underline{M}(s)$ was defined before.

The positive real property of $\underline{M}(s)$ is a necessary condition for stability.

2. The total system non-linearity must satisfy:

$$K_{RH}(x, y, Y_L)(x - x_0) + K_{RV}(x, y, Y_L)(y - y_0) \geq 0$$

for all x, y and Y_L (5.28)

Condition 5.28 is satisfied for all values of x , but only for a limited range of values of y . The allowable range of y values is that for which the Smith chart Y circles, which result from varying x , intersect the K_{RV} -axis. The values of x and y , for which condition 5.28 is satisfied, are plotted in an x - y plane in figure 36, and are denoted by region R . The dimension of region R , along the y -axis, is a function of Y_L

and increases with increasing load reflection coefficient. The equilibrium point (x_0, y_0) , for a particular Y_L , and the point $x=y=0$ always lie within region R.

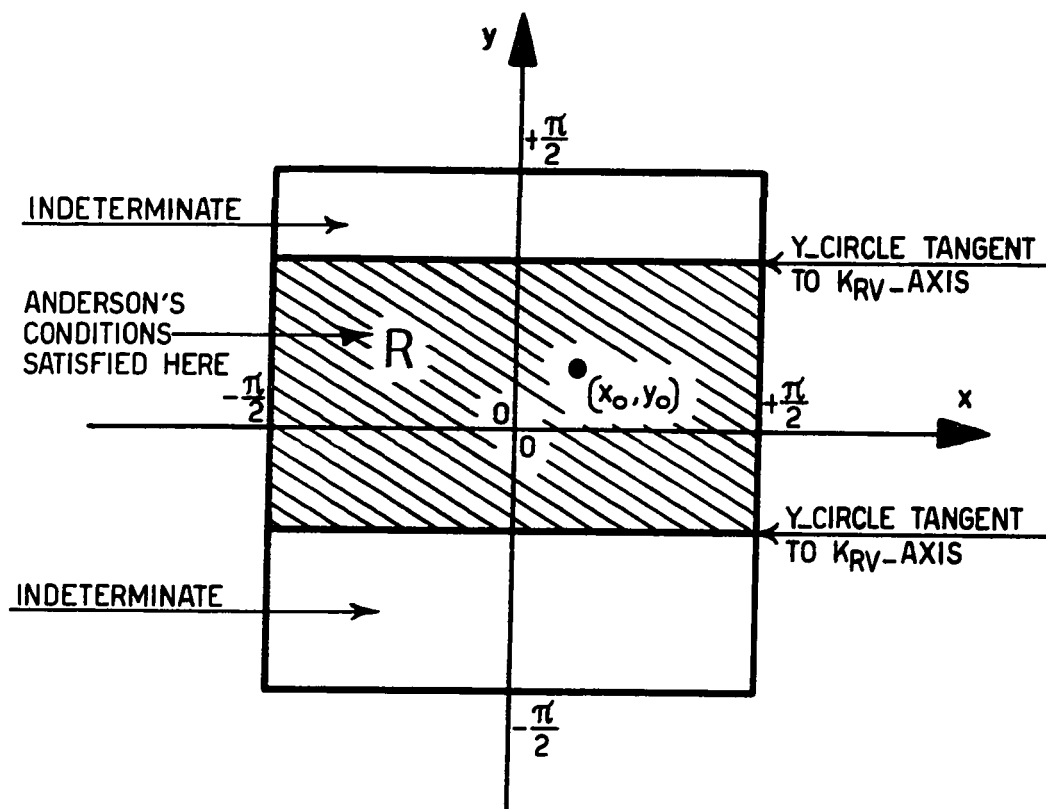


FIGURE 36. DIELECTRIC SLAB POSITIONS SATISFYING ANDERSON'S CONDITIONS ON THE NON-LINEARITIES

Assume that the stub phasemitter slab motion is restricted mechanically in such a way that the stub and series phasemitter slab motions are confined to region R of figure 36. Furthermore, let the mechanical restriction on the stub slab motion be accounted for, in the system block diagram, by means of a saturating non-linearity, inserted into the stub loop, ahead of the non-linearity representing $K_{RV}(x, y, Y_L)$. This is shown in detail in figure 37. The most important

consequence of this procedure is, that condition 5.28 is now satisfied for all x , y , and Y_L . The explanation of the effect of this saturating non-linearity is analogous to that given in section 5-5 for the stub phaseshifter control loop alone.

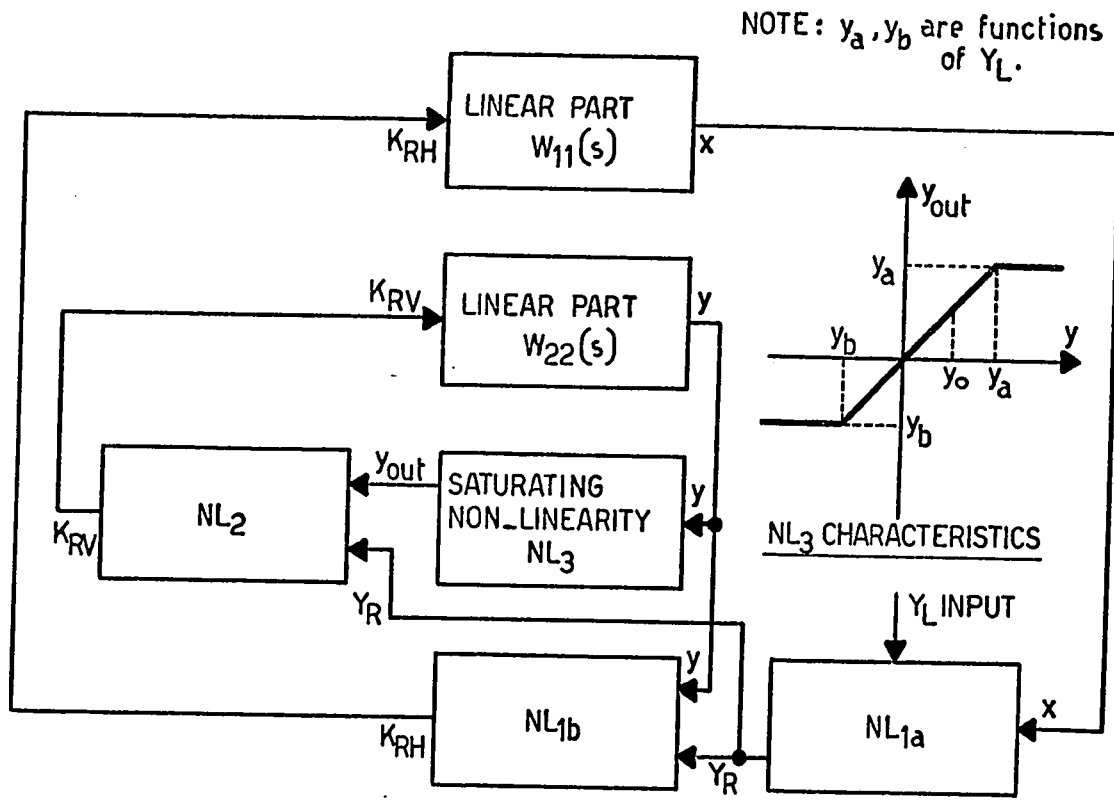


FIGURE 37. THE SYSTEM BLOCK DIAGRAM WITH RESTRICTED STUB PHASESHIFTER SLAB MOTION

The applicable expressions for the Lyapunov function and its time derivative, based on 5.9 and 5.10, are given by:

$$V(z) = z^T P z + 2\beta \int_0^{\infty} z^T H z f^T(\sigma) d\sigma \tag{5.35}$$

$$\begin{aligned} \dot{V}(\underline{z}) = & -[\underline{z}^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}) \underline{W}_0^T] [\underline{L}^T \underline{z} - \underline{W}_0 \underline{f}(\underline{H}^T \underline{z})] \\ & - 2\alpha [\underline{f}^T(\underline{H}^T \underline{z}) \underline{H}^T \underline{z} - \underline{f}^T(\underline{H}^T \underline{z}) \underline{K} \underline{f}(\underline{H}^T \underline{z})] \end{aligned} \quad 5.36$$

where \underline{z} is the state vector of equation set 4.10, and is defined as:

$$\underline{z} = [x, \dot{x}, y, \dot{y}]^T$$

Select conditions 5.23b as the appropriate system operating conditions (some damping included). Then, the matrix $\underline{M}(s)$ is Positive Real, and by Anderson's theorem \underline{P} is positive definite and \underline{L} exists. Also, since the composite system and saturating non-linearities obey condition 5.28 for all x, y , and \underline{Y}_L , there results:

$$\begin{aligned} V(\underline{z}) & \text{ is Positive Definite} \\ V(\underline{z}) & \rightarrow \infty \text{ for } \|\underline{z}\| \rightarrow \infty \end{aligned} \quad 5.37$$

Once again, just as for the separate stub control loop stability proof given in section 5-5, expression

$$[\underline{z}^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}) \underline{W}_0^T] [\underline{L}^T \underline{z} - \underline{W}_0 \underline{f}(\underline{H}^T \underline{z})]$$

is at best positive semi-definite, but with condition 5.28 satisfied and with $\alpha > 0$ from condition 5.23b, one has:

$$\dot{V}(\underline{z}) \text{ is Negative Definite} \quad 5.38$$

Therefore, by 5.37 and 5.38 and by the stability theorem of Lyapunov's second method, the system with restricted stub phaseshifter motion is asymptotically stable in the large.

Note particularly, that the velocities \dot{x} and \dot{y} in the actual

system (assuming it is stable) are bounded. Hence, with sufficiently large damping incorporated into the system, and without the type of mechanical slab motion restriction described above, a state point may enter region R of figure 36, and be captured there subsequently to proceed towards the origin in a stable manner.

Now reconsider the entire system. Insert mechanical motion stops in the stub phaseshifter, as required in an actual system, exactly as was done in section 5-5, such as to limit the obtainable phaseshift to $\pm(\frac{\pi}{2}-\delta)$ radians, where δ is a small positive constant. Any other phaseshifter motion stops have been removed at this point. Consequently, the Lyapunov function of Anderson's theorem needs re-examination.

A summary of applicable system non-linearity properties, based on the work of chapter four and section 5-4, may be given as follows:

1. Condition 5.28 is satisfied only in region R of figure 36. Outside region R the sign of the left hand side of inequality 5.28 is indeterminate.
2. The system non-linearity is periodic in x and y . Since the series phaseshifter slab motion x has not been mechanically restricted (the periods have only been folded as described in section 4-5), the periodicity in x will be apparent.
3. For values of y such that $|y| \geq (\frac{\pi}{2}-\delta)$, the outputs from the non-linearity will be:

$$|K_{RV}| \rightarrow \epsilon; K_{RH} \rightarrow -1,$$

where ϵ is a small constant.

4. For all values of x , y , and Y_L , the outputs from the non-linearity are bounded; i.e.

$$|K_{RH}|_{\max} = 1; |K_{RV}|_{\max} = 1.$$

On the basis of the above summary, a worst case curve and envelope for the normalized values of the left hand side of inequality 5.28 are shown plotted, in figure 38, as a function of the norm of the slab displacement vector. Note that σ is merely a dummy variable. The portions of the curves to the left of line "r" in figure 38 correspond to region R of figure 36. Anderson's conditions are satisfied there, and the normalized product of the non-linearity output and slab displacement vectors is positive semi-definite. To the right of line "r", the worst case value of the same function is -1.

The Anderson type Lyapunov function for the system is still given by expression 5.35. According to this expression, the worst case non-linearity output may be integrated with respect to the slab displacement vector to produce a worst function representing:

$$\int_{\underline{0}} \underline{H}^T \underline{z} \underline{f}^T(\underline{\sigma}) d\underline{\sigma}$$

Consider, now, the term $\underline{z}^T \underline{P} \underline{z}$, where, by Anderson's theorem, \underline{P} is a positive definite matrix. Clearly, this is a quadratic form which is zero at the "origin". A typical range of values of this function, as \underline{z} is varied, is plotted in figure 38.

It is obvious from the final curves of figure 38, that a sufficiently small positive value of β may be selected such that:

$V(\underline{z})$ is Positive Definite

$$V(\underline{z}) \rightarrow \infty \text{ for } \|\underline{z}\| \rightarrow \infty$$

5.39

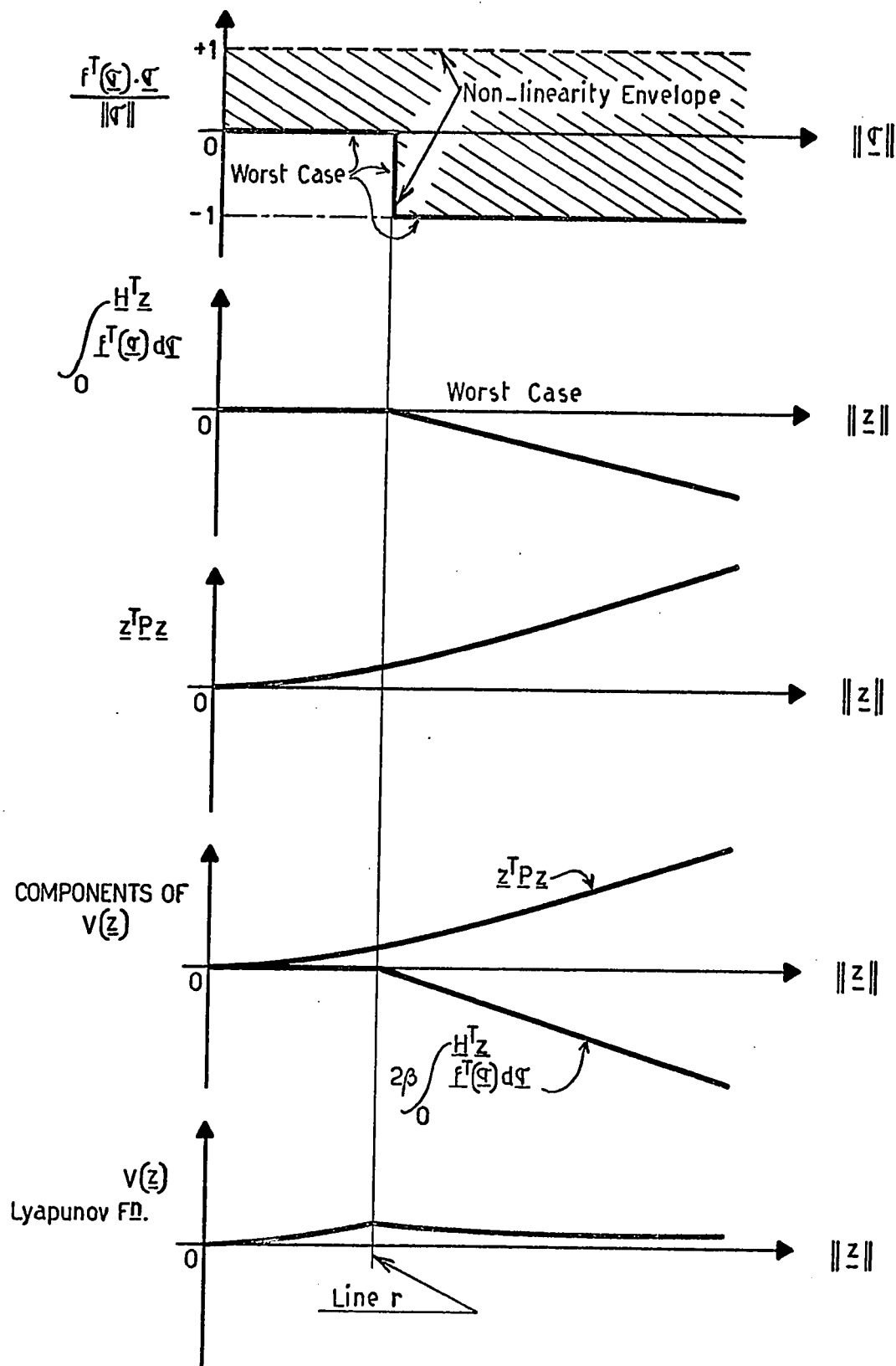


FIGURE 38. POSITIVE DEFINITE NATURE OF THE SYSTEM LYAPUNOV FUNCTION

for the worst case of the non-linearity outputs, and therefore, for all possible outputs from the non-linearity.

To prove stability of the entire system, select conditions 5.23a as the appropriate Anderson conditions. That is:

$$\alpha = 0$$

$$\beta > 0 \text{ but sufficiently small for } V \text{ to be Pos. Def.}$$

$$c_1 \geq 0; c_2 \geq 0$$

$$a_1 \gg 0; a_2 \gg 0 \text{ for considerable damping}$$

Therefore, by Anderson's theorem, the matrix \underline{L} exists and the time derivative of the Lyapunov function becomes:

$$\dot{V}(\underline{z}) = -[\underline{z}^T \underline{L} - \underline{f}^T(\underline{H}^T \underline{z}) \underline{W}_0^T][\underline{L} \underline{z} - \underline{W}_0 \underline{f}(\underline{H}^T \underline{z})] \quad 5.40$$

According to previously presented arguments, one has from 5.40:

$$\dot{V}(\underline{z}) \text{ is Negative Semi-definite} \quad 5.41$$

From Lyapunov's stability theorem, and from 5.39 and 5.41, one may conclude that the entire system is stable. Nevertheless, there remains the possibility of the existence of a stable limit cycle.

It was shown above that, for initial values of x and y in region R of figure 36 (and for any initial velocities), the system is asymptotically stable, if the damping is not zero, and if y is restricted to region R . We selected $a_1 \gg 0$ and $a_2 \gg 0$. Therefore, with y restricted to region R , asymptotic stability prevails. Consequently, even if

y were not so restricted, region R can certainly not support a limit cycle. In fact, for a large range of velocities of entry into region R, an asymptotic approach of the equilibrium point is guaranteed.

From the block diagrams of figure 30 and 37 it is clear that the stub phaseshifter can be considered as being somewhat independent of the series phaseshifter loop, but the converse is not true. More importantly, for values of y outside region R, and as x is varied, the Smith chart Y admittance circles intersect neither the K_{RV} -axis nor possibly the K_{RH} -axis. The non-linearity properties are such that a stable limit cycle, for which y does not enter region R (to be captured there), cannot exhibit symmetry about the K_{RH} -axis on the Smith chart, and will lie mostly in one quadrant of the Smith chart. The projection of this limit cycle on the K_{RV} -axis represents the error signal K_{RV} for the stub phaseshifter control loop. For this limit cycle, the error signal will contain a sizable "DC" component.

On the other hand, it was shown that the stub phaseshifter control loop, separately, is asymptotically stable in the large. This control loop contains a free integrator. The stub control loop cannot, therefore, support an error signal with a lasting DC component. The average magnitude of K_{RV} must decrease, and consequently, the average magnitude of y must also decrease. As $|y|$ decreases, the system state will enter region R of figure 36, resulting in a stable approach of the equilibrium or origin. No stable limit cycle is thus possible.

Asymptotic stability in the large of the entire matching system under consideration has now been demonstrated.

5-7 System Response Speed

Once a system has been proven stable, considerable attention is customarily devoted to analysis of the system response speed. In the case of this system, that task is greatly complicated by the fact that there are two interacting loops, each with its own non-linearity. To make matters worse, the non-linearities are functions of the load admittance Y_L . As a result, some of the normally used methods of estimating the overall system response speed are impractical.

For instance, the method of using the normalized rate of change, $\dot{V}(z)/V(z)$, of the Lyapunov function to estimate the response speed, cannot be used here, because only the general forms of the Lyapunov function and its derivative are known. And, the Lyapunov function changes with load admittance Y_L .

Perhaps the most practical method of investigating the system response speed, is to linearize the system about the equilibrium point (x_0, y_0) , for every load admittance Y_L . The response speed of the linearized system may then be investigated instead. For values of x and y far removed from (x_0, y_0) , some impression of the rates of adjustment of the phaseshifters may be obtained by considering the actual magnitudes of the error signals K_{RH} and K_{RV} , and by regarding these as activating two second order linear loops.

Linearizing of system state equations 4.10, by means of a Taylor series expansion about equilibrium point (x_0, y_0) and considering the first order terms only, while ignoring the constant forcing terms necessary to keep the system at the equilibrium point (which does not change the character of a linear system), results in:

$$\begin{aligned}
 \dot{x}_1 &= x_2 \\
 \dot{x}_2 &= -c_1 G_1 x_1 - a_1 x_2 - c_1 G_2 y_1 \\
 \dot{y}_1 &= y_2 \\
 \dot{y}_2 &= -c_2 G_3 x_1 - a_2 y_2 - c_2 G_4 y_1
 \end{aligned}
 \tag{5.42}$$

where:

$$\begin{aligned}
 G_1 &= \left. \frac{\partial K_{RH}(x_1, y_1, Y_L)}{\partial x_1} \right|_{x_0, y_0} \\
 G_2 &= \left. \frac{\partial K_{RH}(x_1, y_1, Y_L)}{\partial y_1} \right|_{x_0, y_0} \\
 G_3 &= \left. \frac{\partial K_{RV}(x_1, y_1, Y_L)}{\partial x_1} \right|_{x_0, y_0} \\
 G_4 &= \left. \frac{\partial K_{RV}(x_1, y_1, Y_L)}{\partial y_1} \right|_{x_0, y_0}
 \end{aligned}
 \tag{5.43}$$

Based on equations 5.42, a signal flow graph for the linearized system may be drawn as in figure 39. By Mason's rule⁽²²⁾, the characteristic equation for the linearized system is then:

$$\begin{aligned}
 \Delta = s^4 + (a_1 + a_2) s^3 + (a_1 a_2 + c_1 G_1 + c_2 G_4) s^2 + \\
 + (a_1 c_2 G_4 + a_2 c_1 G_1) s + c_1 c_2 (G_1 G_4 - G_2 G_3)
 \end{aligned}
 \tag{5.44}$$

where s is a complex frequency operator (Laplace).

Some general conclusions may first be drawn concerning this system (figure 39 and equation 5.43):

1. For this fourth order system, the s -plane root locations

are markedly affected by all of the constants a_1 , a_2 , c_1 , c_2 , G_1 , G_2 , G_3 , and G_4 .

2. There is a specific range of values that each of the constants G_1 , G_2 , G_3 , and G_4 can assume as the load admittance Y_L is varied. Since none of the constants in the characteristic equation appear merely as gain constants, but rather affect also the open-loop pole and zero locations, it is entirely possible that varying any one of these constants over a large range of positive values will not cause instability, but will only change the character of the system response.
3. Routhian array⁽²³⁾ analysis of the characteristic equation has shown that appropriate ranges of values exist for the constants, such that the system will be stable.

The actual range of values of each of the constants G_1 , G_2 , G_3 , and G_4 , as the load admittance Y_L is varied, may be very approximately

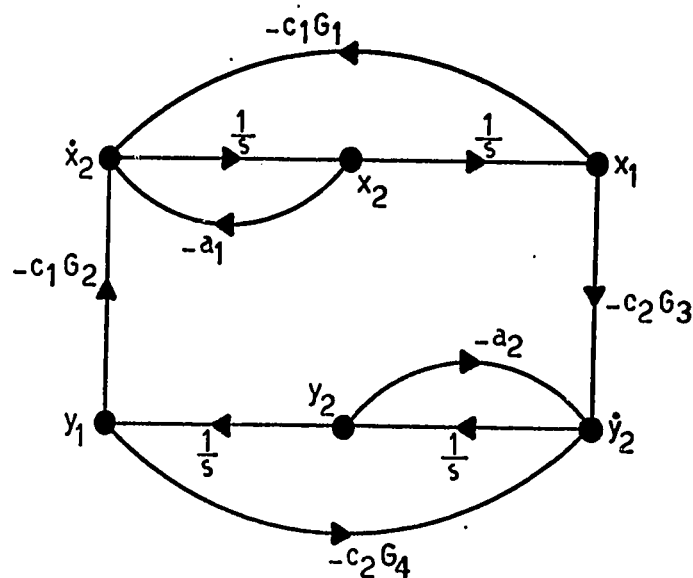


FIGURE 39. SIGNAL FLOW GRAPH FOR THE LINEARIZED SYSTEM

established by consideration of figures 31, 32, and 33, and the material in chapter four and section 5-4. The ranges of values of each of these constants, expressed in terms of the load reflection coefficient magnitude K_{RL} rather than Y_L , are as follows:

1. $|G_1| = 0$ for $K_{RL} = 0$
 $|G_1| \approx 1$ for $K_{RL} \rightarrow 1$
2. $|G_2| = 0$ for all K_{RL}
3. $|G_3| = 0$ for $K_{RL} = 0$
 $|G_3| \approx 1$ for $K_{RL} \rightarrow 1$
4. $|G_4| = 1$ for $K_{RL} = 0$
 $|G_4|_{\max} = g$ for $K_{RL} \rightarrow 1$ with the imaginary part of Y_L large.

g is a large positive constant whose value depends on how closely the stub phaseshift is allowed to approach $\pm\pi/2$ radians.

Clearly, with $G_2=0$, the number of loops in the signal flow graph is reduced and the system now merely consists of two cascaded second order systems. The new characteristic equation becomes:

$$\Delta = (s^2 + a_1s + c_1G_1)(s^2 + a_2s + c_2G_4) \quad 5.45$$

The two separate second order parts of the system are always stable, while, for a given set of values for G_1 and G_4 , the other constants a_1 , a_2 , c_1 , and c_2 may be adjusted to give any desired type of response from overdamped through critically damped to underdamped, at any speed of response. Obviously, for fixed values of a_1 , a_2 , c_1 , and c_2 , the load admittance Y_L , to be matched, will have a profound effect on

the type of response obtained from the system, since G_1 and G_4 depend on Y_L .

The linearized model of the system is highly inaccurate, because even small deviations away from the equilibrium point will greatly change G_1 , G_2 , G_3 , and G_4 . Especially G_2 is seriously affected and will no longer be zero, thus completely changing the nature of the system. However, the fact that the linearized system is unconditionally stable proves that, in the actual system, the origin itself, at least, is asymptotically stable. This supports some of the theory presented in chapter five.

Far away from the equilibrium point, the relative magnitudes of the error signals K_{RH} and K_{RV} determine how rapidly each of the linear parts characterized by:

$$\frac{c_1}{s(s + a_1)} \quad \text{and} \quad \frac{c_2}{s(s + a_2)}$$

will respond and adjust x and/or y . For example, for load admittances close to the K_{RH} -axis on the Smith chart, the series phaseshifter will start to adjust most rapidly. On the other hand, for loads close to the K_{RV} -axis, the stub phaseshifter, initially, will adjust most rapidly. In general, these effects will become more obvious in the material of the next chapter.

The following chapter will consider stability and response speed further, but from an experimental and practical point of view.

CHAPTER SIX

SYSTEM SIMULATION AND PRACTICAL CONSIDERATIONS

An automatic microwave load matching system was analyzed in some detail and the stability of the somewhat idealized closed-loop system was demonstrated. But, the complexity of the system non-linear part prevented the presentation of a stability proof that was mathematically rigorous.

Similarly, the character and speed of response of the system were considered, but no accurate details could be presented. The system response varies a great deal with operating conditions and microwave loading, making detailed analysis impractical.

To show by alternate means that this system is indeed asymptotically stable and that time-varying inputs can be stably handled, a computer simulation was carried out. At the same time, the computer simulation served two other purposes. The types of system response and the response speed could be examined, and experimentation could be carried out to attempt improvement of the system.

This chapter deals in some detail with the above matters.

Additional discussions in this chapter are concerned with the properties of an actual system and its components. Stability of an actual system, particularly, needs further consideration, since in the idealized system analyzed thus far, some important component properties were neglected.

6-1 Computer Simulation of the System

The performance of the proposed microwave load matching system was tested, in detail, by means of a computer simulation.

Generally, computer simulations of control systems may take the form of all digital simulation, all analogue simulation, or analogue-digital hybrid simulation. The latter method was favored for the following reasons:

1. Real-time simulation of a system offers the advantages of easy observation of system variables and system operation, greater freedom to experiment and to examine the effects of modifications immediately, and convenient means of developing a "feeling" for and understanding of system operation.
2. The linear part of the system is easily implemented on an analogue computer, and initial conditions and inputs are easily applied (and varied while the system is running).
3. The non-linearity of the system under consideration is sufficiently complex that it cannot be simulated on an analogue computer. Digital simulation of the same non-linearity, on the other hand, is easy and accurate.

The hybrid computer system used consisted of a standard analogue computer with electronic mode control, operable from the digital computer, analogue to digital and digital to analogue multi-channel conversion equipment, and a small digital computer, with typewriter input and output device. The analogue computer was used only to build up the linear part of the system and to incorporate the necessary loop gain. The required transfer functions for the linear part (two loops) are

given by the elements of the transfer function matrix defined by expression 5.13. The method of implementation was similar to that shown in block diagram form in figure 25. For each loop, only two integrators were needed, one of which had damping feedback applied around it. Complete details of this part of the simulation are shown in figure 40.

The non-linearity could have been simulated by programming formulas 4.26 and 4.27 into the digital computer. However, the digital computation is slightly faster if formulas 4.15, 4.16, 4.18 and 4.19 are used instead. For each run, a particular value of load admittance Y_L was read into the computer via the typewriter. Based on Y_L and current values of the dielectric slab positions x and y , the digital computer was programmed first to calculate G and B , the real and imaginary parts respectively of the admittance at the neutral plane of the H-plane tee (by 4.15 and 4.16). The computer would then convert the values of G and B into current values of K_{RH} and K_{RV} through use of formulas 4.18 and 4.19. Communication between the analogue and digital computers was obtained by means of the A/D and D/A converters. The digital computer would repeatedly update K_{RH} and K_{RV} as the sampled current values of x and y changed.

The simulation of the entire system, shown diagrammatically in figure 40, was quite accurate in that no approximations were made with respect to either the linear part or the non-linearity. However, the simulation was essentially a sampled-data system. To make the simulation appear continuous, the digital computation had to be performed as rapidly as possible, and the linear part had to be made purposely slow. As long as no attempt was made to increase the response speed excessively, the simulation was good.

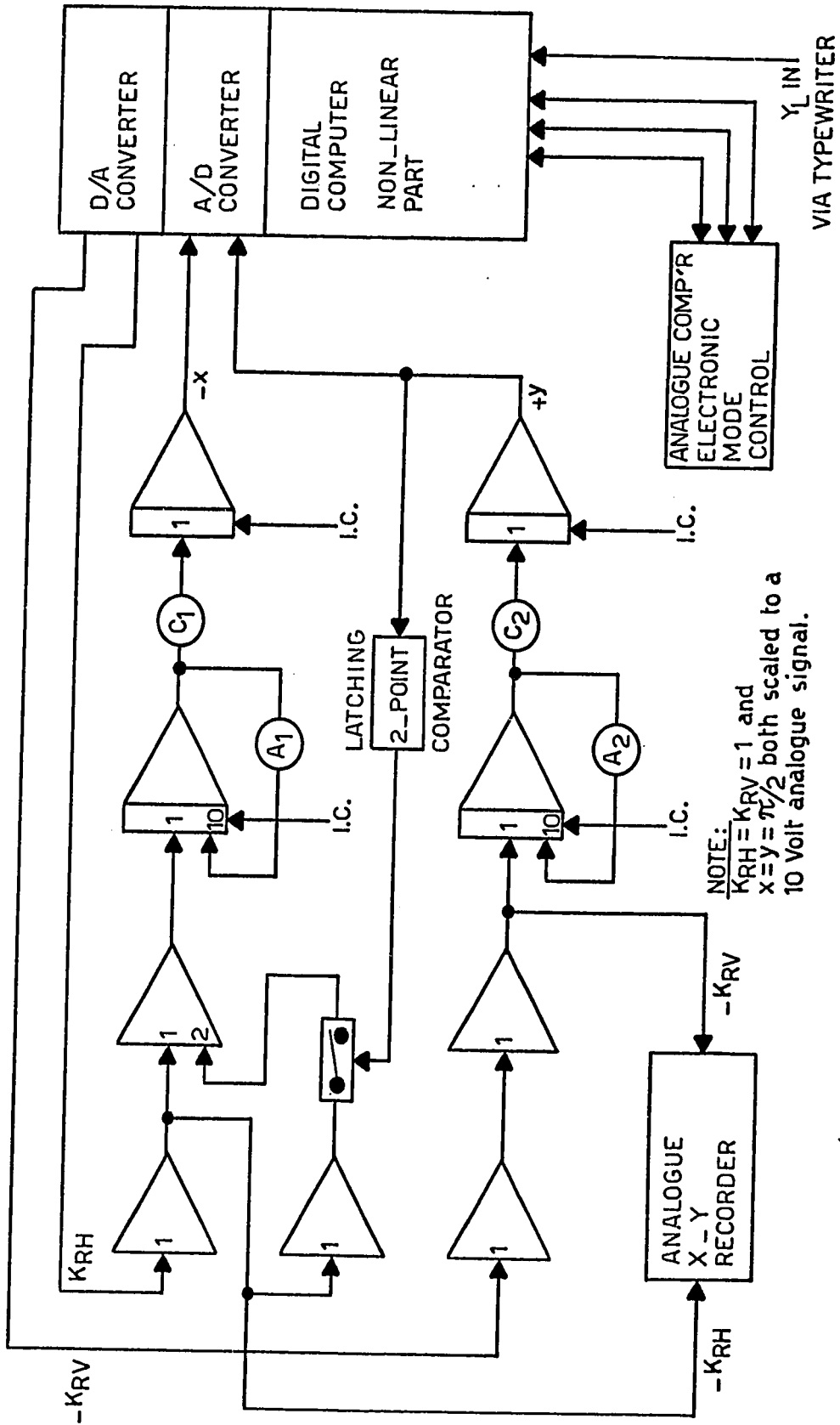


FIGURE 40. SYSTEM SIMULATION ON A HYBRID COMPUTER

The changes in the dielectric slab positions, x and y , could be monitored directly on the analogue computer. More importantly, the analogue values of K_{RH} and K_{RV} could be scaled to be plotted directly on a Smith chart. This produced an accurate record of the Smith chart trajectories of either the reflection coefficient, or of the admittance presented to the microwave power source and its output transmission line, by the load and matching device.

Complete results of the simulation will be presented in the next section.

6-2 System Simulation Results

The main reasons for simulating the system on a hybrid computer were, first, to establish its practical stability, and second, to gain an impression of its response characteristics.

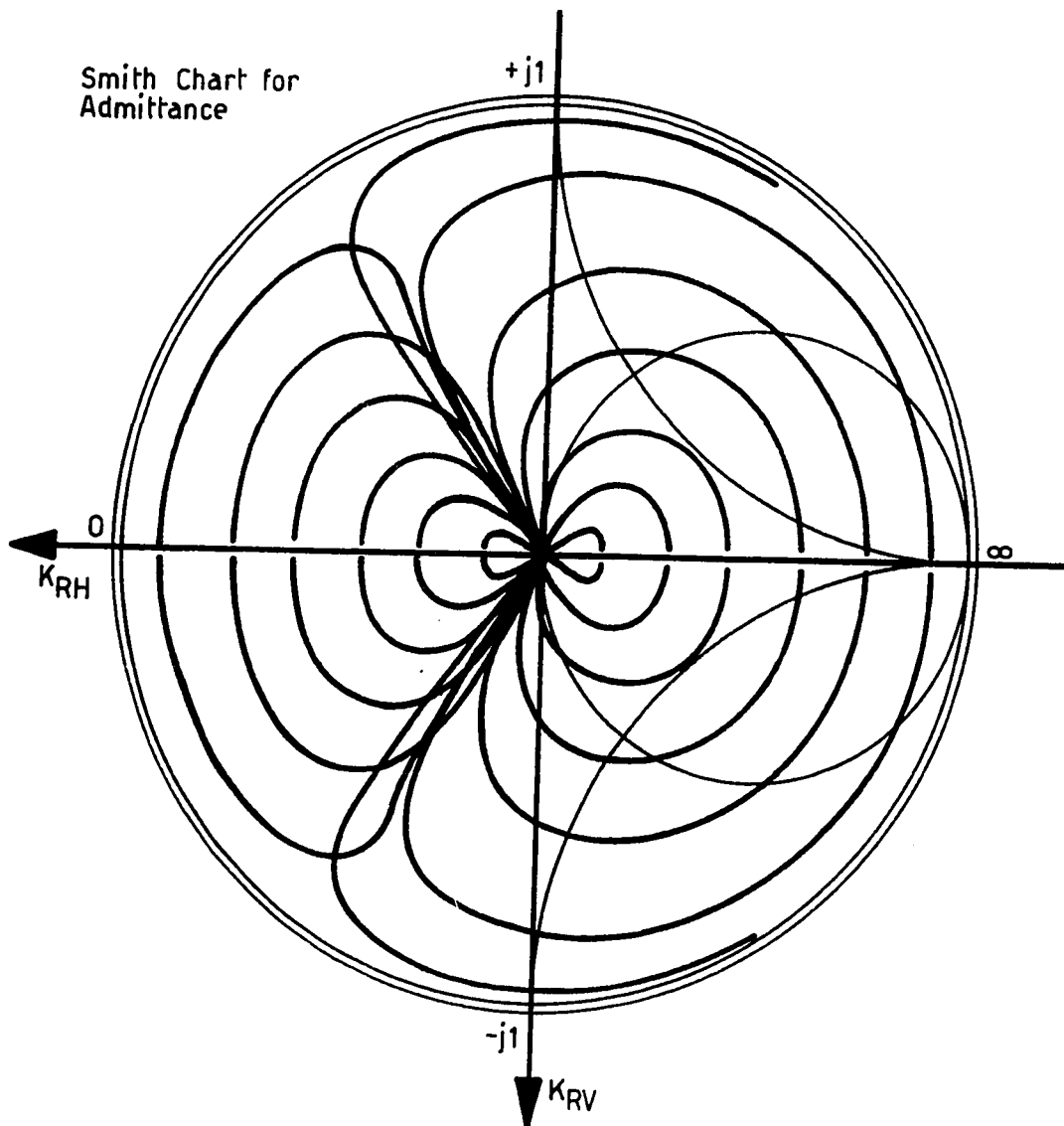
Extensive testing at many loading and operating conditions has shown that the system is indeed stable as predicted in the previous chapter. Moreover, the system was tested for many initial conditions on both position and velocity of the dielectric slabs, such as might occur during operation of an actual system, as well as for forced conditions with sinusoidal inputs. Even in these circumstances, which were not covered in the proof of stability given in the previous chapter, the system reacted in an entirely stable manner.

For the sake of both simplicity and ease of comparison, graphical results are presented only for zero initial conditions on the dielectric phaseshifter slab positions and velocities, and for suddenly applied unmatched loads. These results will now be discussed.

Figures 41, 42, and 43 show sets of Smith chart matching trajectories for, respectively, an overdamped, a critically damped, and an underdamped system. The trajectories shown in these figures all consist of three identifiable sections. These are an initial load rotation, a section of direct approach to the origin, and a trailing portion of rather slow approach to the origin (the latter is not really visible, but can be observed in analogue computer read-out of the rate of change of the dielectric slab positions). During the direct approach of the origin, all the trajectories in one set appear to move towards the origin along one preferred direction line.

The definitions of closed-loop overdamped, critically damped, and underdamped response, as used above, are purely qualitative and refer mainly to the character of the matching trajectories at the preferred direction line, and partly refer to the average (for many load admittances) behaviour of x and y , as functions of time, as matching proceeds. The underdamped trajectories tend to overshoot and oscillate about the preferred direction line, while also exhibiting a barely discernible spiralling into the origin. The critically damped trajectories, on the other hand, move into the origin almost directly with little overlap of the curves. Lastly, overdamped trajectories merely approach the preferred direction line asymptotically, with almost no overlap of the curves.

Considering the two loops individually, the degree of damping of the series loop affects mostly the amounts of overshoot of, and oscillation about the preferred direction line. In contrast to this, the speed of approach to the origin and the spiralling in effect depend mainly on the degree of damping of the stub loop. Nevertheless, there

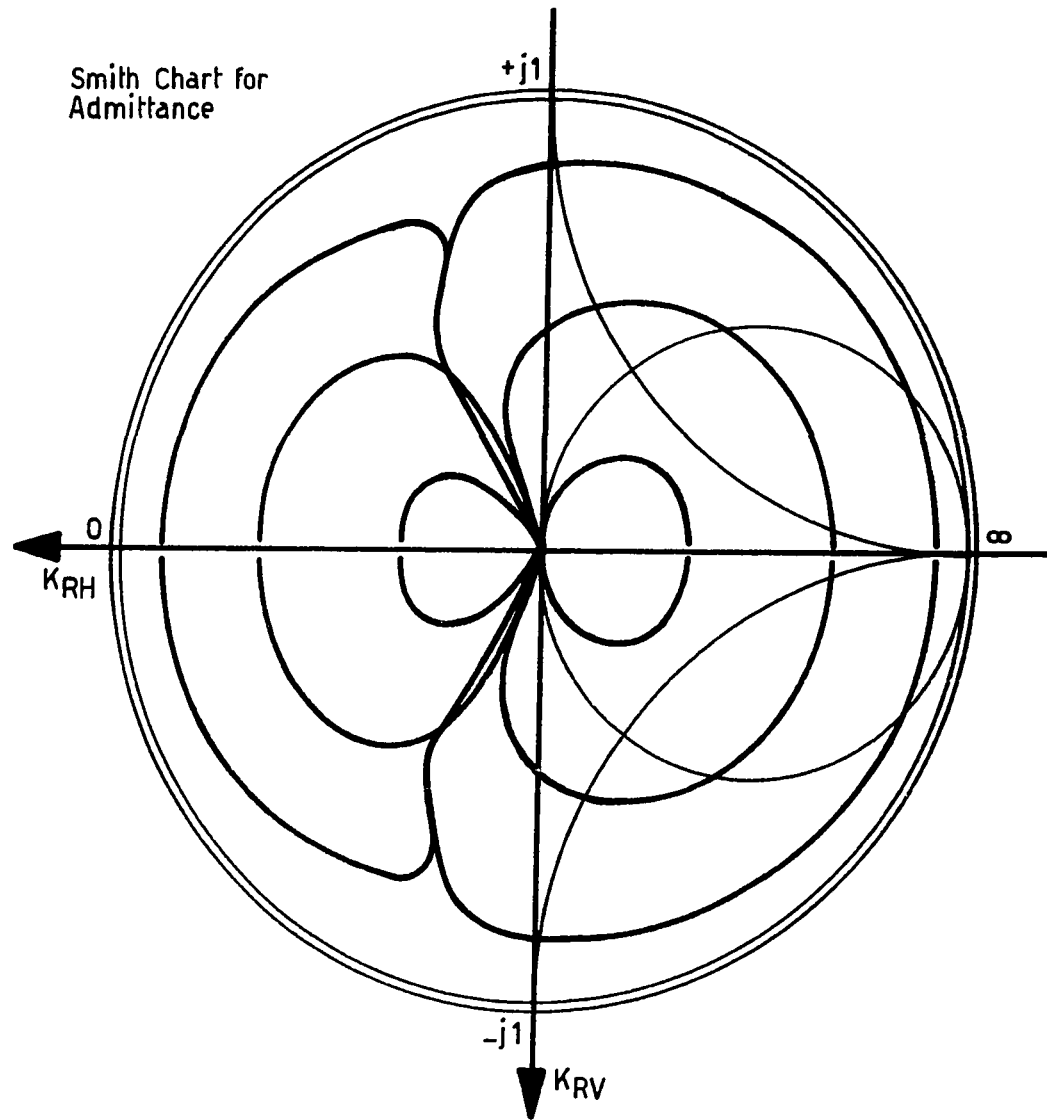


CONDITIONS:

Error Signals: K_{RH} and K_{RV} , normal system.

Operating Conditions: $C_1=C_2=0.5$ into X1 Amp. Gain
 $A_1=A_2=1.0$ into X20 Amp. Gain,
 (see figure 40).

FIGURE 41. COMPUTER SIMULATION:
 MATCHING TRAJECTORIES FOR AN "OVERDAMPED" SYSTEM

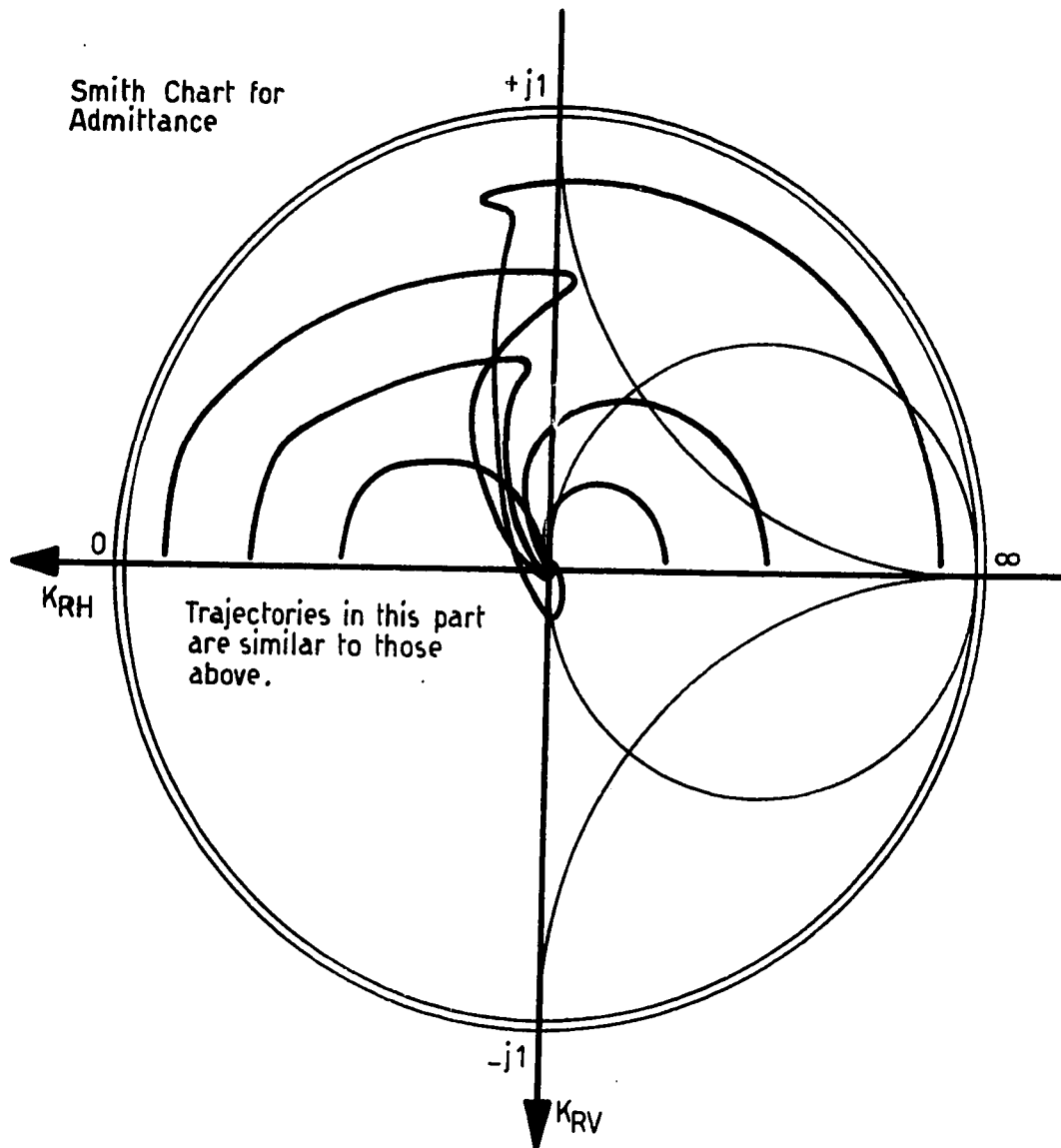


CONDITIONS:

Error Signals: K_{RH} and K_{RV} , normal system.

Operating Conditions: $C_1=C_2=0.05$ into X1 Amp. Gain
 $A_1=A_2=0.4$ into X1 Amp. Gain,
 (see figure 40).

FIGURE 42. COMPUTER SIMULATION:
 MATCHING TRAJECTORIES FOR A "CRITICALLY DAMPED" SYSTEM



CONDITIONS:

Error Signals: K_{RH} and K_{RV} , normal system.

Operating Conditions: $C_1=C_2=0.05$ into X1 Amp. Gain
 $A_1=A_2=0.2$ into X1 Amp. Gain,
 (see figure 40).

FIGURE 43. COMPUTER SIMULATION:
 MATCHING TRAJECTORIES FOR AN "UNDERDAMPED" SYSTEM

is enough interaction between the loops that, to change the overall character of the system response from underdamped to overdamped, the amount of damping of both loops has to be changed appropriately.

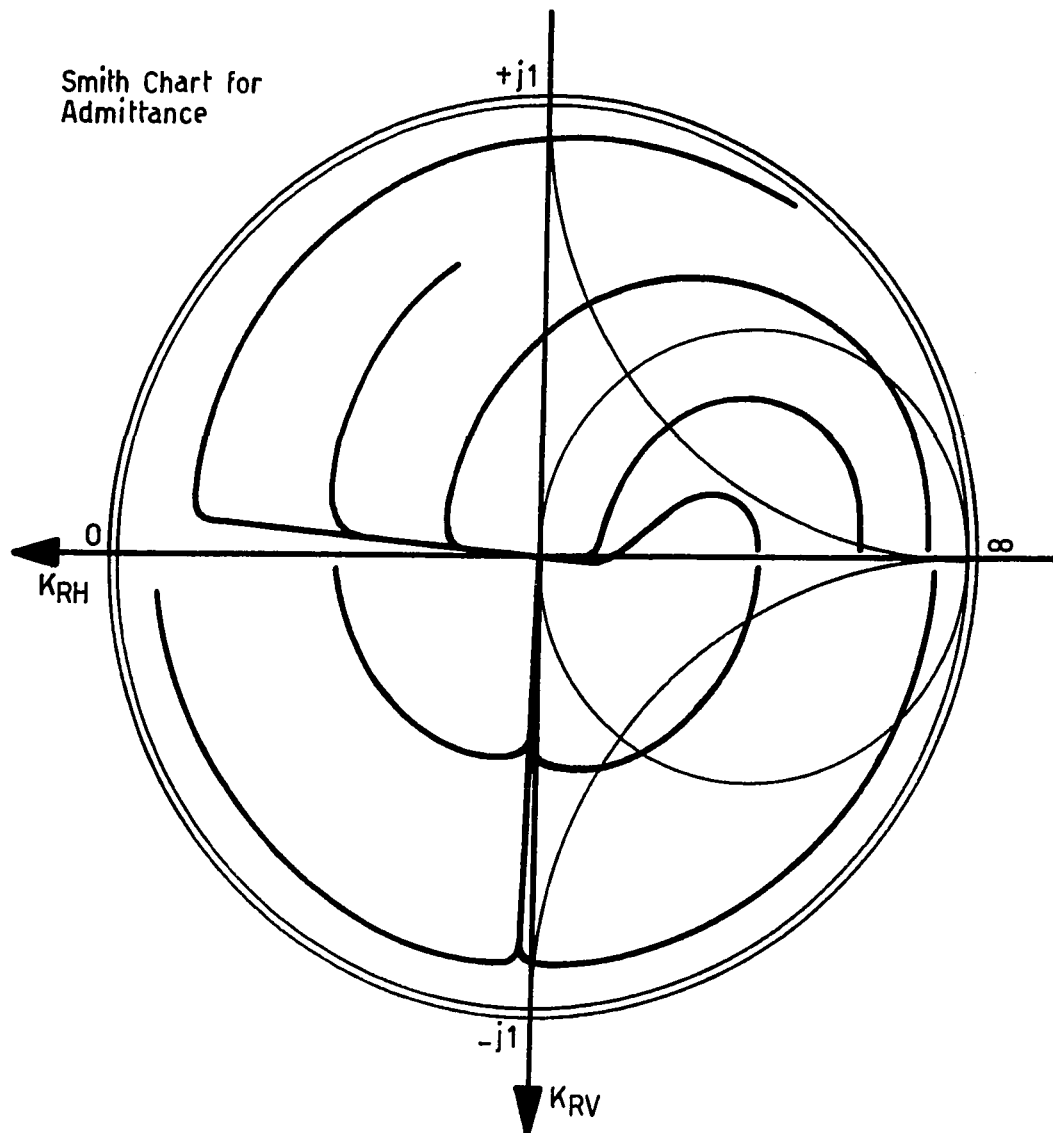
There was but one set of circumstances which did cause instability of the simulated system. If the loop gains were increased excessively, and simultaneously the damping loop gains were drastically reduced, such that the effective response speed was high compared to the number of samples taken per second by the D/A and A/D converters, the system would break into oscillation. The problem appeared to be the additional phaseshift introduced as a result of the sampled-data nature of the simulation. This was further tested by the introduction of an additional lagging time constant into each of the two loops. At low values of gain, and with sufficient damping, the system remained stable. As the gain was increased and/or the damping decreased, the system became unstable, but the instability conditions were affected by the value of Y_L . Loads representing large reflection coefficients caused instability for lower values of gain (or greater damping), than loads with correspondingly lower reflection coefficient magnitudes. In fact, for "borderline" stability settings of the gain and damping, there appeared to be a definite curve that divided the Smith chart into stable and unstable reflection coefficient loads. This last test has some bearing on the stability of a practical system. Hence, it will be discussed again in a later section.

Further discussion of the preferred direction of approach to the origin is required. For the operating conditions shown in figures 41, 42, and 43 (all appropriate operating conditions are marked directly on the figures and refer to figure 40), the angle of the preferred direc-

tion line is about 120° . The direction of this line is a function of the relative response speed of one loop compared to the other. Since the response speed of each loop depends directly on the linear part gain and damping of the loop in question, these same parameters also affect the angle of the preferred direction line. If the stub loop is much faster than the series loop, the preferred direction of approach is almost coincident with the K_{RH} -axis, while for a faster series loop, the approach is nearly along the K_{RV} -axis. These cases are clearly illustrated in figure 44, where the upper half of the Smith chart corresponds to a fast stub loop and the lower half to a fast series loop.

The trajectories can never coincide completely with either the K_{RH} -axis, or the K_{RV} -axis. To show this, let the matching trajectory be the K_{RH} -axis. Motion along the K_{RH} -axis requires a rate of change of the series loop slab position. But, to change the trajectory which, for the series loop alone is normally circular, into a straight line, the stub loop must also be active, and there must be a rate of change associated with its output (non-linear coupling between loops). Since each loop contains only one free integrator, there must be an error signal to maintain the stub loop output rate of change. Hence, K_{RV} cannot be zero while K_{RH} is changing.

The rate of decrease of the reflection coefficient, as matching proceeds, is of great practical importance. In fact, the rate of reflection coefficient decrease is the true measure of system response speed. This matter was, therefore, investigated by having the digital computer calculate the magnitude of the reflection coefficient on a continually sampled basis from:



CONDITIONS FOR UPPER HALF OF PLOT:

Error Signals: K_{RH} and K_{RV} . Stub control loop faster.

Operating Conditions: $C_1=0.05$, $C_2=0.5$ both into X1 Amp. Gain

$A_1=A_2=1.0$ into X20 Amp. Gain (see fig. 40).

CONDITIONS FOR LOWER HALF OF PLOT:

Error Signals: K_{RH} and K_{RV} . Series control loop faster.

Operating Conditions: $C_1=0.5$, $C_2=0.05$ both into X1 Amp. Gain

$A_1=A_2=1.0$ into X20 Amp. Gain (see fig. 40).

FIGURE 44. MATCHING TRAJECTORIES FOR A FASTER STUB LOOP (UPPER) OR A FASTER SERIES LOOP (LOWER)

$$|K_R| = \sqrt{K_{RH}^2 + K_{RV}^2}$$

6.1

Plotting of the reflection coefficient magnitude as a function of time for several typical loads results in a family of curves as shown in figure 45, all of which were obtained for the same over-damped system with nearly equal loop speeds. Each of the curves shows the three typical sections of the matching trajectories discussed before. Especially the slow trailing portion is clearly visible.

Varying of the loop gains or damping did not materially alter the nature of the curves representing the reflection coefficient as a function of time. Amount of damping affected mostly the relative length of three typical periods of each curve one with respect to another, while changing of the loop gains either shortened or elongated the curves. It became apparent that, under all possible operating conditions, large reflection coefficient loads are matched at least twice as quickly as small reflection coefficient loads.

Finally, the integral of the squared reflection coefficient magnitude is of some interest. It is given by:

$$Q = \int_0^t |K_R|^2 d\tau$$

6.2

where t is time and τ is a dummy variable. For a given fixed value of incident microwave power, the quantity $Q|_{t \rightarrow \infty}$ is directly proportional to the total microwave energy reflected from the load and matching device combined during a complete matching operation. Thus, it serves as

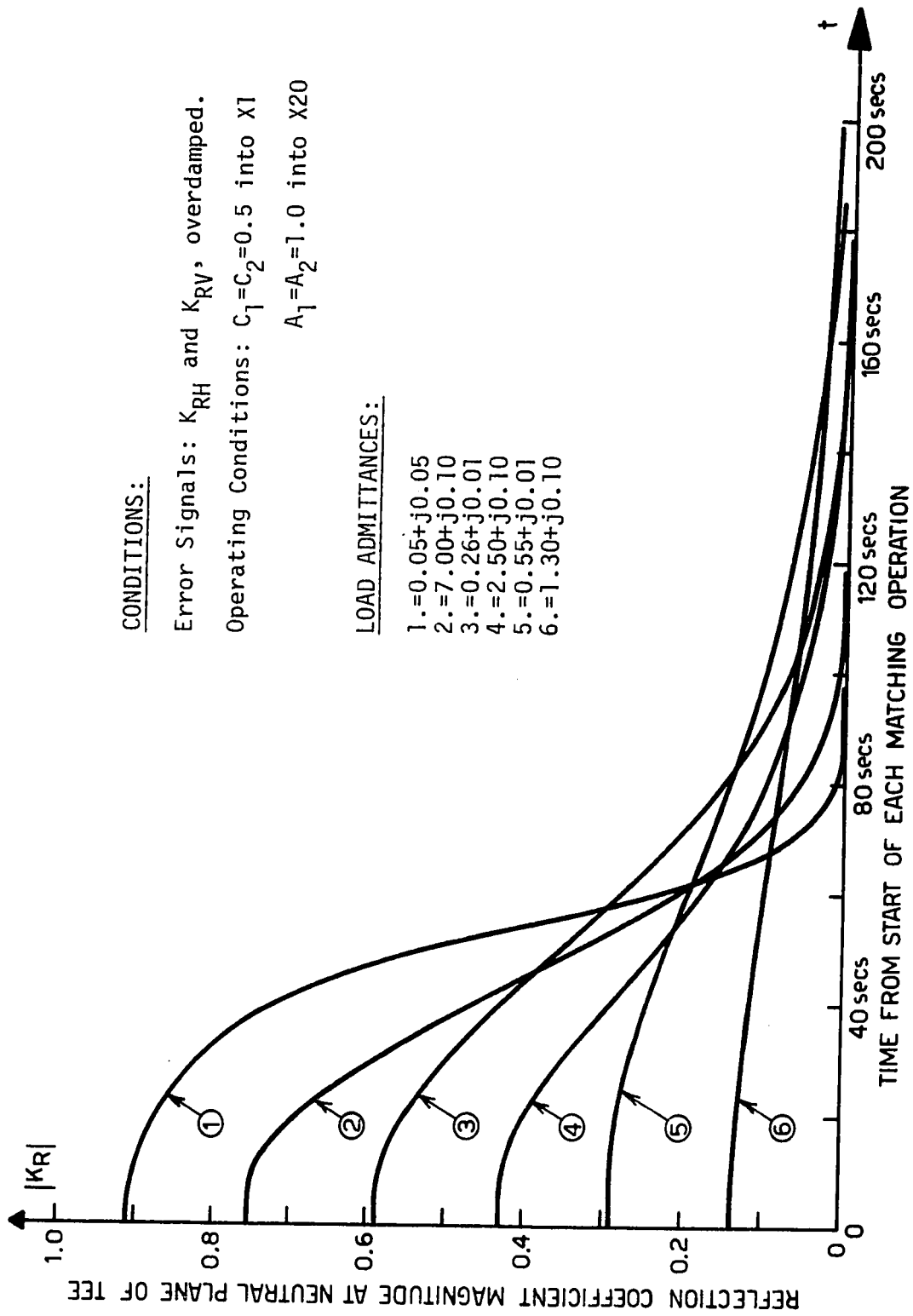
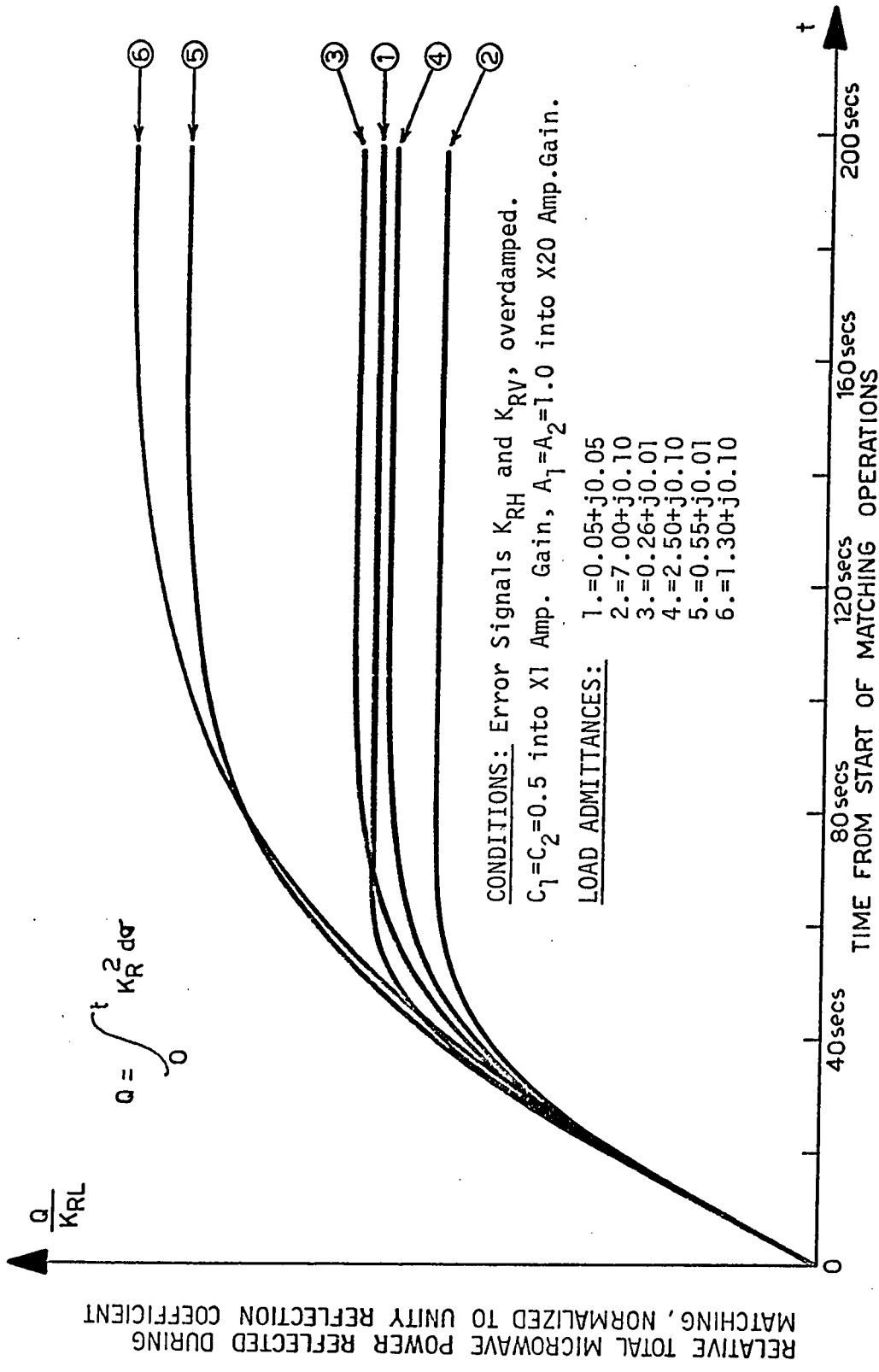


FIGURE 45. CHANGES IN THE REFLECTION COEFFICIENT MAGNITUDE AS MATCHING PROCEEDS



RELATIVE TOTAL MICROWAVE POWER REFLECTED DURING MATCHING, NORMALIZED TO UNITY REFLECTION COEFFICIENT

FIGURE 46. TOTAL MICROWAVE ENERGY REFLECTED FROM START OF MATCHING PROCEDURE NORMALIZED TO UNITY REFLECTION COEFFICIENT

a figure of merit for the automatic matching device, which, ideally, should be minimized for all possible loading conditions.

In figure 46, Q is plotted as a function of time for the same loads and operating conditions as those of figure 45. But, rather than plotting the actual values of Q , the curves were all normalized to an effective load reflection coefficient magnitude of one. This was done to make comparison of the curves easier. For example, if the system were to consist of a single linear loop, the normalized integral of the "squared error" would be the same single curve for all step input disturbances.

The shape of the normalized Q -curves did not change appreciably as operating conditions were varied, but the value of $Q|_{t \rightarrow \infty}$, for each particular load admittance, did depend on the operating conditions. Notably, for a particular value of Y_L , the value of $Q|_{t \rightarrow \infty}$ appeared to be minimum for a slightly underdamped system with approximately equal response speeds for the two loops. Since the system response characteristics are very much dependent upon the value of Y_L , no one set of operating conditions could assure that $Q|_{t \rightarrow \infty}$ was minimum for all loads. This is clearly illustrated in figure 46, where the normalized values of $Q|_{t \rightarrow \infty}$ for small reflection coefficient loads are, in fact, larger by about 50% than the values of the same function for large reflection coefficient loads.

6-3 System Improvement and Optimization

Overall system response speed is a major consideration in designing a satisfactory automatic load matching device, particularly since

the response speed is related to the total microwave energy reflected from the load and matching device during a matching operation. It is clearly desirable to minimize the total energy reflected from an initially mismatched load. Several methods, that can aid in improving the response of the system, will be discussed.

Consider first the linearized system of section 5-7. It was shown there that, very close to equilibrium, the system behaves like two cascaded second order loops, while away from equilibrium, the system characteristics are those of a multi-loop fourth order arrangement. In either case, the characteristic equations (5.44 and 5.45) may be factored into two second order parts and appropriate loop error signals may be defined which are generally not equal to, but which are related to K_{RH} and K_{RV} . An integrated squared error signal similar to Q of expression 6.2 may, thus, be derived, which is equal to the sum of the integrated squared errors of the second order parts. It was shown elsewhere⁽²⁴⁾ that, for a type one, second order, linear system the squared error integrated from zero time to infinity is minimal if the closed-loop damping ratio, ζ , equals 0.5 (underdamped system with about 16% overshoot). Unfortunately, the actual system is so severely non-linear in x and y (the slab positions), and the non-linear characteristics are so greatly dependent upon load admittance Y_L , that no specific damping loop gains to optimize the system can be dictated on the basis of this theory. The above may be used as a guide only to adjust the system parameters in a general sense, "averaged" over all possible loading conditions. This was shown experimentally in section 6-2.

The proposed control system for automatic matching of microwave loads is rather unusual in that no discrete feedback elements or sum-

ming junction are identifiable. The system output, consisting of slab positions x and y , are not measured and compared, as such, to an input signal to produce a system error signal. Rather, the error signal, comprising the real and imaginary parts of the reflection coefficient of the load and matching device combined, is measured and used directly to activate the control elements. This causes peculiar problems if optimization of the system is attempted.

The reflection coefficient as measured is a non-linear function of slab positions x and y and of load admittance Y_L . It contains no explicit information regarding the actual values of x , y , and Y_L at any instant in time. A change in load admittance Y_L is indistinguishable from a change in slab positions x and/or y , in the manner in which the reflection coefficient is affected. Time optimization of the system, which will assure that the microwave energy reflected from an initially mismatched load is minimal, implies that to each set of initial values of x and y and to each value of Y_L for these initial conditions, there corresponds a particular optimal Smith chart matching trajectory. However, these trajectories can be neither computed, nor obtained in an actual system based on information extracted from the reflection coefficient alone.

To force the system to follow certain specific matching trajectories, it is necessary to measure or compute both the state of the system and the complete loading conditions. The required additional information may be obtained either by direct measurement of x , y , and Y_L , or by computation from measurements of the load reflection coefficient, the reflection coefficient immediately after the series phaseshifter, and the admittance added by the stub.

An example of a method which will lead to a time optimal system design is outlined below.

By appropriate, but rather complex calculations performed upon measured data of the load reflection coefficient, the slab positions, and the reflection coefficient at the neutral plane of the H-plane tee, two error signals may be derived, which are directly proportional to the remaining angles through which the phaseshifters have to be adjusted to obtain a matched load condition. In other words, two linear and independent feedback signals may be computed as functions of time, which may be used to adjust the phaseshifters. The system non-linearity has thus been "inverted" and the system has been transformed into a linear system with two second order loops. Once the system has been linearized in this manner, standard, well documented methods⁽²⁴⁾ of time optimizing second order systems may be applied to each loop. This will lead to "bang-bang" controllers for the loops.

The two linear loops will not be completely independent, but will have a common output, namely the reflection coefficient at the neutral plane of the H-plane tee. Since, for time optimal control, both control loops must simultaneously reach equilibrium, and since the Smith chart origin must be approached from a specific direction for each load to be matched, the maximum adjustment rates of the two loops must be coordinated, and altered in relative ratio, according to the actual value of the load to be matched. This implies that the system must contain memory.

The time optimal system, the general form of which was suggested above, has very limited practical value. The complexity of the necessary calculations and the need for memory forces the use of a digital

computer to implement such a system. Moreover, the measurement procedures now required are more complicated and difficult to perform. The attractive advantages of simplicity and reasonable system cost have, thereby, been lost. In view of the above, time optimization of this type was pursued no further.

Provided the coupling between the loops is neglected, another form of system improvement may be attempted, based on the limited information contained in the measured reflection coefficient at the neutral plane of the H-plane tee.

Assume that the angle of the measured reflection coefficient is indicative of the remaining angular adjustment of the series phaseshifter necessary to bring the conductance to unity. Assume further that series phaseshifter adjustment causes pure rotation about the Smith chart origin. The coordinates of the intersection point (only the intersection point in the upper half of the Smith chart is considered) of the unity conductance circle and a particular constant reflection coefficient magnitude circle are given by:

$$K'_{RH} = K_R^2 \quad 6.3$$

$$K'_{RV} = K_R \sqrt{1 - K_R^2} \quad 6.4$$

where the primes indicate that the values of K'_{RH} and K'_{RV} refer to the intersection point coordinates. From figure 47, the desired remaining angle of adjustment is $\gamma = \beta - \alpha$. One has from the diagram:

$$\tan \gamma = \tan(\beta - \alpha) = \frac{K_R K'_{RV} - K'_{RH} \sqrt{1 - K_R^2}}{K'_{RV} \sqrt{1 - K_R^2} + K'_{RH} K_R} \quad 6.5$$

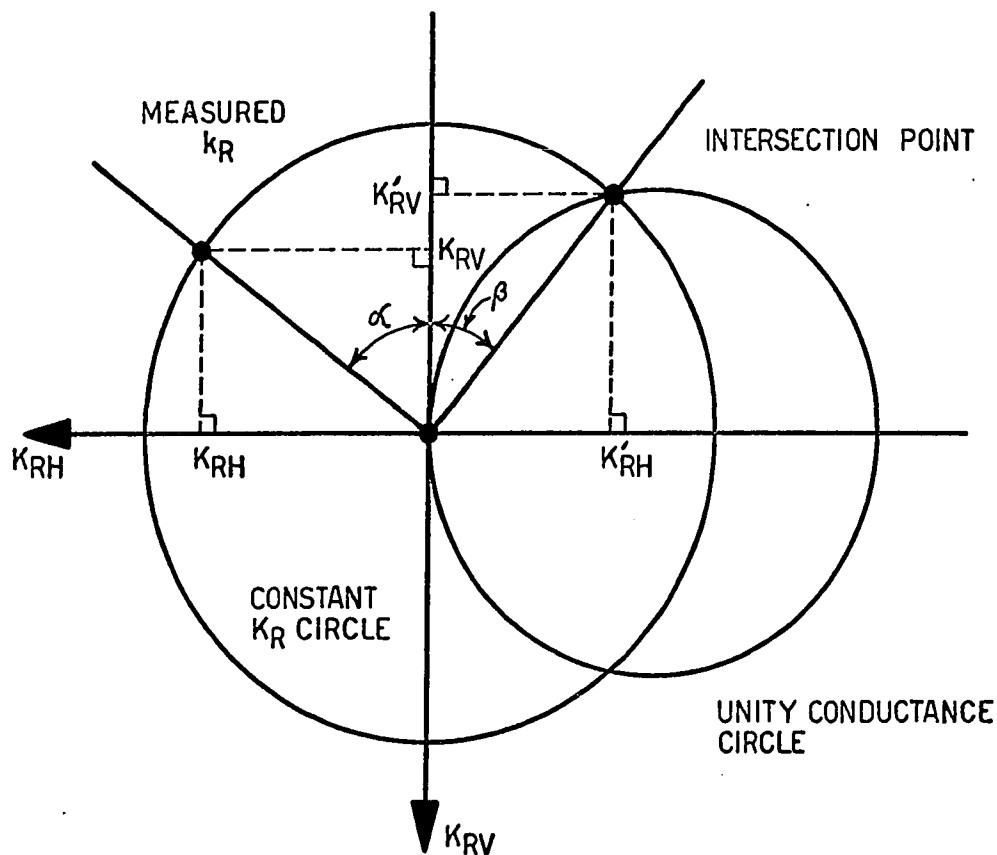


FIGURE 47. SMITH CHART GEOMETRY FOR LINEARIZED ERROR SIGNALS

Since $K_R^2 = K_{RH}^2 + K_{RV}^2$, the error signal for the series phasemitter becomes:

$$e_L = \gamma = \tan^{-1} \frac{K_{RV} \sqrt{K_{RH}^2 + K_{RV}^2} - K_{RH} \sqrt{1 - K_{RH}^2 - K_{RV}^2}}{K_{RV} \sqrt{1 - K_{RH}^2 - K_{RV}^2} + K_{RH} \sqrt{K_{RH}^2 + K_{RV}^2}} \quad 6.6$$

At the same time, from the Smith chart geometry, the remaining reactive admittance that has to be added by means of the stub phasemitter is equal to $(2 \tan \beta)$. The error signal for the stub phasemitter is thus:

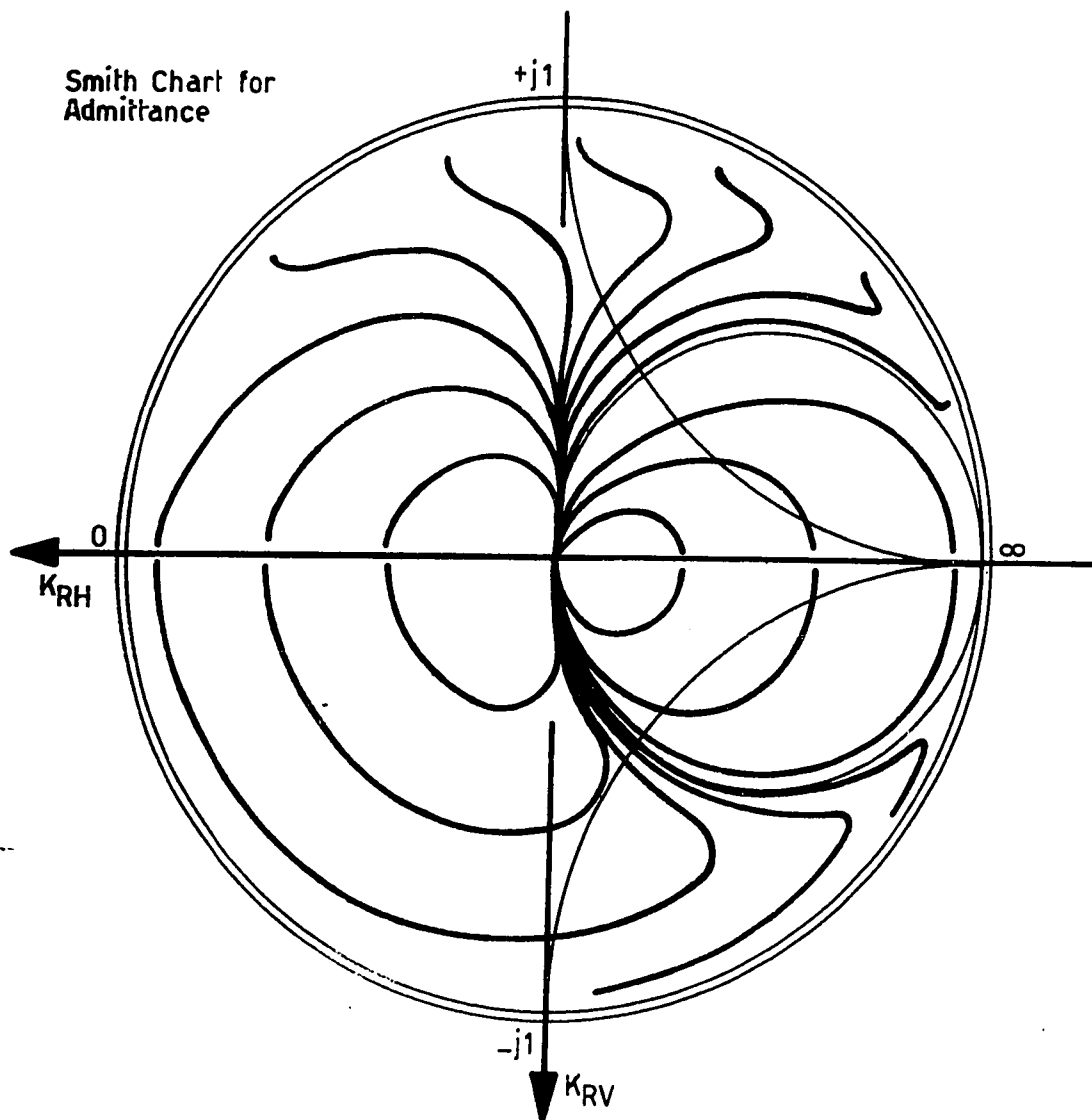
$$e_S = \tan^{-1}(2 \tan \beta) = \tan^{-1} \frac{2\sqrt{K_{RH}^2 + K_{RV}^2}}{\sqrt{1 - K_{RH}^2 - K_{RV}^2}} \quad 6.7$$

In a hybrid computer simulation of the system, the above linearized error signals are easily computed. Rather than feeding back signals K_{RH} and K_{RV} in a scheme as shown in figure 40, signals e_L and e_S may be used for activation of the phaseshifters. If this is done, matching trajectories for various load admittances may again be obtained. Results for two representative sets of static loop gains and damping loop gains are shown in figure 48, where the actual gains used have been marked directly on the diagram.

The matching trajectories shown in the upper half of figure 48 are especially interesting since they represent a set of operating conditions for which each and every load (starting with zero initial conditions on the slab positions) is matched in equal time. If the same relative loop speeds are maintained, it should, therefore, be possible to minimize Q (defined before) with respect to the damping loop gains. This would represent a form of optimization. These same curves also give some idea of the shape of true time optimal trajectories.

The above method of improving the system, though simpler than full time optimization, is still not suitable for analogue implementation in an actual system. No memory is required, but the non-linear functions of K_{RH} and K_{RV} required to calculate e_L and e_S are too complex for anything but digital equipment.

There remains one method of improving the system response (especially for small reflection coefficient loads), which is simple enough to be suitable for direct implementation in an actual system by analogue



CONDITIONS FOR UPPER HALF OF PLOT: Max. Error Signals scaled to 10 V

Error Signals: e_L and e_S replace K_{RH} and K_{RV} , see text for details.

Operating Conditions: $C_1=C_2=0.2$ into X1 Amp. Gain

$A_1=A_2=1.0$ into X10 Amp. Gain.

CONDITIONS FOR LOWER HALF OF PLOT: Max. Error Signals scaled to 10 V

Operating Conditions with $C_1=0.2$ into X1 Gain, $C_2=0.1$ into X1

Error Signals as above: $A_1=A_2=1.0$ into X10 Amp. Gain.

FIGURE 48. MATCHING TRAJECTORIES FOR
A PARTIALLY LINEARIZED SYSTEM

methods alone. This approach will now be described.

As mentioned in several previous sections and as shown graphically in figure 26, the diameters of the circular Smith chart trajectories, obtained for various load admittances Y_L as x is varied (but for constant values of slab position y), are directly related to the magnitude of the reflection coefficient corresponding to the load to be matched. As the load reflection coefficient decreases in magnitude, so does the circle diameter. The average change in K_{RH} , which is the projection of the circular trajectories upon the K_{RH} -axis as x is varied is, thus, also a function of the load reflection coefficient. Therefore, the effective average series phaseshifter loop gain varies with the reflection coefficient of the load to be matched.

Similarly, by examination of figure 31, and by consideration of the theory presented earlier, one finds that the slope of the K_{RV} versus y curves close to equilibrium is a direct function of the reflection coefficient magnitude of the initially unmatched load. The effective stub phaseshifter loop gain, as the system approaches equilibrium is, thus, dependent upon the load reflection coefficient magnitude. This effective loop gain increases with increasing reflection coefficient magnitude.

The above two load dependent effects are responsible for the slower system response for small reflection coefficient loads than for loads of large reflection coefficient. Clearly, to compensate for this effect, the loop gains must be made into functions of the load reflection coefficient magnitude such as to counteract the effects inherent in the primary system non-linearities. Fully in accordance with earlier remarks regarding the necessity of obtaining additional information not

contained in the error signals as measured, the load reflection coefficient must be measured.

Let a directional coupler be used to sample some of the microwave power reflected from the actual load. If a matched crystal detector terminates the coupler sidearm, the output signal from the crystal will be proportional to the reflected power. And, at a constant incident microwave power level, the same crystal output signal will be proportional to the squared load reflection coefficient magnitude K_{RL}^2 .

For the sake of circuit simplicity, rather than extracting the square root of the K_{RL}^2 -signal, the crystal output may be amplified and used directly to modify the error signals K_{RH} and K_{RV} according to the following formulae:

$$(K_{RH})_m = (p - K_{RL}^2) K_{RH} \quad 6.8$$

$$(K_{RV})_m = (p - K_{RL}^2) K_{RV} \quad 6.9$$

where the subscript m indicates the modified values of the error signals and where p is a positive constant. Since the maximum and minimum values of K_{RL}^2 are 0 and 1 respectively, the choice $p=2$ will cause a maximum system loop gain change factor of two; this gain variation corresponds directly to the experimentally determined difference in system response speed to low and high reflection coefficient loads (see section 6-2). In an actual system both p and K_{RL}^2 must, of course, be scaled to the input and output voltage capabilities of the electronic components used. Details of this will be given in the next chapter.

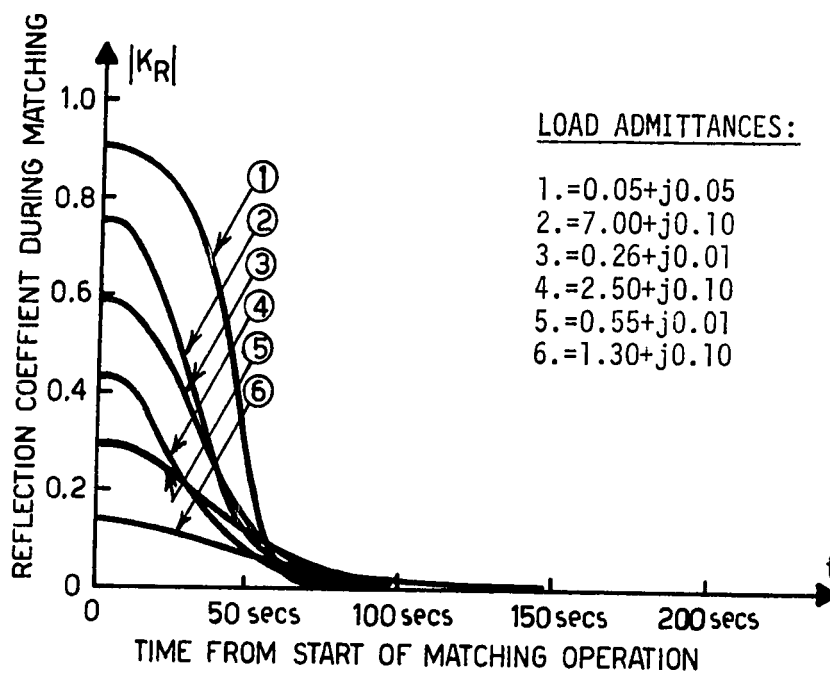
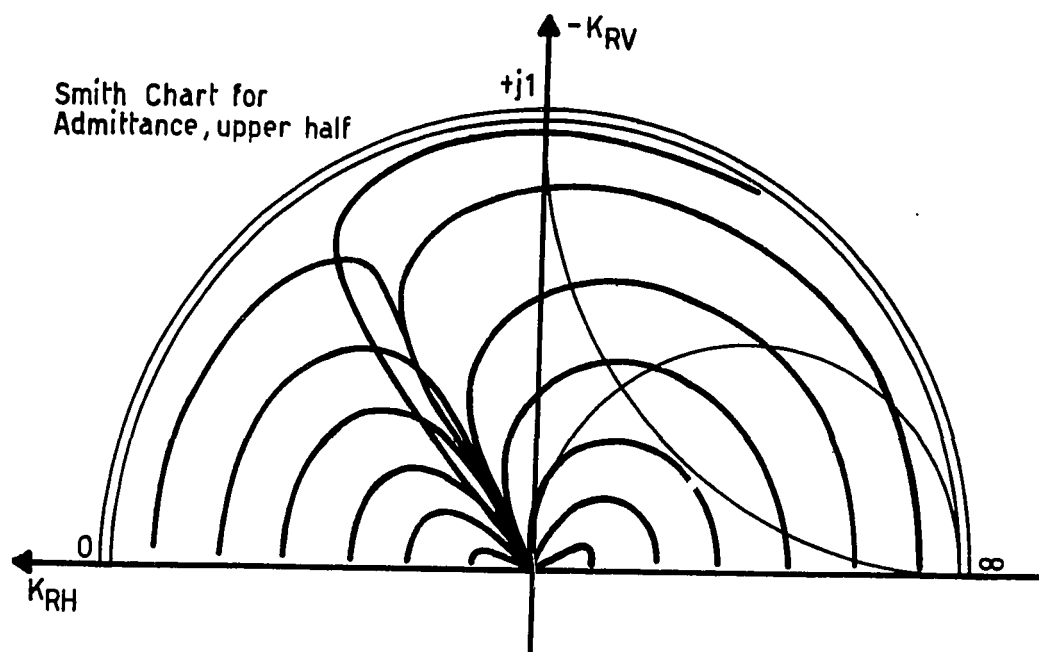
Although this has not been emphasized anywhere in the previous

work, the error signals K_{RH} and K_{RV} , as measured, are not equal to, but are merely proportional to the real and imaginary parts of the reflection coefficient. The magnitude of the proportionality factor is linearly related to the microwave power level at which the system is operated. The effective system loop gains, therefore, also depend on the microwave power level, resulting in changes in response speed as a function of microwave power level. However, since the K_{RL}^2 -signal, as measured, is affected in exactly the same manner, the response speed compensation suggested by formulae 6.8 and 6.9 will partly cancel the power level effects.

In chapter five the idealized system was proven stable for any set of series and stub loop gains, provided the damping was never zero. The adaptive type of gain modulation given by expressions 6.8 and 6.9 is a function of the load admittance only, and does not depend on slab positions x and y . Furthermore, each of the two loops is affected in exactly the same manner. Anderson's conditions on the non-linearities will, thus, not be upset. Therefore, the system will remain stable; only the response speed will be altered.

Once again, the system modifications suggested above were easily incorporated into the digital simulation of the non-linear part of the system. A representative set of matching trajectories is shown in figure 49. For a specific set of operating conditions such that the system was approximately critically damped, the trajectories for small reflection coefficient loads approached the origin noticeably more rapidly.

The suggested response speed compensation is, by no means, ideal or exact, nor is the compensation for microwave power level perfect.



CONDITIONS:

Error Signals: $(K_{RH})_m$ and $(K_{RV})_m$ replace K_{RH} and K_{RV} with $p=2$, see text for details.

Operating Conditions: $C_1=C_2=0.5$ into X1 Amp. Gain

$A_1=A_2=1.0$ into X20 Amp. Gain, (see fig. 40).

FIGURE 49. MATCHING TRAJECTORIES AND REFLECTION COEFFICIENT CHANGES FOR AN ADAPTIVE SYSTEM

Nevertheless, the improvement in system performance is considerable, and the required computations are easily performed by means of operational amplifiers and integrated circuit analogue multipliers.

6-4 Non-Ideal Properties of the System and its Components

In all of the previous work only an ideal system was considered, but the open-loop frequency response of the linear part of the system is, actually, more complex than the double integrator with rate feedback model used. Some of the factors that affect the frequency response and which have, so far, been ignored are:

1. The time constants associated with the inductance and resistance of the phaseshifter drive coils and the output impedances of the power amplifiers.
2. The time constants of the rate feedback coil inductances and resistances, and the input impedance of the rate signal amplifiers.
3. The closed-loop frequency response of the various, individual, operational amplifiers that must be used in the system electronics.
4. The dynamic response of the microwave portion of the system, i.e. the bandwidth and envelope roll-off characteristics of any resonant effects that may occur in the microwave system.

None of the system imperfections mentioned above are likely to cause any severe difficulties in obtaining system stability. If the drive coil and rate feedback coil inductances prove bothersome, the

time constants associated with them may always be decreased by increasing of the power amplifier output impedances and rate signal amplifier input impedances. The closed-loop bandwidth of integrated circuit operational amplifiers is normally at least 100 kHz and is much wider than the expected bandwidth of the automatic matching device, which contains mechanical components. Lastly, the bandwidth of even narrow band microwave resonant effects can be expected to be much greater than the bandwidth of a mechanical adjustment system.

By far the most important imperfection of the system components is the occurrence of mechanical resonance effects in the moving structural parts of the phaseshifters. To test the phaseshifter performance, a positional servo was arranged using the actual phaseshifters. The circuit schematic of the test set-up is shown in figure 50. To measure the position of the dielectric slab, the image of a narrow slit was projected through the phaseshifter waveguide section onto a rectangular solar cell type of photocell. At various positions of the slab a portion of the slit image was shadowed off by the slab. An output signal proportional to slab position could, in this manner, be obtained from the photocell.

The test circuit was first subjected to square wave input signals. Although the damping could be adjusted to produce any type of response desired from underdamped to overdamped, the gain settings had a very pronounced effect on the appearance of 650 Hz ringing superimposed on the normal output. At low gains the ringing was not evident, but as the gain was increased ringing would first occur and as the gain was increased still further, the system would break into oscillation. The amount of damping did not influence this effect at all. An oscilloscope

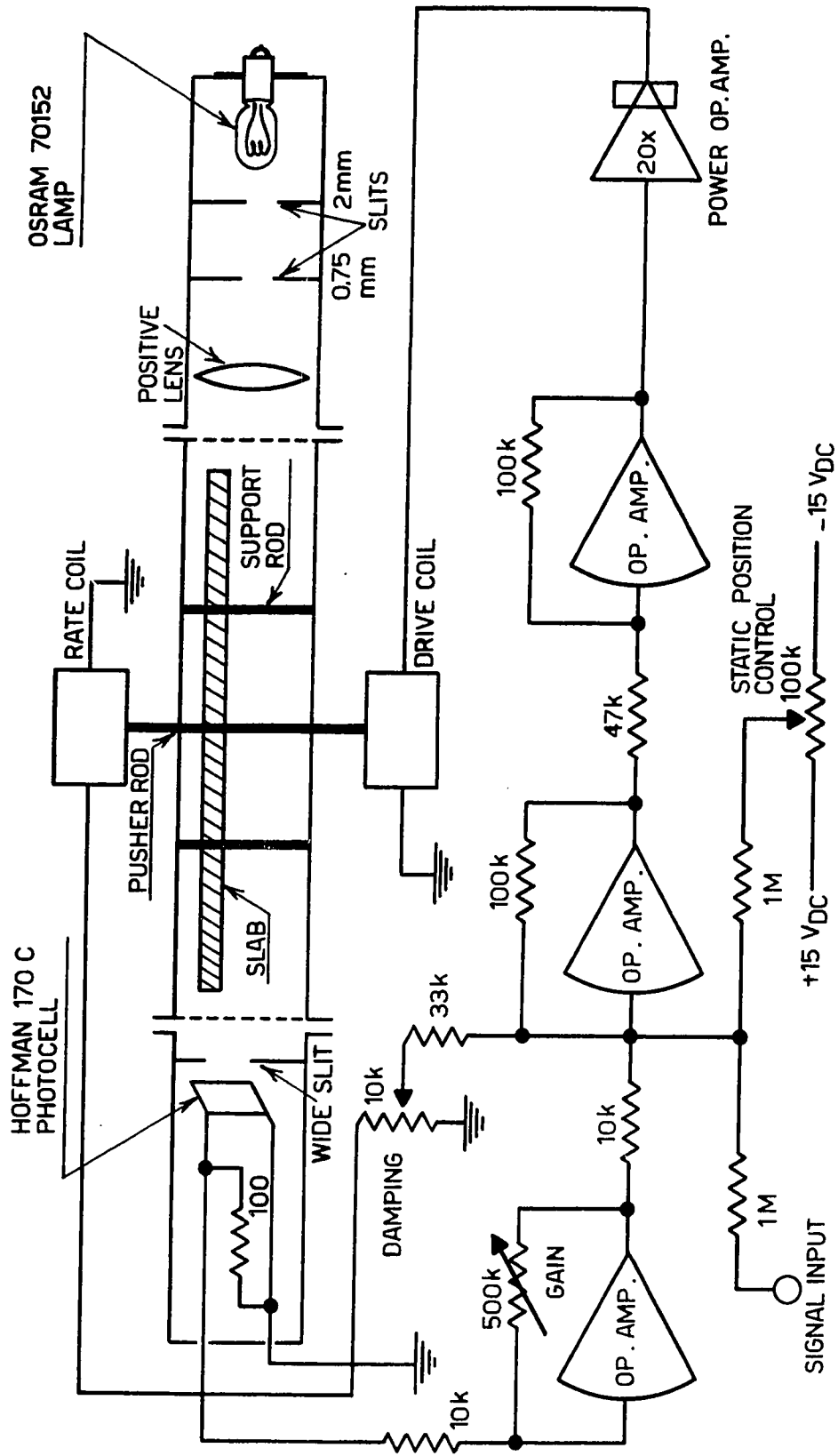


FIGURE 50. SCHEMATIC DIAGRAM OF PHASESHIFTER TEST ARRANGEMENT

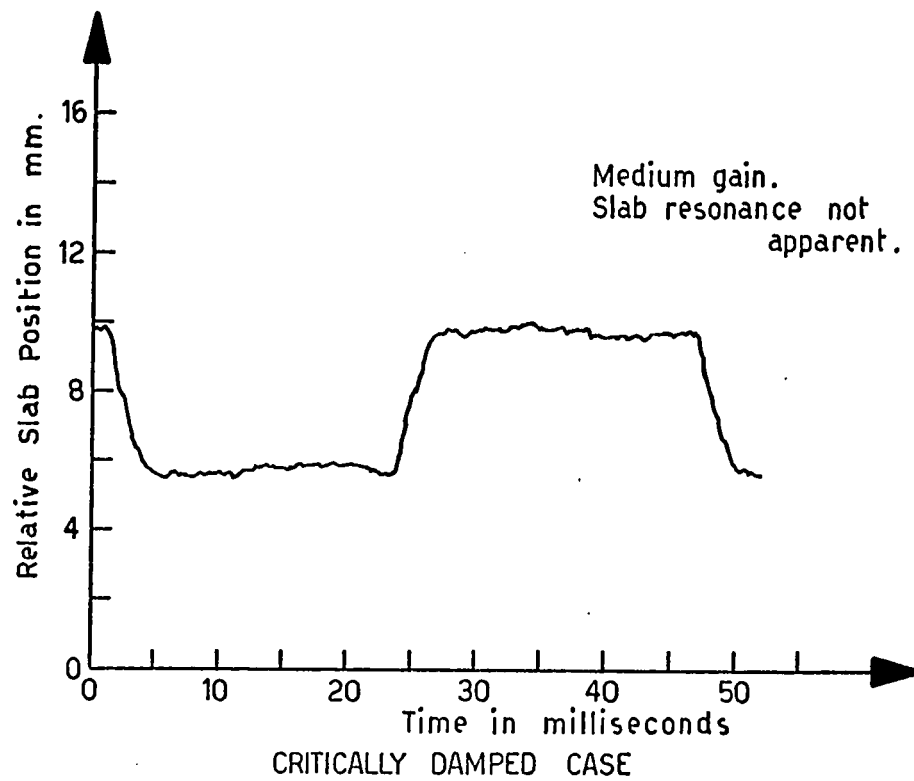
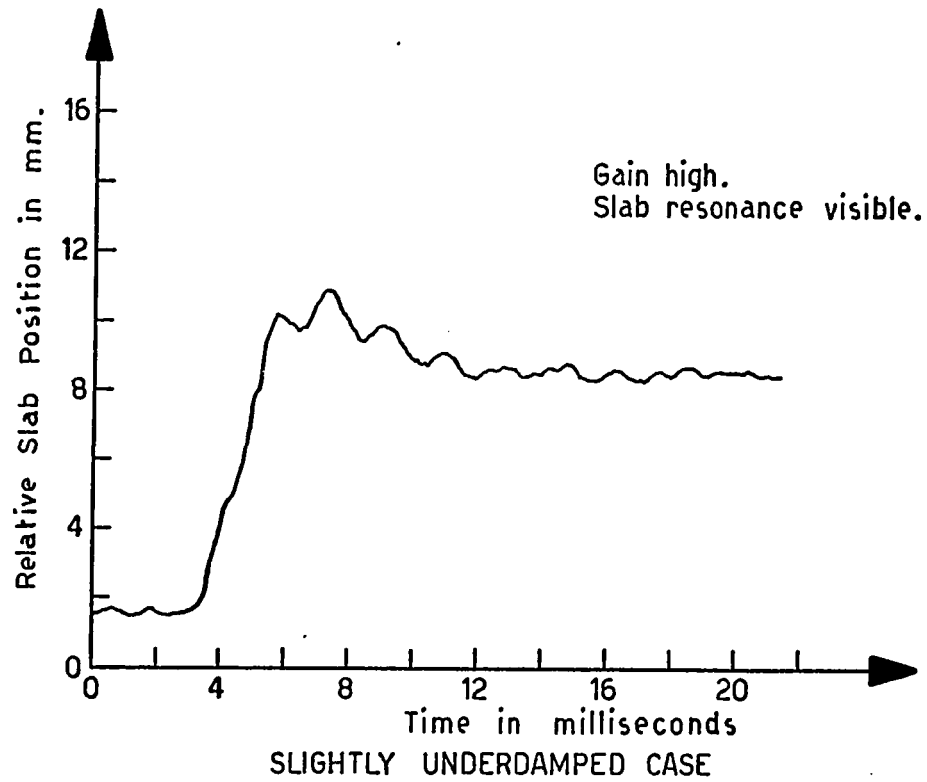


FIGURE 51. PHASESHIFTER TEST CIRCUIT,
SLAB POSITION SQUARE WAVE RESPONSE

record of slab position versus time, for the X-band phaseshifters tested, illustrates the above points. (see figure 50).

In an effort to characterize the frequency response of the phaseshifters, sinusoidal inputs were also applied to the test circuit. By determining the complex ratios of slab position to drive coil input signal, and slab velocity to drive coil signals, it was attempted to obtain Bode plots for the phaseshifters. However, the results obtained were very dependent on signal amplitude, average slab position in the waveguide, frictional effects, and play in the bearings. No Bode plots suitable for the accurate determination of the phaseshifter transfer function were, therefore, obtained. Nevertheless, it became abundantly clear that damped mechanical resonance effects existed, which resulted from the spring-like flexing properties of the dielectric slab itself. Several resonance peaks and dips are present in the phaseshifter response curve, with the main and lowest one occurring at 650 Hz. Efforts to increase the resonance frequencies to more than 650 Hz led only to solutions that were not compatible with proper microwave operation of the phaseshifters.

For purposes of determining stability of an actual matching system, the transfer function of the linear part of each control loop (with the damping loop closed) may be assumed to include at least a complex conjugate pair of poles and possibly a complex conjugate pair of zeros, as well as the poles of the ideal transfer function discussed before. However, the actual s-plane location of the additional poles and zeros could not be determined.

The effects discussed above will be considered again in the next section in connection with stability of the actual system.

6-5 Stability of the Actual System

As for the idealized system discussed before, proof of the stability of the actual system depends, first of all, on the positive real nature of the expression $[\alpha \underline{K} + (\alpha + \beta s) \underline{W}(s)]$ (see chapters 4 and 5), as well as on the properties of the non-linearities. In particular, to prove stability one requires that $[\alpha \underline{K} + (\alpha + \beta s) \underline{W}(s)]$ be positive real for $\alpha \geq 0, \beta > 0$, and $\underline{K} = 0$. Since $\underline{W}(s)$ for the actual system has two diagonal elements (the others are zero) which are complicated and not accurately known, this is virtually impossible to show. Nonetheless, there are several considerations which, although they do not constitute a proof, lead one to expect that the actual system might well be stable for certain operating conditions. These considerations are discussed below.

A. It is conceivable, although not very likely, that $\underline{W}(s)$ may be such that $\underline{M}(s) = \alpha \underline{K} + (\alpha + \beta s) \underline{W}(s)$ is positive real for $\alpha \geq 0, \beta > 0$, and $\underline{K} = 0$. Should $\underline{M}(s)$ not be positive real as such, it may be possible to compensate the system in order to make $\underline{M}(s)$ positive real for the conditions given above. In that case, stability may be proven by methods similar to those used for the idealized system.

B. It may be stated with some degree of certainty that $\underline{M}(s)$ will be positive real, or may be made positive real by compensation, for a limited range of values of α, β , and \underline{K} . The acceptable values of α, β , and \underline{K} thus determined will not likely include $\alpha = 0$ and $\beta \rightarrow 0^+$, as required for a stability proof as given for the idealized system. Therefore, a proper Lyapunov function can only be found if the non-linearities do, in fact, obey Anderson's conditions on the non-linearities. In section 5-4, condition 5.28 was shown to hold true for a limited range of val-

ues of y only, because proper period endpoints for x could not be defined for extreme values of y . Inability to define such period endpoints, however, does not imply that these period endpoints do not exist. There are several endpoint choices which make the non-linearities obey condition 5.28, but they cannot be proven to be the correct ones.

Assuming that the non-linearities obey condition 5.28, the system can be proven stable, but only for a non-zero value of \underline{K} . A non-zero value of \underline{K} restricts the non-linearities to sectors in the first and third quadrants. However, this need not be of any concern, because adjustment of the gains associated with $\underline{W}(s)$ allows these sectors to be opened up. Furthermore, restriction of the load reflection coefficient magnitude to less than one by means of mechanical range of motion stops on the stub phaseshifter will restrict the non-linearities to sectors less than one quadrant wide. One may, thus, expect a relationship between the loop gains used and the maximum allowable reflection coefficient magnitude of the loads to be matched.

C. In the case of linear systems, stability is often determined, both experimentally and theoretically, on the basis of the dominant open-loop poles and zeros only. The method used is that of reducing the loop gains sufficiently such that the troublesome minor poles and zeros occur at frequencies for which the effective loop gain has dropped to considerably less than one. In other words, the minor poles and zeros have been "buried" below the zero dB line in a Bode plot.

Similar methods may, frequently, be applied to non-linear systems. If this were not so, computer simulations of non-linear systems (or linear systems) would be valueless, since any non-ideal computer will add undesirable poles and zeros of its own to the system simulated.

Therefore, it is reasonable to expect that an idealized system, which has been proven stable by both theoretical means and computer study, may be implemented as a stable practical system, despite the presence of additional higher frequency poles and zeros.

D. As discussed in section 6-2, the hybrid computer simulation of the idealized system remained stable, even if additional poles were purposely added to the linear parts of the system. The addition of these poles merely imposed restrictions on the allowable loop gains and unmatched load reflection coefficients, if stability was to be maintained. The simulation study, therefore, supports the theories put forth in the above paragraphs.

The final and most powerful demonstration that an actual system can be stable will be presented in the form of an actual system prototype. The performance of the prototype will be discussed in chapter seven.

CHAPTER SEVEN

THE ACTUAL SYSTEM AND ITS PERFORMANCE

To establish with certainty that an actual system is both stable and practical, an automatic matching system prototype was built and tested. This chapter will deal with the construction of the system prototype and will give details of the required electronics. System performance results will be presented, and suggestions will be made regarding alternate system forms and possible practical improvements.

7-1 The Microwave System Arrangement

For convenience and safety reasons, the system prototype was executed as a 10 mWatt X-Band system. The microwave power source used was a Sylvania type SYA 3200 tuned cavity Impatt diode device, which operated at a frequency of 9.9081 GHz and with a power output of about 12 mW. The Impatt diode was supplied with DC power from an electronic constant current source with built in overvoltage, and "crowbar" type overcurrent protection. To prevent excessive frequency pulling and reflection of major amounts of power from the mismatched power source during the initial mismatched load conditions, the power source was coupled to the transmission waveguide through a standard isolator. The transmission waveguide power level could be adjusted, if desired, by means of a variable attenuator.

The loads used for testing purposes consisted of such varied devices as a precision variable attenuator terminated by a movable short,

an open-ended waveguide radiating towards or through metal or dielectric objects respectively, a horn antenna radiating towards a rapidly movable reflecting metal plate, as well as others.

The matching device was assembled exactly according to the single stub tuner arrangement suggested in chapter one. It consisted of an H-plane tee, a series phaseshifter inserted in the main waveguide between the H-plane tee and the load, and a shorted stub connected to the H-plane tee, comprising a stub phaseshifter terminated by a movable short. The purpose of the movable rather than fixed stub short was to allow the stub length to be adjusted, such that the zero stub admittance position of the stub phaseshifter corresponded to the center of its allowable range of motion. This assured that the largest, but finite, available stub admittance was equal in the positive and negative directions.

The waveguide loading conditions, resulting from the combined load and matching device, were measured with an E-field sampling version of the three detector measuring device discussed in chapter three. Since mounting of E-field, X-band detector probes to an accuracy of a few thousandths of an inch is rather difficult, the probes were located on a slotted section of waveguide. This made it possible to adjust their positions with respect to the standing wave pattern produced by a short at the neutral plane of the H-plane tee. In all other respects, the measuring device construction was similar to that discussed in an earlier chapter.

A directional coupler, with its sidearm terminated in a matched detector crystal, was inserted between the load and the series phaseshifter for the purpose of obtaining a signal that is proportional to the

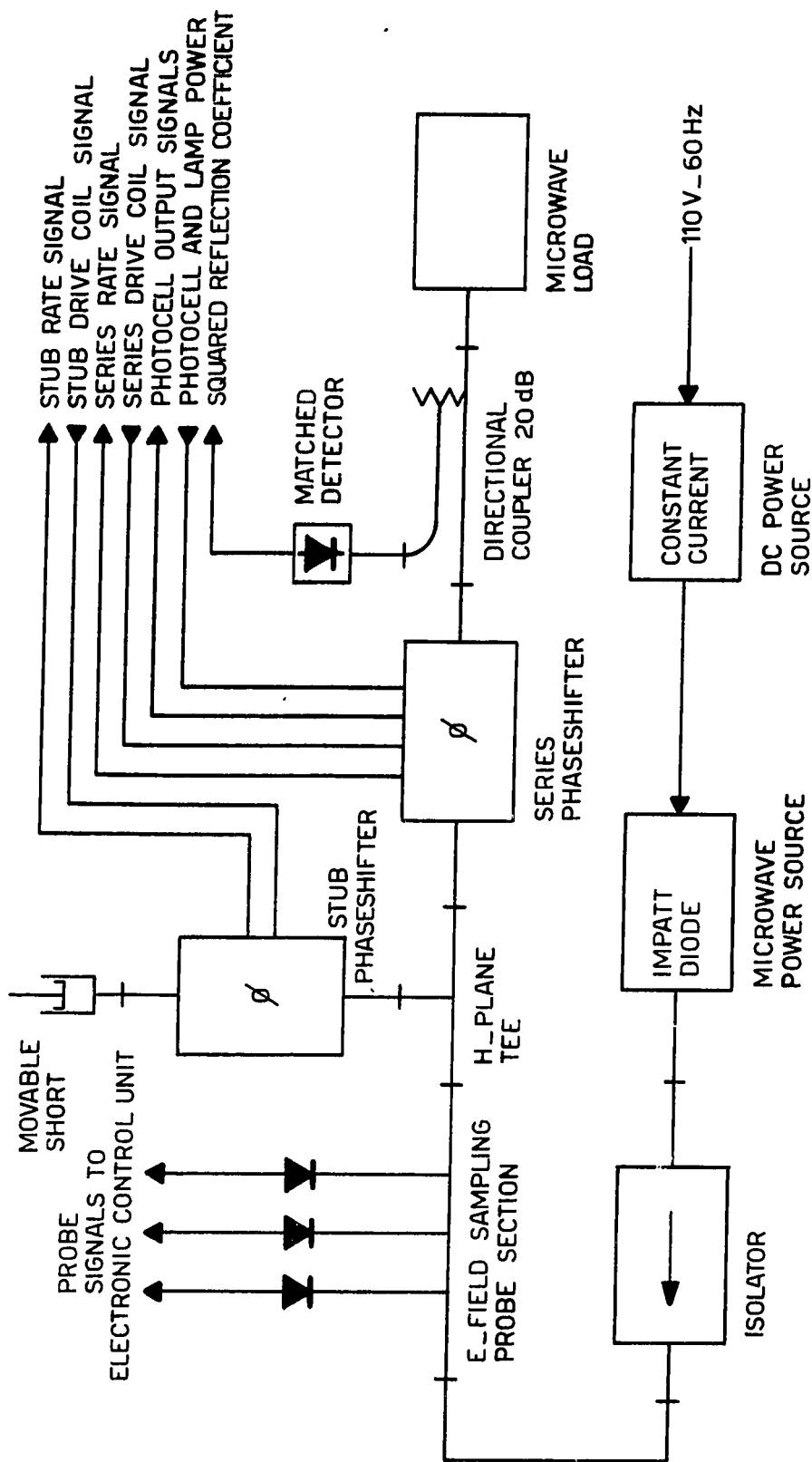


FIGURE 52. THE PROTOTYPE MICROWAVE COMPONENT ARRANGEMENT

square of the load reflection coefficient magnitude.

A schematic diagram of the microwave component arrangement is shown in figure 52. The connections to the electronic control unit are shown, but will be discussed in greater detail in a later section.

7-2 Special Phaseshifter Features

The general manner of construction and mode of operation of the dielectric phaseshifters was discussed in detail in chapter two. Some special features not previously considered or shown in diagrams need mentioning here.

For successful operation of the automatic matching system, the range of motion of the stub phaseshifter slab must be restricted such that the maximum obtainable phaseshift is slightly less than $\pm\pi/2$ radians. This was easily accomplished by simply bolting two slotted angle brackets to the top support plate of the phaseshifter structure. The slots allowed transverse positioning of the brackets with respect to the waveguide. At either adjustment extreme for the slab, the drive coil or rate coil endplates, respectively, would strike the angle brackets and would, thus, restrain the slab from moving further in that direction. Normally, the brackets were positioned such that the slab remained as close as possible to the waveguide sidewall, in order to take advantage of the linear range of the phaseshifter. The cross-sectional view of part of the phaseshifter and one angle bracket, shown in figure 53, further explains the phaseshifter motion stop arrangement.

The series phaseshifter range of adjustment must also be restricted

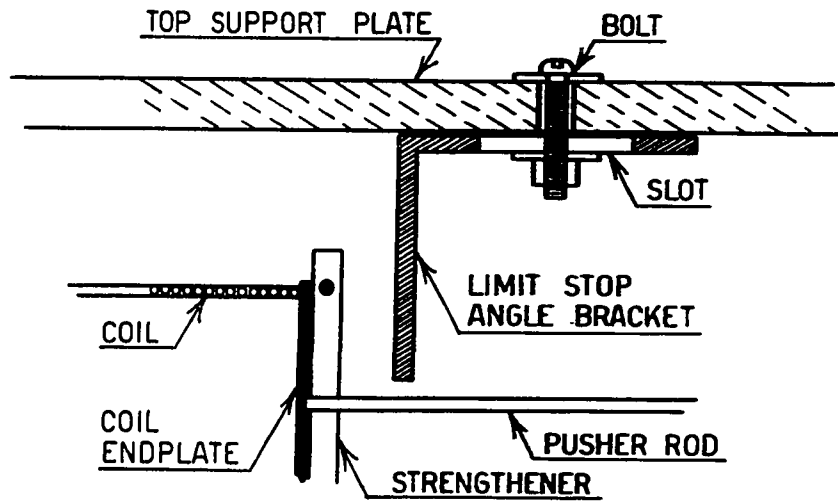


FIGURE 53. STUB PHASESHIFTER MECHANICAL LIMIT STOP

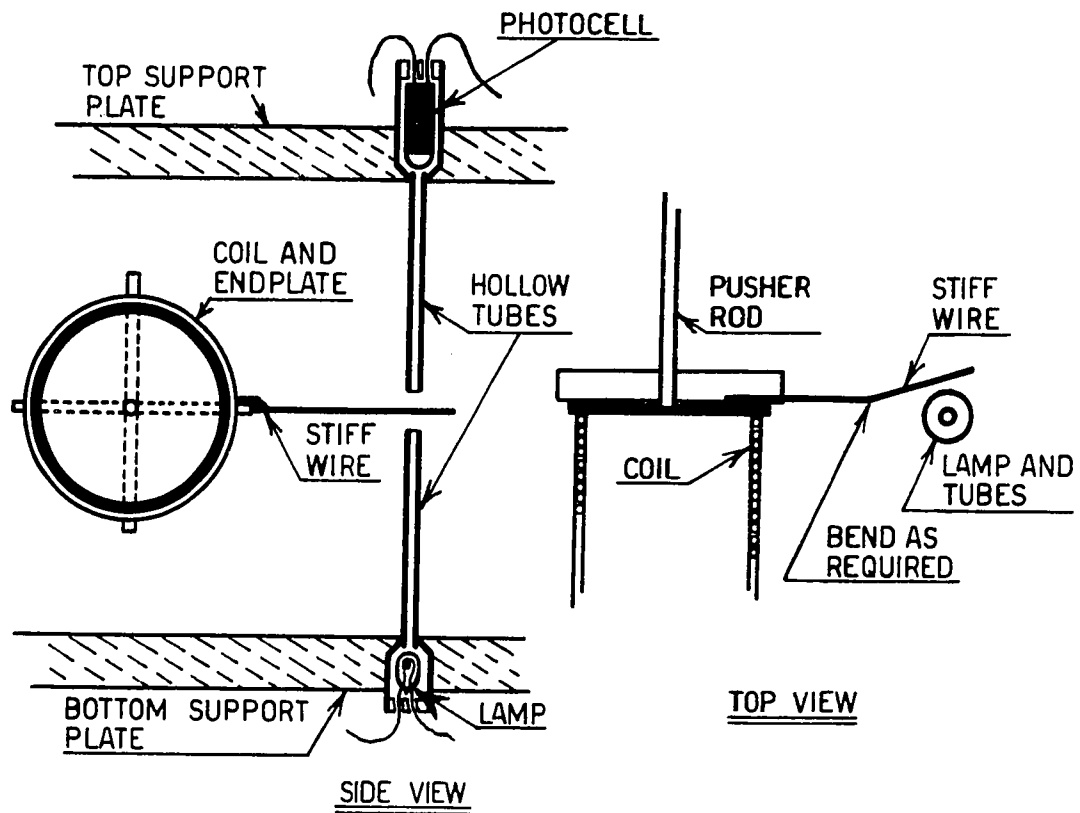


FIGURE 54. SERIES PHASESHIFTER PHOTOELECTRIC LIMIT STOP

to prevent undesirable overshoots and to keep the slab in the linear range of phase angle adjustment. Preferably, the series phaseshifter range should be less than π radians, but, to make matching of all dissipative loads possible, the adjustment range should be at least $\pm\pi/2$ radians. In the actual device, the series phaseshifter motion was restricted in exactly the same manner as that for the stub phaseshifter.

Whenever the series phaseshifter slab reaches one of the motion stops, a signal must be generated, which may be used in the electronic unit to switch the polarity of the series phaseshifter loop gain. Simultaneously, a short pulse must be generated and added to the loop error signal to help drive the slab away from the motion stop. The primary signals to activate the sign reversal and pulse circuits were obtained from lamp and photocell systems placed one at each motion stop. Stiff wires, attached to the drive coil and rate coil endplates, would interrupt light beams that passed from narrow beam sources to narrow acceptance angle photocells mounted close to the motion stops. The light beams and acceptance angles were narrowed without the use of lenses, and ambient light was kept out of the system, by almost joining the lamps and photocells with very small diameter brass tubes. The wire attached to the coils would pass through narrow gaps between the ends of the brass tubes in front of the photocells and light sources, respectively, whenever the slab approached one of the motion stops. The photocell tripoints could be easily adjusted by bending of the wires attached to the coils. The actual arrangement of one photocell and lamp system is shown in figure 54. Circuit details will be given in the next section.

7-3 The Electronic Control Unit

The electronic control unit performs a number of important functions, which may be enumerated as follows:

1. The measuring device crystal detector output signals are accepted, amplified equally, and combined arithmetically to yield error signals K_{RH} and K_{RV} .
2. The rate coil output signals are accepted, amplified, and mixed with the appropriate error signals to provide loop damping.
3. The series control loop gain polarity is reversed in response to signals from photocell limit stops mounted on the phaseshifter. Initial velocity pulses are generated during each polarity reversal operation, and are added to the series loop error signal.
4. The directional coupler detector crystal output signal (squared load reflection coefficient magnitude) is amplified, modified, and used to modulate both loop gains.
5. The rate feedback frequency response characteristics of both loops are modified to compensate for phaseshifter mechanical resonance effects.
6. The composite loop signals are further amplified into powerful output signals suitable for feeding of the phaseshifter drive coils.

The electronic control unit front panel provides full electrical and supervisory access to the control circuitry. As such, the control panel contains the following:

1. Two center-zero meters to indicate the static values of the crystal detector output signals, or the relative static values of the error signals K_{RH} and K_{RV} .
2. Dial potentiometers for gain adjustment of the two control loops.
3. Dial potentiometers for damping adjustment of the two control loops.
4. Coaxial cable connectors (BNC) for the crystal detectors (input signals).
5. Multi-pin connectors for the signals to and from the phaseshifter units.
6. Test signal input jacks for each loop.
7. Test signal output jacks for each loop, for purposes of oscilloscope examination of K_{RH} and K_{RV} .
8. AC power on-off switch.

The electronic control unit circuits are all silicon solid state. Wherever possible, integrated circuits were employed to perform the required circuit functions. More specifically, integrated circuit operational amplifiers provide all necessary gain and perform the arithmetic operations. Figure 55 shows the complete electronic control unit circuit diagram. It may be assumed that, where no details of the integrated circuit connections are shown, the manufacturer's recommendations regarding connections were followed. In figure 55, the gain and input resistance values of each amplifier channel are indicated. This provides all necessary information. To conserve space, the actual input and feedback resistors were not drawn. Some electronic control circuit details follow, and refer to figure 55.

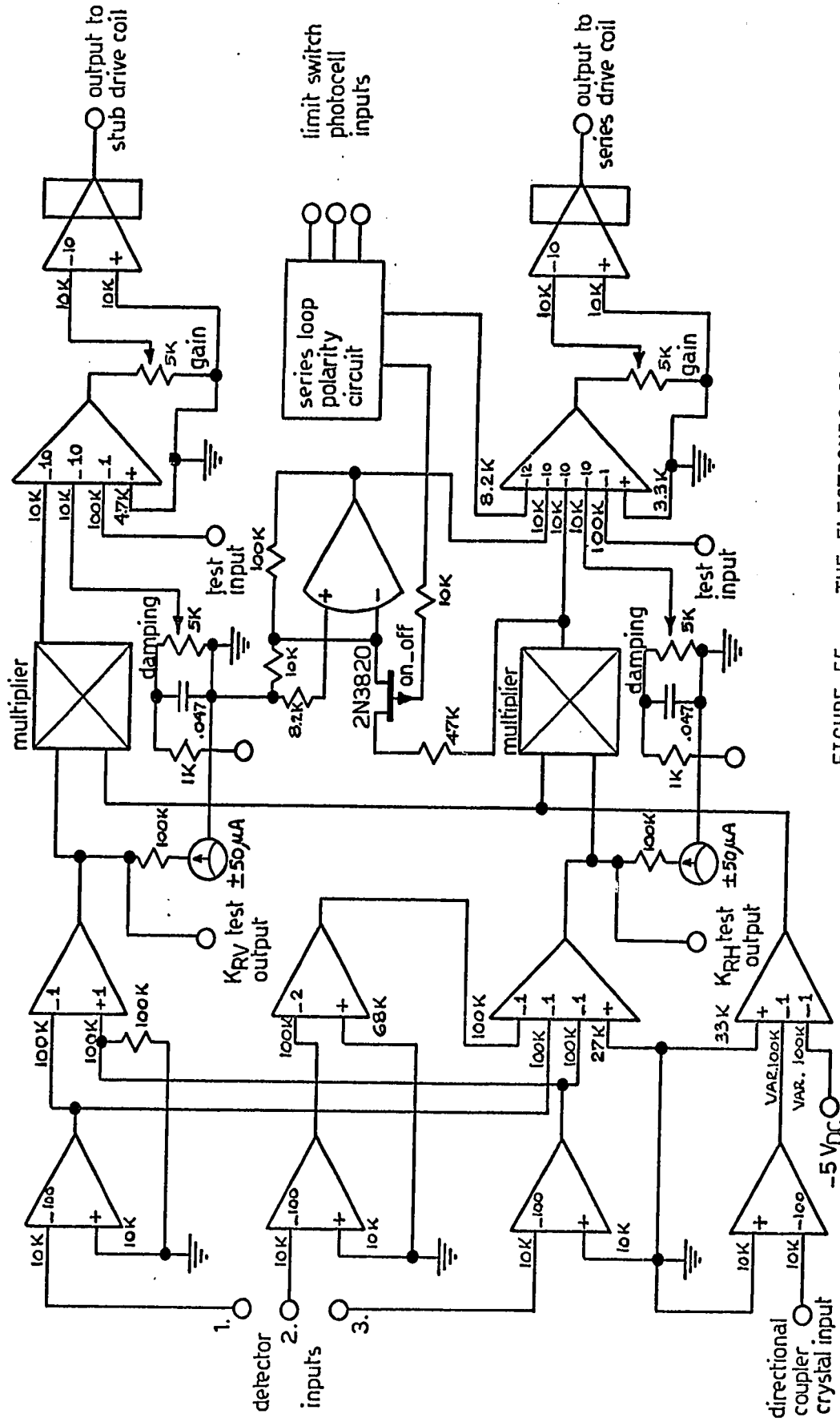


FIGURE 55. THE ELECTRONIC CONTROL UNIT CIRCUIT DIAGRAM

The four input signal preamplifiers are special low noise, low drift units that consist of a Fairchild μ A 709 operational amplifier with a type μ A 726 temperature controlled, integrated, differential input stage added. The combined unit is still used in an operational amplifier mode. A typical drift figure for these combined units is $0.5 \mu\text{V}/^\circ\text{C}$. All other operational amplifiers used are standard μ A 709 types.

The multipliers need not be accurate, because the compensation for load admittance and power level effects is only approximate in any case. Simple integrated circuit multipliers, with direct inputs and output (and automatic scaling to 10 Volts), could thus be employed. The units used are of type MC 1595L (Motorola).

The power amplifiers for the drive coil signals, again, contain a μ A 709 operational amplifier, but with a complimentary compound emitter follower power stage added. The output capabilities for voltage are, therefore similar to those of the operational amplifier, but up to 1.5 Amps of load current may be drawn from them. As in a standard operational amplifier, feedback is applied from the emitter follower output to the summing junction.

The series loop polarity reversal circuit operates in the following manner. The multiplier output signal passes to the next amplifier in two ways, once directly, and once through a switched amplifier with a gain of minus two. With the switched amplifier off, only the direct signal is present at the next amplifier. But, with the switched amplifier on, the direct signal plus twice the negative of the direct signal appear at the input of the next amplifier (to be added there). The net result is a simple sign reversal. On-off switching of the amplifier

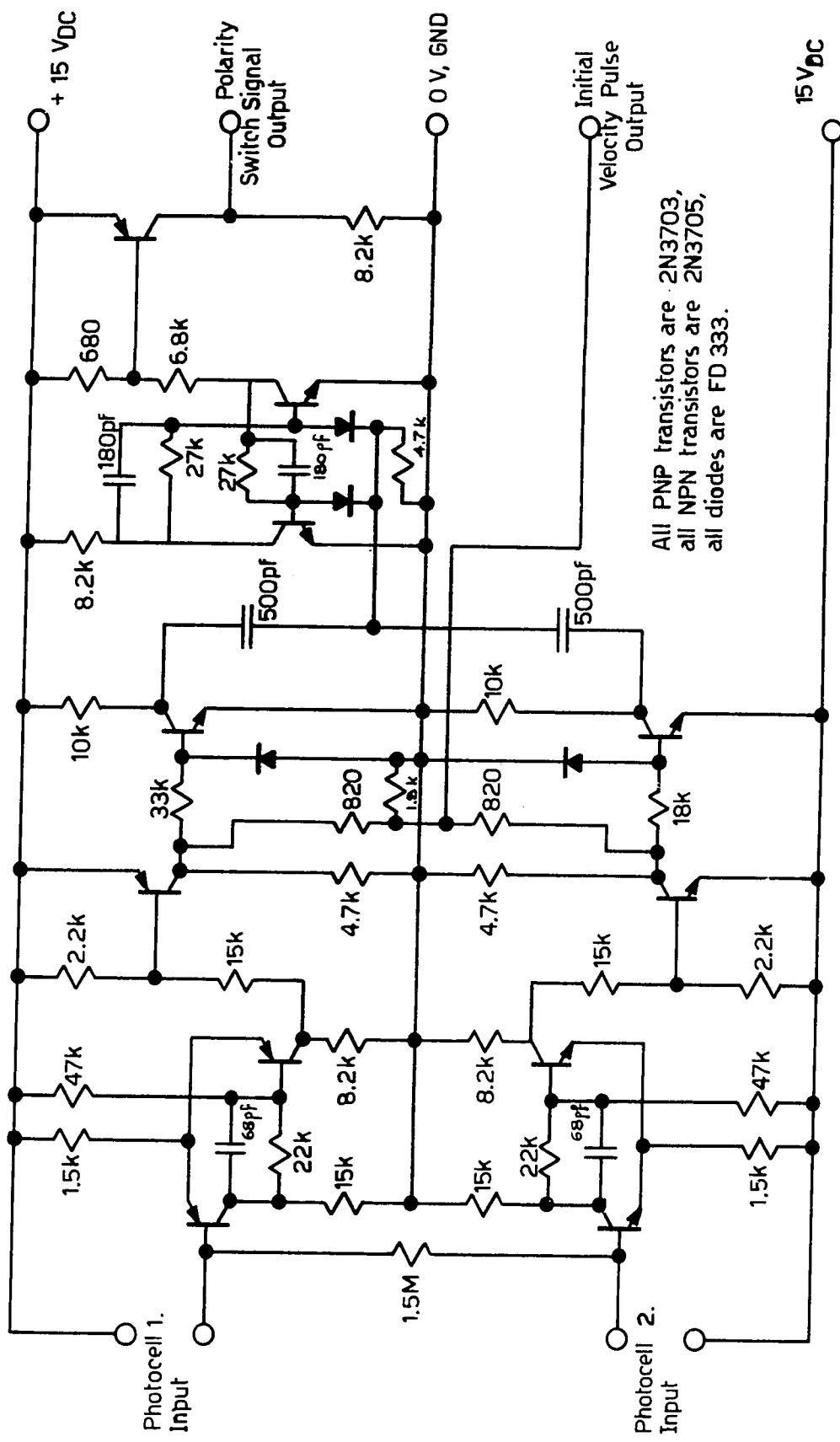


FIGURE 56. THE SERIES LOOP POLARITY SWITCHING CIRCUIT

is accomplished with a field effect transistor in the summing junction. With 0 Volts on the field effect transistor gate, the transistor is on, and the amplifier operates. With 15 Volts on its gate the field effect transistor is off, thus isolating the amplifier from its input. The switching signal for the field effect transistor gate, as well as the initial velocity pulses are derived from the series loop polarity circuit. A complete circuit diagram of this is shown in figure 56. An explanation of the circuit operation follows.

Normally, both photocells are illuminated, but, whenever the series phaseshifter reaches a limit stop, one of the photocells will receive no light. The resulting change in resistance causes a Schmitt trigger, associated with that photocell, to change states. An additional common emitter stage alters the Schmitt trigger output to a signal which is zero whenever the photocell is illuminated, but which becomes a 1.5 Volt signal (of a specific single polarity) during the time that the photocell is dark. The second photocell operates a similar, but complimentary circuit, which also produces an output whenever that photocell is dark. However, the polarity of its output is opposite to that of the first circuit. The outputs are combined in a resistive network and are then fed to the initial velocity pulse output of the circuit. The initial velocity pulse output voltage is, therefore, +1.5 Volts, -1.5 Volts, or 0 Volts depending on whether photocell 1 is dark, photocell 2 is dark, or both are illuminated, respectively. The velocity pulse polarity has been arranged always to drive the phaseshifter slab away from its stops. Note that the pulse lasts only as long as the photocell is darkened. The magnitude of the velocity pulse and the gain of the amplifier channel into which it is fed are such that the pulse

overrides even the largest possible error signal magnitude. This prevents "hang-up" of the slab at one of the limit stops.

The change of states of the Schmitt triggers, as a photocell darkens, is changed to a pulse of appropriate polarity and magnitude (by means of another pair of complementary transistors and differentiating networks), which is fed into a flip-flop. The circuit is so arranged that a pulse originating from either photocell and Schmitt trigger combination will cause the flip-flop to change states. The flip-flop output is fed to the final switch signal output terminal by means of a pulse shaping transistor stage. Since the flip-flop state determines whether the field effect transistor in figure 55 is on or off, darkening of either one of the photocells will cause a loop gain sign reversal which lasts until a photocell is again darkened.

This completes the discussion of the electronic control unit circuit.

7-4 System Adjustments

To assure that the system operates properly, the measuring device probe positions along the waveguide and the probe attenuations must be adjusted correctly. Furthermore, the phaseshifter limit stops must be set. Finally, the loop gain and damping potentiometers must be properly adjusted. The adjustment procedure, which applies to the system prototype in particular (with E-field probes on a slotted guide section), will be given in abbreviated form.

1. Arrange a directional coupler and detector crystal such as to make possible the measurement of the power reflected

- from the matching device.
2. Connect a matched load to the system and set the phase-shifters to their minimum phaseshift positions. Remove the drive coil fuses from the electronic control unit panel.
 3. Adjust the stub movable short until no power is reflected from the matching device.
 4. Connect just one outside probe to the electronic control unit and adjust its attenuator to obtain a stub error meter reading of 40% of full scale.
 5. Connect the second outside probe also and adjust its attenuation such as to return the stub error reading to zero.
 6. Connect the center probe to the control unit and adjust its attenuator until the series error signal is zero.
 7. Readjust the stub movable short to obtain maximum power reflection from the matching device.
 8. Connect only the center probe to the electronic control unit and locate the probe at a convenient minimum (using the stub error meter for indication).
 9. Connect only an outside probe to the control unit and adjust its position to the nearest minimum away from the center probe.
 10. Connect the other outside probe only and move it to the nearest minimum on the other side of the center probe. Then, increase its distance from the center probe until the outside probe in question is at the next maximum.
 11. Connect both outside probes to the control unit. Maintain the distance between the outside probes as established in

steps 9 and 10 (use a micrometer) and slide the probes along until the stub error meter reads zero. The direction of motion to accomplish this must be the one which tends to center the middle probe between the two outside ones.

12. Repeat steps 2 to 6 inclusive if necessary.

Probe adjustment is complete at this point. For a measuring device with pre-positioned probes (such as the H-field sampling section of chapter three), only steps 1 to 6 are required. Alternate probe adjustment methods by measurement and/or calculation are, of course, possible.

The stub phaseshifter range of motion and the movable short are adjusted as follows:

1. Make the system fully operational, but remove the drive coil fuses. Adjust the series phaseshifter to its minimum phaseshift position.
 2. Connect a matched load to the unit.
 3. With the stub phaseshifter set to its minimum phaseshift position, adjust the movable short until a reading of plus 5% of full scale is obtained on the stub error meter.
 4. While holding the phaseshifter against its maximum phase-shift limit stop, adjust this limit stop until a reading of minus 5% of full scale is obtained on the stub error meter.
- For these adjustments to be correct the stub error meter reading must behave in the following manner to slow manual adjustment of the stub slab position from minimum to maximum phaseshift: the meter reading must increase from +5% to a maximum; it must then decrease (through zero) towards a minimum; from the minimum it must increase again to a

maximum reading of -5%. Only one zero value of stub error must be included within the stub slab range of motion.

The series phaseshifter mechanical stops should simply be adjusted to give a series slab range of motion equal to that of the stub phaseshifter. The photocell arrangements should then be adjusted to obtain complete darkening of the photocells with the series slab about 0.5 mm from its extreme positions.

The gain and damping of each loop should be set by applying a small square wave signal to the test input for the loop in question, and by examining the loop response by connecting an oscilloscope to the loop output test jack. For each loop:

1. Connect a mismatched microwave load to the apparatus. Set the gains to low values and the damping to maximum. Make the unit operational.
2. Adjust the damping for the desired type of response.
3. Adjust the gain to the maximum value which will just leave the slab resonance, of the loop in question, invisible.
4. Repeat steps 2 and 3 several times for both loops, to eliminate the effects of interaction.

The circuit is purposely so arranged that only the damping affects the type of system response. The gain adjustment affects mainly the response speed. This arrangement allows maximum response speed to be obtained without interference from the slab mechanical resonances.

The system adjustment is now complete and should not have to be changed unless the microwave operating frequency or maximum power level is changed.

It should be noted that small errors in positioning of the meas-

uring probes on the waveguide section will have no effect on the system operation. If the spacing between the probes is correct, but the distance to the neutral plane of the H-plane tee is in error, the K_{RV} -axis and K_{RH} -axis will remain orthogonal, but will be rotated with respect to the main Smith chart axis. Should the probe spacing be in error instead, the K_{RH} -axis and K_{RV} -axis will not remain orthogonal. Errors in crystal detector tracking and deviations from square law operation will cause curving of the K_{RV} and K_{RH} axes with respect to the Smith chart. If these errors are small, they are of no importance, provided the detector output signals are equal for matched load conditions. However, large errors could cause instability problems.

7-5 Actual System Performance

The performance of the actual system prototype may be summarized as follows:

A. By displaying one test jack output versus the other on a storage oscilloscope, while changing the load and slowly adjusting the phaseshifter slab positions the theoretical results regarding the error signal and non-linearity derivations of chapters three and four were verified. All of the "Smith chart" trajectories were essentially circular. Some distortions occurred as a result of detector crystal tracking errors and lack of square law operation. However, the errors were not considered detrimental.

B. Loads with reflection coefficient magnitudes as large as 0.95 (VSWR of about 50) could be matched in a stable manner. The maximum allowable load reflection coefficient magnitudes (determined by the

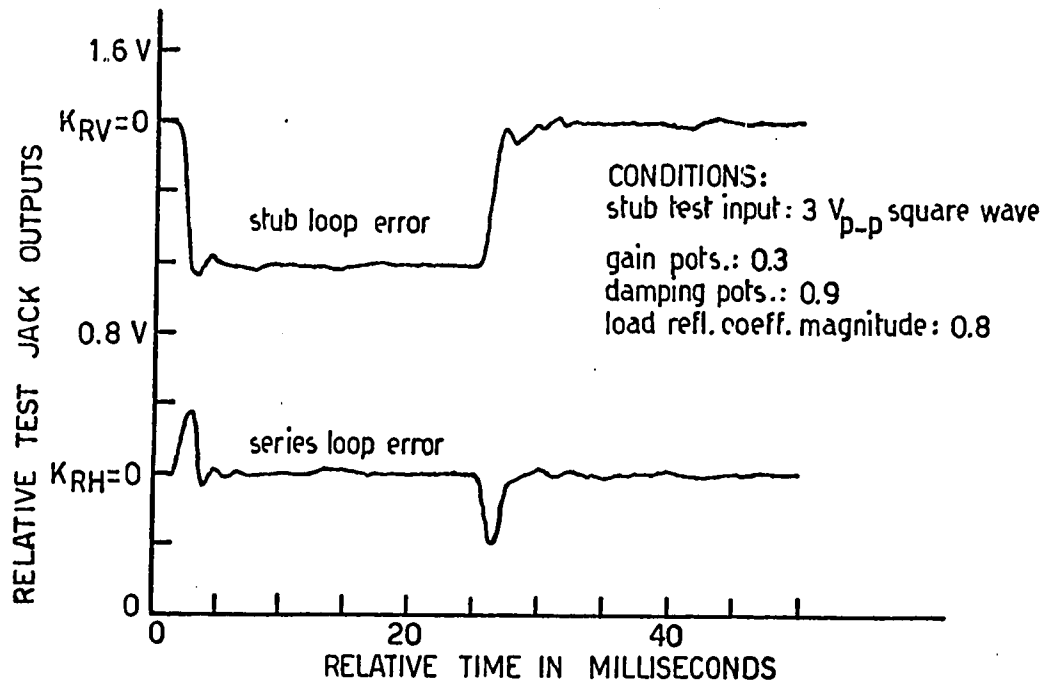


FIGURE 57. SYSTEM RESPONSE TO STUB LOOP SQUARE WAVE FORCING

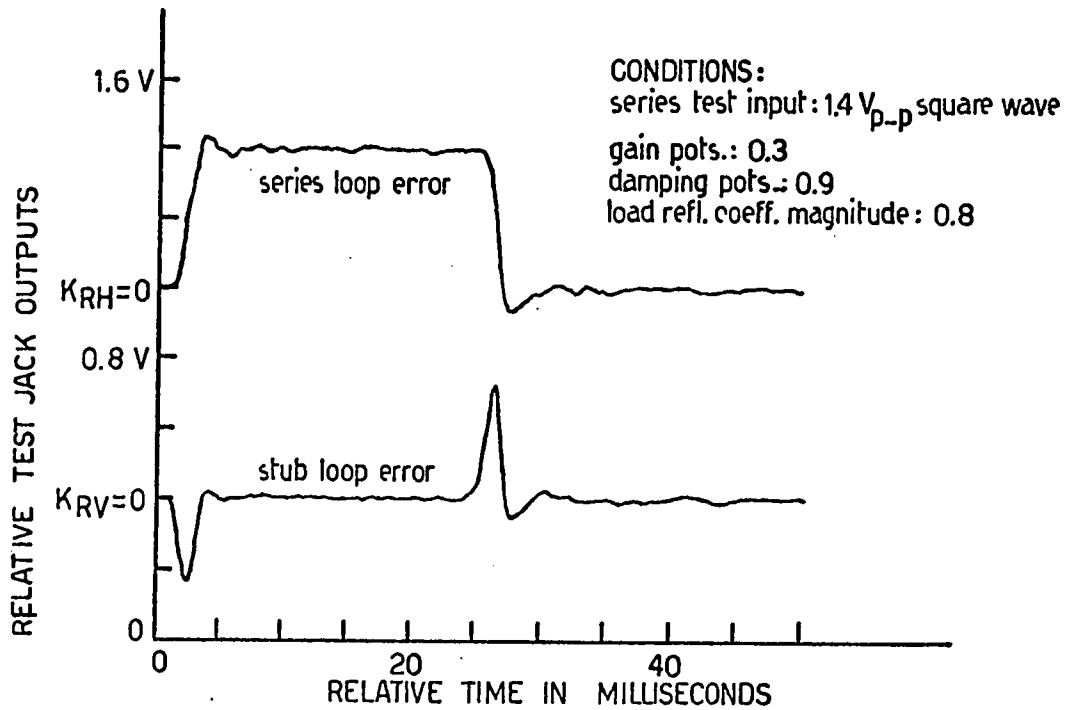


FIGURE 58. SYSTEM RESPONSE TO SERIES LOOP SQUARE WAVE FORCING

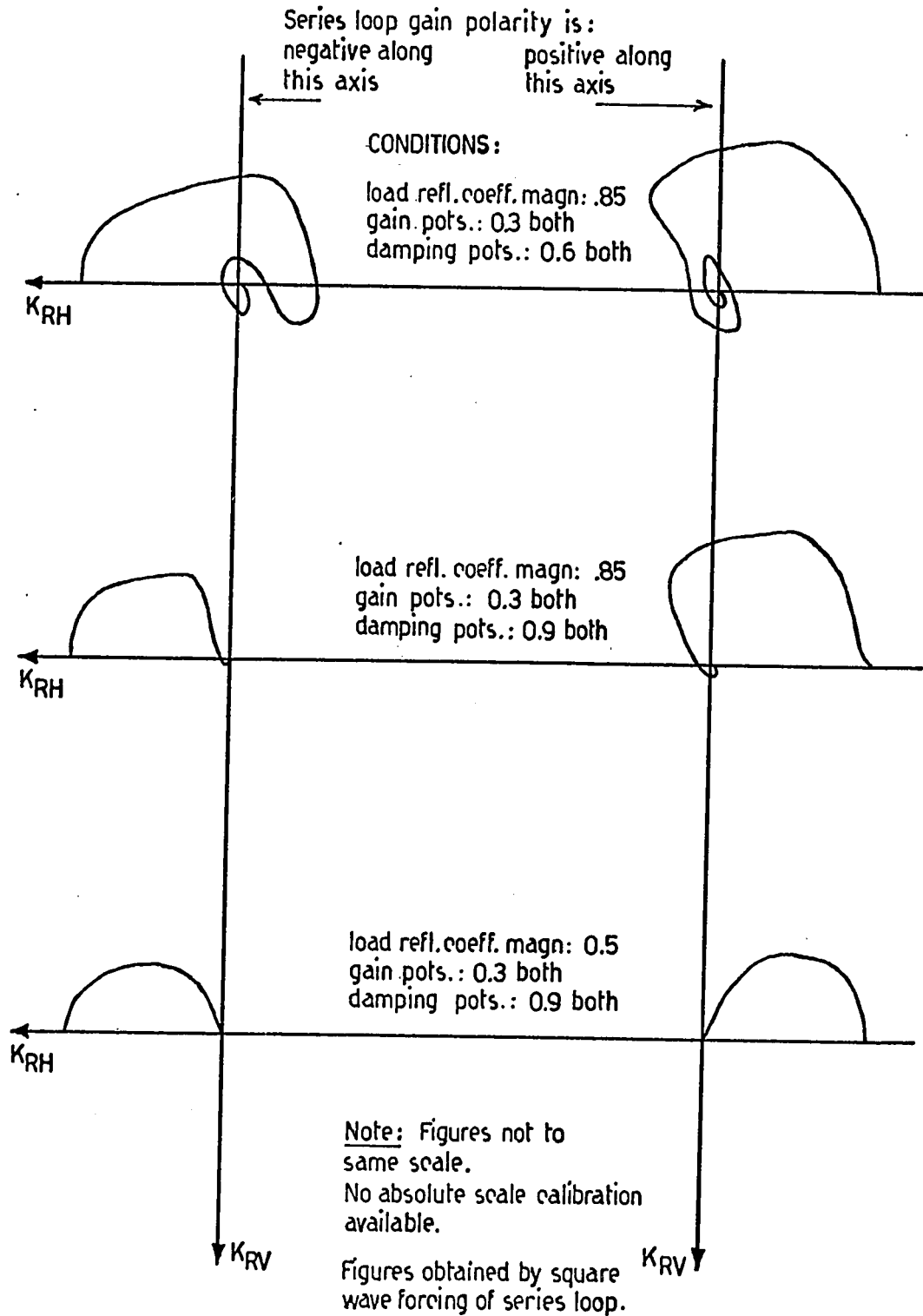


FIGURE 59. ACTUAL SYSTEM MATCHING TRAJECTORIES FOR TYPICAL CONDITIONS

stub phaseshifter range of motion), the maximum allowable gains, and the minimum allowable damping were all found to be interdependent. The matching of larger reflection coefficient loads required the use of less gain and greater damping to maintain stability.

C. After matching of a load the voltage standing wave ratio in the main waveguide was no more than 1.03. At the same time, loads with reflection coefficient magnitudes of less than 0.02 caused no matching device corrective action. These results were affected by the effective loop gains and the phaseshifter frictional effects. To overcome the static friction in the phaseshifter bearings a certain minimum drive coil signal is required. The gain and degree of mismatch, together determine the minimum drive coil signal magnitude available.

D. The maximum stable system response speed obtainable for large reflection coefficient loads was dependent upon the slab mechanical resonance frequencies. This was tested by using several slab designs of differing dimensions. Figures 57 and 58 show some typical tracings of oscilloscope records for a system with slab resonance frequencies of 650 Hz. The obtainable response speed is clearly between 2 and 5 milliseconds per complete matching operation. Small reflection coefficient loads were matched in up to 15 milliseconds. The response speed for small reflection coefficient loads could be improved by incorporation of the load dependent loop gain modulation discussed in chapter six. This was proven by manually adjusting the loop gains in accordance with the formulae given in chapter six. Small reflection coefficient loads were then matched in less than 5 milliseconds also.

E. Typical prototype system matching trajectories, as displayed on an oscilloscope, are shown in figure 59. These trajectories differ

somewhat from those obtained from the computer simulation. The differences were thought to be the result of:

1. Differences between linear part transfer function matrices for the simulated and actual systems.
2. Additional non-linearities caused by slab bearing friction.
3. Crystal detector tracking errors and deviations from square law operation.

In general, however, the performance of the actual system was much as expected.

7-6 System Operation at High Microwave Power Levels

The X-Band system prototype was used at very small microwave power levels. However, S-Band phaseshifters of a similar design were tested at microwave power levels of up to 1.5 kWatts. An H-field measuring section was also operated at this same power level. Although these devices were never assembled into a complete matching system, it was felt that (based on slab temperature and leakage radiation measurements), the S-Band system could be operated at levels of up to 10 KW of microwave power.

Larger systems are thought to be feasible if other than polystyrene dielectric slabs are used. Materials such as beryllia and titania appear especially promising. They are hard ceramics that exhibit high dielectric constants. As such, use of these materials would result in phaseshifters with slabs of smaller dimensions and much higher mechanical resonance frequencies. Although these materials are more lossy than polystyrene at microwave frequencies, they are far more tolerant

of temperature increases. Therefore, their use at high power levels is recommended.

If the necessarily slow response speed and much greater mechanical friction effects can be tolerated, the series and stub phaseshifters may be replaced by devices that contain motorized movable shorts. Careful design of such a modified system can result in a matching device that is useful even at very high microwave power levels. For example, the stub phaseshifter and short may be replaced by a stub terminated directly by a motorized movable short. Similarly, the series phaseshifter may be replaced by a motorized microwave "line-stretcher". One form of line-stretcher consists of a short-slot hybrid (a form of 3-dB coupler) with two adjacent ports terminated by symmetrically coupled, tracking, movable shorts.

Automatic matching system modifications, as suggested above, do not alter the basic principles of operation of the matching device. The non-linear functions remain the same, and even the transfer function matrix for the linear part is altered very little. The system analysis presented in this thesis is, thus, applicable and stability of such a system is guaranteed.

Several other types of matching devices, incorporating two or three movable shorts, and applicable to use at high power levels, are known. These alternate devices match loads in a manner which is different from that of the single stub tuner. Hence, the real and imaginary components of the reflection coefficient may no longer be useful as system error signals, and adjustment may not be unique. In any case, the system nonlinearities would be vastly different from those of the single stub matching device. Complete re-analysis, to establish modes of operation

and to determine stability, would certainly be required. No further consideration was given, at this time, to systems of this type.

CONCLUSIONS AND RECOMMENDATIONS

The automatic microwave load matching system, described in this thesis, employed special adaptations of standard microwave components. The matching system was arranged as a single stub tuner with the shorted stub and series line lengths made adjustable by means of dielectric slab phaseshifters. Short stroke linear motion DC servomotors (of brushless design) were directly coupled to the phaseshifter dielectric slabs and all moving parts were suspended on very low friction bearings. The resulting electrically actuated load matching device was found to be ideally suited for use in an automatic microwave load matching system.

The main feeder waveguide loading conditions were sampled with the aid of three crystal detector probes. Probe spacing was selected as one-eighth of a guide wavelength with the center probe positioned at an E-field or H-field minimum with respect to a waveguide short at the stub tee connection. The signals from the crystal detector probes were arithmetically combined to produce two error signals that, respectively, were proportional to the real and imaginary components of the combined load and matching device reflection coefficient. This simple device has proven remarkably effective as a means of extracting accurate error signals suitable for use (after amplification) as actuating signals for the phaseshifters.

The actual, automatic system became practical only after means were devised of dealing with the periodic nature of the matching device characteristics. The particular selection of appropriate error signals, the limiting of the phaseshifter ranges of adjustment, and the incorporation of automatic series phaseshifter loop gain polarity reversal

circuits all combined to make the matching device adjustment unique. Although periodicity of the device adjustment was not eliminated, these effects were successfully modified such that the control loop action was not adversely affected.

The complete closed-loop system was analyzed in detail for stability and performance. Theoretical results were tested experimentally by means of computer simulation and the building of a prototype system. Good agreement was obtained between theoretical and experimental results. In particular, the X-band prototype was found to be capable of matching all loads with reflection coefficient magnitudes of less than 0.95 to a residual reflection coefficient magnitude of 0.02 within 15 milliseconds. Incorporation of a circuit which modulated the control loop gains, in response to actual load reflection coefficient changes, reduced the matching time required for all loads to less than 5 milliseconds. Furthermore, the system operated in a continuous fashion and would follow rapid load changes in a stable manner.

The system, as devised, is of reasonably simple and economical design and has a response speed that is more than adequate for industrial applications. As a special feature, the system is readily adaptable to use at higher microwave power levels, without there being a need for complete re-design or re-analysis.

Although the work carried out for this thesis was detailed, it has, by no means, been exhaustive. Attractive areas for investigation remain. Particularly interesting is the possibility of developing a very fast automatic matching device, incorporating electrically actuated ferrite phaseshifters. A fast matching device of this type may well prove useful in some specialized research applications of microwave power, where

response times of microseconds rather than milliseconds are required. The methods of analysis and stability determination used in the work of this thesis would be directly applicable to the development of a ferrite phaseshifter automatic matching device.

It is hoped that the work, as reported in this thesis, will contribute substantially to the more economical and convenient industrial utilization of microwave power.

BIBLIOGRAPHY

1. WINTER, R.A.: Self Balancing Capacitance Bridge, M.Sc. Thesis Report, Department of Electrical Engineering, University of Alberta, Edmonton, 1968.
2. BEITMAN, B.J.Jr. et Al.: Automated Coupling and Impedance Matching Network, U.S. Patent No.3160832, December 8, 1964.
3. ROZA, J.E.: Impedance Matching System, U.S.Patent No.3443231, May 6, 1969.
4. LANCE, A.L.: Introduction to Microwave Theory and Measurement, McGraw-Hill, New York, 1964.
5. COLLIN, R.E.: Foundations for Microwave Engineering, McGraw-Hill, New York, 1966.
6. ALTMAN, J.L.: Microwave Circuits, D. Van Nostrand Co., New York, 1964.
7. SMITH, P.H.: Transmission Line Calculator, Electronics, January 1939, p.29.
An Improved Transmission Line Calculator, Electronics, January 1944, p.130.
8. COLLIN, R.E.: Field Theory of Guided Waves, McGraw-Hill, New York, 1960.
9. MONTGOMERY, C.G., Editor; Technique of Microwave Measurements, Radiation Laboratory Series, McGraw-Hill, New York, 1947.
10. JOHNSON, W.C.: Transmission Lines and Networks, McGraw-Hill, New York, 1950.

11. LYAPUNOV, A.M.: Problème Général de la Stabilité du Mouvement, Ann. Fac. Sci. Univ. Toulouse, Vol. 9, 1907, pp.203-474.
12. LUR'E, A.I.: Nektorye Nelineinye Zadachi Teori Avtomaticheskogo Regulirovaniya, In Gostekhizdat, Moscow, 1951.
13. POPOV, V.M.: Absolute Stability of Non-linear Systems of Automatic Control, Translation, Automation and Remote Control, Vol. 22, No. 8, pp.857-875, March 1962.
14. KAHLMAN, R.E.: Lyapunov Functions for the Problem of Lur'e in Automatic Control, Proc. Nat. Acad. Sci., Washington, Vol. 49, pp.201-205, 1963.
15. BROCKETT, R.W. and WILLEMS, J.L.: Frequency Domain Stability Criteria - Part 1, IEEE Trans. on Automatic Control, Vol. AC-10, No. 3, pp.255-261, July 1965.
16. IBRAHIM, E.S. and REKASIUS, Z.V.: A Stability Criterion for Non-linear Feedback Systems, IEEE Trans. on Automatic Control, Vol. AC-9, No. 2, pp.154-159, April 1964.
17. JURY, E.I. and LEE, B.W.: The Absolute Stability of Systems with Many Non-linearities, Transl. from Avtomatika i Telemekhanika, J. of Automation and Remote Control, Vol. 26, Part 1, pp.943-961, 1965.
18. KU, Y.H. and CHIEH, H.T.: Stability of Control Systems with Multiple Non-linearities and Multiple Inputs, J. of the Franklin Institute, Vol. 282, pp.357-365, 1966.

19. ANDERSON, B.D.O.: Stability of Control Systems with Multiple Non-linearities, J. of the Franklin Institute, Vol. 282, pp.155-160, 1966.
20. ANDERSON, B.D.O.: A System Criterion for Positive Real Matrices, SIAM Journal on Control, Vol. 5, 1967.
21. DeRUSSO, P.M., ROY, R.J. and CLOSE, C.M.: State Variables for Engineers, John Wiley & Sons Inc., New York, 1966.
22. KINGMA, Y.J.: Signal Flow Graphs, Department Monograph, Department of Electrical Engineering, University of Alberta, Edmonton, 1965.
23. D'AZZO, J.J. and HOUPIS, C.H.: Feedback Control System Analysis and Synthesis, McGraw-Hill, New York, 1960.
24. DORF, R.C.: Modern Control Systems, Addison Wesley Publishing Co., Reading, Mass., 1967.