CR 86017

MODULATION TECHNIQUES FOR RANGE

MEASUREMENTS FROM SATELLITES

Contract Number NAS 12-146

FINAL REPORT

JUNE 1967

National Aeronautics and Space Administration

Electronics Research Center

SIPD

SPECIAL INFORMATION PRODUCTS DEPARTMENT   DEFENSE ELECTRONICS DIVISION

GENERAL ELECTRIC

SYRACUSE, NEW YORK

MODULATION TECHNIQUES FOR RANGE

MEASUREMENTS FROM SATELLITES

Contract Number NAS 12-146

FINAL REPORT

JUNE 1967

National Aeronautics and Space Administration

Electronics Research Center

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT.)

TABLE OF CONTENTS (CONT.)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (CONT.)

LIST OF TABLES

SECTION I

GENERAL

A. INTRODUCTION

This report contains analyses and evaluations of modulation techniques for range measurements from satellites. The material is presented as an aid in the selection of concepts and in the design of systems using satellites for navigation and traffic control. Satellites are attractive for these applications because they can serve as relays for line of sight transmission between fixed earth stations and distant ships and aircraft.

Present-day ship and aircraft long-range communications use ionospheric reflection in the medium and high frequency bands from 3000 kHz to 30 MHz. Propagation at these frequencies is dependent upon solar activity and the transmissions are not reliable. Satellites have already demonstrated that they can be used to provide superior communications for these craft.

Figure I-1 depicts the transmission links between an earth station, a satellite and a user craft. High signal-to-noise ratios can be achieved between the earth station and the satellite at the information rates necessary for navigation and traffic control because it is practical to use high transmitter power and high gain antennas at the earth station.

The transmission links between the satellite and the user craft present more difficult problems to the system designer. The limitations are set by the energy available in the satellite and by the small aperture of a mobile user antenna. Solar cells are the only currently available practical energy source for satellites that must have a long life in orbit. They are not efficient and the number that can be attached to a satellite of practical size and weight is limited so that the prime power available is on the order of 1 kW.

Fortunately, the state-of-the-art in satellite stabilization has progressed so that it is possible to use the radiated energy efficiently. As indicated by the dotted lines in Figure I-1, it is possible to direct the satellite's transmitted energy towards the earth so that very little is lost in space.

In spite of the advantage gained by using a directive antenna on the satellite, its radiated energy is spread over a very large area, and therefore the power density seen by a receiving antenna is small. The power captured for the receiver is a function of the antenna's physical dimensions. The system designer is faced with the conflicting requirements for large capture area; wide beamwidth; and for aircraft, aerodynamic acceptability. The selection of an operating frequency for a system is involved in this trade-off. For a given capture area, the antenna directivity increases as the frequency increases. At frequencies above VHF it becomes necessary to point the antenna in the direction of the satellite and this introduces design problems, particularly for aircraft application.
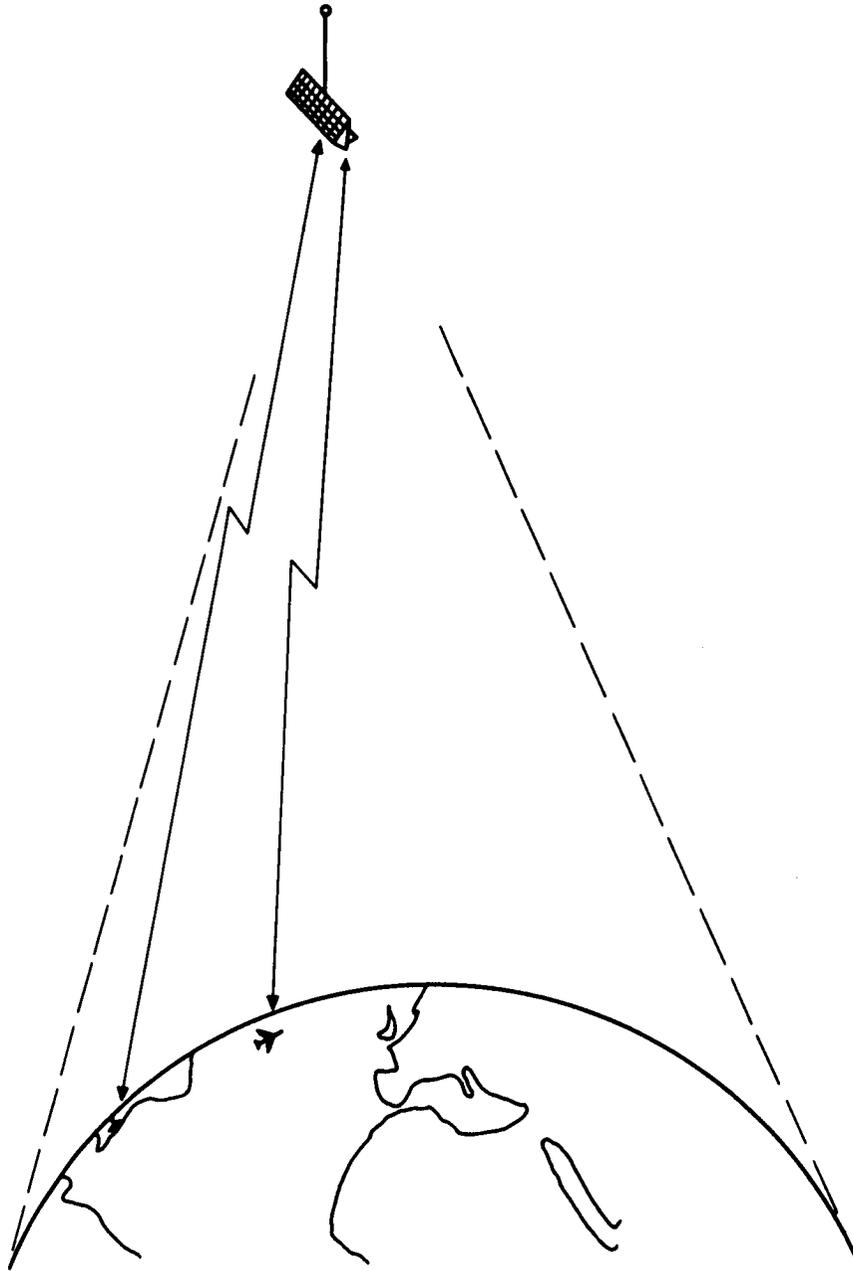
Figure I-1. Satellite Communication Geometry

The transmission link from the user craft to the satellite will usually have a higher signal-to-noise ratio than the link from the satellite to the user, because the transmitter power aboard the user craft can be higher than the satellite transmitter power.

Economic considerations have an important influence on satellite and user equipment designs. A satellite mean time to failure of approximately 5 years is desirable. Reliability of user equipment is also important. While preventative and repair maintenance can be performed, the revenue lost by taking an

aircraft or ship out of service for maintenance must be considered in the maintenance cost. The cost and ease of retrofitting existing aircraft to make them serviceable in a traffic controlled environment must be taken into account in the system design.

Although the communication links between the satellites and the ground terminals require careful engineering, successful transmissions between earth stations and aircraft through the Syncom and ATS satellites have demonstrated the practical advantages of satellites for communication with mobile terminals. The results have matched predicted performance and together with data available from other satellite experiments, they serve as the basis for confident predictions that satellites can be used advantageously for navigation and position fixing.

B. PURPOSE AND ORGANIZATION OF THE REPORT

This report is limited to a consideration of position fixing by range measurements from two or more satellites. It is further restricted to measuring range by measuring the propagation time of a radio frequency signal from the satellite to the user craft and return. Propagation time is measured by introducing a time marker on the radio frequency signal and measuring the time required for the marker to travel from the satellite to the user and back to the satellite. There are several ways in which the time marker can be modulated onto the radio frequency signal. The accuracy of the range measurement does not depend upon the choice of the modulation technique. In theory the energy required for the measurement is very nearly the same for every modulation technique. In practice they differ widely in their means of implementation, and have a greater influence on system design than any other major consideration.

The four transmission links up and down between the earth station and the satellite and down and up between the satellite and the user craft have such different characteristics that the best choice of modulation technique for one link may not always be the best choice for another. The equipment complexity necessary to change from one form of modulation to another must be assessed against the equipment complexity without the change in technique and also against the benefits of such a change.

Section II of the report deals with requirements, constraints and parameters common to all of the modulation techniques. It serves as the basis for the analysis of each technique.

Section III presents the single pulse ranging technique that has been used for many years in simple radars. It is included as a basis for comparing the various modulation techniques. As an aid in making the comparisons, a power budget table and graph are included showing the satellite to user path transmitter power required for range measurements over a wide range of frequencies and system parameter values.

Section IV is devoted to pulse train ranging. Pulse train ranging differs from single pulse ranging in that a coded sequence of pulses is transmitted instead of the single pulse. Section IV presents side tone modulation, including a side tone-single pulse hybrid implementation.

Several appendixes are included to provide additional detail. Appendix A describes a multiple access approach for a large number of users; Appendices B, G and H present additional information beyond that discussed in Section IV regarding the properties and applications of pulse trains.

A number of modulation techniques were studied in detail. Several of them proved to have serious limitations for the application. The analyses made during the study and the resulting evaluations of these techniques are also included as appendices as follows:

Appendix C—FM Pulse Compression

Appendix D—FM-CW Triangular Modulation

Appendix E—Pseudo-Random Code Techniques

Recognizing the voluminous studies that have been accomplished on Pseudo Random Codes, Appendix E examines those aspects of the technique that are unique to this application. The interleaving of ranging information and data and voice information is analyzed. Also a detailed analysis of various techniques to improve the acquisition properties of pseudo random code is presented.

Appendix F develops the evaluation equations for the hybrid assymetrical spectrum presented in Section V.

Appendix J presents a list of symbols employed in each section of the report.

C. CONCLUSIONS

Range measurement accuracy and the energy required for measurement are theoretically very nearly the same for all of the modulation techniques studied. The techniques differ considerably in equipment complexity and compatibility with existing communications equipment and procedures. They differ also in acquisition time, ease of addressing large numbers of users, and the methods for resolving ambiguity. As a result of these differences, the techniques vary widely in the efficiency of their use of radio frequency energy. Specific conclusions about each of the modulation techniques are summarized below.

- Single Pulse Modulation

    1. The concept is simple and performance is easily predicted.
    2. User equipment is simple, but it is difficult to equip the satellite for transmitting the required high peak power.
    3. It is highly efficient in the use of radio frequency energy.
    4. It is efficient in the use of the measurement time available because ambiguities do not have to be resolved.
    5. Some loss of efficiency results because address codes do not contribute to the measuring process.
    6. It is compatible with existing AM receivers.
    7. It requires high peak power.

- Multiple Pulse or Pulse Train Modulation

    1. Each pulse in a multiple pulse address code can contribute to the measurement accuracy.
    2. The basic transmitting and receiving equipment are the same as for single pulse ranging; however, additional pulse processing equipment is necessary.
    3. It is highly efficient in the use of radio frequency energy, because the pulse train can contain the range measurement information and the user's address and can also provide the ambiguity resolution.

4. Flexibility in coding design allows a wide variation of system parameter values.

5. By comparison with a single pulse system, peak pulse power can be reduced by a factor between $\sqrt{n}$ and n times, where n is the number of the pulses in the pulse train.

6. If a stable phase control clock is used in the pulse processing equipment, the return transmission from a transponder can be delayed until after the entire train is received. A single antenna may be used for receiving and transmitting by the use of an electronic coaxial switch, rather than a diplexer. Reception and transmission may be on the same frequency.

● FM Pulse Compression

1. This is primarily a coding scheme for single pulse ranging.

2. The technique permits a wide range of trade-offs in range resolution, measurement time, bandwidth and peak power.

3. When bandwidth is limited, there is a direct ratio between peak power and range resolution.

4. The characteristics of the dispersive elements in the transmitter and receiver must be closely matched.

● FM-CW Triangular Modulation

1. Measurement time is long and the measurement must be made in several discrete steps.

2. Acquisition and address times are long.

3. Bandwidth, accuracy and ambiguity interval are inter-related.

4. The technique is inefficient in the use of radio frequency energy because of the long measurement time. In addition, ambiguity resolution requires the measurement to be made in several steps.

● Multiple Side Tone Modulation

a. Transponded FM Side Tone

1. Acquisition time on the return link from the user is so long that it can seriously reduce multiple access capability.

2. Special purpose phase locked transponders are required.

3. It comparison with other techniques, transmitted power can be low since the receiver bandwidth can be reduced to a few Hertz by means of tracking filters. This advantage is achieved at the expense of long acquisition times.

4. If a phase lock receiver is not used, relatively high transmitter power is necessary of the FM threshold effect in the receiver.

b. Hybrid AM Side Tone

This hybrid technique is described in detail in Section V. FM side tone is used from the ground to the satellite, AM side tone from the satellite to the user, and a single pulse return from the user through the satellite to the ground terminal.

1. A conventional AM receiver may be employed by the user.

2. Signal processing equipment is relatively simple.

3. Post detection filtering contributes to high efficiency in the use of satellite energy.

4. A store-and-forward technique may be implemented in the user equipment to simplify the sharing of the antenna for transmitting and receiving.

5. A large multiple access capacity can be achieved with the use of a block code multiple access concept as described in Appendix A.

6. As with all "CW" types of modulation, initial system acquisition creates an energy penalty unless the system is near time saturation.

● Pseudo-Random Codes

1. Implementation is relatively complex.

2. Acquisition times are long.

3. The technique is relatively insensitive to interference.

4. Multiple addressing of many users is comparatively simple.

5. Data, voice and digital communications may be multiplexed with range measurements.

Pulse modulation systems are efficient in their use of both energy and time. Peak power is comparitively high. However, there are a number of coding techniques that allow variation in pulse system parameter values. In general, the larger the number of pulses in the waveform, the greater is the freedom in the selection of parameter values. The equipment is not complex; however, as the number of pulses increases and the coding becomes more complex, the pulse processing equipment also becomes more complex.

Continuous wave systems usually require low power. They are not efficient in the use of energy or time because the acquisition time becomes long. Efficiency increases as the interrogation rate increases toward a 100 percent use of the channel. Implementation can be simple or complex, depending upon the trade-off between acqusition time and power.

Combination or hybrid systems, in which different modulation techniques are used on each transmission link, should be considered for the design of any system. If the processing equipment to change from one form of modulation to another is simple, the special problems of each link may be solved independently.

Multiple access capability is severely limited by the acquisition time on the links from the user to the satellite to the earth terminal. Since the range from the user to the satellite is not known, there is little a priori knowledge available for reducing acquisition time on the return signal. As a consequence, techniques that require short acquisition times are to be preferred if the user interrogation rate is high.

D. RECOMMENDATIONS

The study did not lead to a recommendation of one modulation technique that would be best for all applications. However, recommendations are made for two important applications. Multiple pulse train modulation is recommended for a navigation and air traffic control system operating in the 118 to 135 MHz VHF band. It is further recommended that a hybrid technique using side tone ranging on the links to the user and pulses on the return link be considered for L-band implementation of a navigation/air traffic control system.

Multiple pulse train ranging was found to have the following advantages for the VHF implementation:

1. The equipment can be compatible with presently used VHF aircraft transmitters and receivers.
2. User equipment cost is low.
3. The equipment may be easily retrofitted to existing aircraft.
4. It is possible to use pulsed data and voice transmissions in the range measurement process so that one channel and one set of equipment can be used for all functions.

While equipment costs may preclude fully coherent signal processing, satellite peak power can be reduced to almost the same level needed for continuous wave ranging through the use of non-coherent processing of the pulse train.

The choice of a modulation system for L-band operation (1540-1660 MHz) is not as clear cut as the choice at VHF. The trade-off between antenna effective area and beamwidth requires either higher satellite power or the ability to point the antenna toward the satellite. The current state-of-the-art will not permit a large enough increase in satellite power for non-directional antennas on the aircraft. L-band operation will require that operationally and aerodynamically acceptable antennas be developed for high performance aircraft. The trade-off between peak power and efficient use of the energy available in the satellite must be examined for L-band operation.

Communication equipment for L-band is not presently available, so that completely new designs will be required for L-band operation. When the application is defined in detail, each transmission link must be examined to ascertain the optimum use of equipment, power and spectrum. Studies directed towards an L-band navigation/traffic control system implementation should include consideration of the hybrid system described in Section V.

Frequencies lower than approximately 100 MHz are not useful because uncertainty in propagation delay and ray path bending in the ionosphere introduce such excessive range errors. Frequencies above approximately 500 MHz require directional antennas on the user craft. Such antennas introduce operational and aerodynamic design problems.

SECTION II

OPERATIONAL PARAMETERS

A. EXPECTED TRAFFIC

Aside from average satellite consumed power, the major effect of traffic density on a satellite navigation system occurs at that point when the simple "one-at-a-time" approach to measurements is exceeded. For a satellite in a 24-hour orbit, an average round trip propagation time of 0.52 second per measurement exists (earth to satellite to earth and return). Thus, for a fix of position to be made (at least two range measurements) a single user essentially ties up the system for 1.04 seconds; allowing approximately 3460 fixes per hour. This assumes that range measurements are completed sequentially. Thus, if the expected traffic requires a fix rate in excess of 3460 fixes per hour, some form of multiple access to the system will be necessary. With parallel channel operation utilizing two satellites, the fix rate without multiple access rises to 6920 per hour.

Several studies have been conducted to prepare estimates of expected traffic for such a system. A result of the "Study of Satellites for Navigation" (NASw-740) predicted traffic similar to that shown as curve A in Figure II-1. Since that time, development has begun on the large or "jumbo" jet aircraft and the super-sonic transport (SST) aircraft. Certainly these aircraft are going to make a large modification in the air traffic market.

A British Ministry of Aviation study, including the effects of the large jets and SST's, has shown an expected peak of aircraft in the air in the late 1970's. Barring a presently unforeseen large increase in the air traffic volume, an actual reduction in numbers of aircraft is anticipated after this due to the gradual take-over of SST's. To relate this study to the NASw 740 study, a fix rate for the SST aircraft of three times that of a subsonic jet was assumed. This assumption was based mainly on an expected air traffic control desire to know an aircraft's position (regardless of type of aircraft) within a certain maximum uncertainty. This assumption, along with the British Ministry of Aviation prediction, was the basis of the additional curve (B) in Figure II-1 which incorporates the effect of the large jets and SST's.

It should be noted that these predicted traffic curves are for the North Atlantic which is used as a model for this modulation study.

The significance of these traffic predictions to the modulation study is that they greatly exceed the 3460 or 6920 fix rate per hour that is the threshold of a multiple access requirement. Thus, either multiple channels or time multiple access, or both, will be needed to fully meet the system needs.
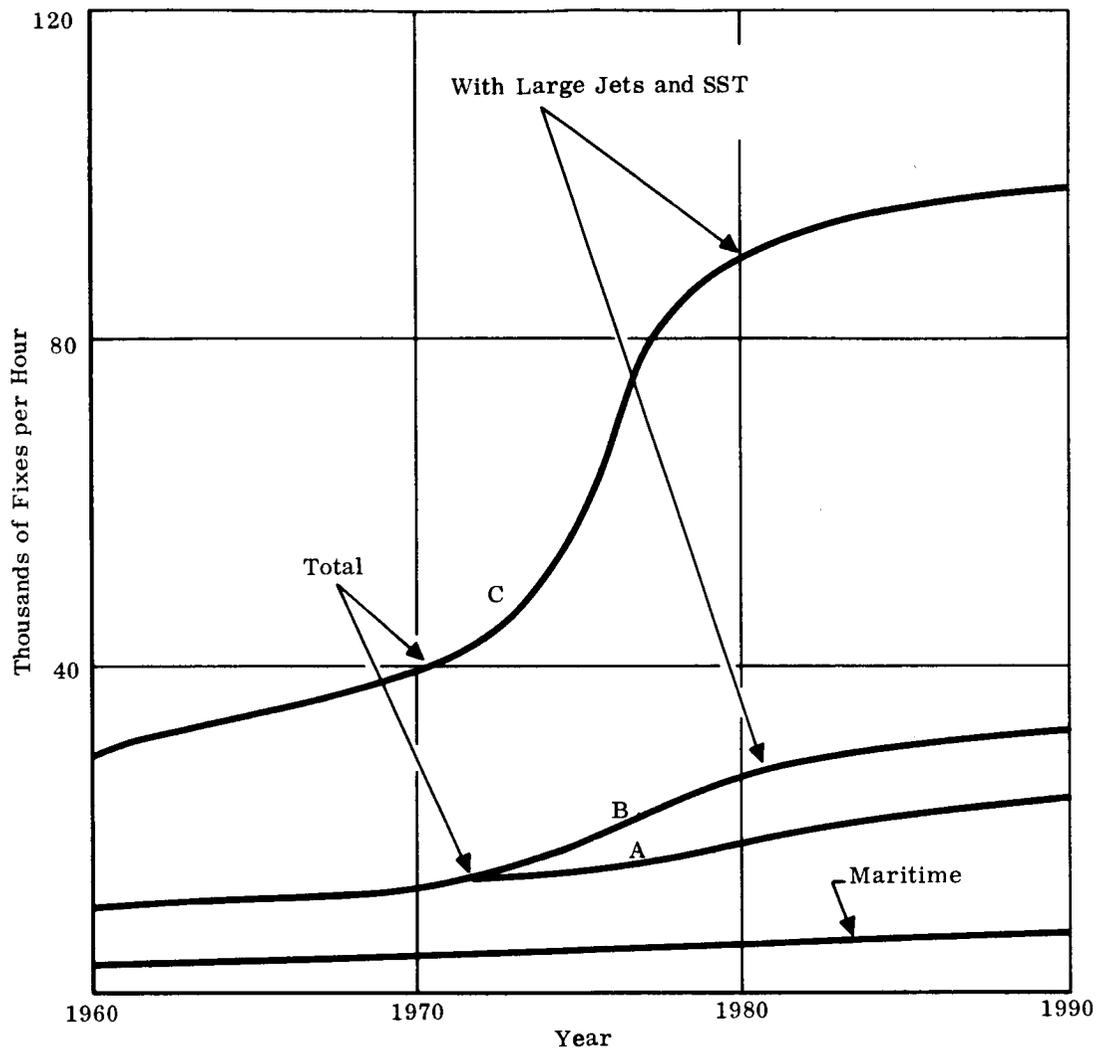
Figure II-1. Expected Traffic

Noting that the traffic model used for the NASw 740 study was based on 25 to 50 percent use by air traffic, the curves of Figure II-1 are suspect if applied to the air traffic control case. Since tight traffic control is already becoming necessary, a more likely curve for fix rate is curve C of Figure II-1. This curve is based upon a 100-percent use by air traffic. There will be times when 100-percent use will not be necessary; however, if weather is not going to seriously disrupt traffic flow, an air traffic control system must be capable of 100 percent use. The curves shown in Figure II-1 thus represent a composite of air traffic density predictions and navigation traffic ground rules. The importance of these curves lies not in their complete accuracy but in that they serve to define the bounds of an engineering study. The data ultimately can be grossly wrong in either direction and not alter the fact that a modulation technique that allows multiple access measurements will be needed.

B. FREQUENCY ALLOCATIONS

At this time it is nearly impossible to predict exactly what allocation for a navigation satellite system may be made in the radio frequency spectrum. Table I represents a survey of the current allocations indicating the frequencies that have similar usage and could be considered for the system without a greater perturbation in international allocations. Except for the VHF and low UHF bands, there is not a great amount of equipment already in use with users. This does not say, however, that the current equipment could not be used. It may well be possible to precede a present VHF receiver, for example, with a simple solid state crystal controlled converter and use the current receiver for a microwave band. This conversion to higher frequencies carries the penalty of decreased antenna capture area. As a result the power requirements will generally rise with increasing frequency. Such simple conversions are, perhaps, not available for user transmitters, though they could likely be used as driver RF sources for higher frequency transmit equipment. Thus, in studying the area of modulation techniques, a specific frequency band does not necessarily rule out the importance of compatibility with present equipment. In a like fashion, it is not mandatory that a user receive and transmit on the same band, even though that is what present VHF systems equipment usually do. Conversely, it may be desirable in some cases to have transmission and reception on the same frequency. In any event, Table II-1 indicates the likely frequency bands for a navigation satellite system.

C. GROWTH CAPABILITY

The purpose of this section is to organize an approach to evaluating the future traffic handling capacity of various modulation techniques. The use of many satellites can undoubtedly increase a given system's capacity, especially on a world wide basis. The amount of increase, per satellite, would be heavily dependent on the system implementation of the satellite, however, the actual modulation technique involved would have very little or no impact on the affect of multiple satellites. As a result, this section of the report will treat the capacity problem on a "per satellite" basis.

In general, there can be two limits to the traffic capacity of a satellite. One restriction on the system would be due to power limitations in the satellite. Another restriction is the time-bandwidth restrictions of the system.

Table II-1. Possible Allocation for Navigation Satellite Use (Table of Frequency Allocation)

Worldwide		Region 2		United States		Federal Communications Commission				
Band (Mc/s) 1	Service 2	Band (Mc/s) 3	Service 4	Band (Mc/s) 5	Allocation 6	Band (Mc/s) 7	Service 8	Class of Station 9	Frequency (Mc/s) 10	Nature of Services of Stations 11
117.975-132	Aeronautical Mobile. (R) (273) (273A)			117.975-121.975 (273) (US26) (US27) (US28) (US85)	G, NG.	117.975-121.075	Aeronautical Mobile. (R)	Aeronautical. Aircraft.	118-121.4 (NG34)	Airdrome control.
									121.5	Aeronautical mobile.
									121.6	Aeronautical search and rescue mobile; aero-nautical utility mobile.
									121.65-121.95 (NG34)	Aeronautical utility land; aeronautical utility mobile.
				121.975-123.075 (US29) (US30) (US31) (US80) (US85) (US102)	NG.	121.975-123.076	Aeronautical Mobile.	Aeronautical. Aircraft.	122.0-123.05 (NG34)	Private aircraft.
				123.075-123.575 (US32) (US33)	G, NG.	123.075-123.575	Aeronautical Mobile.	Aeronautical. Aircraft.	123.1	Flight test; aviation instructional.
									123.15	Flight test.
									123.2	Do.
									123.25	Do.
									123.3	Flight test; aviation instructional.
									123.35	Flight test.
									123.4	Do.
									123.45	Do.
									123.5	Flight test; aviation instructional.
									123.55	Flight test.
				123.575-128.825 (US26) (US85)	G, NG.	123.575-128.825	Aeronautical Mobile. (R)	Aeronautical. Aircraft.	123.6-128.8 (NG34)	Aeronautical mobile.
				128.825-132 (US2) (US85)	NG.	128.825-132	Aeronautical Mobile. (R)	Aeronautical. Aircraft.	128.85-132 (NG34)	Aeronautical mobile.
132-136		132-136	Fixed Mobile. (273A) (276)	132-136 (US2) (US26) (US85)	G, NG. (US3)	132-136	Aeronautical Mobile. (R)	Aeronautical. Aircraft.	132-05-135.05 (NG34)	Aeronautical mobile.
140.0-150.05 (285B)	Radionavigation-Satellite.			149.9-150.05 (US100)	G, NG. (US100)	149.9-150.05	Radio-navigation-Satellite.	Space.		Radionavigation-Satellite.
						156.25-157.45 (287) (US9) (US77)	Maritime Mobile. (NG5)	Coast. Ship.	156.3	Maritime mobile.
									156.35	Maritime mobile. (NG24)
									156.4	Maritime mobile. (NG24)
									156.45	Maritime mobile. (NG24)
									156.5	Maritime mobile.
									156.55	Maritime mobile. (NG24)
									156.6	Maritime mobile.
									156.65	Do.
									156.7	Do.
									156.8	Maritime mobile. (Calling and Safety)

Table II-1. Possible Allocation for Navigation Satellite Use (Table of Frequency Allocation) (Cont.)

Worldwide		Region 2		United States		Federal Communications Commission				
Band (Mc/s) 1	Service 2	Band (Mc/s) 3	Service 4	Band (Mc/s) 5	Allocation 6	Band (Mc/s) 7	Service 8	Class of Station 9	Frequency (Mc/s) 10	Nature of Services of Stations 11
									156.9 / 156.95 / 157.0 / 157.2-157.4	Maritime mobile. / Do. / Do. / Maritime mobile.
390.9-400.05	Radionavigation-Satellite. (311A)			399.9-400.05	G, NG.	399.9-400.05	Radionavigation-Satellite.	Space.		Radionavigation-Satellite.
960-1215	Aeronautical Radionavigation. (341)			960-1215	G, NG.	•960-1215	Aeronautical Radio-navigation. (341)			
1300-1350	Aeronautical Radionavigation. (346) Radiolocation.			1300-1350	G, NG. (US3)	1300-1350	Aeronautical Radio-navigation. (346)			
1540-1660	Aeronautical Radionavigation. (352A) (352B)			1540-1660	G, NG. (352A) (352B) (US29)	1540-1660	Aeronautical Radio-navigation. (346)			
2450-2550 (357)		2450-2550	Fixed Mobile. Radiolocation.	2450-2500	NG. (US41)	2450-2500 (357)	Fixed Mobile. Radiolocation.		2450	Industrial, scientific and medical equipment.
2700-2900 (366)	Aeronautical Radionavigation. (346) Radiolocation.			2700-2900	G. (346) (366) (US42) (US43)					
2900-3100	Radionavigation. (367) Radiolocation.			2900-3100	G, NG.	2900-3100	Maritime Radio-navigation. Radiolocation. (US44)			
4200-4400 (352A)	Aeronautical Radionavigation.			4200-4400	G, NG. (352A) (US47)	4200-4400	Aeronautical Radio-navigation.	Altimeter.		
5000-5250 (352A) (352B)	Aeronautical Radionavigation.			5000-5250	G, NG. (352A) (352B)	5000-5250	Aeronautical Radio-navigation.			
5350-5460	Aeronautical Radionavigation. (385) Radiolocation.			5350-5460	G, NG.	5350-5460	Aeronautical Radio-navigation. Radiolocation. (US48)			
5460-5470	Radionavigation. (385) Radiolocation.			5460-5470	G, NG.	5460-5470	Radio-navigation. (385) (US65) Radiolocation. (US69)			
5470-5650 (387)	Maritime Radio-navigation. Radiolocation.			5470-5600	G, NG.	5470-5600	Maritime Radio-navigation. (US65) Radiolocation. (US50)			
				5600-5650	G, NG.	5600-5650	Maritime Radio-navigation. (US65) Meteorological Aids. (387) Radiolocation. (US51)			
9000-9200	Aeronautical Radionavigation. (346) Radiolocation.			9000-9200	G, NG.	9000-9200	Aeronautical Radio-navigation. (316) (US54) Radiolocation. (US55)			

To assess the power limitations of the satellite requires a survey of the state-of-the-art in long life satellite power sources. The results of such a survey have indicated that through the next decade, solar cell power sources will most likely remain superior in cost over other systems. For solar cell power sources, typical available powers are summarized in Table II-2.

Table II-2

Typical Available Powers (Synchronous Satellites)

	Satellites/Launch Vehicle	Satellite (weight)	Average Power Available
Gravity Gradient and Flywheel Stabilization	1/Atlas Agena	950 Pounds	10^3 watts—4 years
	2/Atlas Agena	425 Pounds (each)	4×10^2 watts—4 years
Spin Stabilization	Syncom	76 Pounds	28 watts—1 year
	Spinning Global Comsat	250 Pounds	40 watts

A secondary limitation on available power is the earth's shadow interruptions of the input power. These interruptions require the satellite to carry sufficient batteries to handle the system energy requirements during the shadowed time. The solar cell input power must be sufficiently above the system energy requirements during the illuminated portions of the orbit to recharge the batteries.

For a worst case shadow, where the satellite's orbital plane is not inclined with respect to the earth-sun axis, the shadow time can be evaluated as pictured in Figure II-2.
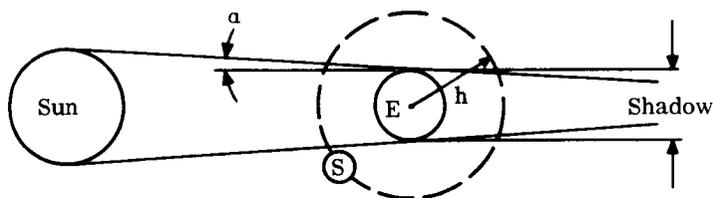
Figure II-2. Earth's Shadow Geometry

The time a satellite spends in the earth's shadow is then:

$$S_T = \frac{(R_e - h \sin\sqrt{h})}{\sqrt{R_e^2 \, g_e}}$$

(2-1)

where:

R_e = radius of the earth

h = satellite altitude (from center of the earth)

g_e = gravitational constant of the earth

$$S_T = (0.458 - 0.615 \times 10^{-6} h) \sqrt{h} \text{ minutes} \qquad (2-2)$$

for h in nautical miles.

Thus, for a geostationary satellite, the maximum eclipse time is 67.4 minutes, occurring once a day near the equinoxes. The satellite is not eclipsed near mid-winter and mid-summer. A medium altitude satellite (6-hour period) is eclipsed for 43 minutes, four times a day. These times, then, represent the worst case for "power off" operation. The peak energy requirements during these dark periods will determine the battery storage requirements for the payload power.

An evaluation of the capacity limit, set by time and bandwidth, requires an extensive study of measurement geometry, channel availability and time delays. If there is no time interleaving of interrogations, the round trip radio propagation time sets a capacity limit for high altitude satellites. A capacity limit due to propagation time may be defined as:

$$C_T = \frac{1}{T_P} \qquad (2-3)$$

where T_P is the round trip propagation time.

In the instance of satellites where this propagation time is long enough to be limiting, a possible way to avoid the full penalty is to intermix several messages during the propagation time. The limit of this mixing is:

$$M = \frac{T_P}{T_M + T_A + T_D + T_G} \qquad (2-4)$$

where

T_M = message length (in time)

T_A = acquisition times that may be necessary for signal reception

T_D = data gathering time (may be greater than T_M if message integration is necessary)

T_G = time guard band between messages.

Actually it does not make sense to intermix fractional messages so this factor may be further defined as some $K \leq M$ where K is constrained to be an integer.

Another factor of importance is the efficiency of a technique for gathering data which may be defined as:

$$\eta = \frac{T_M + T_A}{T_M + T_A + T_D}$$ (2-5)

As a function of these various factors the limit (due to time) of a modulation technique is defined as:

$$C = K \cdot C_T \cdot \eta .$$ (2-6)

Of course, for multiple channel operation this value would be multiplied accordingly. Since the factor C_T is independent of the modulation technique, a time figure of merit for the growth capacity of a modulation technique can be defined as:

$$F_T = K \cdot \eta .$$ (2-7)

A possible implementation of multiple access that takes advantage of interleaving of messages is described in Appendix A.

Thus, the various modulation techniques can be compared (with respect to growth limit) on the basis of two factors:

    1. Power efficiency with respect to both average and shadow power availability

    2. Time figure of merit.

D. ORBIT PARAMETERS

One of the major effects, aside from the great propagation distances involved, of the orbit used for the satellite is the geometrical degradation of the range measurements. As pointed out in the analysis in "Study of Satellites for Navigation" (NASw-740), a simple range measurement from one satellite fixes a user somewhere along a circle on the earth's surface. Ranging from two satellites yields two such circles which usually intersect in two areas. For this system, some other navigation means (dead reckoning, sun or star fixes, etc.) is needed to resolve the intersection ambiguity.

The size of the intersection area is a variable dependent upon the range measurement error, satellite position errors and user position with respect to the sub-satellite point (that point on the earth's surface directly beneath the satellite). Analyses were performed on NASw-740 that were based upon 100-foot errors in range measurements and in satellite position. For the higher frequency bands considered, these assumptions were reasonable. However, when the lower frequencies are considered, the available bandwidth is constrained and the propagation errors due to the ionosphere increase. These factors can force the assumed 100-foot ranging errors to increase greatly—perhaps to as high as 3500 to 4000 feet (as examined in the propagation time delay effects section (II-E) of this report).

Since error results may not simply scale accordingly to the change in range measurement accuracy, a computer analysis (pages II-17 and II-18) has been run to determine the effects of the measurement geometry and the range measurement errors on the user position errors. Figure II-3 is a graph which illustrates the geometrical effects. For example, for a user 76 nautical miles (140.8 KM) from the sub-satellite point of a geostationary satellite, a unit range measurement error would result in a 54 unit error in the calculated user position. However, as the user position becomes far from the sub-satellite point, the geometrical degradation becomes much less (reducing by an order of magnitude when the user is 550 nautical miles (1019 KM) from the sub-satellite point). The graph of Figure II-3 does not assume any position errors in the satellite.

To evaluate, more completely, the expected fix accuracies of the system, another computer program (pages II-11 to II-18) was run. This program was set up for either two or three satellites, with complete freedom as error values and positions of the users and satellites. Five locations were considered:

1. New York to London great circle midpoint
2. New York Coastal area
3. Mid-Atlantic near equator
4. Just south of Canary Island
5. Mid-North Atlantic

As a point of interest the data also shows the effects of altitude errors. As is to be expected, the maximum errors occur near the equator as the user approaches the mid-point of the two satellites. Also this analysis includes in the error figure, the calculated altitude errors. As such, the fix "on the ground" error would be less than calculated. This analysis presents the data as the root-sum-square of the three errors of a rectangular coordinate set centered at the center of the earth.

While this analysis does not include all error sources, nor does it systematically include all coverage areas, it does indicate the relative fix accuracy available from the results indicate that a range measurement error of ±1000 feet (304.8 meters) is sufficient to provide a fix accuracy of one nautical mile throughout a usefully large coverage area.

E.   PROPAGATION TIME DELAY EFFECTS

An experimental investigation of propagation effects is presently being conducted at the GE Radio Optical Observatory near Schenectady, New York. This investigation employs signals from the GEOS I Satellite which is in an orbit inclined 59 degrees to the equator with a perigee of 1115 KM and an apogee of 2275 KM. Passes over Schenectady include all azimuth and elevation angles at changing times of day. Passes have a maximum period of approximately 20 minutes, so that an essentially stationary ionosphere is scanned.
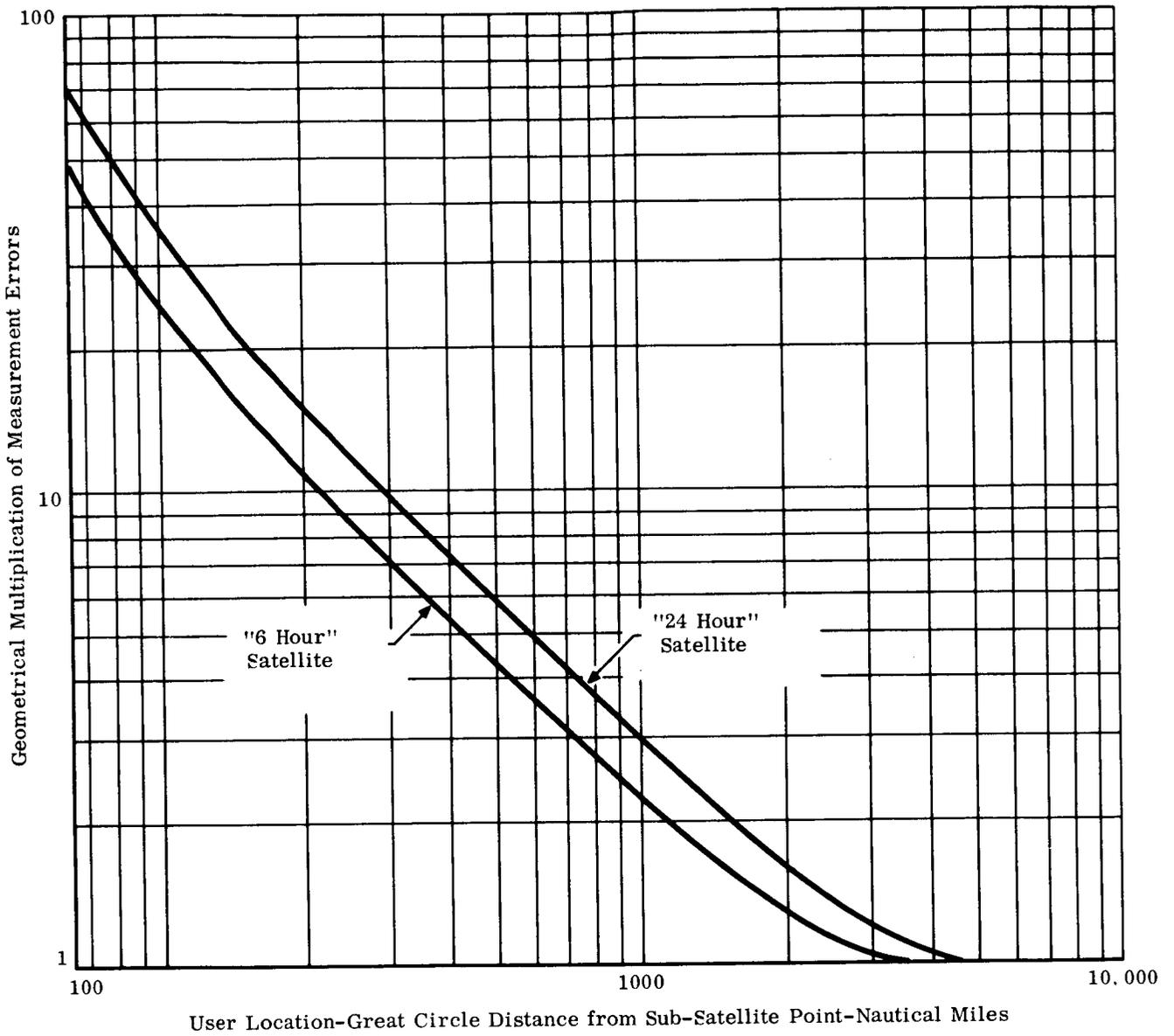
Figure II-3. Geometrical Multiplication of Measurement Errors

CALCULATION OF GEOMETRY EFFECTS ON MEASUREMENT ERRORS FOR RANGING FROM A SATELLITE TO THE EARTH'S SURFACE

FOR A SATELLITE ALTITUDE (NAUTICAL MILES) OF ? 19327

RANGE (NM)	DISTANCE (NM)	ERROR RATIO
19328	76.4256	54.0429
19337	241.751	12.1108
19347	342.003	8.56423
19357	419.007	6.9985
19367	483.99	6.06661
19377	541.299	5.43147
19387	593.163	4.96316
19397	640.905	4.5996
19407	685.387	4.30686
19417	727.208	4.06463
19427	766.803	3.85995
19437	804.502	3.68405
19447	840.559	3.53078
19457	875.177	3.39571
19467	908.522	3.27551
19477	940.728	3.16766
19487	971.91	3.07021
19497	1002.16	2.98157
19507	1031.57	2.90053
19517	1060.19	2.82604
19527	1088.11	2.7573
19627	1337.21	2.27457
19727	1549.41	1.99037
19827	1738.32	1.79897
19927	1910.92	1.65967
20027	2071.34	1.55305
20127	2222.26	1.46851
20227	2365.56	1.3997
20327	2502.59	1.34259
20427	2634.38	1.29446
20527	2761.72	1.25342
20627	2885.25	1.21808
20727	3005.49	1.18742
20827	3122.86	1.16067
20927	3237.73	1.13723
21027	3350.4	1.11662
21127	3461.14	1.09848
21227	3570.17	1.08249
21327	3677.7	1.06841
21427	3783.91	1.05604
21527	3888.96	1.04521
21627	3993.	1.03579
21727	4096.16	1.02765
21827	4198.57	1.02069
21927	4300.34	1.01484
22027	4401.57	1.01004
22127	4502.38	1.00622
22227	4602.85	1.00334
22327	4703.08	1.00136
22427	4803.15	1.00026
22527	4993.16	1.00002

CALCULATION OF GEOMETRY EFFECTS ON MEASUREMENT ERRORS FOR
RANGING FROM A SATELLITE TO THE EARTH'S SURFACE

FOR A SATELLITE ALTITUDE (NAUTICAL MILES) OF ? 5595

RANGE (NM)	DISTANCE (NM)	ERROR RATIO
5596	65.2765	46.1602
5605	206.533	10.3519
5615	292.256	7.32613
5625	358.153	5.99141
5635	413.806	5.19768
5645	462.926	4.65712
5655	507.412	4.25887
5665	548.395	3.94992
5675	586.609	3.70138
5685	622.564	3.49589
5695	656.631	3.32236
5705	689.092	3.17337
5715	720.162	3.04367
5725	750.016	2.92943
5735	778.793	2.82788
5745	806.608	2.73682
5755	833.558	2.6546
5765	859.725	2.5799
5775	885.178	2.51164
5785	909.976	2.44896
7795	934.173	2.39116
5895	1150.97	1.98694
5995	1336.98	1.751
6095	1503.73	1.5935
6195	1657.12	1.47995
6295	1800.63	1.39386
6395	1936.51	1.32629
6495	2066.32	1.27189
6595	2191.21	1.22726
6695	2312.03	1.19012
6795	2429.43	1.15887
6895	2543.96	1.13237
6995	2656.03	1.10975
7095	2766.02	1.09038
7195	2874.2	1.07375
7295	2980.85	1.05948
7395	3086.17	1.04726
7495	3190.36	1.03683
7595	3293.59	1.02801
7695	3396.01	1.02061
7795	3497.75	1.0145
7895	3598.95	1.00958
7995	3699.7	1.00576
8095	3800.13	1.00295
8195	3900.33	1.00109
8295	4000.38	1.00015
8395	4100.39	1.00007

WHERE*   13:49   S1 THU 02/23/67

IF YOU WISH THE USER COVARIANCE MATRIX PRINTED TYPE 1,
IF NOT TYPE 0? 0

HOW MANY SATELLITES ARE THERE IN THE SYSTEM? 3
WHAT IS LATITUDE [DEG], LONGITUDE [DEG] AND ALTITUDE [NM]
OF EACH SATELLITE
? 0,40,19327,0,20,19327,15,20,19327
WHAT ARE THE SATELLITE POSITION ERRORS [FT]
? 100,100,100,100,100,100,100,100,100
WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 54,27,35E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,500

RSS POSITION ERROR [NM] .583697   [1.08101 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 40,75,10E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,500

RSS POSITION ERROR [NM] .537737   [.995889 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 1,34,40E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,500

RSS POSITION ERROR [NM] .935353   [1.73227 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

```
WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 25,25,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,500


RSS POSITION ERROR [NM] .662812                    [ 1.22753 KM ]


DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1


WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 35,45,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,500


RSS POSITION ERROR [NM] .538588                    [ .997465 KM ]


DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1


WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 54,27,35E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,50


RSS POSITION ERROR [NM] .550575                    [ 1.01967 KM ]


DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1


WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 40,75,10E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,50


RSS POSITION ERROR [NM] .511849                    [ .947945 KM ]


DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1
```

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 1,34,40E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,50

RSS POSITION ERROR [NM] .913301   [1.69143 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 25,25,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,50

RSS POSITION ERROR [NM] .583271   [1.08022 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 35,45,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,1000,50

RSS POSITION ERROR [NM] .493927   [.914753 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 0

TIME: 59 SECS.

WHERE*   14:04   S1 THU 02/23/67

IF YOU WISH THE USER COVARIANCE MATRIX PRINTED TYPE 1,
IF NOT TYPE 0? 0

HOW MANY SATELLITES ARE THERE IN THE SYSTEM? 2
WHAT IS LATITUDE [DEG], LONGITUDE [DEG] AND ALTITUDE [NM]
OF EACH SATELLITE
? 0,40,19327,0,20,19327
WHAT ARE THE SATELLITE POSITION ERRORS [FT]
? 100,100,100,100,100,100
WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 54,27,35E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,500

RSS POSITION ERROR [NM] .659525   [1.22144 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 40,75,10E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,500

RSS POSITION ERROR [NM] .855408   [1.58422 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 1,34,40E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,500

RSS POSITION ERROR [NM] 8.44126   [15.6332 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 25,25,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,500

RSS POSITION ERROR [NM] .702258   [1.30058 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 35,45,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,500

RSS POSITION ERROR [NM] .698266   [1.29319 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 54,27,35E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,50

RSS POSITION ERROR [NM] .635435   [1.17683 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 40,75,10E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,50

RSS POSITION ERROR [NM] .828499   [1.53438 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 1,34,40E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,50

RSS POSITION ERROR [NM] 6.26164   [11.5966 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 25,25,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,50

RSS POSITION ERROR [NM] .643123   [1.19106 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 1

WHAT IS USER'S LATITUDE [DEG],LONGITUDE [DEG],AND ALTITUDE[FT]
? 35,45,30E3
WHAT ARE THE RANGE AND ALTITUDE MEASUREMENT ERRORS [FT]
? 1000,1000,50

RSS POSITION ERROR [NM] .66156   [1.22521 KM]

DO YOU WISH ANOTHER USER POSITION? IF SO, TYPE 1;
IF NOT, TYPE 0? 0

TIME: 21 SECS.

The satellite transmits very stable, phase coherent signals at 162 and 324 MHz. These are received on highly stable phase-locked receivers at the Observatory. A 30-foot steerable antenna is used to receive the signals so that the data is not corrupted by noise. The voltage controlled oscillators of the two-phase locked receivers are on the same frequency. If there were no ionosphere, these oscillators would have a constant phase difference between them when locked to the coherent signals from the satellite—i.e., doppler shift is cancelled by this process. At a point in the frequency multiplying chains, where both are at 162 MHz, the signals are applied to a phase comparator and thence to a paper chart recorder. The integral of the phase difference is also recorded. One cycle of phase difference is recorded each time the path delay difference changes by the period of one cycle at 162 MHz (6.07 feet). A portion of a recording is shown in Figure II-4.

Plots of integrated phase difference for three passes are shown in Figures II-5, II-6, and II-7. The first indicates that the ionosphere was uniform in its structure, the second indicates a very irregular structure, and the third shows an interesting phase reversal that was observed frequently but only within narrow azimuth and elevation limits. The reason for the phase reversals have not yet been determined, but it has been noted that they occur when the ray path is through or along the edge of the auroral zone.

The phase difference recordings provide a direct and precise measurement of the difference in path length change at 162 and 324 MHz. Assuming that the effects are proportional to $1/f^2$, a close estimate of the path length change can be made at 162 and, similarly, at 118 to 136 MHz.

Data from 36 passes, taken at various times of data were plotted to determine the path length change due to the ionosphere for each 5-degree change in elevation angle. The results are summarized in Figures II-8 through II-11, which can be used to determine the mean and the ninety percentile values at any elevation angle.

The values in Figures II-8 to II-11 are not to be taken as total range error due to the ionosphere, however, as they do not include the error at the zenith. In addition to phase measurements, highly accurate, high time-resolution measurements were made at both frequencies through the passes. It is believed that these data can be used to determine the zenith angle error for each pass. But the data analysis has not yet been completed to provide that result.

The zenith angle range error can be estimated as follows:

$$\Delta R_g = \frac{40 \times 10^6}{f^2} \int_0^s Ndh .$$

(2-8)

The range of values for the integral is from 10^{12} to 10^{14}, depending on diurnal, seasonal, and solar activity changes. Conditions of the ionosphere are known at any time, so that the range uncertainty can be reduced to much less than the 50-to-5000 feet total variation of ΔR_g at VHF.
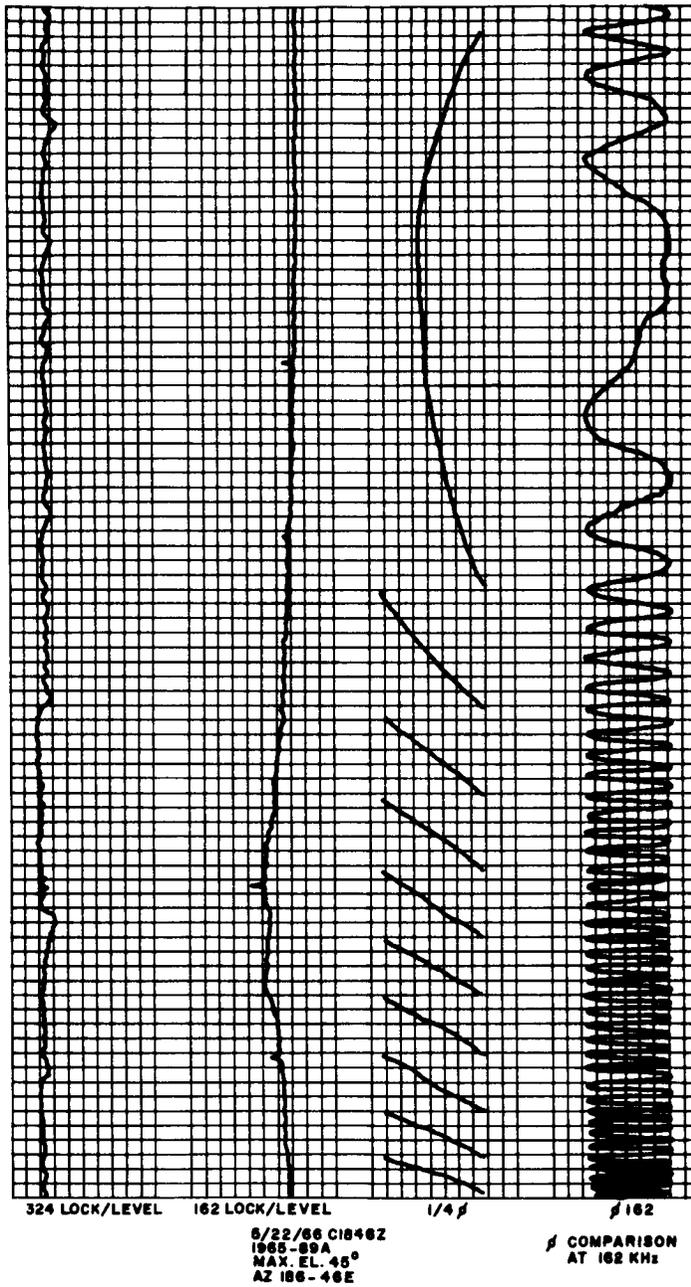
Figure II-4. Recording of GEOS I Satellite Signal
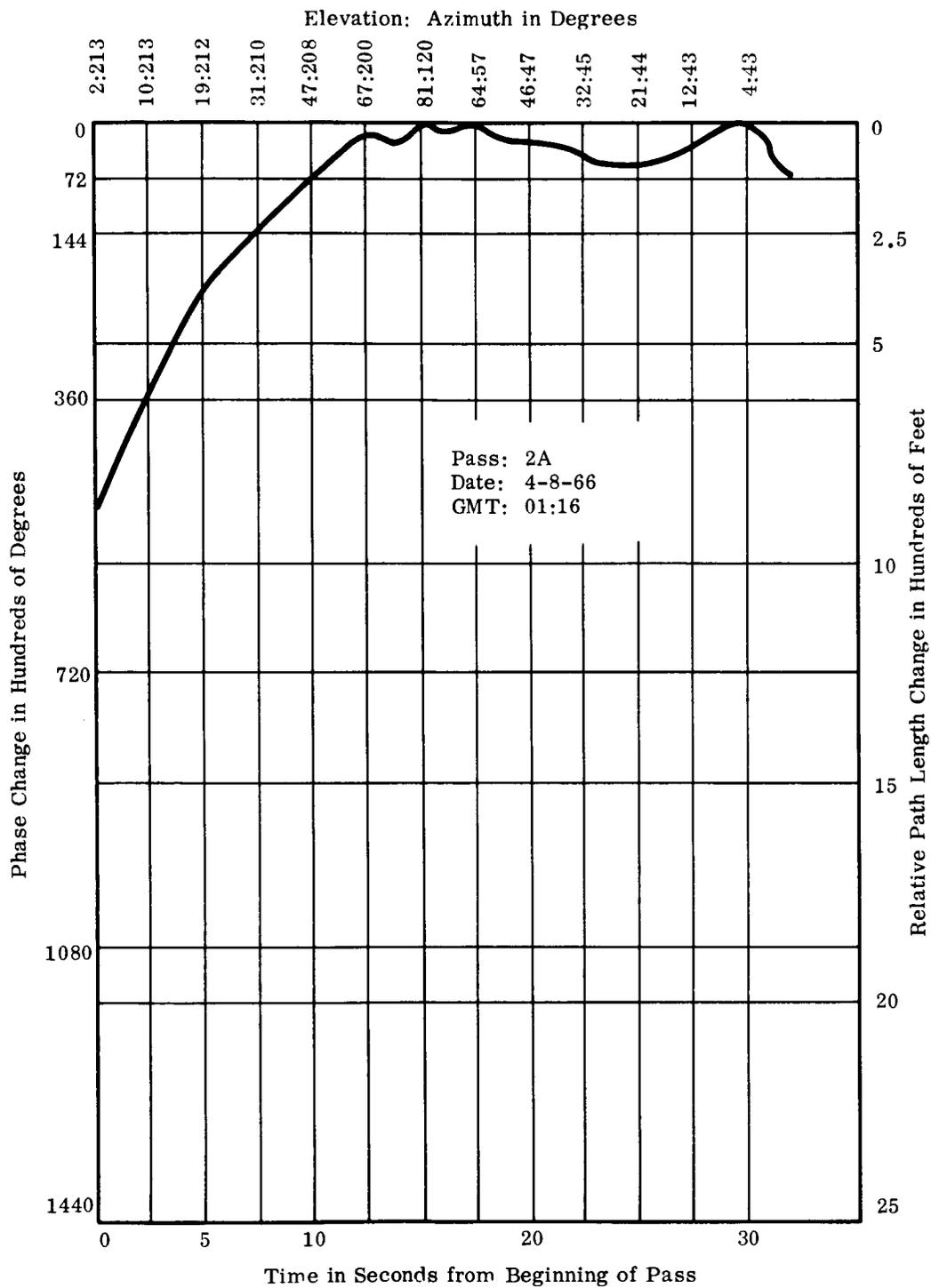
Figure II-5. Plot of Integrated Phase Difference—Pass 2A
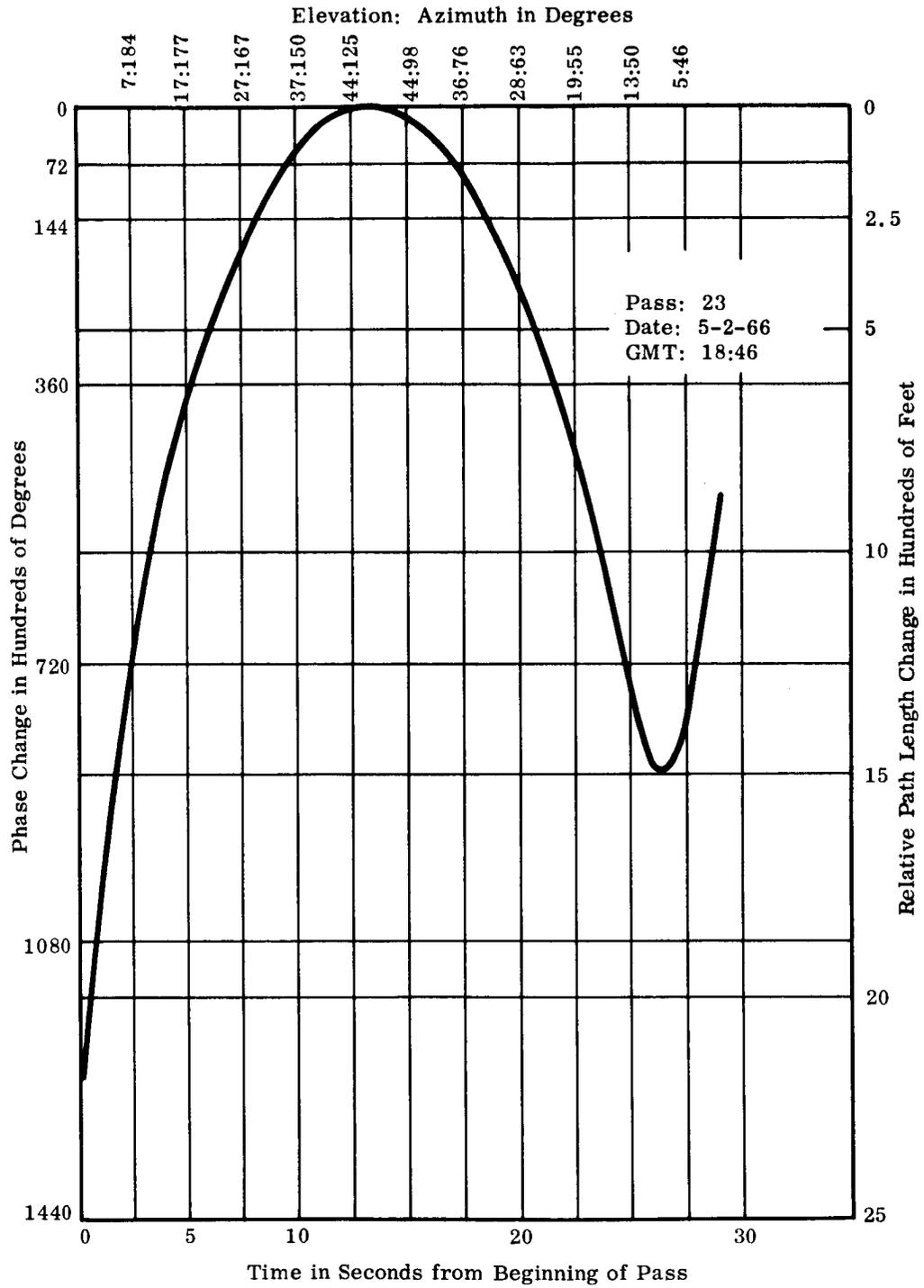
Figure II-6. Plot of Integrated Phase Difference—Pass 23
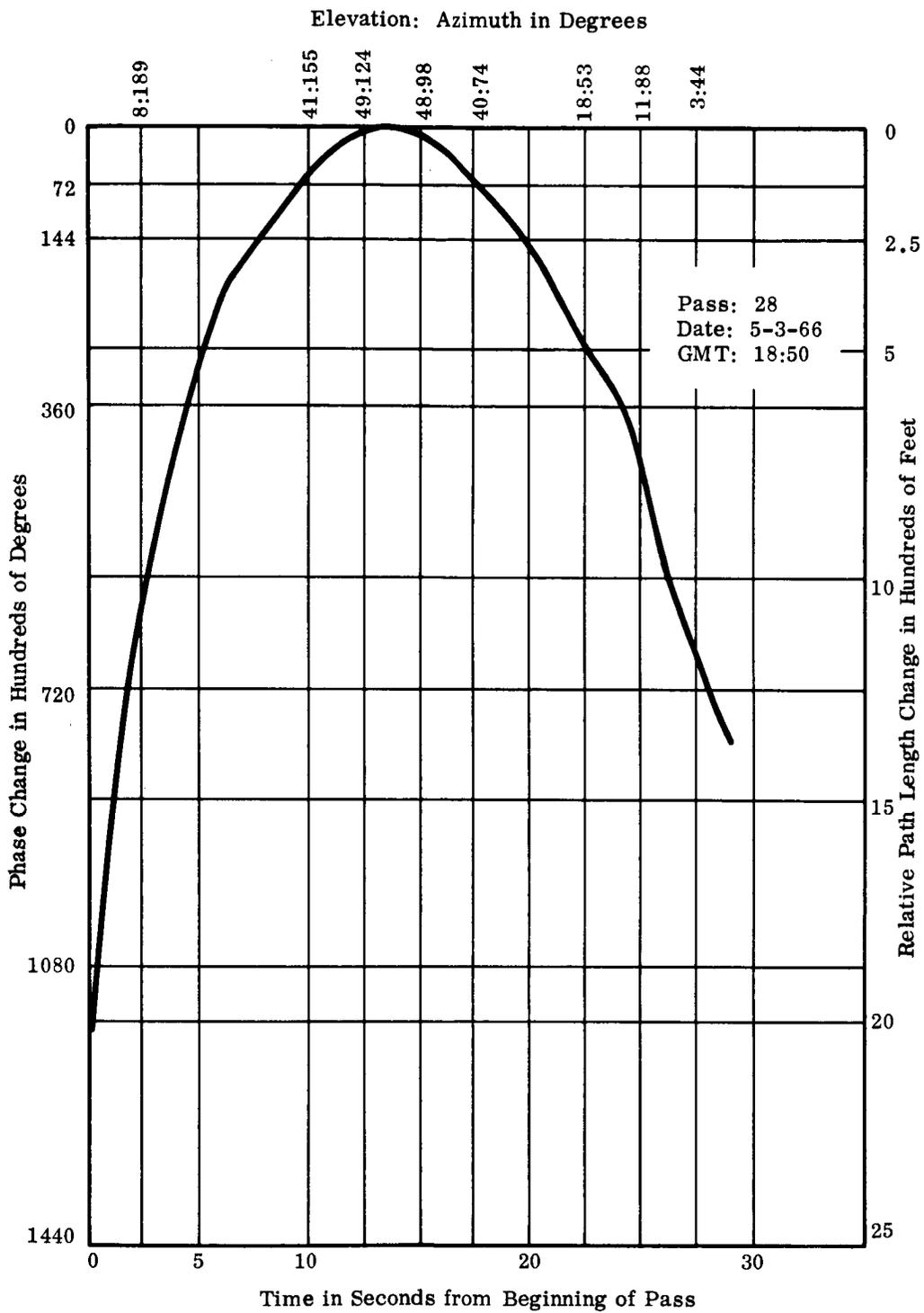
Figure II-7. Plot of Integrated Phase Difference—Pass 28
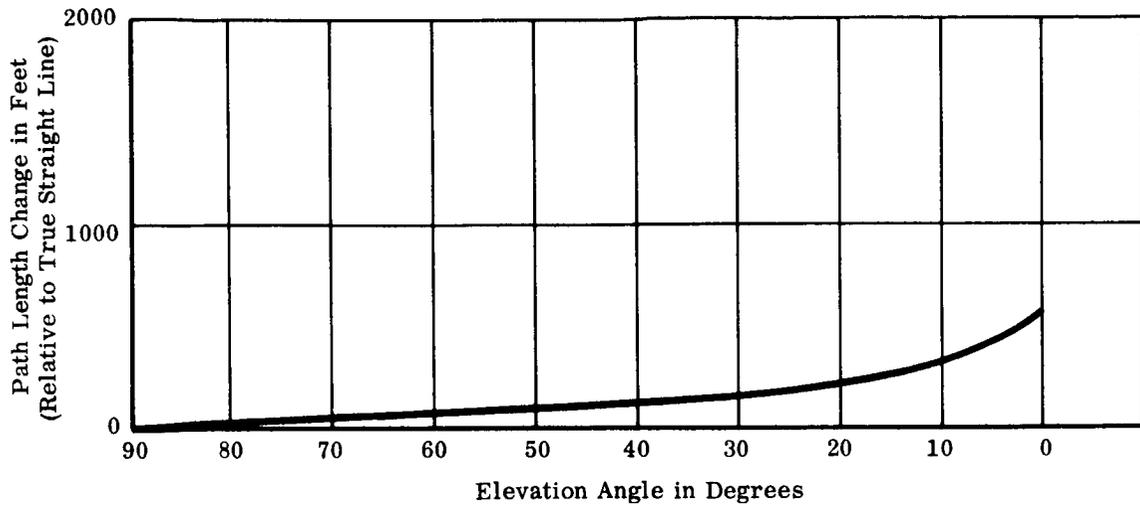
Figure II-8. Total Path Length Change Accumulated from 90° Elevation (GMT:05:00-11:00)
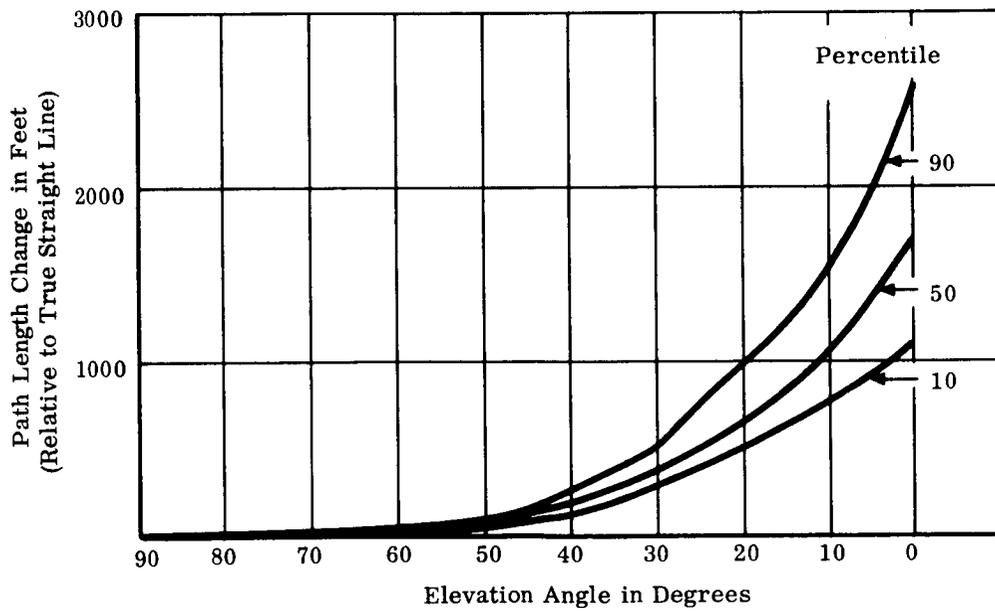
Figure II-9. Total Path Length Change Accumulated from 90° Elevation (GMT:11:00-17:00)
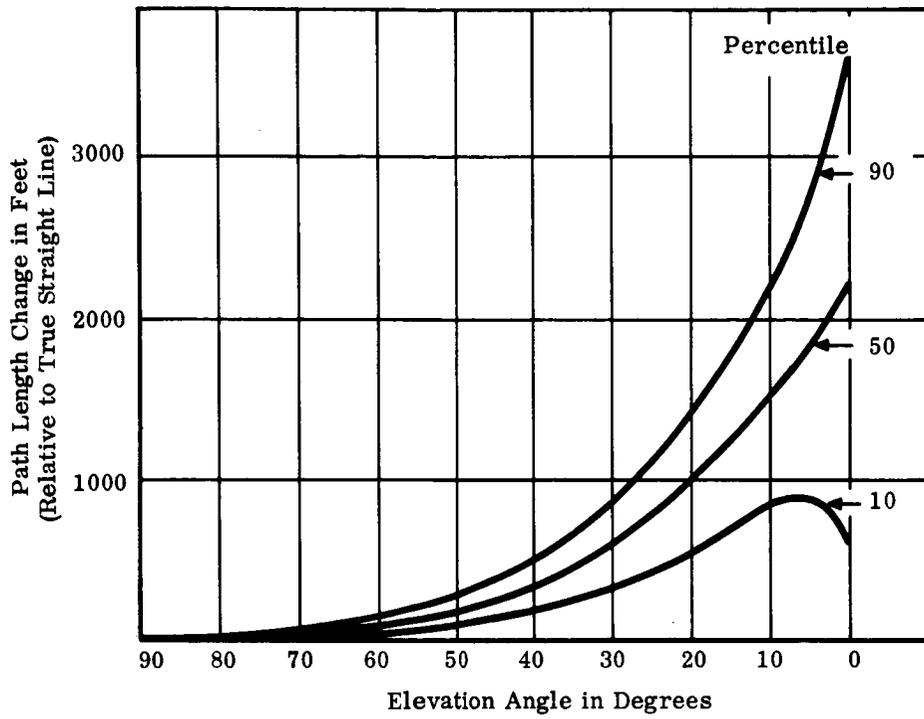
Figure II-10. Total Path Length Change Accumulated from 90° Elevation (GMT:17:00-23:00)
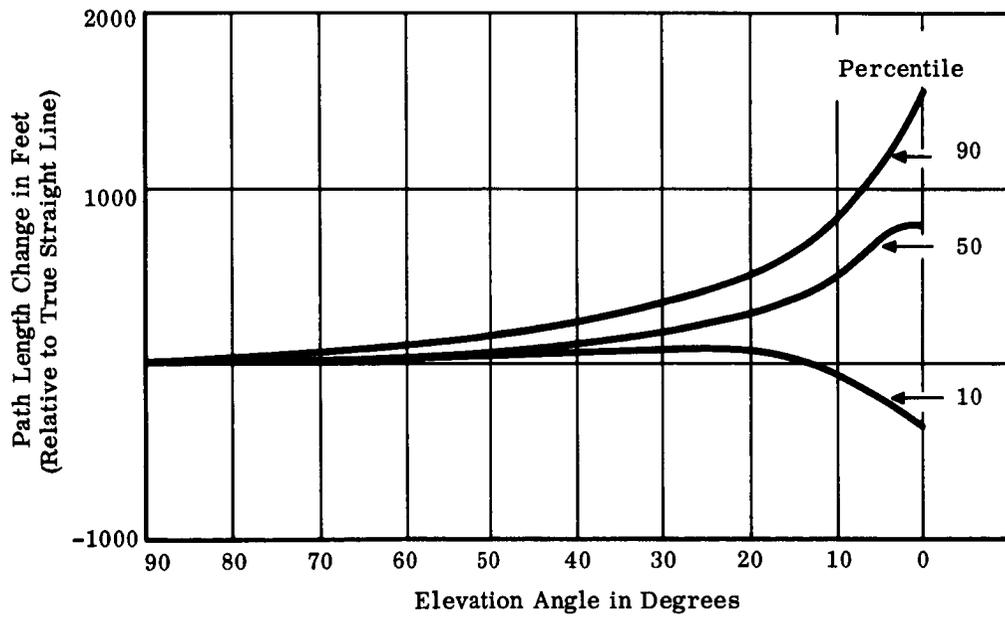
Figure II-11. Total Path Length Change Accumulated from 90° Elevation (GMT:23:00-05:00)

Total range error at VHF may be estimated by combining the propagation errors and those due to noise jitter on the ranging signals. Since each contribution is independent of the others, they may be combined as their root-mean-square to provide a one-sigma estimate of their total contribution. The total range measurement error is used in the computation of fix accuracy by applying the geometrical dilution factor for the region of interest.

Figure II-12 shows estimated position error contours determined theoretically. An independent estimate using the observatory measurements yielded a 3-sigma error of 30, 200 feet for a region of the Atlantic near the theoretical 6-nautical mile, 3-sigma contour line of Figure II-12. The estimate is based on the parameters selected for the system described in Section IV.
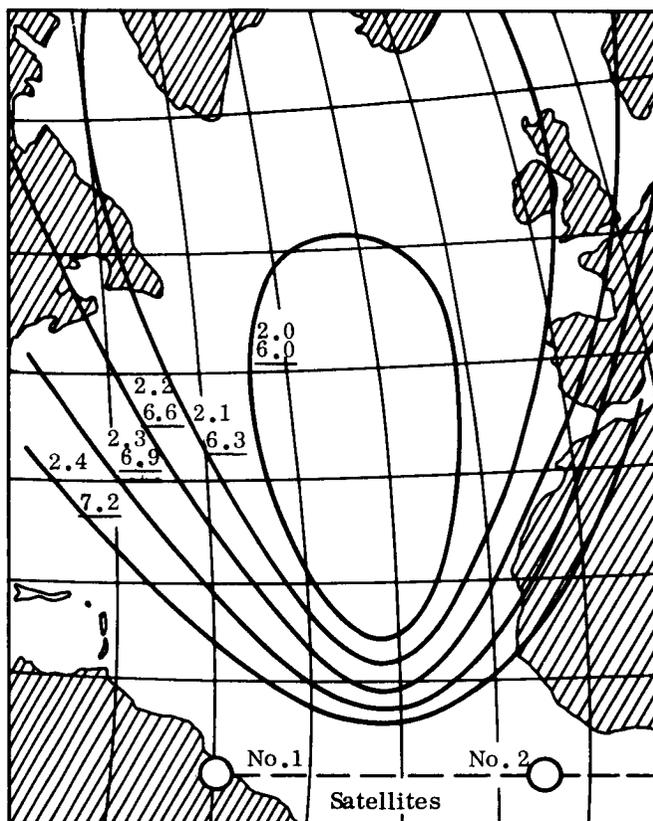
Underlined Numbers are Absolute Accuracy, 3σ
Others are Relative Accuracy, 3σ

Figure II-12. Synchronous Satellite Coverage and Accuracy (VHF Ranging System)

SECTION III

SINGLE PULSE AMPLITUDE MODULATION

A. TECHNIQUE DESCRIPTION

Single pulse amplitude modulation is the simplest concept for ranging. It is of interest as a possible method for implementation, particularly at lower frequencies, and it is of further interest because it can serve as a basis for comparing modulation methods. To aid this comparison, power budget calculations at representative frequencies and bandwidths are included in this section.

A range measurement is accomplished by first addressing the individual user with a digital address sent as a sequence of pulses. A single ranging pulse of short duration follows. The address and ranging pulse are sent from the ground station to the satellite, which repeats the sequence pulse by pulse. After the user's equipment recognizes its own address, it repeats the ranging pulse, which is again transponded by the satellite. The ground station measures the time between the two repetitions of the ranging pulse by the satellite, and thus determines the two-way propagation time between satellite and user plus the response time of the user and satellite transponders. When the known transponder response times are subtracted from measurement, it yields the range from the known position of the satellite to the user.

The pulses are transmitted by keying the transmitter "on" for the pulse duration. They are detected using envelope detectors.

B. PERFORMANCE ANALYSIS

No special requirements are placed on the orbit design. Attitude stabilization of the satellite is not required for the basic measurement since the satellite serves only as a relay for the pulsed signals. Altitude stabilization is desirable, however, to permit the highest possible gain for the satellite antenna. Maximum antenna gain is especially important with this method because it requires the highest peak effective radiated power of all the methods considered. In fact, one of the strongest reasons for considering methods other than this one is this high peak power requirement.

Envelope detection of an amplitude modulated pulse requires that its peak signal be considerably greater than the root mean square noise power in the bandwidth occupied by the pulse. Time can be resolved to a fraction of the duration of the pulse if the signal-to-noise ratio is high. Woodward, gives the following relationship of time resolution, bandwidth, and signal-to-noise ratio:

$$\delta\tau = \frac{1}{\beta(R)^{1/2}} \tag{3-1}$$

where:

β = effective bandwidth of pulse

R = ratio of received signal energy in pulse to noise density

$\delta\tau$ = standard deviation of τ (seconds).

This equation applies if the bandwidth of the receiver is matched to the bandwidth of the pulse, and if R is large. System design constraints are most likely to set limits on range resolution and bandwidth. With these constraints, the signal-to-noise ratio is determined from Equation 3-1 and a power budget drawn to determine peak transmitter power. Equipment state-of-the-art sets the limit on peak power. The intended use of the system and frequency allocation regulations determine range resolution, operating frequency, and bandwidth.

Power budgets were calculated for several postulated implementations of the method and plotted on a chart (Figure III-1) to show required transmitter power for the links between the satellite and user.

Figure III-1 relates peak power to user antenna beamwidth. Beamwidth is an important parameter in a ranging system design because it determines the magnitude of the antenna pointing problem faced by the user and also the user antenna aperture and, hence, antenna gain and physical size. Hemisphere coverage is desirable because it eliminates the pointing problem. In general, a practical upper limit for antenna gain in most mobile applications may be between 10 and 15 db; that is, between approximately 60- and 30-degree beamwidths for practical antennas.

The ordinates of Figure III-1 may be read directly in watts for single pulse ranging, where N is one. The $1/\sqrt{N}$ factor on the left hand ordinate scale gives a lower bound of the reduction in peak power for the average of a number of single pulse measurements or for a pulse train. They may be further reduced if pulse compression is used.

No method of ranging is more efficient than single pulse ranging in terms of the energy required for a range measurement, neglecting the energy required for address, if it is assumed that noise in the channel is Gaussian and that only one user is interrogated at a time. In practice, single pulse ranging may be slightly more efficient than other pulse methods because the equipment is simple. For example, a matched filter for a single pulse is the intermediate frequency amplifier of the receiver adjusted for proper bandwidth. The slightly greater efficiency that may be achieved using single pulse ranging is of minor significance. The high peak power may be an overiding disadvantage in some cases. Furthermore, energy needed for addressing users and furnishing communications is much larger than that required for the actual range measurement itself.

Efficiency of single pulse measurement is emphasized to point out that other methods may reduce peak power but they do not save energy over single pulse ranging. They trade between time, bandwidth, and peak power. They may offer some other advantages as well. For example, pulse train ranging may accomplish both the addressing and the range measurement in one pulse train. Other methods may also provide better protection against impulse noise for a given energy per interrogation.
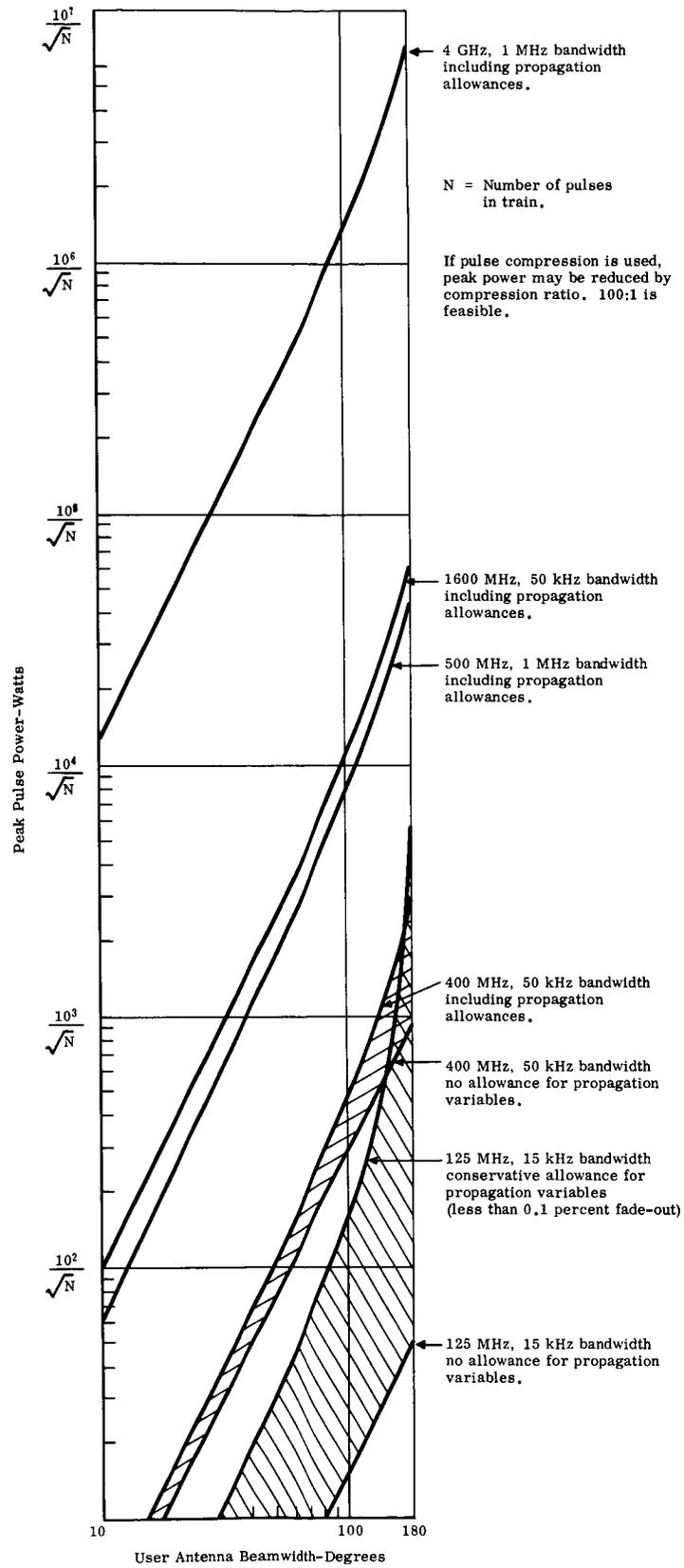
Figure III-1. Satellite Transmitter Power Requirements

Table III-1 includes power budget calculations for several representative frequencies and bandwidths. Values selected for the various parameters are reasonable values, but some are subject to wide variation depending on equipment design and propagation conditions, particularly at lower frequencies. The 125 and 400 MHz cases plotted in Figure III-1 indicate the ranges of these values at the two frequencies.

Tropospheric and ionospheric attenuation and scintillation are small most of the time, so that the 6 db value at 125 MHz is conservative, except for infrequent periods lasting from minutes to approximately an hour, when short period fading fluctuations much greater than 6 db may take place.

Multipath fading occurs when the signal reflected from the earth's surface arrives out of phase and tends to cancel the direct signal from the satellite. Reflection coefficient varies with ground conductivity, sea state and polarization. The ratio of the direct and reflected signals depends on the beamwidth and pointing direction of the user's antenna. As the user moves, he may pass through regions where the direct and reflected signals alternately add in phase and cancel, so that the received signal level varies between enhancement and degradation in comparison with the direct signal above. Under some circumstances multipath can cause complete cancellation. When a directional antenna is pointed upward toward a satellite, multipath fading may be small. For this reason, in plotting Figure III-1, it was not considered important for user antenna beamwidths smaller than approximately 90°. Although it can be important at any antenna beamwidth if the antenna is pointed close to the horizon.

Satellite antenna peak to full coverage allowance is affected both by antenna beam shape and satellite stabilization. The earth subtends 17.5 degrees at synchronous altitude. A 17 db antenna has a beamwidth of 24 degrees between its half-power points. The power budget assumed a satellite stabilization of approximately ±3 degrees so that a user who sees the satellite on the horizon would suffer the 3 db loss, if the satellite were at maximum tilt in the direction away from him.

Polarization loss is affected by Faraday rotation and antenna design. Faraday rotation can cause more than a 90° change in the polarization of a plane polarized wave at 500 MHz, and much greater rotation at lower frequencies. Signal reduction due to Faraday rotation may be avoided by using circularly polarized antennas on both satellite and user craft. Circular polarization on the satellite and linear polarization on the user craft can eliminate polarization fading, but cause a loss of 3 db in received signal. The power budget calculation assumed the use of circularly polarized antennas at both ends of the link. The -1 db allowance is for ellipticity of the polarization.

Figure III-1 extends the power budget calculations to narrower antenna bandwidths. The calculations and plots are presented without regard to whether they can be implemented. Peak power may be reduced at least in proportion to square root of the number of pulses averaged, and in proportion to the pulse compression ratio, as explained in Appendix C. Hence, if 100 pulses were averaged, and each pulse were transmitted with a 100:1 compression ratio, the peak power would be reduced by a factor approaching $\sqrt{100} \times 100 = 1000$. Such reduction requires matched filter techniques, however. Simple envelope detection of pulses far below the noise level is not possible. These considerations are discussed further in Section IV and Appendix C.

Table III-1

Power Budget

		125 MHz 15 kHz BW Note 1	125 MHz 15 kHz BW Note 2	400 MHz 50 kHz	1600 MHz 50 kHz BW	4 GHz 1 MHz BW	500 MHz 1 MHz BW
I. POWER DELIVERED TO USER RECEIVER							
Sat. Ant. Coupling Loss	-db	0.5	1.0	0.5	0.5	0.5	0.5
Path Loss (Geometric + Ant. Apert)	-db	167.0	167.0	177.0	189.0	197.0	179.0
Tropospheric + Ionospheric Attenuation and Scintillation	-db	0.0	6.0	1.0	0.0	0.0	0.0
Multipath Allowance (See Note 3)	-db	0.0	10.0	3.0	3.0	3.0	3.0
Sat. Ant. Peak to Full Coverage Allow.	-db	3.0	3.0	3.0	3.0	3.0	3.0
User Ant. Peak to Full Coverage Allow.	-db	3.0	3.0	3.0	3.0	3.0	3.0
Polarization Loss	-db	0.0	1.0	1.0	0.5	0.5	0.5
User Antenna Coupling Loss	-db	0.5	1.0	0.5	0.5	0.5	0.5
Sat. Antenna Gain	+db	17.0	17.0	17.0	17.0	17.0	17.0
DBW/Watt Radiation (into isotropic antenna)	-dbw	157.0	175.0	172.0	182.5	190.5	172.3
II. RECEIVER POWER REQUIRED							
Receiver Noise Figure	+db	5.0	5.0	3.0	5.0	5.0	3.0
Receiver Noise Temperature	°K	916.0	916.0	293.0	916.0	916.0	293.0
User Ant. Noise Temperature	°K		730.0	200.0	100.0	100.0	100.0
Total Noise Temperature	°K		1646.0	493.0	1016.0	1016.0	393.0
Total Noise Figure (db/Hz)	+db	29.6	32.1	26.9	30.0	30.0	25.9
Receiver Bandwidth Factor	+db	42.0	42.0	47.0	47.0	60.0	60.0
Noise Density/1°K (Boltzman's Constant)	-dbw	228.6	228.6	228.6	228.6	228.6	228.6
Receiver Noise Power	-dbw	157.0	154.5	154.7	151.6	138.6	142.7
Desired C/N Ratio	+db	20.0	20.0	20.0	20.0	20.0	20.0
DBW Required at User Receiver	-dbw	137.0	134.5	134.7	131.6	118.6	122.7
III. BALANCING THE BUDGET							
DBW/Watt Radiated less DBW Req. at User Receiver	dbw	20.0	40.5	37.3	50.9	71.9	49.6
User Antenna Gain		3.0	3.0	3.0	3.0	3.0	3.0
Satellite Transmitter Power	+db	17.0	37.5	34.3	47.9	68.9	46.6
Satellite Transmitter Power	watts	50.0	5600.0	2700.0	6.2×10^4	7.8×10^6	4.6×10^4

NOTES

1. No allowance was made for variable propagation effects, or for user antenna noise temperature, which are dependent on antenna beamwidth, pointing elevation, side lobes and geographical location.

2. Probable values for variable propagation effects and antenna noise temperature are included for a wide beamwidth antenna. It is assumed that all effects occur simultaneously in a direction to reduce the signal. Accurate statistical data are not available, but it is estimated that this combination of effects will occur much less than 0.1 percent of the time.

3. In preparing Figure III-1, multipath allowance stated was used for 180° user antenna beamwidth, reducing to near zero at 90° beamwidth. Narrow beam antennas aimed at low elevation angles can experience multipath fading.

Envelope detection of the amplitude modulated pulses is not affected by doppler shift of the carrier frequency.

The ultimate limit on interrogation rate is set by the differential propagation time to the various users. Interrogations may be interleaved so that several can be made in the propagation time from the ground station through satellite to user, and back. Various procedures for address coding and use of a priori knowledge of user ranges may be used to program interrogations. This was discussed in Section II-C and an example is presented in Appendix A.

Although single pulse ranging is well adapted to time division multiple access, it is not well adapted to frequency division multiple access. The bandwidth of the range measuring pulses is wide compared to some other ranging methods so that relatively small number of channels may be accommodated within a bandwidth assignment.

Growth capacity is determined by the energy required for a range measurement and by multiple access capability. The energy for the range measurement itself is small. The energy for addressing users and transmitting their positions and other information is much greater than that for the measurement, and, thus, is one of the limitations on total capacity. Multiple access capability is a limitation on growth capacity because it determines the ability to use available channel assignments through multiplexing the transmissions.

Equation 3-1 relates bandwidth, signal-to-noise ratio, and resolution. If the noise in the channel is Gaussian, white noise, the timing resolution can be a fraction of the pulse duration. For a 20 db signal-to-noise ratio, the timing resolution can be one-tenth the pulse duration. For the examples of Table III-1, a 20 db receiver output signal-to-noise ratio was assumed. Various bandwidths were used as examples. The selections were made on the basis of fixed accuracy requirements as they are now understood for various services.

The range resolutions of Table III-2 apply to each transmission link. There are a total of four links in the measurement; ground station to satellite, thence to user, back to satellite, and from there to the ground station. Because the ground station antenna can be large, thirty feet in diameter or larger, and the transmission frequency and bandwidth can both be high on the links between the ground station and satellite, range resolution can be sufficiently high that the links give a very small error contribution compared to the user-satellite links. Range jitter on the links is not correlated and hence, the total error has a one sigma standard deviation equal to the root sum square of the four link errors.

Table III-2

Bandwidth—Range Resolution

Service	Bandwidth	Pulse Duration	Prop. Time Resolution	Range Resolution
Air Traffic Control	15 kHz	0.67×10^{-4} sec.	6.7×10^{-6} sec.	6700 feet
Air and Marine Position Fixing	50 kHz	0.2×10^{-4} sec.	2.0×10^{-6} sec.	2000 feet
High Accuracy Marine Position Fixing	1 mHz	1.0×10^{-6} sec.	0.1×10^{-6} sec.	100 feet

Lapsed time for a measurement of a user's range is essentially the propagation time from the ground station to the satellite, thence to user and return; that is, approximately 0.5 second. Time for the signal to pass through the satellite and user transponders must be added to the propagation time, but it is negligibly small by comparison. For position fixing, measurements must be made from two satellites, but these may be done essentially simultaneously so that both measurements are made in the same half-second. The results of a previous study[2] indicated that the fix computation time was only about 3 ms in a large computer. Consequently, the range measurements and fix computation can all be completed in a lapsed time of approximately one-half second.

Equipment for single pulse ranging is conceptually simple. Important in the design of the user equipment is the stability and uniformity of circuit delay between reception and retransmission of the pulses. The delay must be constant and known to approximately one-tenth the duration of the pulse. It must be the same for all user equipments. It is also critical to design the pulse detector for a consistent decision on the timing of the pulse leading edge. Pulse shape must be maintained, and amplitude linearity must be maintained over a dynamic range sufficient to accommodate the full range of signal levels. An automatic gain control must be provided to set and hold the gain from one received pulse to the next with a proper time constant to accommodate changes in signal level. Fortunately, the dynamic range and gain control problems are minimized because the total change in signal level due to the users' changing range from the satellite is only about 1.5 db. While signal level changes due to atmospheric scintillation and multipath may be considerably greater at lower RF frequencies, a dynamic range of 30 to 40 db should be sufficient for the receiver. Figure III-2 shows typical block diagrams for the ground equipment, satellite repeater and user transponder.

Single pulse ranging is susceptible to impulse noise interference. Noise pulses that could be mistaken for ranging pulses could cause a disruption of the system. Means to discriminate against noise pulses that differ in amplitude or duration, or expected time of reception, can be provided to minimize the problem. But the interfering pulses are similar to the ranging pulses and if they can occur within expected reception periods, they could cause confusion in the measurements.

C. IMPLEMENTATION CONSIDERATIONS

The energy in a pulse is its peak power multiplied by its duration in seconds. For the examples calculated in Table III-1, the energies are:

Table III-3

Frequency—Pulse Energy

Frequency	BW	Peak Power (watts)	Pulse Duration (seconds)	Energy (watt-seconds)
125 mHz	15 kHz	1.00×10^3	67×10^{-6}	0.067
400 mHz	50 kHz	2.5×10^3	20×10^{-6}	0.05
1600 mHz	50 kHz	6×10^4	20×10^{-6}	1.2
4 gHz	1 mHz	7×10^6	10^{-6}	7.0
500 mHz	1 mHz	4×10^4	10^{-6}	0.04
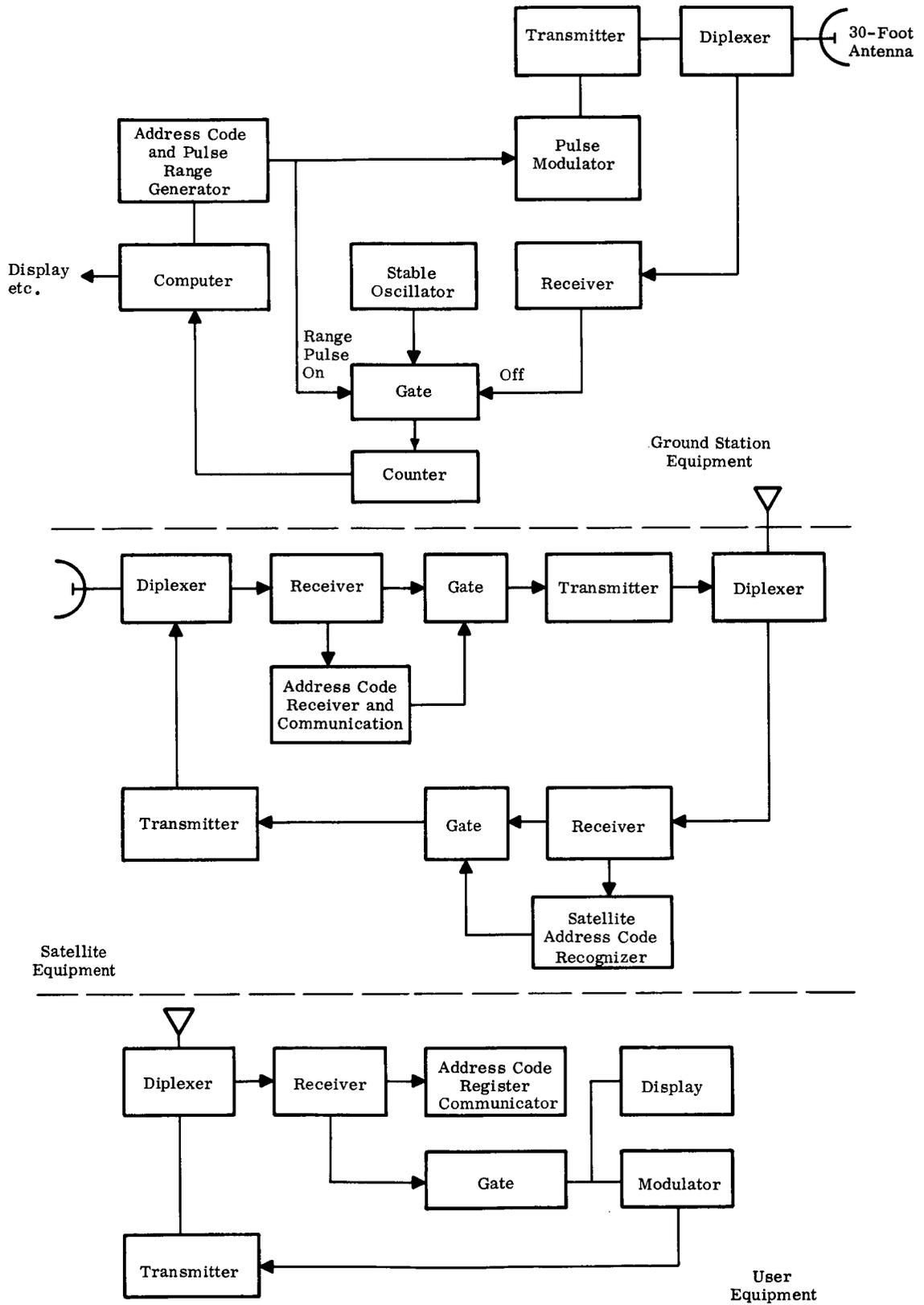
Figure III-2. Equipment Block Diagrams

The energies shown in Table III-3 are selected as being within conservative design range for wide angle antennas. They can be reduced in proportion to the curves of Figure III-1 for user antennas of larger aperture—i.e., narrower beamwidths.

Satellites of a cost and size considered to be economically attractive for the navigation functions can supply bus power in excess of 150 watts. Assuming a transponder bus to RF efficiency of 33 percent, 50 watts is available for the navigation functions. The number of range measurements per second may be determined by dividing energy in watt-second into the 50 watts available. For the values shown in Table III-3, seven to several hundred range measurements could be made each second.

As noted above, the energy needed for addressing and transmitting fix information is much larger—in the range of fifty to one hundred times the energy needed for the ranging above. Practical systems may, therefore, be limited to the order of ten position fixes per second by the energy available in the satellite.

SECTION IV

MULTIPLE PULSE TRAIN RANGING

A. INTRODUCTION

In this technique, the ranging waveform transmitted is a group of pulses much as that indicated in Figure IV-1.

Figure IV-1. Envelope of Typical Pulse Train

Because the accuracy of a range measurement depends primarily upon the ranging waveform bandwidth and the energy it contains, pulse trains have several desirable features. One is that the energy can be increased by increasing the number of pulses without increasing the peak power. This is important because many applications impose peak power limitations.

A second advantage exists when addressing or interrogating one of many users. In this case the pulse train can be coded, for example, by coding the pulse spacing to make it serve as both the ranging waveform and the address.

Figure IV-1 shows a simple pulse envelope of pulses separated in time and implies a constant carrier frequency. This is probably one of the most interesting cases. However, it should be noted that the pulse train ranging technique need not be restricted to the use of such pulses. It is possible to utilize pulse to pulse variation of carrier frequency and phase in addition to the staggering of the interpulse spacing as indicated in Figure IV-1. Further, within the pulse envelope the frequency or the carrier phase can be varied. These various forms of frequency and phase coding are particularly useful when it is necessary to use minimum time for the ranging waveform, when severe peak power limitations exist, or when amplitude limiting amplifiers are to be used in the system. Phase and frequency coding permit the time and bandwidth allotted to the ranging waveform to be used most efficiently. When time duration is to be minimized, phase or frequency coding can be used to construct waveforms of nearly constant envelope amplitude. Hence the energy is maximized; yet, because of the coding, the bandwidth required for ranging accuracy and the possibility of constructing various address codes is retained. Similarly, maximum bandwidth and energy can be obtained when the peak power limitations are stringent. The use of amplitude limiting amplifiers require phase or frequency coding to avoid signal suppression in the limiter. In general the detection of the pulse train ranging waveform utilizes matched filters followed by detectors. The pulse train is most efficiently detected through the use of a filter matched to the pulse train as a whole. However, considerations of

equipment complexity, stability requirement on oscillators and matched filters, the effects of Doppler frequency shifts and repeater design may make the use of filters matched to the pulse train as a whole impractical. In this case, it is possible to detect the pulse train by parts, utilizing filters matched to the parts. In the limit, the smallest parts of the pulse train are their individual pulses. When these are simple pulse amplitude modulated carrier waveforms, the receiver IF amplifier can be a good approximation of a filter matched to the individual pulses.

Individual pulse detection has the advantage that repeaters and transponders need only preserve the envelope and phase relationship within a pulse; whereas detection of a pulse train by a filter matched to the train as a whole requires that transponders and repeaters preserve the phase relationships between pulses as well as the spacing of pulses. In the limit of simple amplitude modulated carrier pulses, individually detected, the transponder can be a pulsed oscillator because phase coherence between pulses is not required.

The use of long coherent processing times is limited by frequency stability requirements, in addition to equipment complexity. A serious loss of signal occurs for frequency shifts equal to the reciprocal of the processing time. However, this sensitivity to frequency shift can possibly be used to advantage in frequency division multiplexing.

B.   TECHNIQUE DESCRIPTION

This technique is based on the use of matched filters in detection of the pulse train. These filters, discussed in Appendix G, coherently process the waveform to which they are matched. Generally, this gives a time compression of the waveform, resulting in a pulsed waveform at the filter output which has a peak signal-to-noise power ratio of about E/N_o; where E is the energy of the pulse train and N_o is the single sided noise power density.

Appendix H introduces the concept of waveform space dimensionality and considers a method for specifying waveforms in that space. Figure IV-2 shows diagrams illustrating waveform spaces of interest. The largest space is that assigned to the complete pulse train and it has dimensions T_s and W_s. As discussed in Appendix H, in a space of this size waveforms with $2T_s W_s$ independent parameters may be specified. The independent parameters were taken as amplitudes of the waveform envelope spaced $1/2W_s$ seconds apart, in the case of low pass waveforms, and simultaneous samples of the inphase and quadrature components of the waveform spaced $1/W_s$ apart, in the case of bandpass waveforms. The smallest region of the space to which a waveform can sensibly be confined has unit area in the passband case and considering the inphase and quadrature components of the waveform, this allows $2T_s W_s$ elementary waveforms to be defined in the space. This number is defined as the dimensionality of the space, D_s.

It is convenient when considering pulse trains to divide T_s into N intervals and define a pulse space dimensionality $D_p = 2W_s T$, where $T = T_s/N$. Each of the pulse waveforms with dimensionality $D_w \leq D_p$ is fitted into one of the pulse spaces. Figure IV-2 shows the pulse waveform made up from three short pulses which may, for example, be phase shifted versions of each other. The resulting pulse waveform is shown at various positions in the pulse space. Some of the many other possible ways of utilizing the space to define pulse train waveforms are shown in Figure IV-2b, IV-2c, and IV-2d.
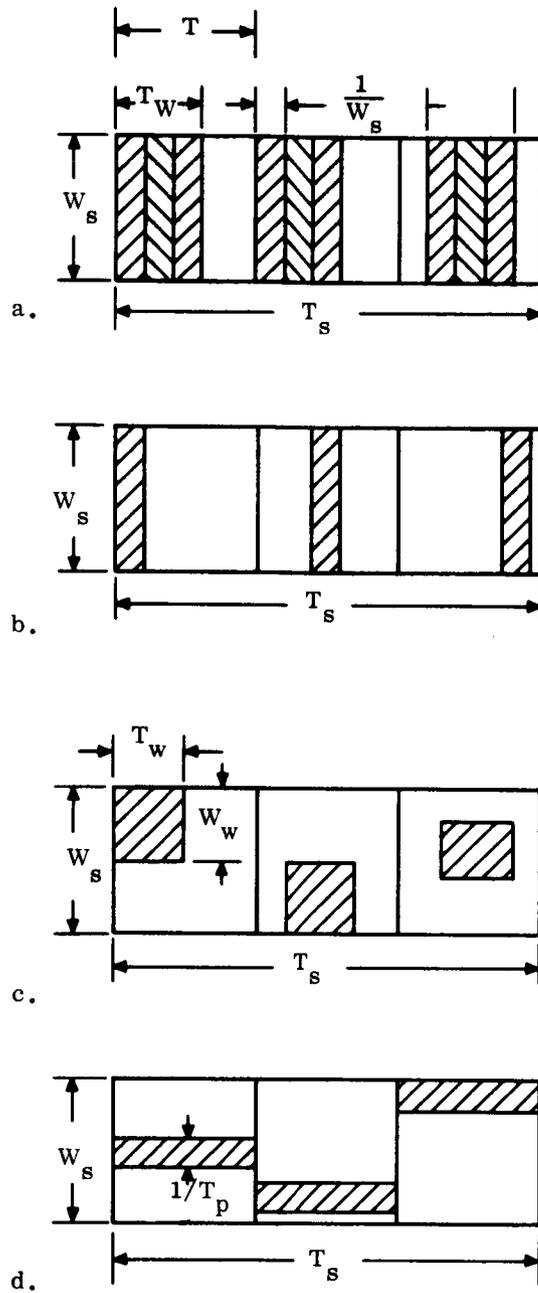
Figure IV-2. Illustration of Pulse Train Waveform Space

The method considered gives approximately orthogonal waveforms; the number being

$$M_p = D_p + 1 \qquad\qquad (4\text{-}1)$$

where

M_p = number of waveforms in pulse space

D_p = dimensionality of pulse space = $2W_s T$.

Which can be written as

$$M_p = 2TW_s + 1 . \qquad\qquad (4\text{-}2)$$

In order for a detector to distinguish between all these waveforms, it must have knowledge of the carrier phase, in addition to the phase structure of the pulse waveform. In the system to be described, this information is fairly easy to acquire on all links, except that, from user to satellite which would require an oscillator synchronizing signal from the user. To avoid this complication and to give somewhat more conservative results, the number of waveforms available is taken as

$$m = TW_s \qquad\qquad (4\text{-}3)$$

$$\approx \frac{1}{2} M_p \text{ for large } TW_s.$$

As mentioned in the appendix, the waveforms, specified by the method described there, have optimally low crosscorrelation coefficients, but not necessarily good, uniformly low, crosscorrelation functions. It is probable that a systematic transformation of the waveforms could be found, which would yield better correlation functions. But since the transformation is not known, the number of available waveforms specified is a reasonable approximation. References (3) and (8) give data useful in designing waveforms.

Figure IV-3 shows one way in which matched filters can be used in pulse train detection. It is applicable only when the pulse waveforms have the same bandwidth and center frequency. It is not directly applicable to the waveforms in Figure IV-2c and IV-2d. In Figure IV-3 the pulse train is applied to a parallel bank of matched filters. Each filter is matched to one of the N pulse waveforms. The response of each matched filter is delayed to bring it into time coincidence with the other filter responses and detected. The detected output may be applied either to a coincidence circuit or to a summing circuit and threshold detector. A particular group of N pulses represents a user address. Coincidence of the N pulses or the sum of the detected pulses exceeding a threshold notifies the user that he is being addressed.

It is apparent that the delay elements and matched filter can be interchanged assuming both are linear. Alternately the detectors can precede the delays. When the pulse train is detected as a whole, only one matched filter is used which incorporates the necessary delays and the detected pulse is applied to a threshold circuit.

Matched Filters   Delays   Envelope Detector

```
┌─────────┐      ┌─────────┐     ┌─────────┐      ┌────────────┐
│  No. 1  │─────▶│  No. 1  │────▶│  No. 1  │─────▶│            │
└─────────┘      └─────────┘     └─────────┘      │ Coincidence│
                                                  │  Detector  │
                                                  │            │
                                                  │     or     │
┌─────────┐      ┌─────────┐     ┌─────────┐      │            │
│  No. 2  │─────▶│  No. 2  │────▶│  No. 2  │─────▶│  Summing   │
└─────────┘      └─────────┘     └─────────┘      │ Amplifier  │
       •                                          │            │
       •                                          │    and     │
       •                                          │            │
┌─────────┐      ┌─────────┐     ┌─────────┐      │ Threshold  │
│  No. N  │─────▶│  No. N  │────▶│  No. N  │─────▶│  Detector  │
└─────────┘      └─────────┘     └─────────┘      └────────────┘
```

Pulse Train Input
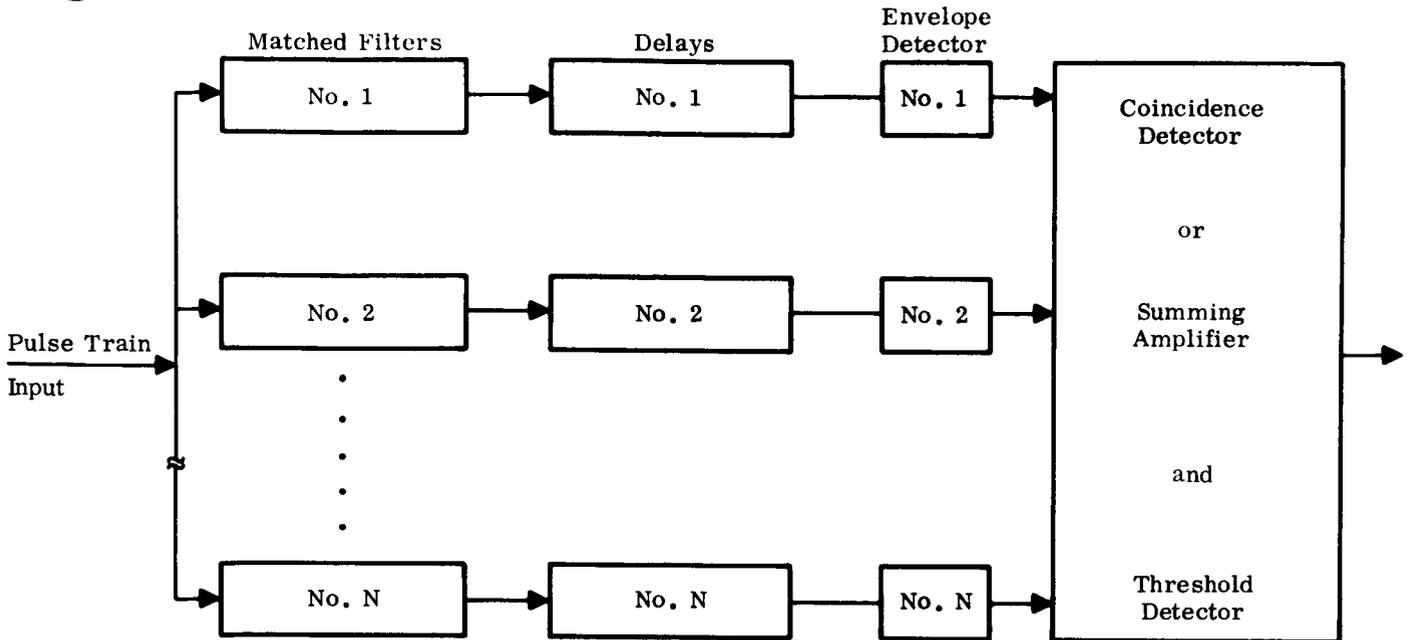
Figure IV-3. Pulse Train Detector

The matched filter technique of Figure IV-3 can be used with equal bandwidth pulses in the train because each pulse has the same range resolving power as the train as a whole, although it contains less energy. However, pulses of different center frequency such as shown in Figure IV-2c and IV-2d do not individually have the bandwidth required for range resolution—the required bandwidth is contained in the pulse train as a whole. Figure IV-3 can be adapted to such waveforms by removing the detectors and summing the delayed responses before detection.

Detection of pulse trains as a whole has the primary advantage of minimizing both the peak and average power. The power is minimized because the pulse resulting from compressing the whole waveform has the maximum possible amplitude, thereby giving least possibility for suppression by noise. If each of the N pulses are individually detected, the energy is $1/N^{th}$ of the train energy and greater suppression of signal in the detector is possible. The advantages of long processing time are paid for by complexity in the receiver in the form of the required matched filter and tighter tolerances on repeater performance. Also, because pulse trains are comparatively long, the matched filter response is more sensitive to doppler and other frequency shifts than is the response of a filter matched to the individual pulses. Appendix G points out a serious loss of signal strength at the matched filter output for frequency shifts of about $1/t_o$, where t_o is the coherent processing time defined in Appendix B. t_o is not necessarily equal to the duration of the signal as conventionally defined. The relationship can be found from the equations in the appendix. However, it is convenient and in most cases adequately accurate, to take t_o as equal to the conventional duration of the waveform to which the filter is matched.

As illustrated in Appendix G, the sensitivity of pulse trains to frequency shifts can be utilized for frequency division multiplexing. That is, with long pulse trains coherently processed by matched filters, only small frequency offsets are required to greatly reduce filter response. Ambiguity functions of Appendix G illustrate the effects.

One possible implementation of the pulse train ranging technique is illustrated in Figures IV-4, IV-5, and IV-6. The implementation is chosen to show general principles rather than to show an optimum configuration. Figure IV-4 illustrates the operations to be performed by the system. Figures IV-5 and IV-6 provide more detail on the user and satellite equipment. Pulse trains are used for both ranging and identifying the user. The N different pulses making up the pulse train are taken as phase coded pulses chosen from a set of k mutually orthogonal, equal energy, equal duration pulses, where $N < k \leq m$. Each of the N pulses in the train is detected through the use of a matched filter, followed by a square law detector. Pulse time compression takes place in the matched filters and each of the N compressed pulses is detected and applied to a digital matched filter which provides the delay and summing features described earlier.
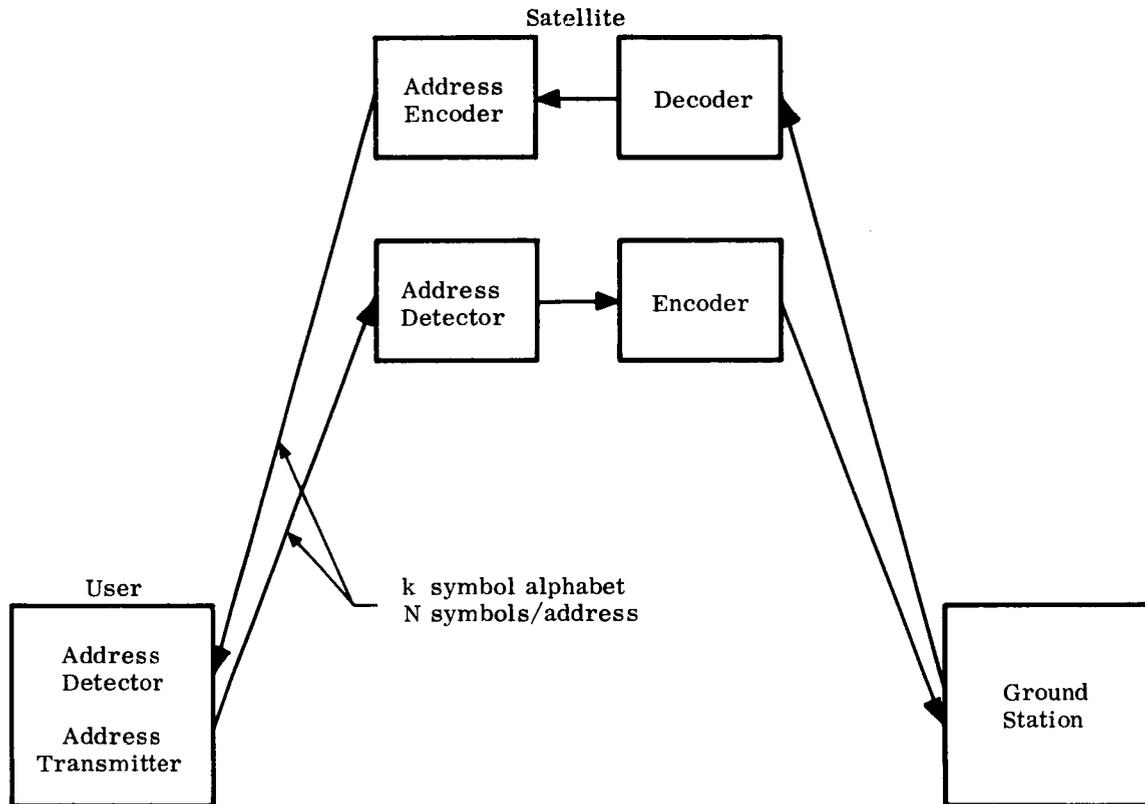
Figure IV-4. Multiple Pulse System

For illustration, the matched filters for the phase coded pulses are shown in Figure IV-5, as made up from tapped delay lines. A single tapped delay line with N sets of tap weights, in a tap weight matrix, is used to construct the N required matched filters. It is important to note that the pulse train detection is done in two levels; first level is the pulse detection; the second level is the detection of a specific pulse pattern.

An interrogation from a ground station via a satellite contains N pulses. The delay line, one set of tap weights and a summing amplifier form a filter matched to a specific pulse waveform.
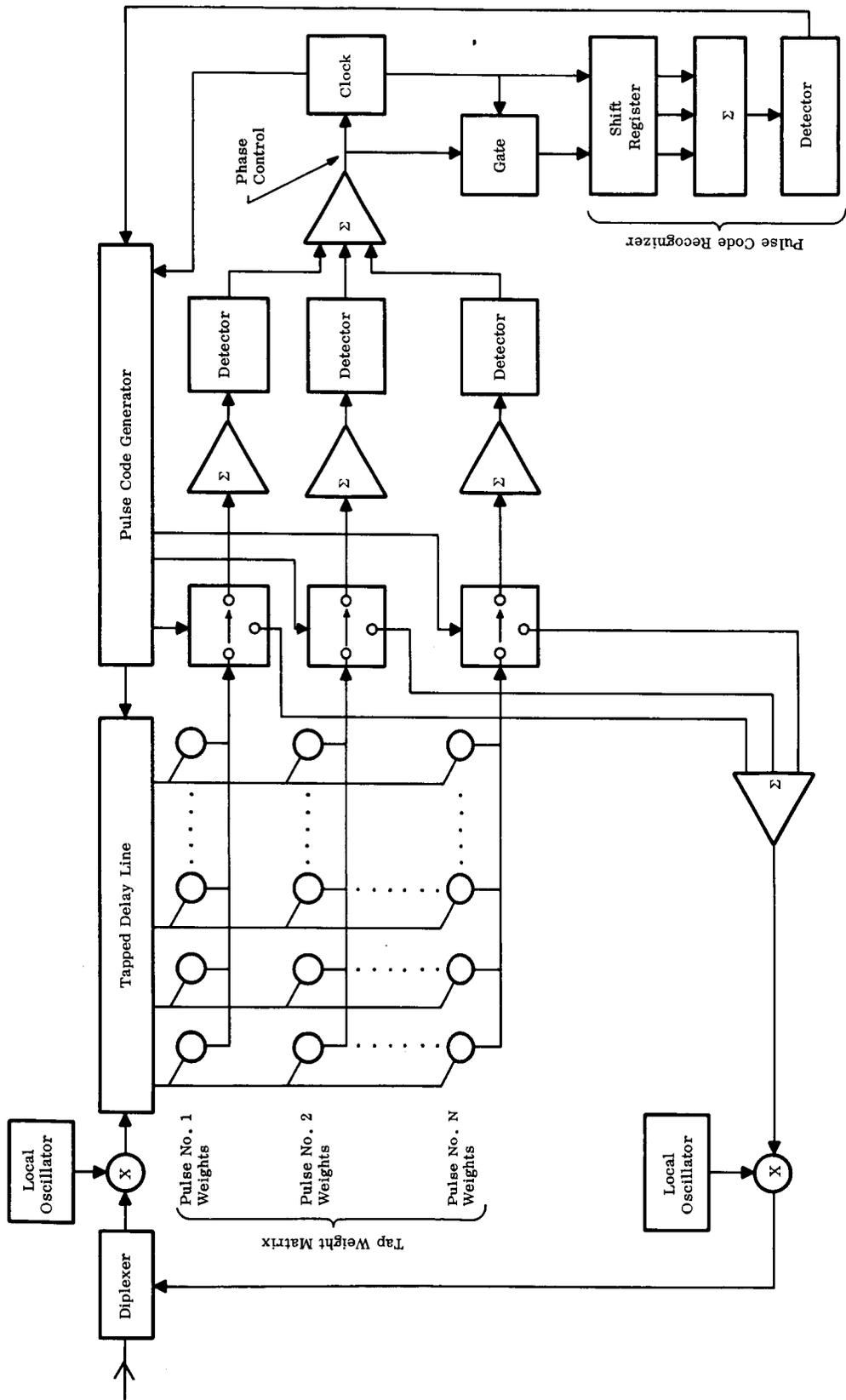
Figure IV-5. User Equipment
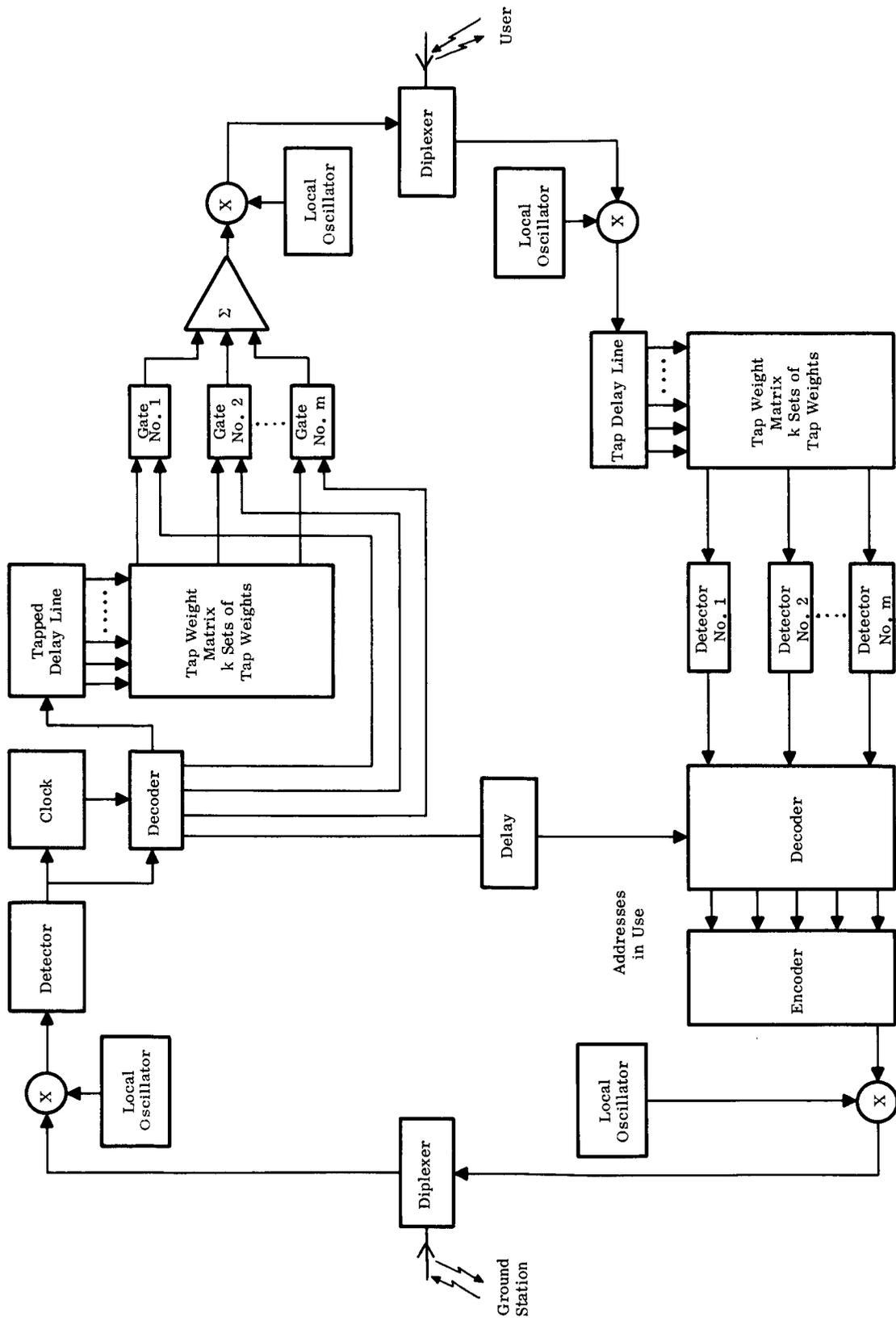
Figure IV-6. Satellite Equipment

It is possible to have one or more of the pulse waveforms the same, except, for a time displacement within the pulse space. In this case, only one set of tap weights is required for each pulse waveform.

The matched filter responses after detection are summed, used to phase control a clock and also applied to a gate. Since the pulses making up the address are used in other user addresses, though not in the same pattern, many pulses are used in controlling the phase of the clock. It is assumed that the ground station interrogations are time quantized by a clock. The gate, driven by the synchronized clock, samples the detected matched filter responses and applies them to the input of a shift register, which acts as a delay line. N stages of the shift register are tapped to a summing amplifier. The taps correspond to the positions of the detected pulses, which are sent to a threshold detector.

If some preselected level is exceeded at the threshold detector—meaning some preset number of pulses in the train have been detected—a signal is generated to trigger a pulse code generator. The pulse code generator forms an N pulse group corresponding to the time position of the pulses in the user pulse train address. Each set of tap weights generates, in response to each pulse from the code generator, the waveform to which the corresponding filter is matched. The pulse code enters the delay line at the opposite end from the signal. Switches controlled by the pulse code generator select the time pattern of the phase coded pulses for transmission. After summing the N pulses, they are frequency translated and transmitted to the satellite. IF, RF and power amplifiers, required in operational equipment, are not shown in the figure. An important feature of this technique is the use of many received pulses to control clock phase. This means the arrival time of an interrogation is known much better than could be determined from a single pulse train. The function of the pulse train is to identify which of the clock pulses represents the arrival time.

A variety of satellite equipment can be used with pulse train ranging. If the satellite-ground station link is wideband and good quality, very little data processing is required at the satellite. For example, pulse trains made up of phase or frequency coded pulses can be transmitted directly through amplitude limiting amplifiers in the satellite. In this case, user address generation for interrogation and matched filter detection of the user response can be done at the ground station with little loss in system performance.

When high energy pulse trains are used, it may be more convenient to generate them in the satellite. Figure IV-6 shows a possible implementation. The satellite equipment has separate receivers and transmitters for communicating with the ground station and the user. It is assumed that power and receiver sensitivity are not problems on the ground station to satellite link. Consequently, a modem different from that on the user-satellite link might well be used to reduce the time required to send interrogations and responses. Since this link is not critical, it has not been analyzed. However, the ground station transmits user addresses in some arbitrary format and the transmission contains timing information, so that transmissions to users can be accurately timed. The ground station transmission is detected and a pulse group generated to excite the tapped delay line. Since the N pulses in the user pulse train address can be selected from k possible waveforms, the tap weight matrix on the tapped delay line must contain k sets of weights. Otherwise, the satellite equipment generates user addresses in the same manner as the user generates a response.

Reception of user responses also requires k matched filters plus the delay and add function. Since the user address is transmitted to the satellite by the ground station, this information is used to simplify the receiving equipment.

Detected matched filter outputs are fed to a decoder circuit which provides the delay and add function through the use of a digital delay line. As the system has been described, no timing information is sent from the user except that contained in the responded address. Because the arrival time of the user response must be derived from the response, the detector output pulses must be sampled at a rate such that several samples are taken of each pulse. This is necessary to achieve accurate definition of the pulse leading and/or trailing edge. The shift register delay line clock must be at the sampling rate. Since the total delay required is constant and equal to the address duration, the number of shift register stages required increases linearly with clock rate. This may make the technique unattractive when long address duration are used.

One alternative to digital delay lines is the use of passive delay lines as the delay mechanism. A passive delay line does not require sampling of the detector output. A second alternative is use of one of the possible pulse trains exclusively for timing information. N detectors would be permanently wired to the input of a passive delay line, tapped to detect the timing train pulse pattern. The address train arrival time then would need only to be known accurately enough to gate the timing pulse to the encoders. The sampling rate required for the other delay lines, if they are digital, is much lower than that needed to define the beginning and end of pulses. Gating of the timing pulse is required only if user responses are allowed to overlap in time at the satellite.

Since many user responses must be detected, it is cumbersome to have one delay line properly tapped and connected to the detectors for each of the users. This is avoided through the use of the information available at the satellite about the addresses currently in use. Specifically, each detector drives a bank of parallel gates—the number of gates required at each detector is equal to the maximum number of addresses that will be in use in any 43 microseconds interval. The 43 microseconds is the uncertainty in the user response delay and is equal to the two way propagation time over the earth radius. One gate on each detector is connected to a delay line, which is tapped at intervals equal to the minimum, discrete, position the pulses can occupy in the pulse train. Each tap is gated to a summing amplifier.

For each address in use, the appropriate N detector gates are energized to pass the detected pulses to a delay line. The delay line tap gates corresponding to the user pulse groups are also energized. Thus, a detector for each of the addresses in use is set up. The detector remains set up for a maximum of 43 microseconds and then is readjusted to detect a new address.

The address information is delayed between the ground station decoder and the user response decoder by a time interval equal to the fixed system delays, such as the propagation time from satellite altitude to the earth and back.

C. PERFORMANCE ANALYSIS

1. ORBIT AND STABILIZATION EFFECTS

These factors uniquely affect this technique only through doppler frequency shifts produced. Doppler effects are discussed below.

2. FREQUENCY EFFECTS

The carrier frequency determines the space loss, antenna gain for a fixed aperture, scintillation due to ionosphere and magnitude of doppler shifts. These effects are common to all techniques. Only doppler has a unique effect on this system. So the discussion will be confinded to doppler.

The doppler frequency shift, one way, is given by

$$\nu_d \cong \frac{\dot{R}}{c} f_c \ \text{Hz} \tag{4-4}$$

where

$\dot{R}$ = range rate

c = propagation velocity

f_c = the carrier frequency.

For a 5,600 mile orbit, the maximum range rate is about 4,200 miles per hour. A mach three aircraft can contribute about 2100 miles per hour to the range rate. These range rates give maximum doppler shifts of

$$\nu_{d,\,sat} = 6.3 \, \dot{f}_c \ \text{Hz}$$

$$\nu_{d,\,ac} = 3.15 \, f_c \ \text{Hz}$$

for f_c in megahertz. Over the frequency range of 118 to 1660 MHz, the maximum contributions to doppler shifts go from $\nu_{d,\,sat}$ = 744 to 10,460 Hz and $\nu_{d,\,ac}$ = 372 to 5,230 Hz. As discussed previously, the detector output is low, if the waveform is mismatched by $1/t_o$ Hz from the matched filter center frequency. Consequently, uncompensated doppler limits the coherent processing time to

$$t_o < \frac{1}{\nu_{d,\,sat} + \nu_{d,\,ac}} \tag{4-5}$$

Automatic frequency control can be imcorporated into the system to compensate for doppler. Since $\nu_{d,\,sat} + \nu_{d,\,ac}$ can reach a maximum of 15,690 Hz, the system would be limited to coherent processing of simple pulses, if the system bandwidth is similarly restricted. Simple pulses require high peak power. The peak power reduction possible is proportional to the dimensionality of the pulse waveform. Consequently, if the bandwidth is restricted by assignment and doppler is large, high peak power is required.

Figure IV-7 shows maximum doppler shifts and the corresponding restrictions on processing time as a function of carrier frequency.
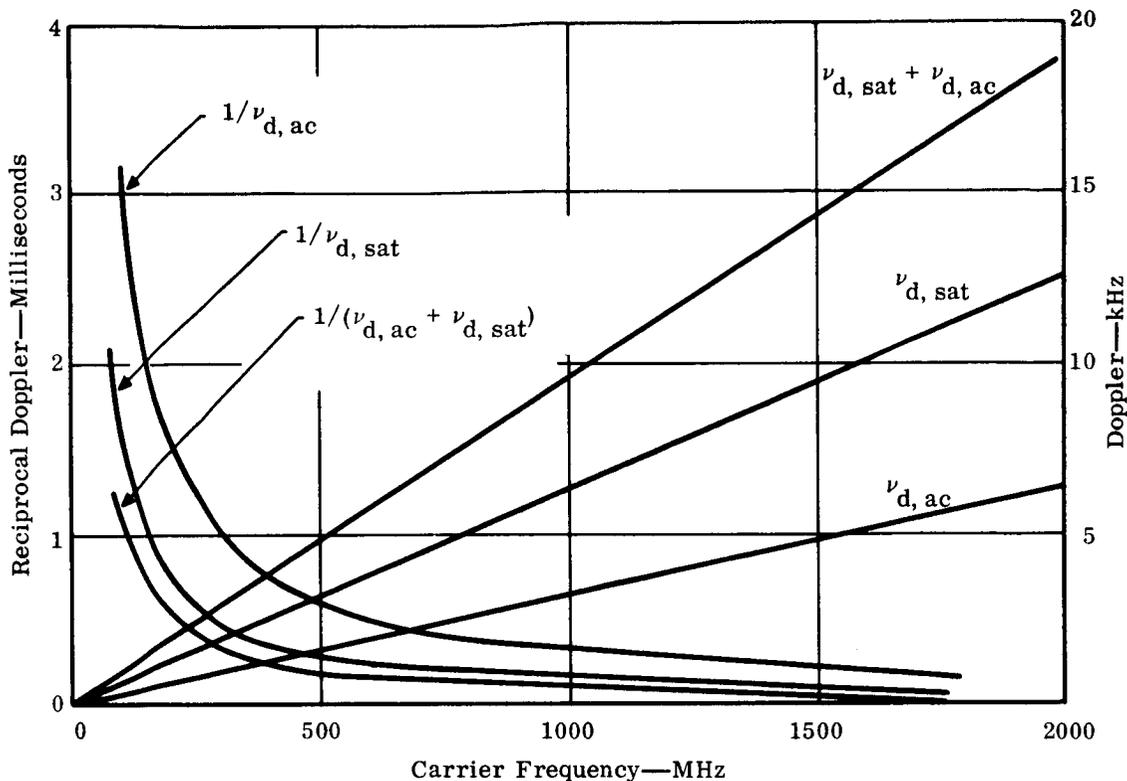
Figure IV-7. Maximum Doppler Frequency Shifts and Reciprocal Doppler Due to 5600 Miles
Orbit Satellite and 2100 Miles per Hour Aircraft

3. MULTIPLE ACCESS CAPABILITY

The multiple access capabilities can be estimated as follows. Assume pulse trains made up of N coded pulses. The coding is such that the pulses are nearly mutually orthogonal to eliminate considerations of partial contributions. This may be achieved by a combination of time and phase coding. In any case, the pulses are spaced by a nominal interval of T seconds. Each of the N places can contain one of k orthogonal pulses. For example, if the pulses are simple, uncoded pulses may lie in one of k non-overlapping time slots in the pulse space.

The limit of the multiple access capability technique is taken as the number of user response pulse trains that may be simultaneously handled by the satellite receiver without the probability of falsely generating an address currently in use exceeding some prescribed level. False addresses generated in satellite transmission will with high probability be rejected by the satellite, if responded to by a user. It is assumed that each of the N pulses making up a train is chosen randomly from the set of k waveforms. The random choice is for convenience in analysis and would not be used in practice, because of the possibility of duplicating addresses. However, such a choice results in a conservative estimate of the multiple access capability.

Let L be the number of user responses simultaneously arriving at the satellite. L is assumed to be constant*. Let

p_1 = probability that in the jth pulse period one of the L users has one of the k waveforms matching a given user address.

= 1/k

p_2 = probability that none of the L users has the particular waveform in the interval

= $(1-p_1)^L$ = $(1 - 1/k)^L$

p_3 = probability of one or more of the L users having the waveform in the given interval

= $(1 - p_2)$ = $\left[1 - (1 - 1/K)^L\right]$

p_4 = probability of one or more of the L users having a given waveform in each of the N intervals

= probability of falsely generating a given user address

= p_3^N = $\left[1 - (1 - 1/k)^L\right]^N$

$$\left.\begin{array}{c}\\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array}\right\} \quad (4\text{-}6)$$

If q is the number of addresses in use at a given time, the probability that one of them is falsely generated is approximately

$$p_5 = q\, p_4 \qquad\qquad (4\text{-}7)$$

Figure IV-8 shows a plot of Equation (4-6) for N = 5, 10 and k = 5, 10, 20. It shows that the multiple access capability is heavily influenced by the degree of security against false address generation required. Note that the curves strictly apply only when the pulse waveforms are chosen randomly. Practically, the would be chosen by design and this would decrease the probability of generating a false address, when L is small, from that shown on the curves. For example, when L = 1, there is only one user transmitting to the satellite and that waveform would be different by design from all other users. That would make the probability of false address zero. Therefore, the curves should go to infinity at L = 1. From a few calculations wherein the non-random choice of waveforms is taken into account, it appears that the curves show a probability of false address that is large by a factor of 3—i.e., the curve are low by 1/2 division, at L = 2 and the error is smaller for larger L.

In addition to the limitation on multiple access caused by false address generation, there is a limitation due to interference between pulses. This is due to the fact that the k pulses cannot be strictly orthogonal for all relative time displacements. This results in a low level noise-like response of a matched filter to a waveform to which it is not matched. This point is discussed in Appendix G.

4.   GROWTH CAPACITY

The example of paragraph D, which follows, shows that the technique is capable of handling more than the expected traffic.

---

*L is actually not constant because the users are distributed in range. A more accurate result requires averaging over the distribution of L. See Reference (9).
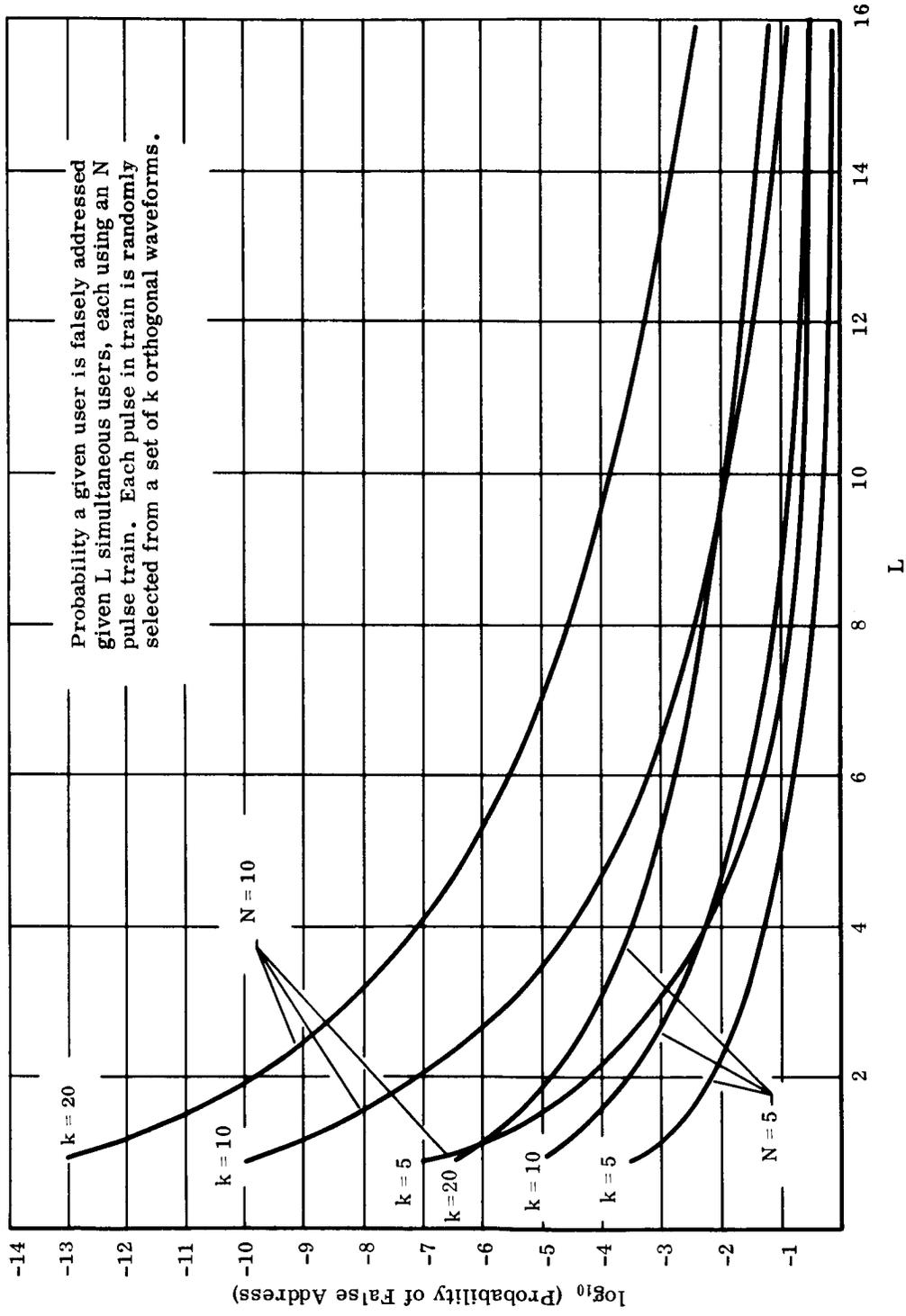
Figure IV-8. Probability of False Address

5. ERROR ANALYSIS

The basic limitation on accuracy of the range measurement is given by the error in measurement of range delay. For thermal noise the error is given by Equation B-7, discussed in Appendix B, which is

$$\sigma_\tau = \frac{1}{\beta_0 \sqrt{\dfrac{2E}{N_0}}} \; \frac{1}{\sqrt{1 - \dfrac{a^2}{\beta_0^2 t_0^2}}} \quad \text{second.} \tag{4-8}$$

If the frequency modulation is small $a \approx 0$ and Equation 4-8 becomes

$$\sigma_\tau = \frac{1}{\beta_0 \sqrt{\dfrac{2E}{N_0}}} \quad \text{second.} \tag{4-9}$$

If E is the total energy of the pulse train, Equation 4-8 applies when the pulse train is detected as a unit. Under large signal-to-noise conditions it also applies (Reference 6 , Chapter 8) when: 1) a range delay measurement is made on each of the pulses in the train and the range delay taken as the average; or 2) the detected outputs of the matched filters for the several pulses are appropriately delayed and added, as discussed in reference to Figure IV-3, and the range delay measurement made on the sum. The second method is potentially more efficient because each of the pulses need only contain enough energy so that the matched filter responses are not suppressed by noise in the detector following the matched filter. If E_p is the energy of the pulse waveform ($E = NE_p$), this requires $E_p/N_0 \approx 3$, assuming the train contains enough pulses so that NE_p is large enough to give reliable detection, about 50. The effects of dividing the pulse train energy into various numbers of pulses and detecting each individually, relates to the discussions of Reference (10). The data presented there is a useful guide to effects of energy distribution in the pulse train. The energy to noise density required in each pulse in the first ranging method is only a few db less than that required in the whole pulse train when the second method is used. As N becomes small, the efficiency of the two techniques becomes nearly equal.

When waveforms, overlapping in time and with almost completely overlapping spectra, are applied to a receiver, as occurs in waveform and some forms of frequency division multiplexing, the quantity N_0 needs some interpretation. In this case in addition to the contribution to N_0 from thermal noise and other sources, each of the waveforms contributes interference of energy density about $1/\beta_0 t_0$ down from the energy of the pulse train, as discussed in Appendix G. If there are L pulse trains of the types described, all of equal energy, simultaneously applied to a receiver in addition to the one intended for the receiver, then the effect on ranging accuracy is about the same as that due to noise of power density $L/\beta_0 t_0$. This puts a limit on the number of multiplexed pulses.

In addition to errors due to noise and interference, there are systematic errors. These come from variations in the delay of amplifiers, clock phase errors and so on. When the bandwidth is small, an important possible source of error is variation in the pulse detection level. A narrow bandwidth results in a slow pulse rise time and variations of detection level can give large delays. This means that good pulse leveling or pulse splitting techniques are required when the bandwidth is small.

In the system shown in Figures IV-5 and IV-6, it is expected that the user to satellite link would contribute the major part of the ranging errors because timing and frequency are continually shifting as different users respond.

D. AN ILLUSTRATION OF SOME PARAMETERS OF A PULSE TRAIN RANGING SYSTEM

This section is an example worked out to show some typical results for the VHF case and the application of equations presented earlier.

The general parameters of the VHF "probable" case, shown in the power budget of Table III-1 in Section III, will be used.

Assume a useable system bandwidth of 15 kHz. Frequency instability is not important and will be discussed later. The number of user addresses is one million.

From Reference (10), a pulse train energy to noise density ratio of 17 db appears appropriate and perhaps somewhat conservative. If L + 1 pulse trains are allowed to overlap—meaning that each pulse train is subjected to interference from L others, the noise level includes the interference. Using the approximation discussed previously—namely, that the interference between pulse trains is similar to thermal noise at a level of $E/\beta_o t_o$ for each interferring pulse train, the approximate total noise is

$$N_t = N_o + N_i \tag{4-10}$$

where N_i is the interferring pulse train noise and N_o is the thermal noise. Equation (4-10) can be written

$$N_t = N_o + \frac{LE}{\beta_o t_o} \tag{4-11}$$

giving the signal-to-noise ratio

$$\frac{E}{N_t} = \frac{E}{N_o + \frac{LE}{\beta_o t_o}} = \frac{1}{\frac{N_o}{E} + \frac{L}{\beta_o t_o}} \tag{4-12}$$

where equal energy and $\beta_o t_o$ product is assumed for all waveforms.

Given a required E/N_t the necessary E/N_o can be determined from Equation (4-12) as

$$\frac{E}{N_o} = \frac{1}{\frac{N_t}{E} - \frac{L}{\beta_o t_o}} \quad . \tag{4-13}$$

Equation 4-13 shows that, since $E/N_o > 0$, the following must hold

$$\frac{\beta_o t_o}{L} > \frac{E}{N_T} \quad . \tag{4-14}$$

From the power budget of Table III-1, the attenuation from satellite to user is 175 db and the thermal noise power density N_o is $-(154.5 + 42) = -196.5$ db. Since E/N_o of 23 db is required, the received energy must be

$$E = N_o - 23 = -196.5 + 23 = -173.5 \text{ db}$$

referenced to one watt-second. The transmitted pulse train energy is $E_T = 175 - 173.5 = 1.5$ db referenced to one watt-second or $E_T = 1.42$ watt-second. This energy is spread over T_s seconds which, from the assumption of $W_s = 15$ kHz and Equations (4-18) and (4-21)

$$T_s = \frac{m}{W_s} N \tag{4-22}$$

$$= \frac{20}{15 \times 10^3} (10) = 13.3 \text{ msec.}$$

The pulse space duration is $T = m/W_s = 1.33$ msec. Assuming the energy E_T is spread uniformly over the pulse train—which means that the pulses are phase coded and 1.33 msec long, the peak transmitted power per pulse train is

$$P_p = E_T/T_s = 1.42/13.3 \times 10^{-3} = 107 \text{ watts} \quad . \tag{4-23}$$

Individual pulse detection has been assumed and the pulse duration taken as 1.33 msec. This means that the pulse matched filter output will be very small for a frequency offset of

$$\nu = \frac{1}{T} = \frac{1}{1.33 \times 10^{-3}} = 750 \text{ Hz} \quad . \tag{4-24}$$

Practically the frequency offset would have to be restricted to a few hundred Hz. This is a stringent requirement. It can be relaxed by making the pulse waveform dimenionality smaller, while retaining pulse space dimensionality, and using time translations of the pulses to generate the m different waveforms. The sensitivity to frequency offset decreases directly with pulse duration. But peak power increases as the pulse is shortened.

It has been assumed that $L = 3$. This means that a receiver has $L + 1 = 4$ pulse trains applied to it. If one of the pulse trains is its address, three other pulse trains cause interference. The previous results have shown some of the pulse train requirements for the receiver to operate properly in this case. To those receivers not addressed, all the pulse trains represent interference and in applying Figure IV-8, L should be taken as 4. From the figure, for the present case, $N = 10$, $k = m = 20$ and $L = 4$, it is seen that the probability that a user is falsely addressed is about 10^{-7}.

Approximating $\beta_o t_o$ as $W_s T_s$, Equation (4-14) can be written

$$\frac{W_s T_s}{L} > \frac{E}{N_t} \tag{4-15}$$

which means that the pulse train waveform dimensionality must be larger than L times the required signal to total noise ratio. It is convenient to put the equation in terms of pulse space parameters, W_s and T. Since $T_s = NT$, Equation (4-15) gives

$$W_s T > \left(\frac{L}{N}\right) \frac{E}{N_t} \ . \tag{4-16}$$

Choosing $W_s T = 20$ and using $E/N_t = 50$ (17 db) as specified earlier, gives the requirement that

$$\frac{N}{L} > 2.5 \tag{4-17}$$

From Equation (4-3) the number of waveforms available in the pulse space is

$$m = W_s T = 20 \ . \tag{4-18}$$

Writing user addresses with k symbols in N places gives as the maximum number of addresses

$$U = (k)^N \ . \tag{4-19}$$

In the present case $k \leq M$ and it is required that $U = 10^6$. Choosing $L = 3$, Equation (4-17) gives

$$N > 7.5 \ . \tag{4-20}$$

This gives $U \gg 10^6$ if $k = m = 20$ is used in Equation (4-19). Let

$$N = 10 \ . \tag{4-21}$$

Using the approximation $W_s T = \beta_o t_o/N$, in Equation (4-13), the signal energy to the thermal noise ratio can be determined:

$$\frac{E}{N_o} = \frac{1}{\dfrac{N_t}{E} - \dfrac{L}{N W_s T}}$$

$$= \frac{1}{\dfrac{1}{50} - \dfrac{3}{10(20)}} = 200 \ (23 \ db) \ .$$

Because most of the noise is due to N_i, system performance will not be very sensitive to thermal noise N_o.

If no pulse trains are allowed to overlap, the total noise is thermal noise and the restriction of Equations (4-11) through (4-17) do not apply. Consequently, a great deal more freedom is available in the design of pulse trains. For comparison, assume the pulse train parameters remain the same, except overlap is not permitted. From Equation 4-12 it is seen that $E/N_t = E/N_o = 17$ db which is 6 db less than that required with four overlapping. Therefore, the peak power transmitted is 1/4 of that shown in Equation (4-123), or about 28 watts per pulse train.

The addressing rates for the parameters chosen are 300 per second when 4 trains overlap and about 75 per second with no overlap.

The range delay measurement accuracy is given by Equation (4-8) which is for $a = 0$

$$\sigma_T = \frac{1}{\beta_o \sqrt{\frac{2E}{N_o}}} .$$

Taking $\beta_o = W_s$ and the assumed $E/N_t = 50$ gives

$$\sigma_T = \frac{1}{15 \times 10^3 \sqrt{100}} = 7 \times 10^{-6}$$

which gives a 1σ range error of about 3500 feet.

SECTION V

SIDE TONE MODULATION

A. TECHNIQUE DESCRIPTIONS

To develop a frame of reference for the multiple side tone (MST) scheme, first look at the "conventional" application, which has been highly developed in many systems. The basis of the technique is to use the time or phase delay of a low frequency signal to measure the propagation path length. Modulation is used to allow effective transmission of the low measurement frequencies. Because the cyclic nature of a tone creates measurement ambiguities every cycle of the tone, it is common to use multiple tones to allow ambiguity resolution of the higher frequency tones. By extension of this, at least in theory, any given range can be measured to any given accuracy.

The normal MST system transmits to a phase locked transponder an FM (or PM) signal which contains the range tones. The use of a phase locked transponder is quite common due to its good sensitivity and highly accurate tone phase control. In general, the output signal of the transponder is used as the first local oscillator for the transponder receiver. In this manner, the receiver compares the phase difference between the received tones and the transmitted tones. This forms the basis of a servo control loop allowing the normal tone phase errors to be reduced by the magnitude of the control loop gain.

The transponder transmitted signal is then received at the ground station and the received tones are phase compared with the tones originally transmitted by the measuring station.

In phase comparison for data extraction, system bias phase shifts are generally removed. Through system calibration and checkout, both with and without the use of a sample transponder, the values for bias errors can be determined. Adjustable phase shifters, similar to that shown in Figure V-1 in the ground station, can be then used to compensate for the bias errors and effectively remove them from the data. The same results can be obtained in computer processing of the data by including bias effects in the computer programming. Figure V-2 shows a block diagram of a transponder for such a system.

To assess the ultimate multiple access capability of an FM MST system, the lock-up time for the receiver must be calculated. Considerable work has been done at General Electric on several past systems to evaluate the minimum acquisition time for a phase locked loop.

For the past several years General Electric has been designing and building a large number of phase locked receivers for use in missile guidance and instrumentation systems. As a result a large effort has been applied to determining the acquisition characteristics of phase locked receivers. This effort has resulted in a determination of the maximum rate that frequency uncertainty can be searched. This rate, as shown in Figure V-3, is a function of signal-to-noise ratio in the phase locked loop, loop noise bandwidth, and first time acquisition probability.
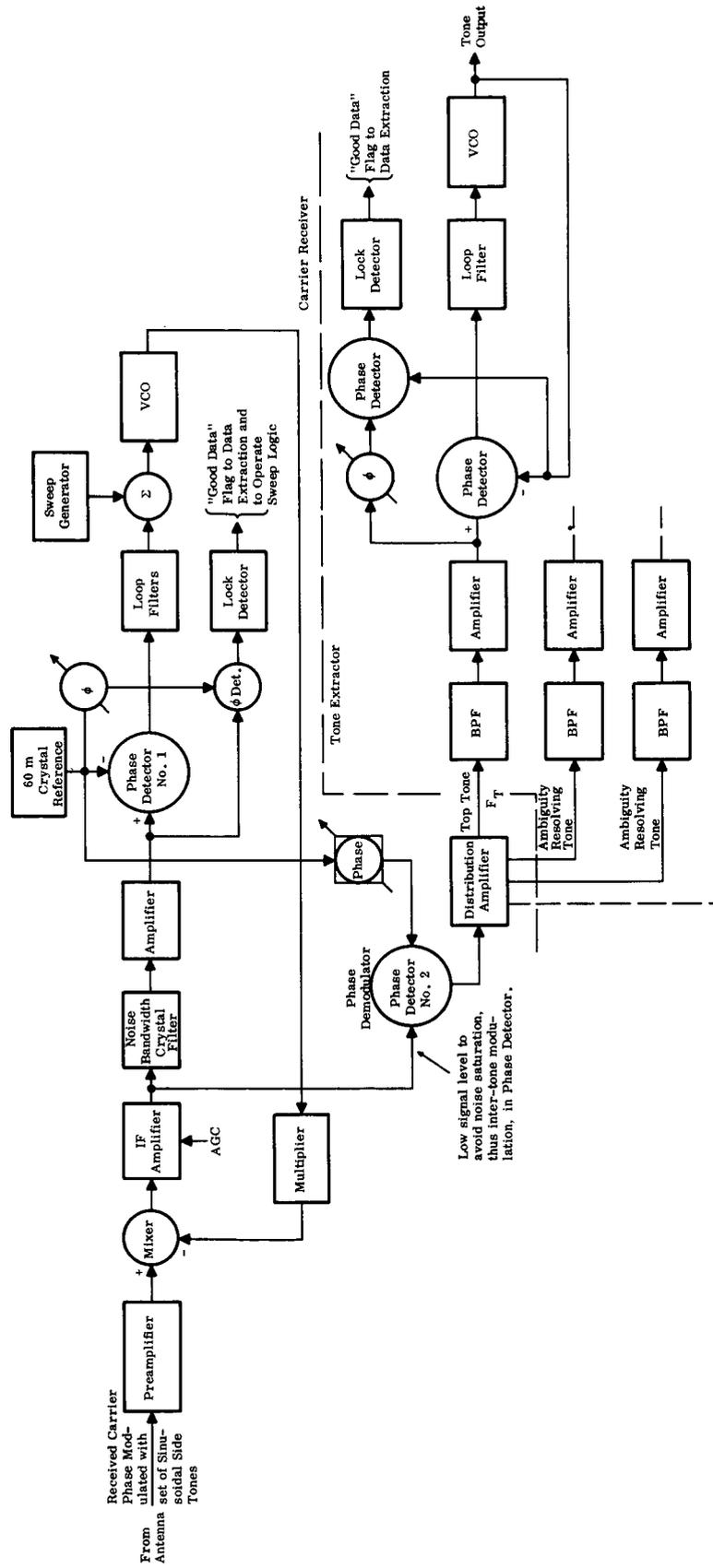
Figure V-1. Ground Station Typical Receiver and Tone Extractor
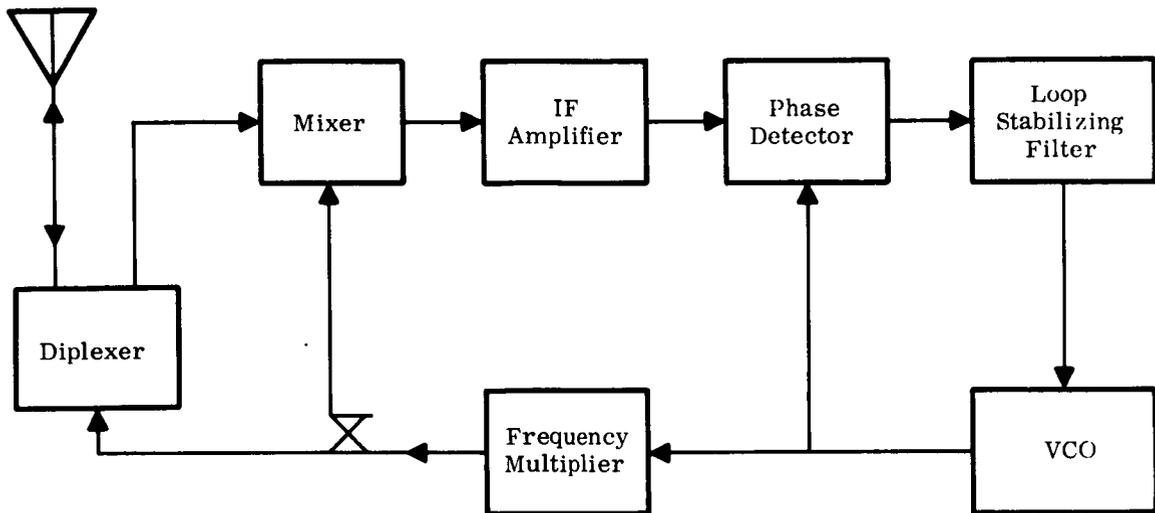
Figure V-2. FM Side Tone Transponder

For a practical system, it should be possible to keep the user receivers locked (if phase locking receivers are used) at all times; thus, the acquisition time of importance is that of the ground receiver. A deep fade would momentarily disrupt the system until the user receivers relocked. This, however, should only cause occasional data interruptions. If phase coherent transponders are used, the returned signals will have the doppler shift over the propagation path as the only uncertainty. On the other hand, if the return signal transmitters are free running, their frequency error must be added to that of the doppler shift.

Based upon the system parameters, the ground receiver signal-to-noise conditions could be assessed and the maximum tolerable bandwidth of the receiver carrier and tone loops could be determined. The carrier loop should see a minimum output signal-to-noise ratio of 10 db. While the "threshold" of a phase locked receiver is generally considered to be 6 db, the probability of continuous lock at this signal-to-noise ratio is small. Recently some empirical and analytical studies have been done to predict the mean time to unlock.[11,12] Condensing these studies, for a 6 db signal-to-noise ratio, a prediction of mean-time-to-unlock can be made as:

$$T_P \simeq \frac{1200}{B_n} \text{ sec} \qquad (5\text{-}1)$$

where, B_n is the two-sided noise bandwidth of the loop in Hz. For example, for a one kHz noise bandwidth loop, at the 6 db threshold, a 1.2-second mean time to unlock is to be expected. As can be seen, this could easily interfere with data gathering.

From the previously mentioned relation:

$$S/N = \frac{1}{2(\phi_\epsilon)^2} \qquad (5\text{-}2)$$
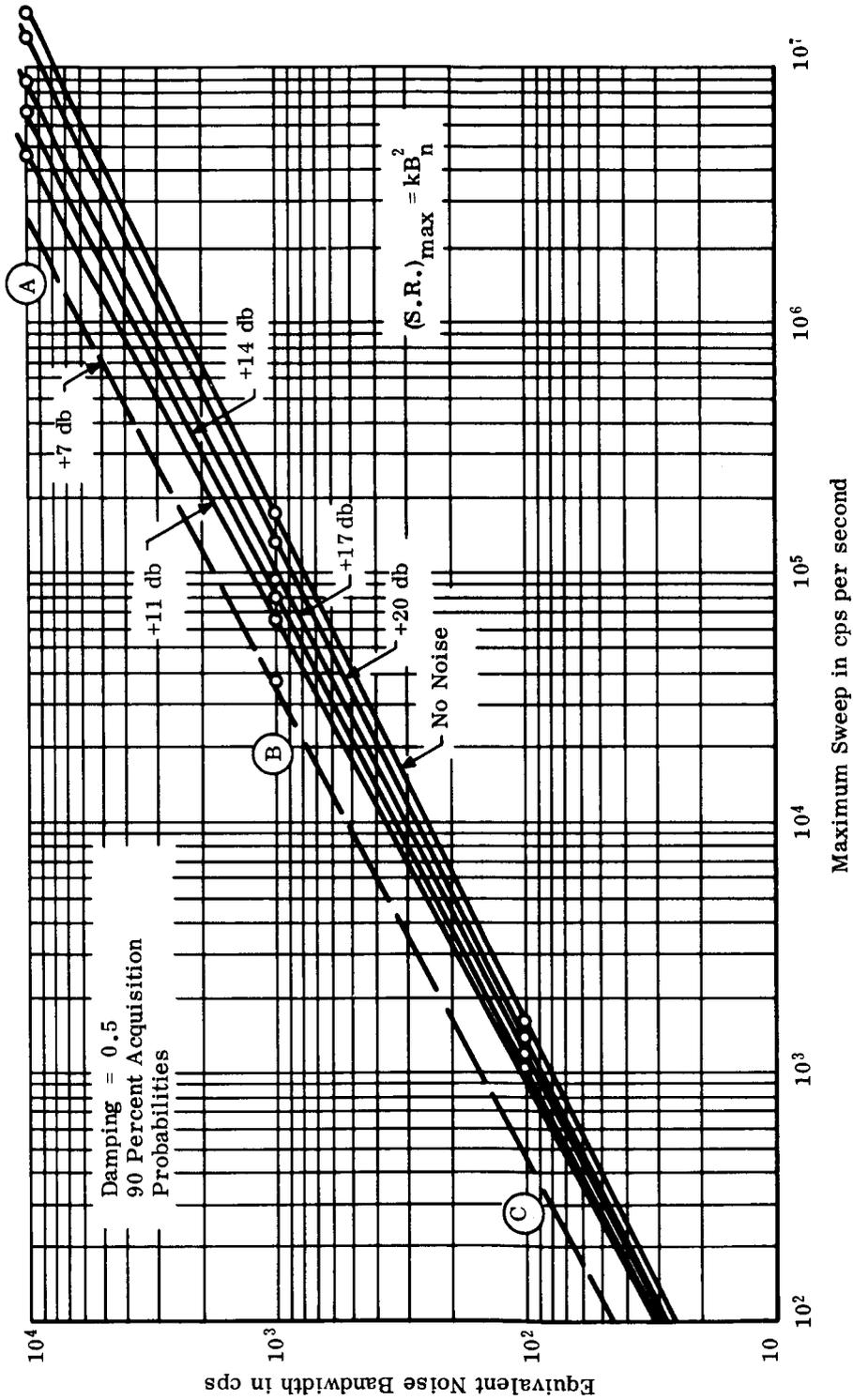
Figure V-3. Maximum Sweep Rate versus Noise Bandwidth with Proportional Plus Integral Control Filter

the output signal-to-noise ratio of the tone loops can be calculated. With the system error budget established, the tone loop bandwidth can be determined based upon the signal-to-noise improvement:

$$(S/N)_o = \frac{B_{CL}}{B_{TL}} (S/N)_i$$ (5-3)

where

$(S/N)_o$ = output signal-to-noise ratio

$(S/N)_i$ = input (to tone loop) signal-to-noise ratio

B_{CL} = noise bandwidth of carrier loop

B_{TL} = noise bandwidth of tone loop.

With all the bandwidths set, the acquisition time can be calculated from:

$$T_A = T_C + T_T$$ (5-4)

where:

T_A = acquisition time

T_C = carrier loop acquisition time

T_T = tone loop acquisition time.

The acquisition time of a loop can be calculated from:

$$T = \frac{K}{B_n} \left(\frac{\Delta f}{SR}\right)$$ (5-5)

where:

Δf = total frequency uncertainty

SR = sweep rate from Figure V-3.

The factor of K/B_n in the time equation allows for loop settling time prior to data measurements. Where B_n is the loop noise bandwidth and K can range from 3 to 5 typically.

From the expressions of Section II,

$$M = \frac{T_P}{T_M + T_A + T_D + T_G}$$ (2-4)

and

$$\eta = \frac{T_M + T_A}{T_M + T_A + T_D} ,$$ (2-5)

the time figure of merit for a MST scheme can be determined. For this system no "guard time" per se is needed. Also the data gathering does not take significant time in addition to the message time. Therefore,

$T_M + T_D \cong T_M$, and:

$$M = \frac{T_P}{T_D + T_A} \quad ; \quad \eta = \frac{T_M + T_A}{T_M + T_A} \cong 1 \quad . \tag{5-6}$$

To resolve the range ambiguities, the tone system must contain a tone whose period is at least equal to the possible time variations in the return signals. This constrains the "message time" to be in the order of 45 milliseconds minimum. To keep the data gathering time to this level requires the data readout to be capable of measuring phase on one cycle. Under these conditions:

$$M \cong \frac{0.5}{0.045 + T_A} \tag{5-7}$$

and

$$\eta \cong \frac{0.045 + T_A}{0.045 + T_A} \quad . \tag{5-8}$$

Thus, from these relations, and considering that the acquisition time can be a large part of a second, or even greater than one second in some cases, the time Figure of Merit for a full FM MST system could become small. In fact, as acquisition time becomes large with respect to propagation time, the acquisition begins to seriously reduce multiple access. The multiple access capability of a FM MST system, therefore, can become a direct function of the power received at the measuring terminal. If low power is used, the loop bandwidths must be small, requiring long acquisition times. With a small amount of received power, the acquisition time could become dominant with respect to the round trip propagation time. The net result of these relations being that the multiple access and growth capacity of this type of system are a function of the transmit power on the link to the data gathering terminal. This relation holds until the acquisition time is small with respect to the range ambiguity time.

The accuracy of a MST system will be tied directly to the frequency of the highest tone used, and thus is bandwidth limited. To achieve the most accuracy per unit of bandwidth, it is imperative that the application of the information (tones) to the carrier be as efficient as possible. This gives rise to serious questions regarding the use of FM (or PM) as a basic modulation, in that FM is basically a spread-spectrum noise improvement technique. As long as it is reasonable to generate such a signal, it appears that the most efficient technique to use would be a "compatible single side band" (CSSB) AM technique. This technique has two attractive points:

1. Highest bandwidth efficiency
2. Compatible with current AM receivers.

B. PERFORMANCE ANALYSIS

Assuming a CSSB-AM modulation, the frequency of the highest usable tone can be calculated by knowing the bandwidth available and the frequency errors to be expected. Considering first the VHF case (125 MHz), the frequency errors will be:

1. Doppler: 750 Hz
2. Satellite Transmitter (0.5 ppm stability) ±62 Hz
3. Receiver L.O. (20 ppm stability) ±2500 Hz

the total frequency error will, therefore, be (as a worst case) the sum of the parts or 5874 Hz. For a 25 kHz bandwidth, subtracting the drift, yields a usable bandwidth of 19.12 kHz maximum. The following example uses 18 kHz.

From the frequency-period of a sine wave relation, the measurement error in feet per degree phase measurement error was calculated. This relation is plotted in Figure V-4. From this graph, a figure of approximately 70 feet per degree phase error is found. This represents then, the measurement error introduced by phase changes (in degrees) in the receiver.

To estimate the phase errors produced by a receiver, a computer program was written to calculate the phase shift as a function of frequency for a typical double tuned transformer circuit. At this time, it is unknown just what the current VHF receivers are equivalent to, in terms of double tuned circuits, since it is the phase non-linearities that are important. As an attempt at a worst case, it is assumed that a current VHF receiver looks equivalent to five cascaded double tuned stages with respect to phase non-linearities. Using a 18 kHz spectral line separation and a total of 6 kHz drift, the differential phase error calculates to be 2.58 degrees per stage or 12.90 degrees total. This error is for a 25 kHz bandwidth circuit with center frequency at 250 kHz. As a comparison, using a center frequency of 500 kHz and a 25 kHz bandwidth, the resultant phase error is 1.32 degrees per stage or 6.61 degrees total. Taking the worst case of 12.90 degrees—at 18 kHz; this would result in a measurement error of 903 feet, which is getting very close to the 1000-foot "goal" indicated in Section II. This indicates that a current VHF receiver, if used, should have better frequency control with respect to its IF circuits than the 20 parts per million assumed.

A swept side tone modulation, such as examined in "Feasibility Study of Satellites for Range Instrumentation" (AF19(628)-4200) would be affected by the total phase change across the IF bandpass, not just by non-linearities in phase. As such, unless this technique were used in a full phase locked loop, the change of phase of the receiver would be completely intolerable.

The susceptibility of the receiver and tone processing equipment to interference will be primarily in two modes. Since the technique is capable of operating with such a weak predetection signal-to-noise ratio, an obvious form of interference would be that of a relatively strong signal, within the passband of the IF, developing an AGC voltage which would suppress the desired signal. If, however, the interfering signal were not strong enough to suppress the desired signal by AGC action, then, to cause trouble, this interfering signal would have to significantly affect the output of the tone filters. If a two degree criteria were put on this "noise" for example, then the AM spectral components of the interference would have to equal or
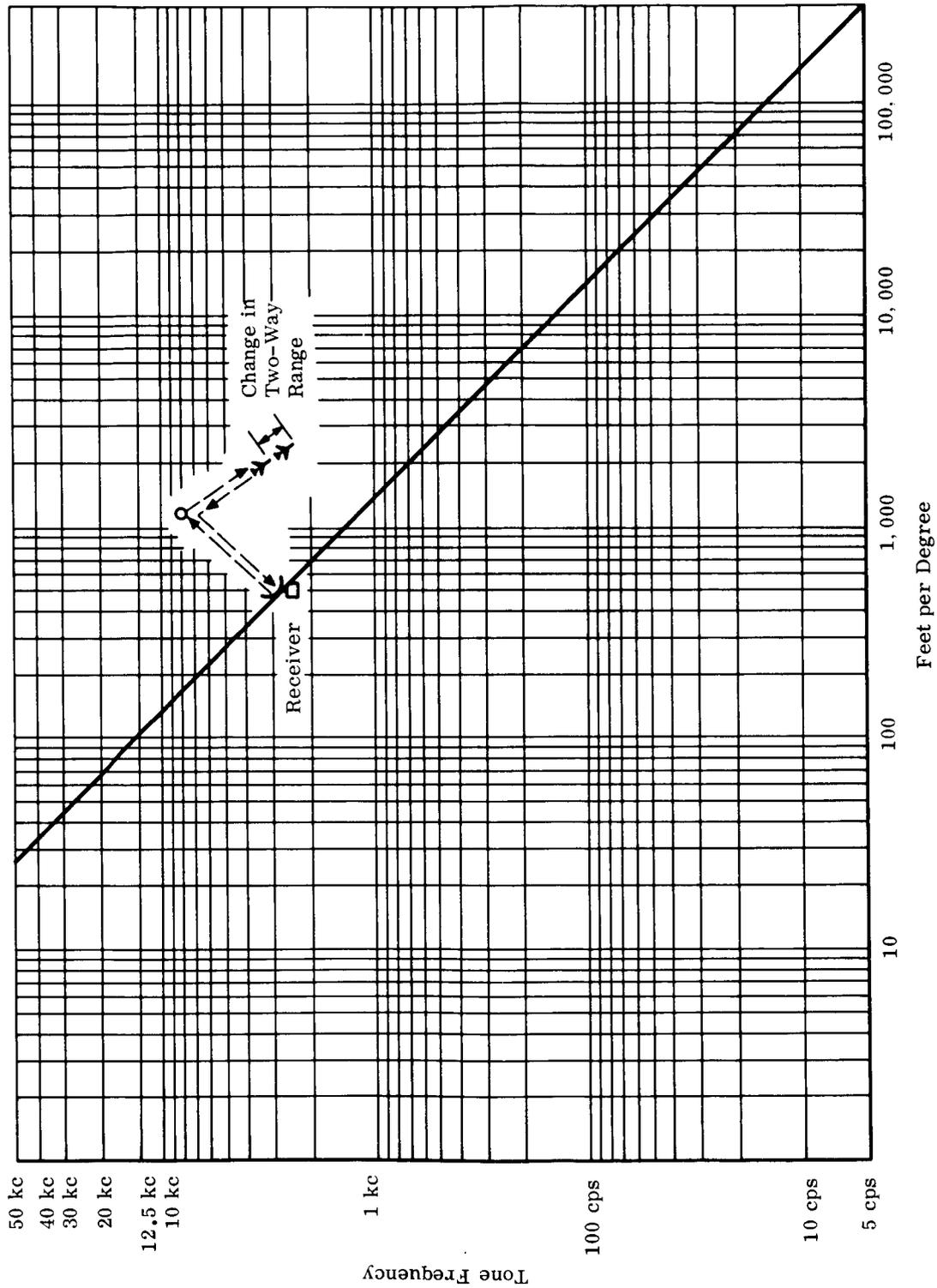
Figure V-4. Change (feet) in One-Way Range Resulting in One-Degree Phase Change at Receiver

exceed 0.035 of the desired tone voltage. This would be due to sufficient amplitude of the interfering signal within the bandpass of the tone filter for a sufficient length of time. To arrive at an explicit expression for the effect of an interfering signal, the frequency-time statistics for the signal would have to be known. In any event, given the ratio of the interfering tone to the desired tone to be A, and the response of the tone filter to a signal within its bandpass to be f(t), then:

$$A \cdot f(t) \leq 0.035 \qquad (5-9)$$

must hold for the two degree arbitrary requirement. The performance of the system with regard to the power budget depends heavily upon the post-detection processing or filtering of the tones. Figure V-5 shows the processing steps used on the measurement tones, if a conventional AM receiver is used. The input signal, at the IF point, is an asymmetrical signal not centered in the IF bandpass. Instead the "carrier" is displaced to one side with the highest measurement tone spectral line displaced equally to the other side of the IF bandpass. If the input carrier-to-noise ratio to the square law detector is low, there will be noise suppression of the output tone due to the non-linear detection process. Overriding this noise suppression, is the process gain of the final tone bandpass filter. As an example for a power budget and for a guide in a typical implementation, the case of a one db carrier-to-noise input to the square law detection was picked. In Appendix F, the output tone-to-noise ratio for such a detection-processing system has been calculated. A 25 kHz predetection bandwidth and a 90 Hz post-detection bandwidth were assumed. The assumed carrier-to-tone spectral power ratio was 6:1. This resulted in a 15.2 db tone output-to-noise ratio.

Input → Bandpass Filter → Square Law Detector → Low Pass Filter → Bandpass Filter → Tone Output
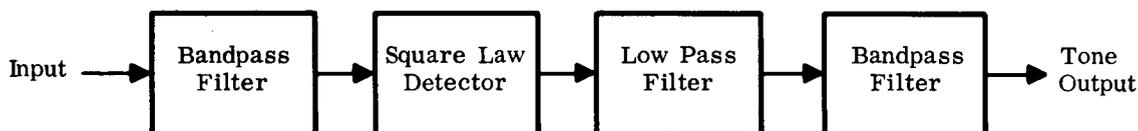
Figure V-5. Tone Signal Processing

Using the relationship:

$$S/N = \frac{1}{2(\phi_\epsilon)^2} \qquad (5-10)$$

where ϕ_ϵ is phase jitter in radians. The resultant phase jitter can be calculated.

$$\phi_\epsilon = \frac{1}{\sqrt{2 \times 33}} = 0.123 \text{ radian} \qquad (5-11)$$

$$\phi_\epsilon = 7.05 \text{ degrees rms.} \qquad (5-12)$$

This value is jitter due to receiver noise only and must, of course, be combined with other error sources to arrive at an overall error performance. At this point, it is, however, significant that a one db input

carrier-to-noise ratio will result in acceptable performance. As such this can serve as a typical value for range-power budget calculations. For this case, a power budget can be established similar to the budget for the single pulse technique. Using the same factors, except where indicated, results in the budget of Table V-1 for the carrier power required.

Thus, by applying the indicated correction factors of

1. -17 db
2. -20 db
3. -22 db
4. -22 db
5. -35 db
6. -35 db

to the ordinate of the curve of Figure III-1 the power required for a MST system can be achieved.

The frequency error on the 18 kHz tone out of the detector, due to doppler, is in the order of 0.054 Hz for the geostationary case and about 0.33 Hz for the six-hour case. For the geostationary satellite, the doppler due to a SST transport looking at a satellite "on the horizon" represents the worst case doppler shift. For the six-hour orbit satellite case, the worst case doppler is calculated from the sum of the satellite and the user vehicle velocities. Using the familiar

$$f_d = \frac{f_o V}{c} \tag{5-13}$$

approximation to the doppler shift, the doppler can be calculated. For the carrier doppler in the cases represented by the various frequency bands chosen as representative, the Table V-2 can be calculated.

C. HYBRID MODULATION IMPLEMENTATION

The advantages and disadvantages of a multiple side tone ranging scheme have been highlighted. It is apparent from a study of the four links of the system:

1. ground station-to-satellite uplink
2. satellite-to-user downlink
3. user-to-satellite uplink
4. satellite-to-ground station downlink

that the various links have different problems. In fact, a good solution for one link can create near intolerable problems for another link, if that same solution is applied there also. For this reason, the hybrid system of Figure V-6 is described here. Some of the factors that influenced the major selections in this implementation are:

1. Large antenna gains available at the ground stations
2. Satellite equipment that is as self-correcting as possible without undue complexity
3. Minimizing the satellite peak power requirements
4. Minimum user equipment cost
5. Multiple access capability
6. Relative freedom from propagation effects.

Table V-1
Power Budget

	Units	125 MHz 25 kHz BW Note 1	125 MHz 25 kHz BW Note 2	400 MHz 25 kHz BW	1600 MHz 25 kHz BW	4 GHz 25 kHz BW	500 MHz 25 kHz BW
I. POWER DELIVERED TO USER RECEIVER							
Sat. Ant. Coupling Loss	-db	0.5	1.0	0.5	0.5	0.5	0.5
Path Loss (Geometric + Ant. Apert)	-db	167.0	167.0	177.0	189.0	197.0	179.0
Tropospheric + Ionospheric Attenuation and Scintillation	-db	0.0	6.0	1.0	0.0	0.0	0.0
Multipath Allowance (See Note 3)	-db	0.0	10.0	3.0	3.0	3.0	3.0
Sat. Ant. Peak to Full Coverage Allow.	-db	3.0	3.0	3.0	3.0	3.0	3.0
User Ant. Peak to Full Coverage Allow.	-db	3.0	3.0	3.0	3.0	3.0	3.0
Polarization Loss	-db	0.0	1.0	1.0	0.5	0.5	0.5
User Antenna Coupling Loss	-db	0.5	1.0	0.5	0.5	0.5	0.5
Sat. Antenna Gain	+db	17.0	17.0	17.0	17.0	17.0	17.0
DBW/Watt Radiation (into isotropic antenna)	-dbw	157.0	175.0	172.0	182.5	190.5	172.3
II. RECEIVER POWER REQUIRED							
Receiver Noise Figure	+db	5.0	5.0	3.0	5.0	5.0	3.0
Receiver Noise Temperature	°K	916.0	916.0	293.0	916.0	916.0	293.0
User Ant. Noise Temperature	°K		730.0	200.0	100.0	100.0	100.0
Total Noise Temperature	°K		1646.0	493.0	1016.0	1016.0	393.0
Total Noise Figure (db/Hz)	+db	29.6	32.1	26.9	30.0	30.0	25.9
Receiver Bandwidth Factor	+db	44.0	44.0	44.0	44.0	44.0	44.0
Noise Density/1°K (Boltzman's Constant)	-dbw	228.6	228.6	228.6	228.6	228.6	228.6
Receiver Noise Power	-dbw	155.0	155.5	157.7	154.6	154.6	158.7
Desired C/N Ratio	+db	1.0	1.0	1.0	1.0	1.0	1.0
DBW Required at User Receiver	-dbw	154.0	154.5	156.7	153.6	153.6	157.7
III. BALANCING THE BUDGET							
DBW/Watt Radiated Less DBW Req. at User Rcvr.	dbw	3.0	20.5	15.3	28.9	36.9	14.6
User Antenna Gain		3.0	3.0	3.0	3.0	3.0	3.0
Satellite Transmitter Power	+db	0.0	17.5	12.3	25.9	38.9	11.6
Satellite Transmitter Power	watts	1.0	56.0	17.0	388.0	2460.0	14.5

NOTES:

1. No allowance was made for variable propagation effects, or for user antenna noise temperature, which is dependent on antenna beamwidth, point elevation, side lobes and geographical location.

2. Probable values for variable propagation effects and antenna noise temperature are included for a wide beamwidth antenna. It is assumed that all effects occur simultaneously in a direction to reduce the signal. Accurate statistical data are not available, but it is estimated that this combination of effects will occur much less than 0.1 percent of the time.

3. In preparing Figure III-1, multipath allowance stated was used for 180° user antenna beamwidth, reducing to near zero at 90° beamwidth. Narrow beam antennas aimed at low elevation angles can experience multipath fading.

Table V-2

Doppler Shift

Frequency	Geostationary	Six Hour
125 MHz	±438	±1200
400 MHz	±1400	±3840
1600 MHz	±5600	±15400
4000 MHz	±14000	±38400

Satellite To Ground
50 Watt Peak Power
Pulse 1 μsecond
Repeated from User

Satellite to User
CSSB-AM Side Tone
Tones Repeated From
Ground Station

Ground Transmission
FM Side Tones
18kHz Highest Tone

User 10^5 watt
Peak Power Pulse
1 μsecond. Pulse derived
from Received Tones
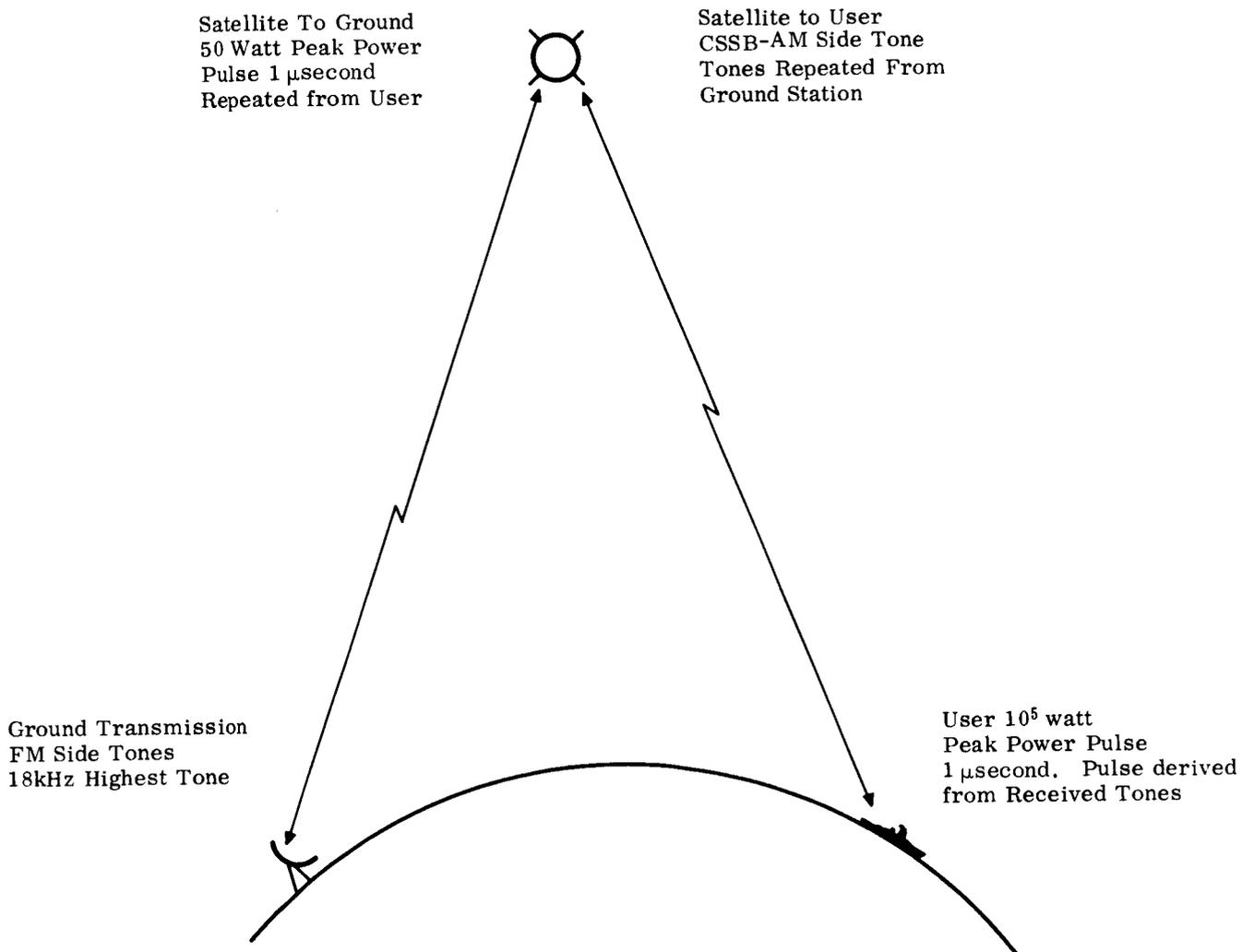
Figure V-6. Hybrid Ranging System

In the light of these factors, the system, as shown, utilizes a multiple FM side tone technique from the ground station to the satellite. From the satellite to the user, the modulation is multiple AM sidetones. On the return path, single pulse transmission is used to increase multiple access as much as possible. As discussed in Section II-E, it appears that the 1600 MHz frequency band is about the lowest frequency that will allow ionospheric disturbances without exceeding the total system error budget. Thus, as an illustrative example, it is assumed the system will operate within the 1540 to 1660 Aeronautical Radio Navigation band.

It is desired that the signal at the user be detectable by a standard AM receiver. As such, a basic amplitude modulation scheme is indicated. To achieve, within the receiver bandwidth, as high a modulation frequency as possible, the AM spectrum on one side of the carrier only (and the carrier) will be transmitted. This then, is a so-called compatible-single-sideband transmission.

Due to the difficulty of achieving high efficiency in linear amplifiers in the microwave region, the satellite transponder will receive, detect and reconstitute a transmitted signal to the user. Perhaps the easiest way to generate the CSSB AM signal is to modulate a signal in the conventional AM fashion and then filter off the unwanted sidebands. The signal is then translated (if necessary) to the transmission frequency and amplified.

The ground transmitter signal will be received in the satellite as an FM signal in a FMFB receiver. By using a feedback receiver, the phase shifts on the tones due to that receiver can be very well controlled. This implementation takes advantage of the high, effective radiated power possible from a ground station. To control the phase shift errors in the satellite receiver, an assumed open loop gain of 40 db at 18 kHz is used. With a 60 db/decade slope on the receiver loop characteristic (and a properly stabilized loop) the minimum noise bandwidth of such a receiver would be 80.5 kHz. Thus, with the gain of the satellite antenna, and considering a 28-foot tracking parabola at the ground station, a power as low as 200 milliwatts at the ground station transmitter would provide an acceptable signal at the satellite.

The demodulated tones are then used to create a transmit signal for the satellite. For illustrative purposes, it is assumed that no linear amplifier with the desired life and efficiency at these frequencies is available. Thus, to generate the signal to the user, the "tone modulated" signal will be made up of five separate signals. These signals can be derived from five, relatively small, separate solid state transmitters, (Figure V-7). The output of the five signals (after they are added on to one line) is sampled by a simple diode detector. The resultant demodulated tones will be compared with the demodulated tones from the ground in phase detectors to generate control signals to control the frequency and phase of the "tone" transmitters. Thus the phase error of the transmitted tones from the satellite can be controlled as a function of loop gain.

At the user, the received 1600 MHz signal will be fed into a tunnel diode preamplifier and then to a mixer, where the signal is heterodyned down to a channel in a receiver like the current VHF-AM receivers (Figure V-8). This receiver would be modified (if necessary) to have the last IF signal available as an output.
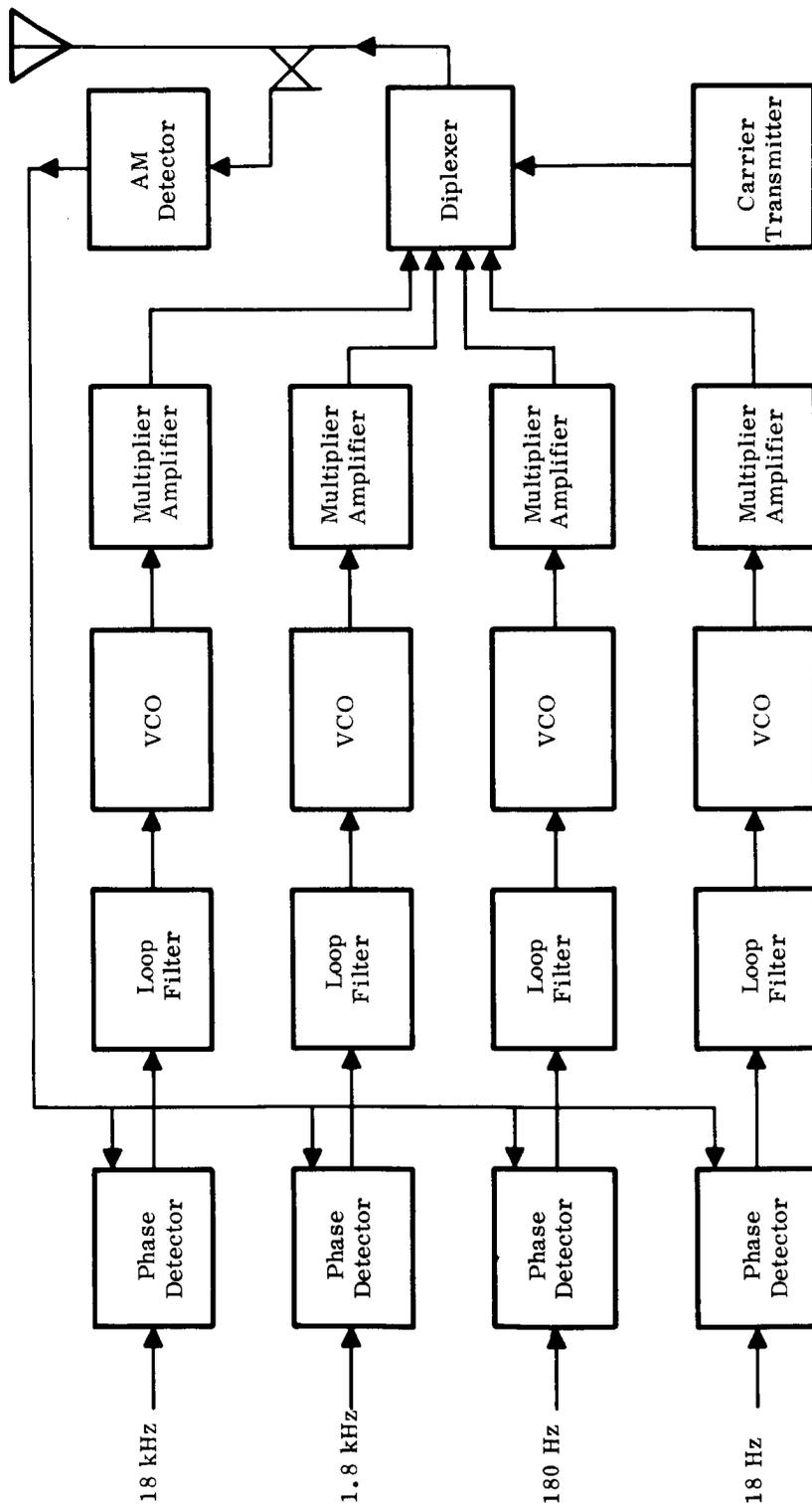
Figure V-7. Satellite to User Transmitter
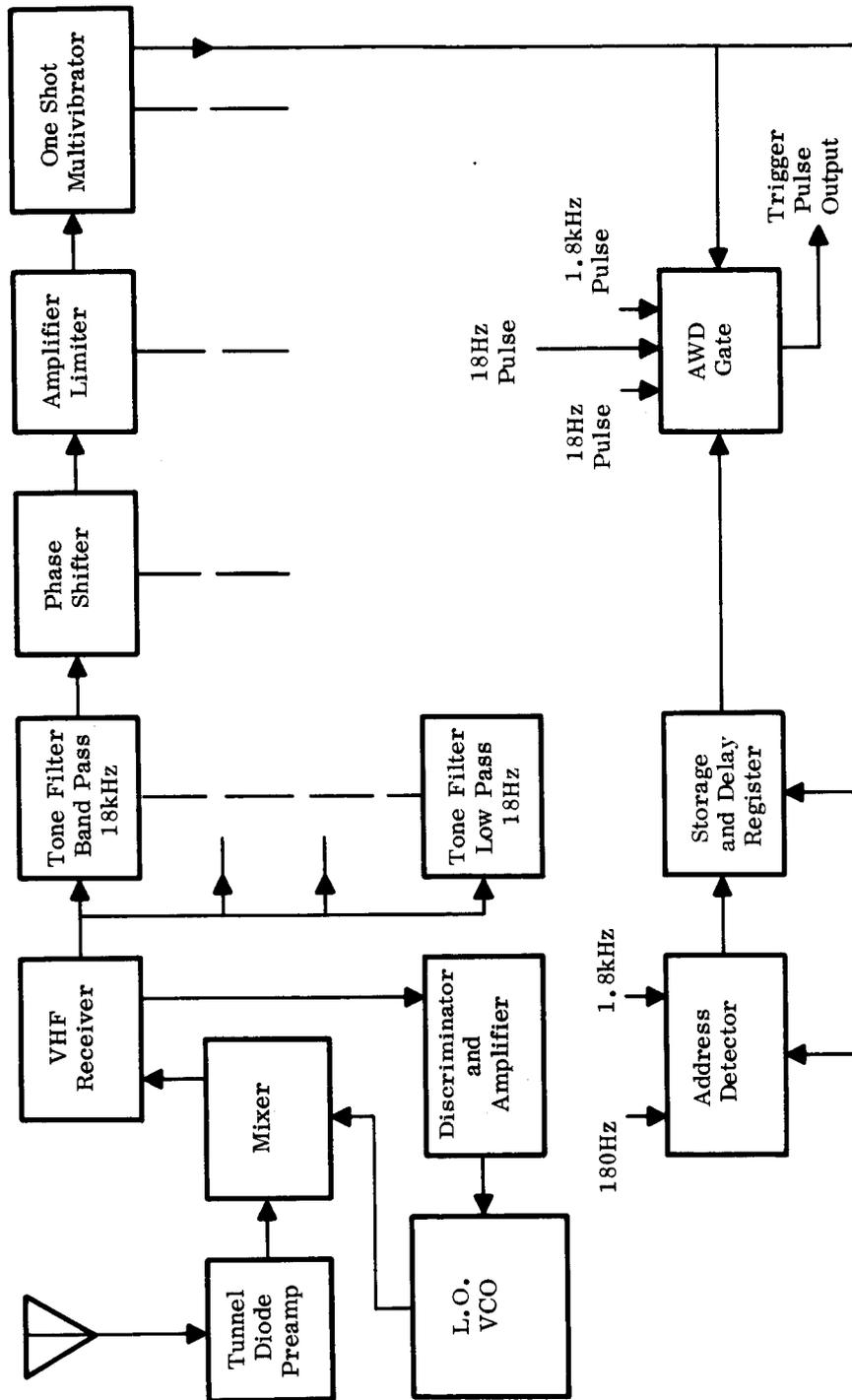
Figure V-8. User CSSB-AM Receiver

The IF signal from the receiver is used to provide an input to a discriminator. As was pointed out earlier, the doppler shift at 1600 MHz plus the possible drift of the receiver could introduce far too much shift (from the receiver IF) in the tones, to permit accurate ranging. To counteract this, the frequency shift and drifts are sensed in the discriminator and the output is used to control the frequency of the local oscillator for the front end mixer. This is the same sort of AFC used on conventional FM receivers. The net result of this frequency loop would be a greatly reduced frequency shift with respect to the receiver IF system. Without the loop, the doppler plus drift could range from 10 to 40 kHz (geostationary—"6 Hr" satellites). A 20 db gain loop could reduce this to 400 Hz shift maximum. This would, (in the previously used worst case) result in one to two degrees of phase shift error (maximum) or a range error of 70 to 140 feet.

The output tones of the VHF receiver would be processed, as shown in Figure V-8, by bandpass filters, adjustable phase shifters, amplifier-limiters and astable multivibrators to generate an output pulse for retransmission purposes. For example, the pulse could be generated when all received tones simultaneously cross the zero volts axis in a positive direction.

It is envisioned that this user receiver could be calibrated quite easily. Figure V-9 indicates a "test set" that would generate an input signal and a time base sweep for a CRT. The vertical axis of the CRT would be fed from the outputs of the multivibrators. Thus, by means of the adjustable phase shifters, the output pulse of the user receiver could be adjusted to a standard time, relative to the phases of the input signal and its modulated tones. This would provide the system with a simple method of achieving a "standard user transponder."
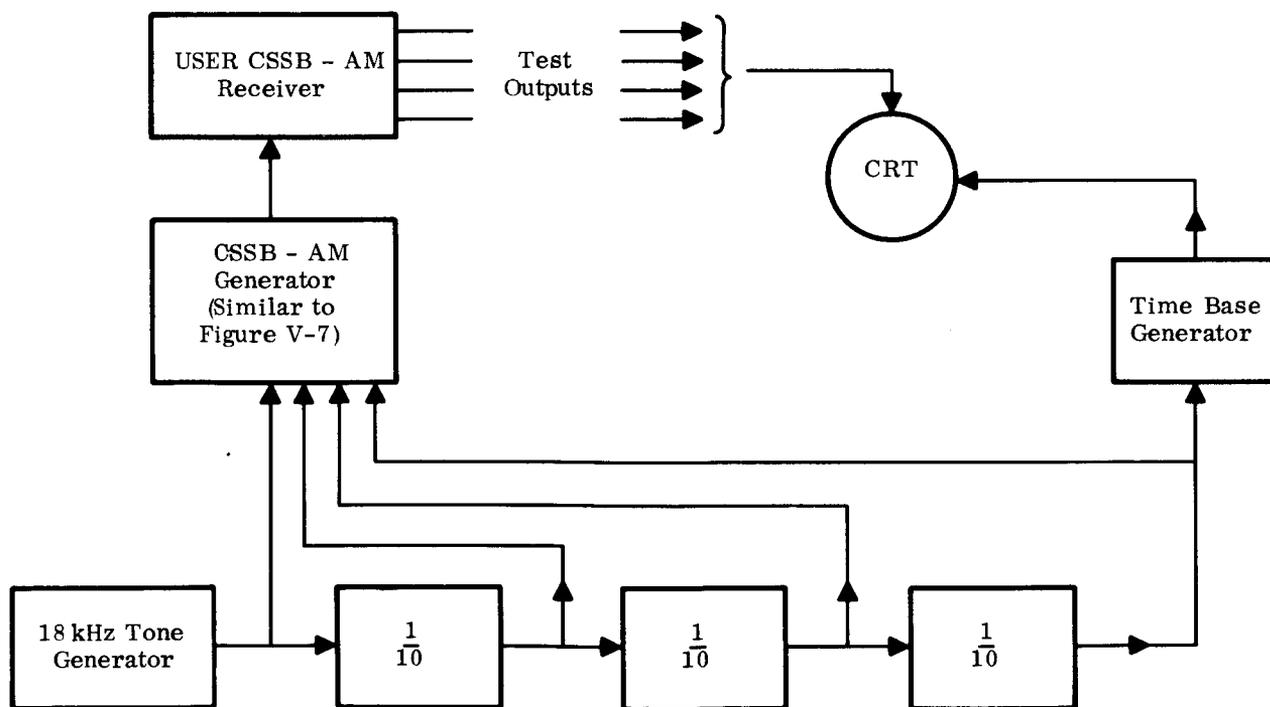
Figure V-9. Test and Calibration Set

With regard to multiple access, the major disadvantage of the tone system is the acquisition time for the returned ranging tones from many users. Therefore, the return signal will be, instead of tones, a single pulse. Since the receiver can be receiving the satellite signals at all times, it does not contribute to acquisition time. With a single pulse return the satellite receiver does not have an acquisition factor. Therefore, this factor has been completely removed from multiple access consideration. Thus, during a span of time covering the possible range differences, the tones will be amplitude modulated to interrogate a block of users. When the user interrogation code matches, a fixed delay after the interrogate code (counted 18 kHz pulses) is received, the "and" gate to the transmitter will be activated, causing a return pulse to the satellite. By keeping the block code and delay time length to within the range difference time, it can be sent on its way while the satellite is listening to the last block of users; thus no extra time is inserted in the multiple access consideration.

As an illustrative example of the address possibilities for the hybrid modulation approach, consider the following implementation, designed to accommodate 1024 blocks of users. Immediately after a range return is made from a block of users, the address is started for the next block. During the first 1/36 of a second, the 1.8 kHz tone will be moved to one of 64 points, 18 Hz wide between 1836 Hz and 4140 Hz. During this time, the user equipment will search for address activation by sensing the output of a tunable filter in the receiver. When the user "reports in" to the system, he will be given an address to set into his equipment. One part of this address is a tunable filter, which is manually set to the proper channel much as one tunes a radio. The second part of the address is made up of an effective "on-off" modulation of the 18 kHz tone during the first 4/90 of a second following the previous range return. This modulation can be accomplished by either moving the 18 kHz tone more than 90 Hz or not, so that output of the receiver filter looks like an on-off keyed signal. By transmitting a "bit" every 1/90 of a second, this modulation can send 4 bits or 16 unique combinations in the 4/90 of a second allocated. This part of the address will be selected by a switch or switches. A diagram of this modulation is shown in Figure V-10. The 1.8 kHz tone will be returned to its nominal frequency and phase for the last half cycle of the 18 Hz tone. For the last 1/90 of a second of that half cycle the 18 kHz tone will also be returned to its nominal frequency and phase. Thus at the end of the 18 Hz cycle, the ones will be in their nominal relationships, so that they all cross zero at the same time triggering a range return from that block of users addressed. Immediately sequence for the next block of users is started.

The user equipment, by definition, is in an environment, as compared to the satellite, that is much more tolerant of maintenance and power needs. This factor, along with the multiple access factor, has influenced the choice of a simple, single pulse transmitter in the user equipment. It is assumed the bandwidth at the satellite receiver will be 200 kHz. Based upon Section III, this would require 2×10^4 watts peak power. For a 5-microsecond pulse, even if interrogated once per second, the average output power would be only 0.1 watt.

At the satellite, the return pulses will be detected and used to trigger a return pulse to the ground. Since this link has approximately a 33 db advantage over the satellite-user link, the peak power necessary will be 50 watts. Here, of course, the duty cycle will be much higher and the average power can approach 8 watts. (This is based on a 6 μseconds one nautical mile time uncertainty of the return pulses; requiring a spacing at least that great in the return block of pulses. This also assumes an ideal interleaving situation.)
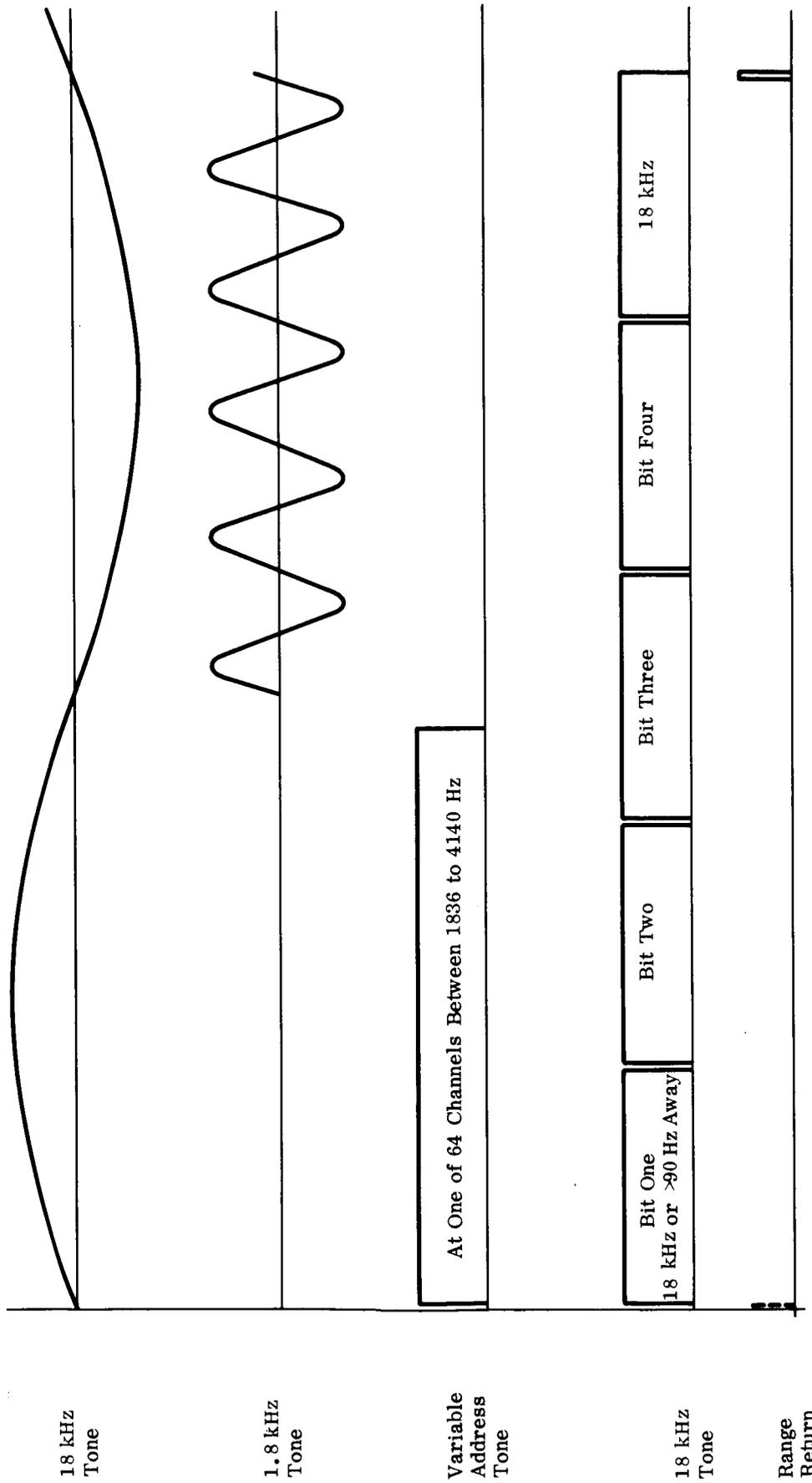
Figure V-10. Hybrid Modulation System Address Timing

18 kHz
Tone

1.8 kHz
Tone

Variable
Address
Tone

At One of 64 Channels Between 1836 to 4140 Hz

18 kHz
Tone

Bit One
18 kHz or >90 Hz Away

Bit Two

Bit Three

Bit Four

18 kHz

Range
Return

The system just described is, with respect to modulation, a hybrid system. It appears, however, that solving the special problems attendant to each leg of the link results in such a system being indicated. To assess the total error of such a system the following error budget is felt to be reasonable for the individual cases specified.

1. GROUND-TO-SATELLITE LINK

Tone control ground transmitter	0.3 degree
Tone control satellite receiver	0.3 degree
Noise jitter	1.0 degree

In both of these cases, phase locked loops are used to control errors, so that errors are decreased by the value of loop gain. In the transmitter case, this includes control of a time reference with respect to the tones. This time reference is used to "measure" the return pulses from the system.

2. SATELLITE-TO-USER LINK

Tone control satellite transmitter	0.1 degree

In this case an overall phase locked loop is used to create the modulated signal.

User receiver tone error	6.0 degrees
Tone filter phase errors	2.0 degrees
Tone and trigger pulse processing	1.0 degree
Tone harmonic distortion	1.7 degrees
Frequency errors	2.0 degrees
Noise jitter	7.05 degrees
RSS TOTAL	9.9 degrees

3. USER-TO-SATELLITE LINK

Trigger-to-RF pulse jitter	20 nanoseconds
Propagation jitter	50 nanoseconds
Satellite receiver noise jitter	500 nanoseconds

4. SATELLITE-TO-GROUND LINK

Trigger-to-RF pulse jitter	20 nanoseconds
Propagation jitter	50 nanoseconds
Ground receiver noise jitter	100 nanoseconds
(1) Ground-to-satellite link propagation jitter	50 nanoseconds
(2) Satellite-to-ground link propgation jitter	50 nanoseconds
RSS TOTAL	511 nanoseconds

Range Errors

9.9 degrees tone phase error	695.0 feet
511 nanoseconds time jitter	511.0 feet
TOTAL RSS RANGE ERROR	862.8 feet

User equipment cost is a critical factor in the system parameters. As such, the equipment shown in Figure V-8 should be as inexpensive to implement as possible. This configuration has been picked to meet this critieria. The use of an AFC loop will compensate for the drifts such that less stable oscillators may be used. For reasonable values of loop gain (up to 100 db), the trade-off of frequency loop gain versus frequency stability rests heavily in favor of loop gain. This, of course, is all the more true when doppler is independent of stability. The system power budget has been based upon reasonable parameters obtainable from passive tone filters. Tone loops, of course, could further decrease the satellite power required. However, the resultant user cost would be quite high.

With these trade-off areas as described, it is estimated that the equipment of Figure V-8 could be implemented for $1,600 each, excluding the costs of the VHF receiver and the basic pulse transmitter. The VHF receiver and pulse transmitter have been excluded from this price, inasmuch, as they are "standard current components." In many cases these components will already be "on-board." This is especially true of the VHF receiver. The price estimated for the signal processing and frequency translation equipment is based on current catalog prices. With the exception of the monolithic digital parts, this equipment is not generally in high production and, as such, the prices tend to be somewhat premium.

A summary of this implementation, providing a list of "specifications" on the various equipments, given below.

Ground Station

Transmitter

Carrier Frequency ~	1600 mHz
Power	10 watts
Tone Frequencies	18 kHz, 1.8 kHz
	180 Hz and 18 Hz

Modulation:

18 kHz—FM Mod Index	1.2
1.8 kHz—FM Mod Index	0.7
180 Hz—FM on 18 kHz	
Tone Mod Index	0.7
18 Hz—FM on 18 kHz	
Tone Mod Index	0.7

Modulation applied progressively as satellite receiver locks up on carrier and individual tone loops.

Address initialed 50 ms prior to all tone simultaneous positive going zero crossings.

Antenna—28-foot tracking antenna Parabolic Reflector.

Receiver—1540 to 1660 MHz

    Bandwidth—matched to one μsec pulse

    Noise Figure $\leq$ 5 db

Satellite

    FM Receiver

        Frequency—same as ground transmitter

        Bandwidth

Carrier Loop	1 kHz
Tone Loop	10 Hz

        Noise Figure $\leq$ 5 db

        Pulse Receiver

            Frequency—same as user transmitter

            Bandwidth—matched to one μsec pulse

            Noise Figure $\leq$ 5 db

    AM Transmitter

        Individual phase and frequency controlled solid state CW transmitters

        Carrier Frequency (f_c) in 1540-1660 band

        Carrier Power—68 watts

        Tone Frequency f_c + 18 kHz

        Power—28.3 watts

        Tone Frequency f_c + 1.8 kHz

        Power—5.65 watts*

        Tone Frequency f_c + 180 Hz

        Power—5.65 watts*

        Tone Frequency f_c + 18 Hz

        Power—11.3 watts*

    Pulse Transmitter

        Carrier Frequency in 1540-1660 band

        Peak Power—50 watts

        Duty Cycle $\leq$ 0.16

        Pulse Width 1 μsec

        All Solid State

    Antenna 17 db gain—earth oriented (Fixed)

User

    Receiver

        "Standard" VHF AM receiver with additional IF output

---

*These powers have been calculated on the basis of 7.5 degrees total rss jitter which is quite conservative for ambiguity resolution.

Converter

    Frequency—same as Satellite AM Transmitter

    RF Preamp   Tunnel Diode

        Noise Figure $\leq$ 5 db

    Local Oscillator—Solid state VCO

                  4 mw output power

AFC Loop

    Gain   > 40 db

    Bandwidth ~ 1 Hz

Tone Processing

    Filters—Passive

Bandwidth			
18 kHz	–	90 Hz	Bandpass
1.8 kHz	–	18 Hz	Bandpass
180 Hz	–	18 Hz	Bandpass
18 Hz	–	36 Hz	Lowpass

    Output Pulse Width   1 μsecond

Transmitter

    Carrier Frequency in 1540-1660 Band

    Peak Power   2 x 10^4 watts

    Duty Cycle < 10^{-6}

    Pulse Width   5 μsecond

Antenna

    140-10-10-degree beamwidth switchable to cover entire hemisphere. Two 10-degree lobes from horizon to +10 degrees, from 10 degrees to 20 degrees above the horizon and a 140-degree lobe across the zenith.

T-R switch receiver protection during transmission.

APPENDIX A

A MULTIPLE ACCESS APPROACH

As pointed out in Section IIC, a method of intermixing of messages could provide a multiple access capability, which would improve the time figure of merit. A possible means of implementing this approach is presented in this appendix. This discussion uses, as an example, a pulse coded system; however, through modifications in the basic techniques, the same principles can apply to other modulation techniques.

For a satellite at 33 degrees elevation relative to a ground terminal, the slant range is 125×10^6 feet. Slant range from a user to the satellite varies from 117×10^6 to 138×10^6 feet over the useful service area. Thus, the one-way range sum may be from 242×10^6 to 263×10^6 feet.

The round trip travel time is between 490 milliseconds and 535 milliseconds. Therefore, if it were necessary to complete the interrogation-response cycle of a specific user before beginning the next, the minimum interrogation rate would be less than two per second, or about 6750 per hour.

Two levels of improvement are possible.
1. Interrogations can proceed without completing the cycle, allowing a "guard interval" for each response of 50 milliseconds (the amount the transit time may vary plus some margin for time to actually execute the interrogation and response). The resulting rate may be as high as 20 interrogations per second, or 72,000 per hour.
2. The next level of improvement exploits the facts that
   a. the actual time to receive a response can be significantly less than the 50 millisecond guard interval, and
   b. due to differences in range, the transmit time between interrogation and response is different for various users. Ordering the interrogation sequence so as to prevent interference requires control of the sequence by a computer, using available data to predict actual time of arrival of responses. The remainder of this section addresses this factor and describes a scheme for ordering the interrogation cycle to achieve maximum efficiency.

The ground station transmits to all users, via the satellite, blocks of pulse-coded data. Each block consists of an address word, an advisory data word, and a sync word.

The repetitious sync word is utilized by all users to generate bit synchronization for decoding the address and data words. When a user equipment recognizes its address, it responds by coherently repeating back, via the satellite, the next following sync word, which upon receipt is used by the terminal station to determine range sum.

At any point in time, each user bears a priority assigned by the computer. Also some interrogation addresses will have been transmitted, with the upcoming times of arrival of the corresponding responses predictable by the computer from previous position data. These we designate as "committed" responses.

The computer also predicts the time of arrival of responses of each of the top n users (n is to be determined) that are to be interrogated at the next opportunity, and it designates for interrogation the highest priority user, who can respond without interfering with a committed response.

As soon as a response has been received from a specific user, his priority is dropped to last, from which he works back up toward the highest. (In some cases, certain users may require a higher average measurement rate than others. In that case, such "blue chip" users can be reinserted higher up in the standings upon completion of a measurement, rather than dropped to last.)

The maximum overall interrogation rate depends upon the data block rate, which of course depends upon the brevity of the data block.

Typical operation is illustrated in the example which follows. The assumed system parameters are for illustrative purposes only, and do not constitute recommendations.

Assumptions:

1. Bit Rate $= 12 \times 10^3$ bits per second

2. Round trip transmit time:
   - minimum $=$ 490 milliseconds
   - maximum $=$ 535 milliseconds

3. Ground Terminal Transmission:
Address	$=$	12 bits
Data	$=$	12 bits
Ranging word	$=$	16 bits
Total	$=$	40 bits
Block length	$=$	3.33 milliseconds

4. User Transmits (upon demand):
Range word	$=$	16 bits
Data (altitude)	$=$	12 bits
Total	$=$	28 bits
Block length	$=$	2.33 milliseconds

Refer to Figure A-1. The time-axis indicated by (T) indicates transmissions at the ground terminal; (R) indicates received events at the same ground terminal.

Prior to t_o, the computer must determine which user address A_o to transmit at time t_o in order to get a suitable response keyed to R_o, the range word. (All R are identical, while A and D are not.)
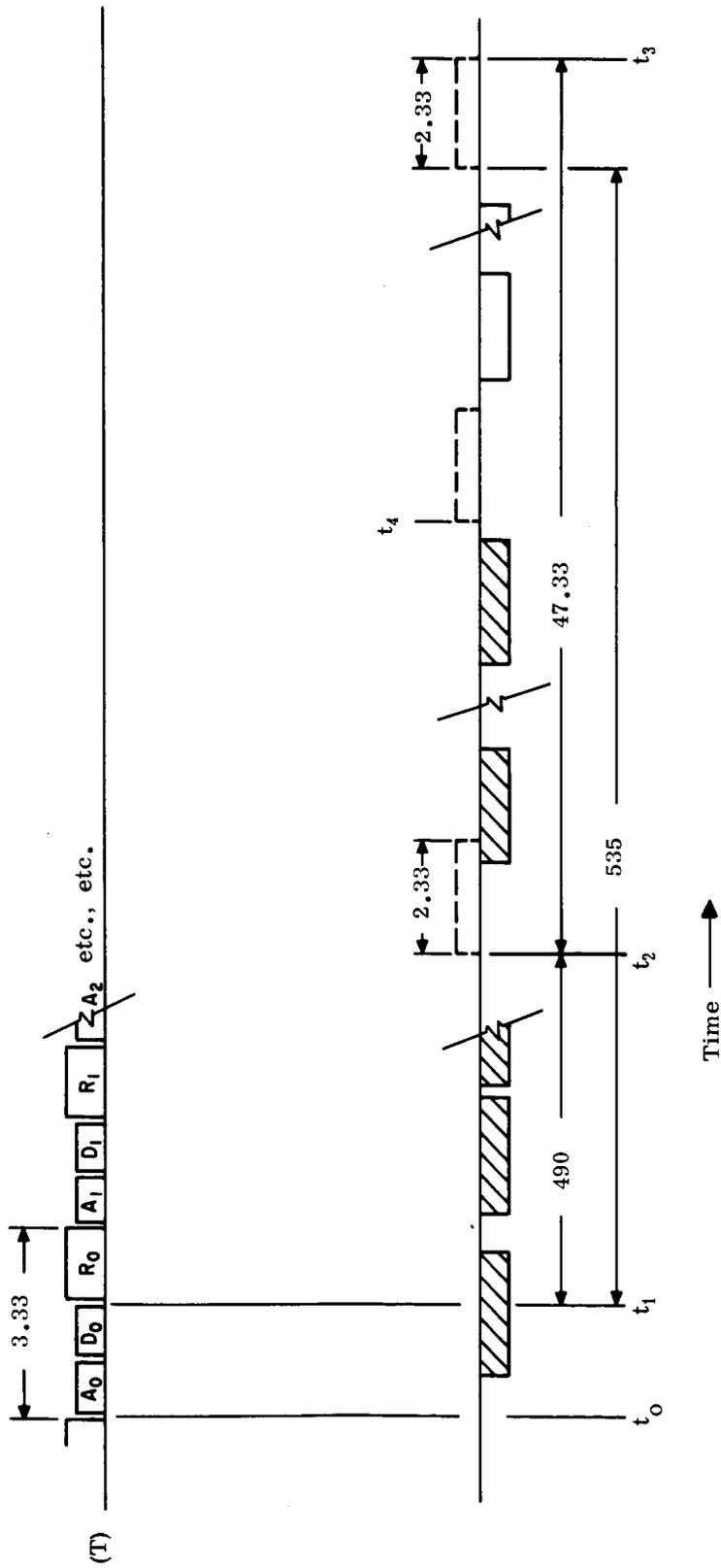
Time ———▶

Time intervals indicated in milliseconds.

Figure A-1. Multiple Access Timing

A number of responses (cross-hatched) are already committed—the appropriate addresses having been sent. Using last measured position (plus flight history, if necessary), the computer has predicted the arrival times of the cross-hatched data blocks.

It is clear that any response to R_o (suggested by dotted blocks) must occur during the interval between t_2 and t_3, so attention can be directed to this interval.

The computer starts with the top priority user, and from last measured data predicts when (during the $t_2 - t_3$ interval) his response to R_o would occur. If it is in a clear interval, such as t_4, that user's address is sent as A_o, and the computer turns its attention to the next block, keyed to R_1. If a response to R_o from the top priority user would interfere with a committed (cross-hatched) response, the computer tests the next priority user, and so on, until compatibility is found.

If it should happen that no match can be made, a dummy address (for example, all zeros) is sent as A_o, which results in no response, and attention is directed to A_1.

When a high priority user is passed by, he maintains his position in the standings. Similarly, if a response is missed, that user remains near the top. (A prolonged series of missed responses from a given user would be cause for alerting the system operator.)

A user accidently responding when not addressed, would probably interfere with another response, causing it to be missed, in which case it would be treated as above.

During periods of low system utilization, many of the address and data words may be omitted (all zeros). If a PAM system is selected, this would conserve power throughout the system. The repetitious range word would be sent, however, to permit all users to remain synchronized and ready to receive when addressed.

In the foregoing example, the 12-bit address allows for approximately 4000 users. Since the data block for responses was assumed shorter than the interrogation data block, the latter establishes the system interrogation rate, which is 300 interrogations per second, or over a million per hour, for this example.

If the full complement of 4000-plus users were active, the average interval between interrogations is 13.5 seconds.

The number of users the system can accommodate can be doubled by adding one bit to the address word (at the expense of the data word). Since the system interrogation rate remains the same, it is clear that the average interval between interrogations for a given user would also be doubled to 26 seconds.

APPENDIX B

WAVEFORM DESCRIPTIONS AND PARAMETERS

This appendix outlines the use of complex notation in waveform analysis and gives definitions of bandwidth, time duration and phase structure parameters of waveforms. The relationship of these parameters to the ranging accuracy is also given.

Although physical signals are real, much of the analytical work with matched filters utilizes complex signals for mathematical convenience.

A physical signal $s_p(t)$ can be written in terms of complex waveforms as

$$s_p(t) = \text{Re}\left[\mu(t)e^{j\omega_o t}\right] \tag{b-1}$$

where $\mu(t)$ is called the complex envelope and $\text{Re}[\ ]$ means the real part of the bracketed quantity. The complex envelope can be written as:

$$\mu(t) = a(t)e^{j\phi(t)} \tag{b-2}$$

where

$\quad a(t) \quad = \quad$ envelope of $s(t)$

$\quad\quad\quad\quad = \quad$ real quantity > 0

$\quad \phi(t) \quad = \quad$ phase modulation function of complex envelope

$\quad\quad\quad\quad = \quad$ phase modulation function of carrier.

Using the above relationships, $s_p(t)$ can be written:

$$s_p(t) = a(t)\cos\left[\omega_o t + \phi(t)\right] \quad . \tag{b-3}$$

This is the observable waveform. The important point in the above description of $s_p(t)$ is that all the necessary information about the waveform is contained in the complex envelope $\mu(t)$. All that is missing is the carrier frequency and that is important primarily in calculating the magnitude of doppler shifts.

It is convenient to describe the performance of pulse trains and other waveforms in ranging applications in terms of general definitions of bandwidth, time duration and phase structure. Such definitions are given below, along with the relationship of these parameters to the attainable ranging accuracy. In general, these definitions give values different from more conventional measures of the parameters, such as 3 db bandwidths and time between half power points. Helstron (6) gives a good treatment of the derivation of the following equations.

Bandwidth $= \beta_0$

$$\beta_0^2 = \frac{(2\pi)^2 \int_{-\infty}^{\infty} f^2 \, |M(f)|^2 \, df}{\int_{-\infty}^{\infty} |M(f)|^2 \, df} \quad \text{(radians}^2 \text{ per second}^2\text{)} \quad . \qquad \text{(b-4)}$$

Time Duration $= t_0$

$$t_0^2 = \frac{(2\pi)^2 \int_{-\infty}^{\infty} t^2 \, |\mu(t)|^2 \, dt}{\int_{-\infty}^{\infty} |\mu(t)|^2 \, dt} \quad \text{(seconds}^2\text{)} \quad . \qquad \text{(b-5)}$$

Phase Structure $= \alpha$

$$\alpha = \frac{2\pi \int_{-\infty}^{\infty} t\phi'(t) \, |\mu(t)|^2 \, dt}{\int_{-\infty}^{\infty} |\mu(t)|^2 \, dt} \quad . \qquad \text{(b-6)}$$

Range Delay Measurement Error $= \sigma_\tau$

$$\sigma_\tau = \frac{1}{\beta_0 \sqrt{\dfrac{2E}{N_0}}} \frac{1}{\sqrt{1 - \dfrac{\alpha^2}{\beta_0^2 t_0^2}}} \quad \text{(seconds)} \quad , \qquad \text{(b-7)}$$

where

$\mu(t)$ = complex envelope

f = frequency in Hz

$M(f)$ = Fourier transform of $\mu(t)$

$\phi(t)$ = phase function of $\mu(t)$

$\phi'(t)$ = $d/dt \left[\phi(t)\right]$

E = energy of ranging waveform

$2E = \int_{-\infty}^{\infty} |\mu(t)|^2 \, dt$

N_0 = white, Gaussian noise power density in watts/Hz = kT

σ_τ = rms error in the measurement of range delay, τ.

Some comment on these formulas is needed. Equations (b-4) and (b-5) compress all the information about the pulse train spectra and envelope into two numbers, β_o and t_o. While these are analytically convenient, far more detail may be required in a specific application. For example, the maximum value of the spectrum as a function of frequency may be prescribed. It is assumed in Equations (b-4) and (b-5) that the frequency origin and time origin are chosen such that the mean frequency of the waveform and mean time are both zero, i.e., the frequency origin is in the "center" of the waveform spectrum and the time origin is at the "center" of the pulse train.

The constant α from Equation (b-6) is a measure of the weighted phase structure within the waveform or pulse train. $\phi'(t)$ is the instantaneous frequency as a function of time. Equation (b-6), therefore, can be thought of as the weighted correlation between time and phase. α tends to be large, for example, when large linear frequency shifts take place within a waveform.

Finally, Equation (b-7) relates those quantities to the rms error in the range delay measurement. Equation (b-7) must be applied carefully. It is applicable when the energy in the ranging waveform, e.g., a pulse train, is appreciably larger than the noise power density N_o. Note that it is the total waveform energy (not peak or average power as such) that is important. Also, it does not take into account errors introduced by spurious responses at a matched filter output. Such responses are illustrated by the ambiguity surfaces, discussed in Appendix G, which provides a convenient means of evaluating the possibility of error.

In words, Equation (b-7) states that the range error decreases directly with bandwidth, only as the square root of energy, and that large values of α will cause an increase in range error. This means that energy is minimized, for a fixed range error, by making the bandwidth large and the exchange of bandwidth for energy is very favorable. Doubling bandwidth decreases the required energy by four (when $\alpha = 0$).

APPENDIX C

FM PULSE COMPRESSION

A. TECHNIQUE DESCRIPTION

This modulation technique is a means of achieving trade-offs between the parameters of a single pulse of RF energy. By transmitting a pulse that has been specially shaped in time and frequency and using a receiver that is matched to this special shaping, a radar can overcome some device characteristics that would ordinarily limit its performance.

As an example, a common application is that of achieving very high time resolution in a radar signal. In most microwave tubes, there is a limit to the narrowest pulse that can be transmitted due to the tube characteristics. Where shorter video pulses are required, a swept frequency pulse and its associated matched filter receiver can accomplish the desired results. Figure C-1 shows in block form such operation. Even though the transmitter cannot transmit a pulse as short as that generated by the pulse generator, advantage can be taken of the known frequency-time relations of the transmitted and received signals to re-create the initial short pulse for the ranging circuits. As might be expected, the better the match between the transmitted pulse characteristics and the matched filter, the better the recreation of the initial trigger pulse. For this reason pulse compression systems very often use the same element in a time shared fashion to shape the transmitted signal and to recreate the short pulse for the received signal. Thus, in Figure C-1, the "Pulse Time-Frequency Shaping" element and the "Matched Filter" element become one and the same. These networks can be designed so that no matter which form is used for the input of the network, the output will be the other form of the reciprocal pair. In this way, a very high degree of match between the received signal and the matched filter can be achieved.

The general method of implementation for FM Pulse Compression is based upon the network shown in Figure C-2. When applied to the transmitter, a short pulse is used to excite a band of filters with adjacent bandpasses. The filter outputs are then applied to a delay line, which inserts progressively longer delays across the total band of the adjacent bandpass filters. The output of the delay line is, therefore, a stretched pulse progressively rising or falling in a step fashion in frequency.

Applying the same dispersive element in the receiver will invert the process. Assuming the network of Figure C-2 is designed to yield an approximation to the time frequency plots of the transmitted signal of Figure C-1, then the highest center frequency filter will have the longest delay. To invert the process, the local oscillator in the receiver is placed on the high side of the received signal. This will cause an inversion of the frequency-time relation of the signal out of the receiver. In terms of Figure C-1, the slope of the frequency-time plot for the received signal will be inverted to a negative slope. Thus, the earliest frequency applied to the filter bank of the dispersive element of Figure C-2 will be the highest frequency.
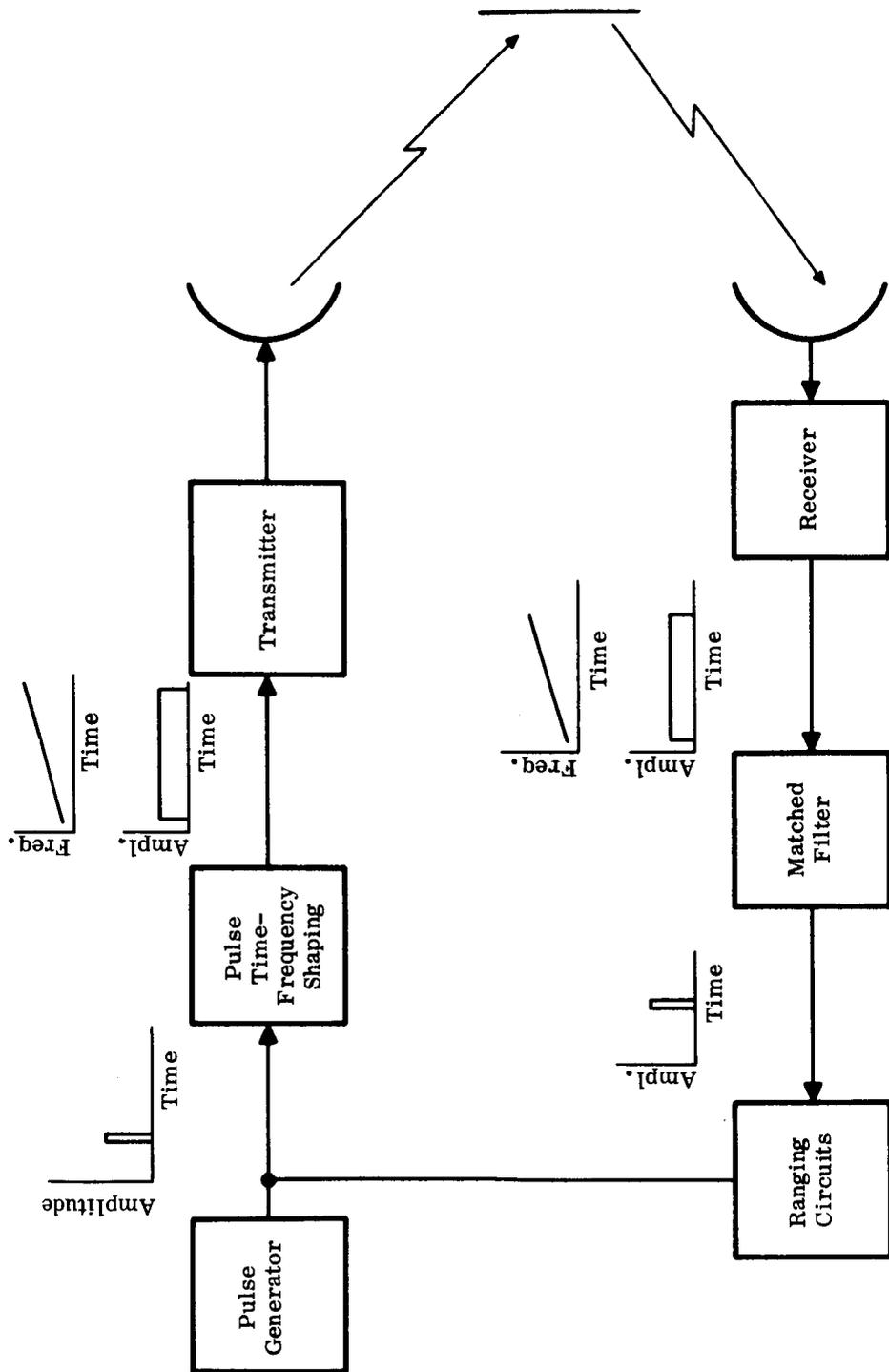
Figure C-1. Pulse Compression Radar
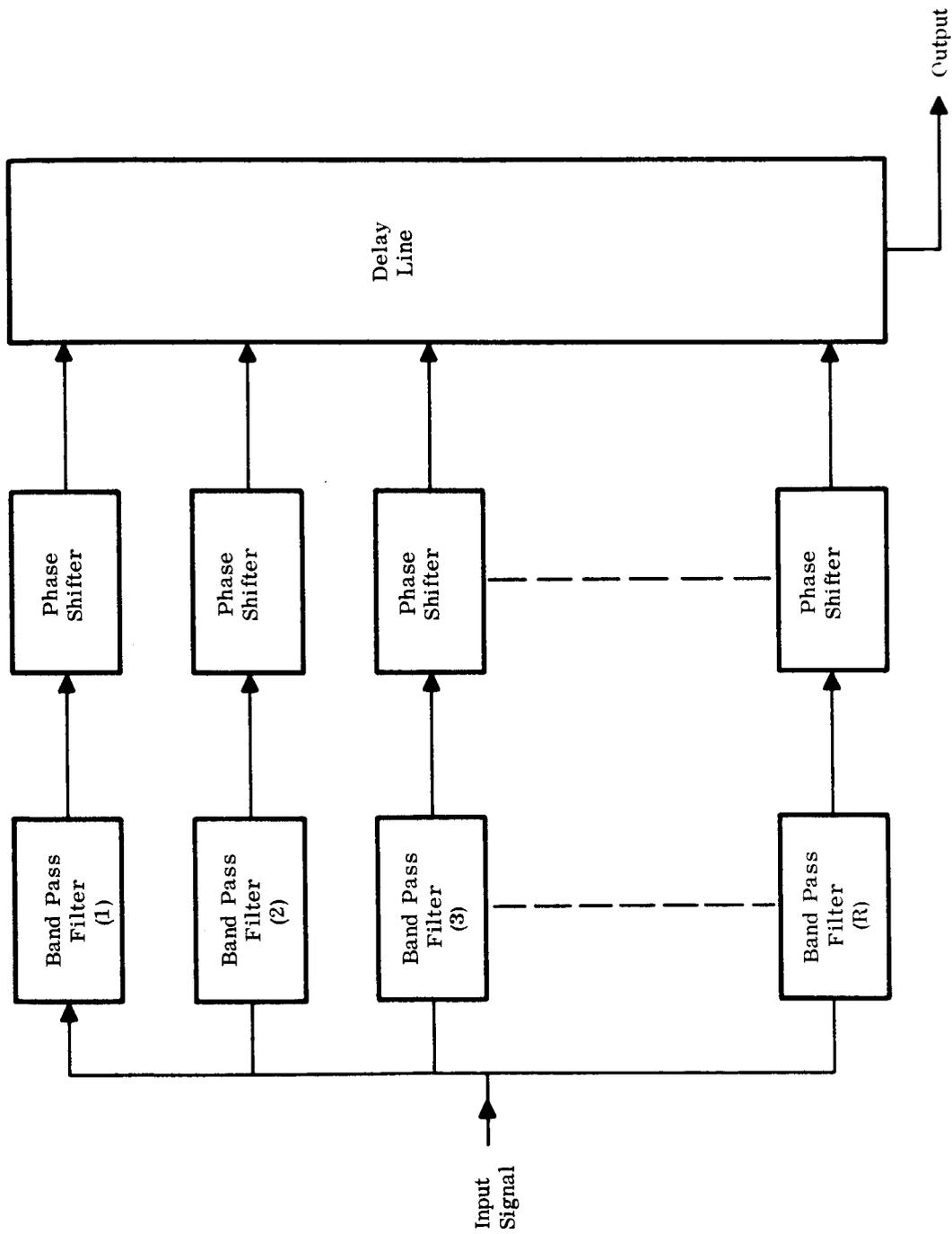
Figure C-2. Dispersive Element

When the entire received signal has been applied to the filter band and the output arrives at the delay line output, the energy will have been recombined into a short pulse of IF energy. The phase shifters, shown in Figure C-2, are used to make fine adjustments of the delay line length to assure that the energy at the delay line output adds in phase. It is necessary that the ratio of the center frequency of the filters to their band-pass be such that, over the time of the trigger pulse width, there is not a significant loss of phase coherence between the highest and lowest frequency.

B.   PERFORMANCE ANALYSIS

To evaluate the trade-off performance of the system, some relationships must be defined. Let:

B_b =   Total bandwidth occupied by the transmitted signal

B_f =   Bandwidth of an individual filter, as shown in Figure C-2

R =   Compression ratio = B_b/B_f

t_t =   Width of trigger pulse (transmit mode)

t_p =   Width of transmitted pulse

t_D =   Time required for detection of received signal.

It is assumed, here, that the delay line is so arranged, such that:

$$\frac{t_p}{t_t} = R \quad . \tag{c-1}$$

In consideration of the lower frequency allocations (VHF) and/or low cost user equipment, the signals used for ranging are going to be bandwidth limited. As a result, in considering FM Pulse Compression, this constraint requires the technique be considered in the light of limiting the total bandwidth for the transmitted signals to be comparable to the bandwidth required by other modulation techniques. This is an unusual con-straint for FM Pulse Compression.

The signal-to-noise ratio of the output of the FM Pulse Compression receiver can be assessed by examining the dispersive elements. Assuming, for the moment, constant peak power transmitted, both with and without compression, the signal voltage out of one of the filters should be the same as the signal voltage out of a conventional receiver of the same gain and antenna capture area. The output noise of the filter, however, will be diminished by 1/R. This assumes the same noise density in both receivers. Thus, the voltage output signal-to-noise ratio of one of the filters is:

$$(S/N)_V = \frac{V_s}{\dfrac{V_n}{\sqrt{R}}} = \frac{\sqrt{R} \; V_s}{V_n} \quad . \tag{c-2}$$

Further, it is assumed the FM Pulse Compression receiver is built so that the signal voltages from the filters all add in phase at the end of the delay line, when it is used as a receiver element. It is also assumed that the noise voltages from the filters arrive at the end of the delay line in a totally uncorrelated fashion. Under these conditions, the output of the delay line with respect to signal will be:

$$V_{s_o} = R\, V_s \tag{c-3}$$

Also, the uncorrelated noise voltage at the output of the delay line will be:

$$V_{n_o} = \sqrt{R}\, V_{n_f} \tag{c-4}$$

where, V_{n_f} is the noise output voltage of any one of the filters. Combining these terms, the total output signal-to-noise voltage is:

$$(S/N)_{T_V} = \frac{R\, V_s}{\dfrac{\sqrt{R}\, V_{n_f}}{\sqrt{R}}} = R \cdot \frac{V_s}{V_n} \tag{c-5}$$

Therefore, the output power signal-to-noise ratio is:

$$S/N = R^2\,(S/N)_P \tag{c-6}$$

where, $(S/N)_P$ is the output power signal-to-noise ratio of the hypothetical single pulse receiver. Thus, the peak power transmitted can be reduced by the factor R^2 for the same output signal-to-noise performance.

The FM Pulse Compression scheme, under the constant bandwidth constraint, also has its drawbacks. Due to the matched filter processing, Woodward's relation can be used:

$$\delta\tau = \frac{1}{\beta\,(S/N)^{\frac{1}{2}}} \tag{c-7}$$

(which holds over a reasonable range of practical signal-to-noise ratio) and the factor $\delta\tau$ can be evaluated. This factor is important, as it directly represents the measurement error achievable in the system. The factor, β, in the relation, is the proportional to B_f of the FM Pulse Compression receiver; just as β is proportional to B_b of the "conventional" receiver.

Thus for the case of equal output signal-to-noise power ratio, the measurement error of the FM Pulse Compression system will be R times that of the "conventional or single pulse" system. Within the range of signal-to-noise ratios, where Woodward's (1) relation holds, a trade-off between signal peak power and measurement accuracy can be made. It must be recognized, however, that the error improves with the square root of the power increase.

Another factor of system importance is the detection time. Since traffic handling capacity can be of importance to the overall system, the time required for a detection (and range measurement) is an important factor in the modulation technique performance.

Assuming the predetection filter shown in Figure C-2 acts in a fashion similar to a single section filter, the output voltage will behave as:

$$e_o = K' \left(1 - e^{-\omega_f t}\right) .$$ (c-8)

For the case where the previously derived power advantage is not compromised in the power-accuracy trade-off, the time a received signal must remain within the bandpass of a given filter is that time required to allow the output voltage to rise to the detection threshold. Thus,

$$t_D = \frac{K_t}{\omega} .$$ (c-9)

The required time is inversely proportional to the bandwidth involved. From this, a comparison of FM Pulse Compression with the single pulse technique can be made:

$$t_{D_c} = \frac{B_b}{B_f} t_{D_P} = R \cdot t_{D_P} .$$ (c-10)

In addition to the "dwell time" within a filter, the signal must traverse R filters; thus, the total detection time is:

$$T_D = R \cdot R t_{D_P} = R^2 t_{D_P} .$$ (c-11)

In this instance, if the accuracy penalty is held constant, a trade-off between dwell time and peak power can be made. Since this trade-off does not affect the delay line length, the trade-off involves a decrease in detection time, T_D, with the square root of the peak power increase. Conversely, if the peak power is kept high, a direct accuracy-detection time trade-off can be made. It is important to note that the full advantage of reduced peak power comes only at the full sacrifice of accuracy and detection time.

The time figure of merit for the modulation technique, with respect to a single pulse technique, can be evaluated as a function of R. For both the single pulse and FM Pulse Compression techniques, the message length and detection times are very nearly one and the same. Thus, the intermixing factor:

$$M = \frac{T_P}{T_M + T_A + T_D + T_G}$$ (2-4)

reduces to

$$M = \frac{T_P}{T_M + T_A + T_G} .$$ (c-12)

C-6

Furthermore, the acquisition time is nearly zero. Also, in both cases, the guard time, T_G, probably reaches a minimum near the value of one to three times the message length. So:

$$M \cong \frac{T_P}{J\,T_M} \qquad\qquad (c\text{-}13)$$

where the factor J, accounting for message length plus guard time, probably ranges from two to four. From the previous analysis, we can say that:

$$M_C = \frac{M_P}{R^2} \qquad\qquad (c\text{-}14)$$

where M_C is the FM Pulse Compression factor and M_P is the factor for the single pulse technique. In both techniques, there is not a significant amount of integration performed, in addition to the message length, itself. Therefore:

$$\eta = \frac{T_M}{T_M + T_A + T_D} \cong \frac{T_M}{T_M} \cong 1 \quad . \qquad\qquad (c\text{-}15)$$

Thus, the data gathering efficiency is independent of the compression. The overall time figure of merit for FM Pulse Compression is approximately:

$$F_{TC} = \frac{F_{TP}}{R^2} \quad . \qquad\qquad (c\text{-}16)$$

C. IMPLEMENTATION CONSIDERATIONS

In the usual implementation of a FM Pulse Compression modulation, the circuits of Figure C-2 are inserted sequentially first in the transmit trigger channel and second in the receiver detection channel. Toward the antenna end of the equipment, the signals in both the transmitter and receiver are amplified and heterodyned to and from the proper operating frequency (Figure C-3). Thus, except for the somewhat unusual transmitter technique of heterodyning to the operating frequency and amplifying to the desired power level, the "extra" equipment necessary is as pictured in Figure C-2. In most cases, however, the FM Pulse Compression technique has been used on "skin trackers." Here, however, a transponder at the user site at least, is inserted in the transmission path. Thus, the earlier assumptions of the received signals adding in phase at the output of the delay line must be examined carefully. One of the major reasons for using the same circuits in the transmit channel and in the receive channel is to keep the received signal as well controlled as possible in amplitude and phase. There are three possible cases to examine: (1) The user equipment detects the received signal and constructs a similar transmit signal, (2) The user equipment detects the received signal and transmits a different type of signal and, (3) The user equipment merely frequency offsets the received signal and amplifies it for retransmission.
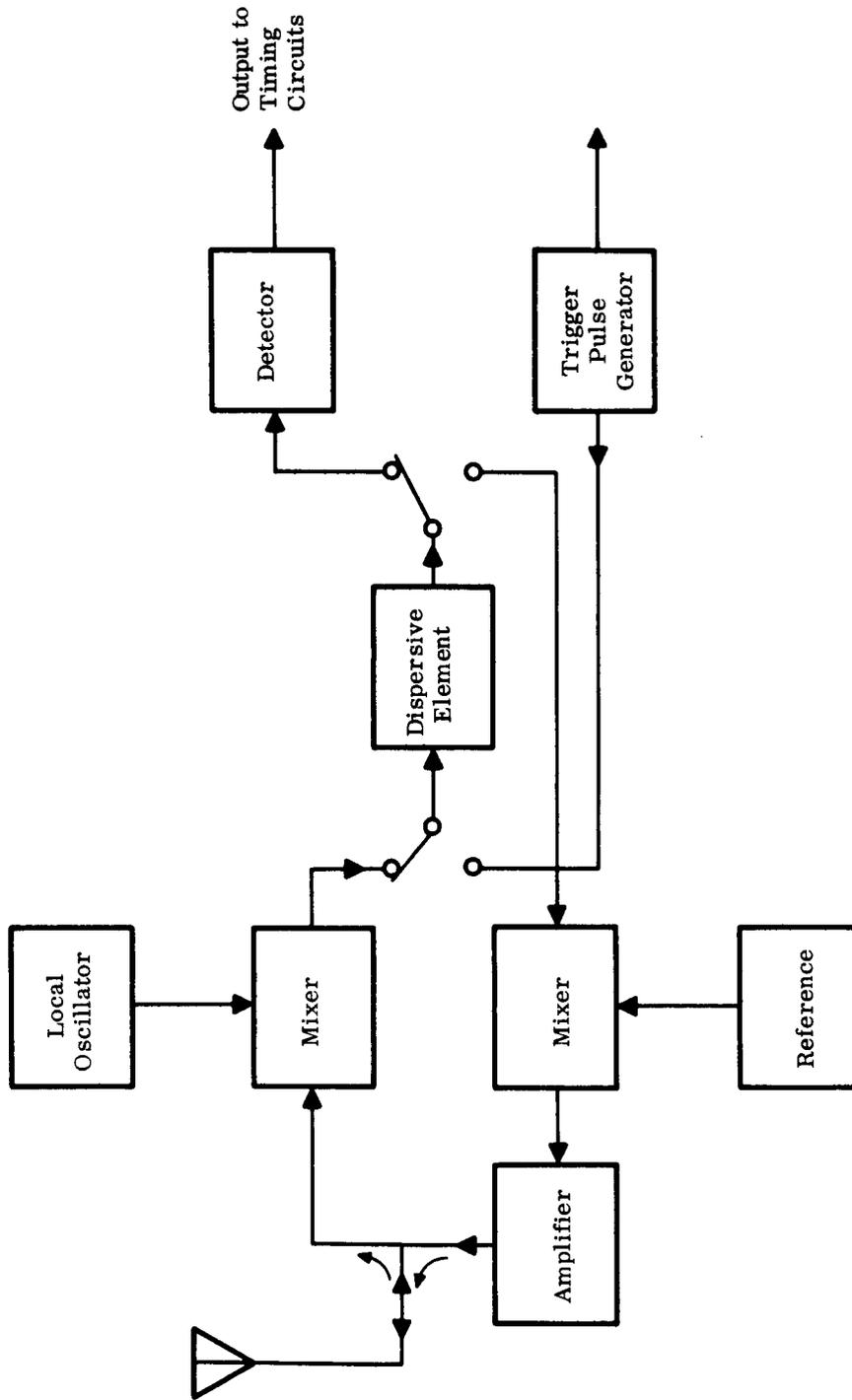
Figure C-3. FM Pulse Compression Radar

In the first case, insomuch as the characteristics of the transmit dispersive element (Figure C-2) of the satellite can be duplicated at each user, the full advantage of the trade-off functions can be realized at the user receiver. While no data has been found on this particular aspect of the problem, the realization can probably be near theoretical. With the small number of satellites that would be used, it is reasonable to talk of having all satellite transmit dispersive elements very closely matched. Under these conditions, the user receiver, when installed, could be phased properly when receiving satellite transmission and thus, be properly phased for all satellites by a screwdriver adjustment.

On the return path, however, the satellite receiver must operate with literally thousands of user transmitters. To expect that these user transmitters, in addition to being low cost, would all be precisely matched to a common standard is unreasonable. This very real problem which would definitely degrade performance, along with the realization that high peak power is not of as crucial a concern with the user transmitter compared with the satellite, leads one to case two.

In this case the receiver situation for the user would be exactly as in case one. However, the received pulse at the end of the delay line would be used to fire a simple pulse transmitter for return trans-mission to the satellite. Thus, the FM Pulse Compression would only be used on the satellite-user down link.

In the third case, the user equipment would consist of a non-detecting receiver that would be used to drive a mixer-amplifier for signal re-transmission. This is, of course, an attempt to make the user return look like a frequency offset "skin return." In order to avoid degradation of the phase characteristics of the signal, the receiver would probably have to be rather broad band. Unfortunately, if the receiver bandwidth is broadened to minimize degradation of the transponded signal, the peak power saving is compromised. Thus, the implementation trade-off here would be received bandwidth, satellite peak power and user return transmitted noise. In this case, it may well be worth while to have the user transmitted signal not completely noise-free in order to ease the peak-power-bandwidth problem. Here, the satellite receiver processing is relied upon to free the resultant ranging signal from noise.

From a preliminary look at the implementation problems, it appears that the second case is the most desirable case.

With the second case of FM Pulse Compression on the satellite-user down link and single pulse on the up link, it appears that some reasonable compromises have been made. The circuits of Figure C-2 are additional items for the user receiver. The satellite peak power can be reduced. The user transmitter and the user-satellite up link receiver can be simple pulse-type equipments. The user receiver and the satellite-user down link transmitter will need to be R times better in frequency stability than for a single pulse scheme. This stability criteria, however, with present state-of-the-art techniques is not a severe constraint. For example, current typical specification, for aircraft receivers, require 20 parts per million frequency stability, while one-half part per million oscillators without ovens, are readily available today.

One aspect of the use of FM Pulse Compression that must be recognized, however, is that the accuracy suffers as peak power is reduced. This is true due to the fixed total bandwidth constraint. At VHF and low UHF frequencies, the fixed total bandwidth is a real constraint due to the limited spectra available. It is, also, at these bandwidths and frequencies that a degradation in accuracy is least acceptable. Thus, the application of FM Pulse Compression is most likely in the areas of the spectrum where bandwidths of 500 kHz and more are available. This is, fortunately, where peak power reduction is most needed.

APPENDIX D

FM-CW TRIANGULAR MODULATION

A. TECHNIQUE DESCRIPTION

Frequency modulated-continuous wave triangular modulation radar is used in some radio altimeters. It can provide high resolution at close range. As an example, the AN/APN-22 radar altimeter provides an accuracy of ±2 feet from 0 to 40 feet and an accuracy of ±5 percent of altitude from 40 to 20,000 feet. (Reference [13].) Altitude of the aircraft is measured by sweeping a transmitted frequency back and forth between two values. The frequency received back from ground reflection thus differs from the frequency being transmitted by an amount proportional to the propagation time and, thus, the altitude of the aircraft.

The same concept may be applied to ranging from a satellite. A ground station transmits a signal that is continually swept between two radio frequencies. The satellite translates the signal to a different carrier frequency and retransmits it. The signal is transponded in a similar manner by the user, thence, back through the satellite to the ground station.

The phase of the swept frequency modulation being received from the satellite is compared with that being transmitted to yield the total distance from the ground terminal through the satellite to the user. The known distance from the ground station to the satellite is subtracted to give the range from the satellite to the user.

The relationships of the transmitted and received waveforms are shown in Figure D-1. As range increases, the received signal is correspondingly delayed in time. Range is measured by heterodyning the transmitted and received signals to produce a beat frequency, as shown in the lower part of Figure D-1. The propagation time from the ground terminal to the user craft and return is so long that it is necessary to sweep the frequency back and forth between its limits several times during the propagation interval. This is necessary in order to keep the bandwidth of the transmission within reasonable limits. The sweep rate must be sufficiently high to give adequate range resolution. An ambiguity exists for each sweep cycle. If the duration of the cycle can be made long enough, the ambiguity may be resolved by "a priori" information about the approximate location of the user. However, the range resolution required, the sweep rate, and hence the bandwidth of the signal are inter-related in such a way that the ambiguity interval may be too short for a priori resolution. Two approaches to ambiguity resolution are recognized. In one apporach, it is resolved in steps, starting with a long period for the sweep to provide a long ambiguity interval, but with low range resolution. In subsequent steps, the period of sweep is shortened, locating the user in progressively smaller range intervals, until the desired resolution is achieved. In the other method, a second modulating waveform is superimposed on the frequency swept signal.
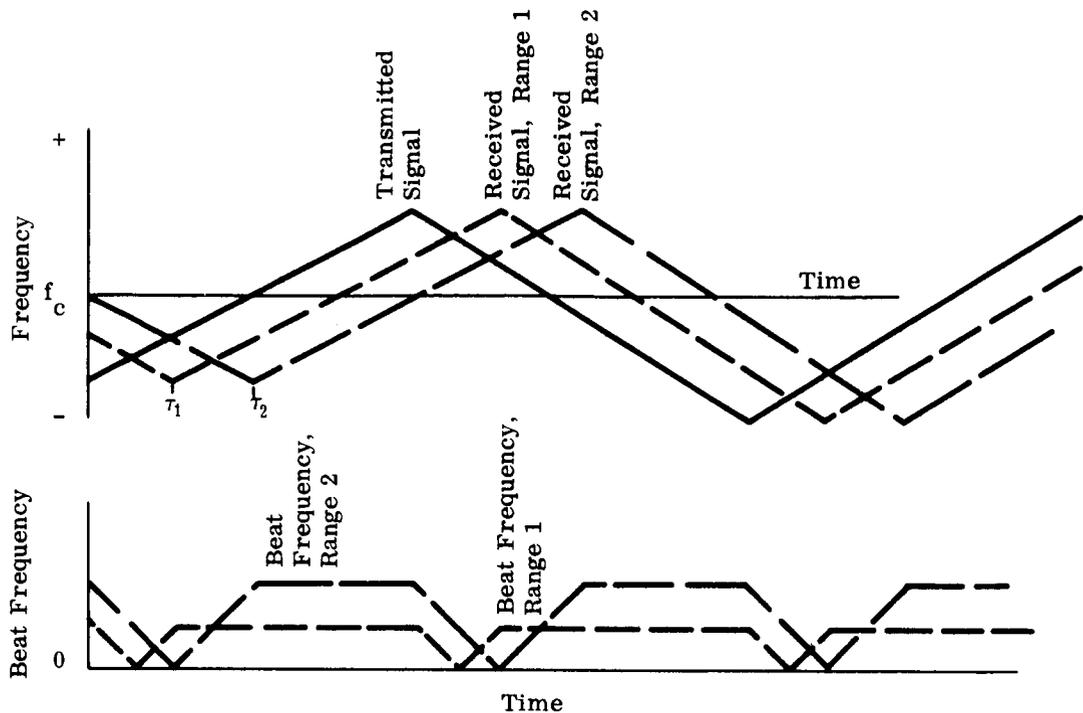
Figure D-1. Transmit-Receive Waveforms

B. PERFORMANCE ANALYSIS

The FM-CW triangular modulation method does not impose any unique requirements on orbits or satellite attitude stabilization. The satellite serves only as a signal relay at a reference point in space. Its attitude does not affect the measurement, except, insofar as control of the attitude makes it possible to improve the antenna gain and thus, reduce transmitter power both at the satellite and at the user.

1. POWER BUDGET

Power budgets are calculated for minimum signal-to-noise ratio and bandwidth. These are conveniently adaptable from those of Section IIIB for single pulse amplitude modulation. Referring to Table III-1, Part I, "Power Delivered to User Receiver" is unchanged, as the same satellite configuration is assumed. Only a few values of Part II, "Receiver Power Required" are changed, as the same noise figures are assumed.

The minimum signal-to-noise ratio in the BF bandwidth for the detection threshold of an FM receiver using a hard limiter and no feedback is 9 db, in accordance with the following expression (Reference [14], page 55):

$$\sqrt{2}\, S_t = 4\, k_1\, (\beta f_D)^{1/2} \qquad\qquad \text{(d-1)}$$

S_t = rms carrier threshold

k_1 = rms noise voltage per unit bandwidth

βf_D = bandwidth of an FM signal

hence:

$$P_t = 8 k_1^2 \beta f_D \; ; \qquad\qquad (d-2)$$

where, P_t = signal power at threshold.

$$\frac{P_t}{k_1^2 \beta f_D} = 8 = 9 \text{ db} \; . \qquad\qquad (d-3)$$

In the example at the end of this section (Appendix D), the minimum bandwidth for 1000-foot range resolution is calculated to be 4900 Hz, if doppler shift is negligible and the counting resolution is assumed to be 4 Hz. For 100-foot resolution, using the same assumptions, the bandwidth is 15,500 Hz.

The FM-CW triangular modulation power budgets for 100-foot range resolution are as shown in Table D-1. The values allow 6 db above the detection threshold. Average power transmitted by the satellite is the value indicated because the transmission is continuous.

Table D-1

Power Budget

		125 MHz 5 kHz Bandwidth No prop. var. allowance	125 MHz 5 kHz Bandwidth Allowance for prop. var.	400 MHz 15.5 kHz Bandwidth	1600 MHz 15.5 kHz Bandwidth	4 GHz 15.5 kHz Bandwidth
I. Power Delivered to User Receiver dbw/watt radiation (into isotropic antenna)		157	175	172	182.5	190.5
II. Receiver Power Required Receiver Bandwidth Factor (Change from Table III-1)		-4.8	-4.8	-5	-5	-18.3
Receiver Noise Power (change from Table III-1)		-4.8	-4.8	-5	-5	-18.3
Desired C/N Ratio (change from Table III-1)		-5	-5	-5	-5	-5
DBW Required at User Receiver (actual -dbw)		146.8	111.3	144.7	141.6	142.1
III. Balancing the Budget DBW/watt Radiated less dbw Required at Receiver	dbw	10.2	30.7	27.3	40.9	48.4
User Antenna Gain	db	3	3	3	3	3
Satellite Transmitted Power	dbw	7.2	27.7	24.3	37.9	45.4
Satellite Transmitted Power	watts	5.3	590	270	6200	54,000

Power values may be reduced in proportion to bandwidth for poorer range resolution. For 1000-foot resolution, the power in watts may be reduced by the ratio 4,900/15,500. Power budgets in Tables III-1 and D-1 assume user antennas with hemispheric coverage. If the beamwidth of the user's antenna is reduced, the aperture becomes larger, and the satellite transmitter power requirement is reduced in accordance with the curves of Figure III-1.

Other detection methods may permit a reduction in power requirements. Feedback does not appear to offer a large saving, however, as the noise bandwidth of a frequency modulation with feedback (FMFB) receiver is essentially that of the modulating waveform. As will be shown later, the modulating waveform bandwidth is one-half the RF bandwidth when, P, the sweep period, is chosen for minimum bandwidth. Hence, the saving with feedback would be less than 3 db.

Matched filter techniques might offer a substantial saving in noise bandwidth and a proportional saving in power. A technique using comb filters was considered briefly. It would resolve range in discrete steps and require matching the sweep period to the filter characteristics for each measurement. It did not appear attractive because of its complexity and, therefore, it was not studied in detail.

2. DOPPLER EFFECTS

Frequency changes due to relative motion of the user, the satellite, and the ground station can cause a frequency difference between the transmitted and received signals and, hence, a beat frequency. Figure D-2 shows the effect for doppler shifts to be smaller than $\tau \, df/dt$ (where τ is the time delay between transmitted and received waveforms, and df/dt is the rate of change of frequency).
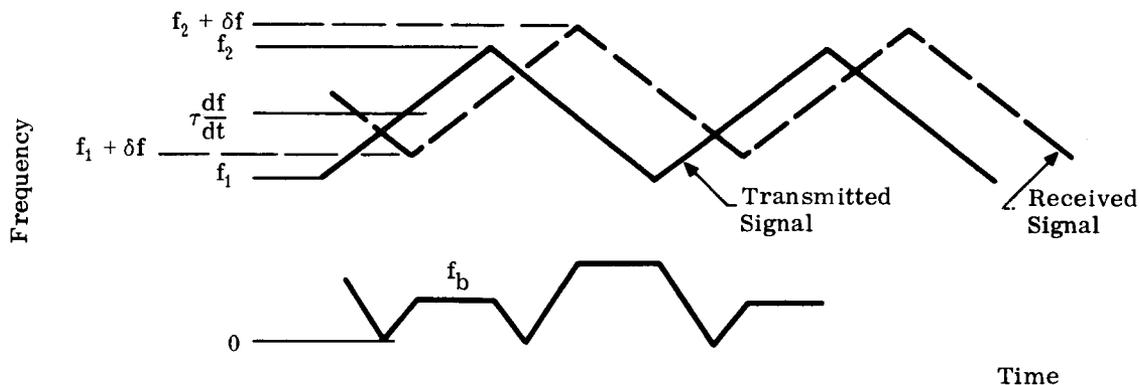
Figure D-2. "Small" Doppler Effect

The average value of f_b, the beat frequency between the transmitted and received waveforms over an integral number of frequency sweep periods, is unchanged from the value without doppler shift.

On the other hand, if the doppler shift is greater than τ df/dt, the average value is changed, as shown in Figure D-3.
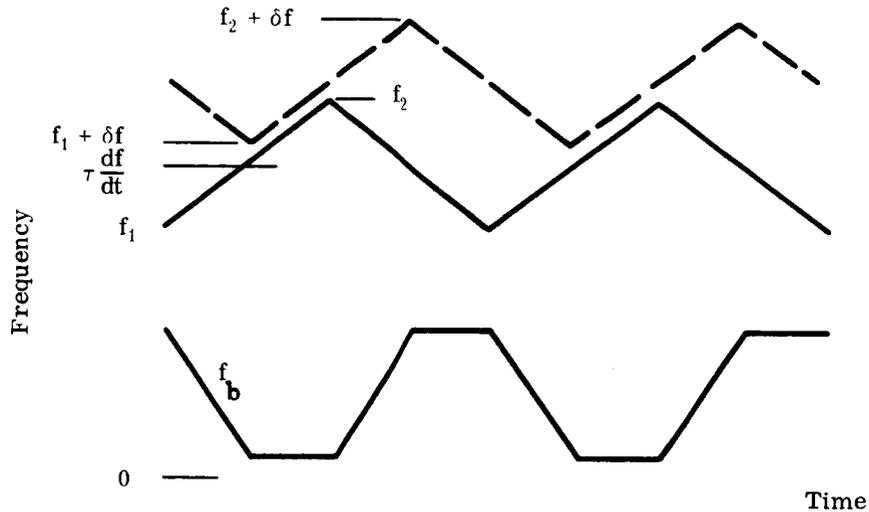
Figure D-3. "Large" Doppler Effect

Doppler shift therefore, restricts the minimum τ in the range interval so that

$$\tau \;>\; \frac{\delta f}{\dfrac{df}{dt}} \tag{d-4}$$

where

δf = frequency shift due to doppler effect.

A convenient rule of thumb for one-way doppler shift is

one Hz/MHz/mach

(if "mach," the speed of sound, is considered as 670 mph). Two-way doppler shift, as in an active ranging system, will double the value.

Extreme cases for doppler shift occur when supersonic aircraft are on the horizon for the satellite and moving directly toward or away from it. In the VHF aircraft band, the doppler shift of the signals for a mach three aircraft can be as large as 820 Hz and in the 1540-1660 MHz band, as large as 10 kHz.

Doppler effects are a function of carrier frequency. Frequency sweep must be approximately twice the maximum doppler in order to insure that a range measurement can be made without the doppler shift exceeding τ df/dt.

Other factors determining bandwidth are independent of carrier frequency. But doppler shift is not. Hence, the bandwidth at higher carrier frequencies may be determined primarily by doppler shift, if fast aircraft are to be accommodated.

3. MULTIPLE ACCESS

As with other methods of range measurement, there are two factors that affect the multiple access capability. These are:

a. Time for measurement and the capability for interleaving user interrogation during measurements.

b. The bandwidth required for the measurement.

The time for measurement is the sum of the address time, lock-up time, measurement time and guard time between measurements. The total bandwidth required for the measurements must include the guard bands between channels.

Users of the system may be interrogated in sequence or on separate radio frequency channels.

Propagation time from ground station through a synchronous satellite to a user and return is approximately one-half second. Several sweep cycles are needed to resolve the user's position within each ambiguity interval. If no a priori information about the user's range is available, the initial sweeps are 40 milliseconds long. Between five and ten sweeps may be necessary to resolve the ambiguity, since several different slight variations of the period must be sent in order to insure that the position of the interval in space places the user in the favorable portion of the interval. (This is further discussed later in this section and illustrated by Figure D-11.) The initial measurement takes at least several tenths of a second. While subsequent steps in the position measurement use shorter sweep periods, the number is such that the total measurement time will be between one and two seconds per user. Thus, the method does not compare favorably with some other methods for time sequential multiple access.

The simultaneous measurement of a number of users on separate radio frequency channels is a possibility. The bandwidth required for the measurement can be minimized, if the ambiguity interval is made relatively short. However, an optimum system design may require a relatively large bandwidth and, thus, limit the number of simultaneous measurements that can be made using frequency division multiple access.

4. GROWTH CAPACITY

Growth capacity is affected by both factors described for multiple access, and in addition, it is affected by the energy required per measurement. Energy per measurement limits growth capacity because of the limited amount of energy available in the satellite to supply additional channels. Since the transmissions are continuous, the satellite transmitted power calculations in Table D-1 may be used to estimate satellite bus power requirements for each channel. Bus to RF efficiency is in the range of 20 to 40 percent, depending on frequency and transponder design. FM-CW triangular modulation is low in overall efficiency in comparison with other methods; thus, it has a relatively poor growth capacity.

5. EQUIPMENT CONSIDERATIONS

There are no unusual requirements on equipment performance. Bandwidth and equipment stability requirements are similar to those for other ranging methods. Signal propagation times through the various equipments must be constant and known. It must not vary from one user's equipment to another by more than the delay time allowed for the range measurement error attributable to that cause. Frequency stability requirements for the transponders in the satellites and users are not unusually stringent. It is important that the frequency stability be such that no more than one cycle of frequency change shall occur during the time of measurement. As this measurement time is approximately 0.2 of a second, stability would be five parts in 10^8 at 100 MHz or five parts in 10^9 at a thousand MHz. While this requires good frequency stability, it is reasonable for operational equipments.

It is essential that the period of the triangular swept modulation be held very stably and be known very accurately, since the modulation envelope is essentially a measuring scale in space. A frequency error will cause a cumulative error along the length of the scale. It is reasonable to hold the error to approximately one foot, so that it will not contribute appreciably to the total error of the system. Such accuracy requires that the frequency not change by more than one part in 65 million. Oscillators with a stability of one part in 10^{10} are readily available, so that this equipment requirement is easily met.

6. PROPAGATION EFFECTS

There are no propagation effects unique to the FM-CW triangular modulation method of ranging. The atmospheric effects are as described in Section IIE.

7. INTERFERENCE EFFECTS

Continuous wave interference within the passband of the satellite transponder will tend to capture a portion of the satellite transmitted energy. Interference of this type was experienced during the initial ground to air tests on VHF using the ATS-1 satellite. Transmissions from aircraft to the satellite on MHz were subject to interference from unidentified, but probably legally operating sources. At times, the level of interference was large enough to capture enough of the 135.6 MHz down link power, so that the aircraft transmissions were not relayed effectively. Interference of this type can be troublesome for a hard limited transponder on a synchronous satellite because the satellite is in view of a large portion of the earth and, therefore, of many potential interfering sources. On the other hand, the hard limited transponder is relatively insensitive to impulsive type interference.

8. DETAILED PERFORMANCE ANALYSIS

The remainder of this section, on the performance analysis of FM-CW Triangular Modulation techniques, is intended to provide the analysis and derivations upon which the previous statements have been based. It also discusses the ambiguity problem inherent in this type of modulation scheme.

The maximum difference, or beat frequency, between the transmitted and received frequencies is

$$f_{b\ max} = \frac{df}{dt}\ \tau \tag{d-5}$$

where τ is the round trip delay time between transmitted and received signals. The maximum frequency increases linearly with delay until the round trip delay time is equal to the time of the linear frequency sweep in one direction. It then decreases linearly, becoming zero at the end of the triangular waveform period. Ambiguity can be resolved within the triangular waveform period, P, by observing that, if τ is less than P/2, the increasing received frequency follows the increasing transmitted frequency. But, if τ is greater than P/2, the decreasing received frequency precedes the decreasing transmitted frequency.

The duration of the maximum frequency difference becomes a shorter portion of the sweep cycle, as range is increased, until the round trip delay time equals P/2—the duration of the frequency excursion, i.e., one-half the sweep cycle. At that point, the maximum difference frequency is reached, but the time available for its measurement becomes zero. It is then more desirable to observe the average difference of the transmitted and received frequencies over an integral number of sweep periods. Maximum and average frequency differences as a function of range are plotted in Figure D-4.

τ = Two-way propagation time, or delay of received signal behind transmitted signal.

P = Sweep period.
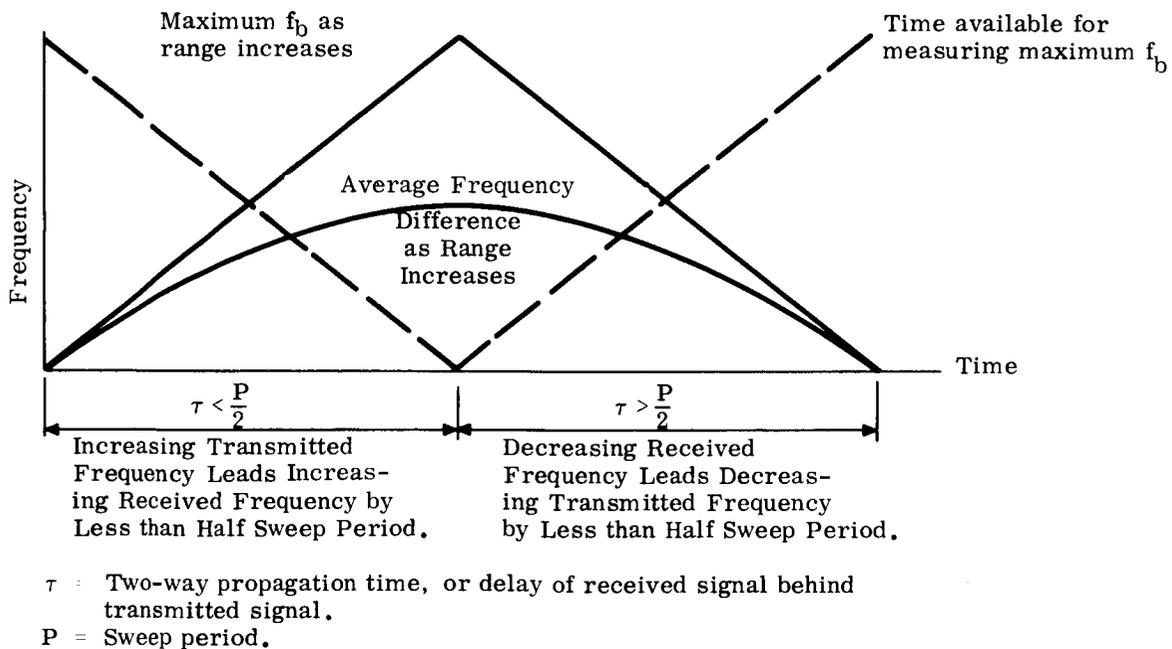
Figure D-4. Frequency Difference Variation with Range

Average frequency difference, $|f_b|_{avg}$, follows a parabolic curve as range is increased, as shown in Figure D-4.

The pattern of average frequency versus range is repeated for every interval of range corresponding to the two-way travel time in period P. For free space, there is a range ambiguity for every Pc/2 range increment, where c is the velocity of light. Range measurement requires that the ambiguity increment must not be shorter than the uncertainty in a priori knowledge of the user's range, or limits of range that can be determined by another measurement, such as an additional modulation waveform on the signal.

When ranging from satellites, the desired accuracy of the range measurement may be high, and the range differences between the nearest to the most distant users may be very great. For example, a synchronous satellite system may require that ranges be measured to users anywhere between the subsatellite point and the horizon. The total range from the satellite to the nearest and most distant users is, thus, between 19,300 and 22,500 nautical miles, respectively. The ambiguity interval is 3200 nautical miles. The two-way propagation time for the ambiguity interval is 40 milliseconds; thus, requiring the P in the triangular waveform modulation be at least that long if no other means is used to resolve the ambiguity.

Range accuracy of a practical navigation system must be a very small part of the total range ambiguity interval. One way to achieve desired accuracy is to measure the range in steps, first locating the user to a reasonably small part of the total interval, then within the smaller interval. The process is repeated in several steps until the range has been determined to the desired accuracy.

It would be feasible to build a system in which the measurement resolution in units of P, $\Delta_n P$, would be in the range of 0.01 to 0.1 P. The first measurement will determine the range to $\Delta_1 P$, the second to $\Delta_2 P$, etc.

The measurement of $\Delta_n P$ requires that average frequency over one modulation period be determined to high accuracy. The resolution required in the measurement of average frequency is highest near the mid-range because of the inflection in the slope of the parabolic relationship of average frequency and range or period, as shown in Figure D-6. For a change of ΔP centered at the midpoint of P, the change in average frequency has a magnitude corresponding only to the change for $\Delta P/2$ because of the inflection. It is also the smallest value of change in average frequency for ΔP anywhere in P. It thus determines the change in the measurement of average frequency that must be resolved in order to measure range to the desired resolution. The expression for frequency counting resolution as a function of range, or period, is derived below.

Given a signal which varies in frequency with time, as shown in the sketch below (solid line),
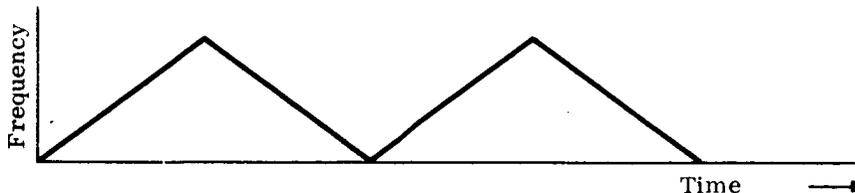
Figure D-5. Sweep Frequency

the signal is to be compared with another signal with exactly the same frequency-time characteristics, but delayed by a time τ which is less than one period, P, of the original signal. This second signal is shown by the dotted line in the following sketch:
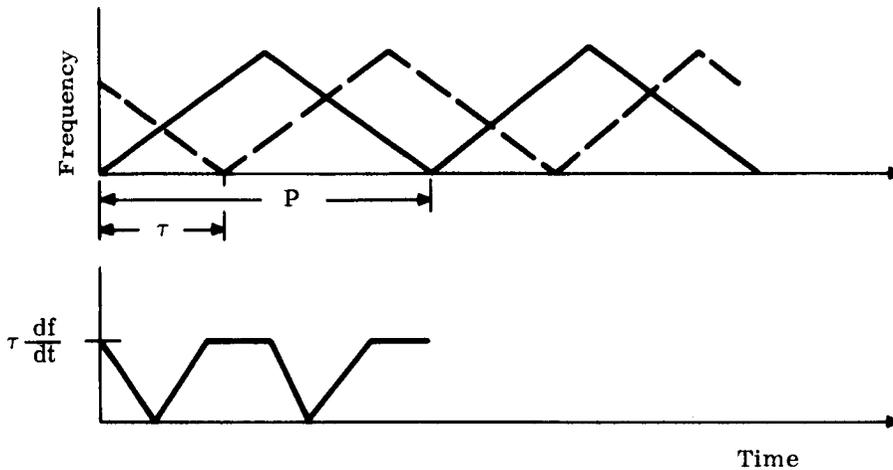
Figure D-6. Frequency Difference

The above figure shows the frequency difference of the two signals expressed as a positive number. The average of the absolute value of the frequency difference $|f_b|_{avg}$, is the area under the curve in the figure, divided by the base, P, or:

$$|f_b|_{av} = \frac{\tau \frac{df}{dt} \left[\left(\frac{\tau}{2}\right) 2 + (P - 2\tau) \right]}{P} \qquad \text{(d-6)}$$

simplifying:

$$|f_b|_{av} = \tau \frac{df}{dt} \left(1 - \frac{\tau}{P}\right) . \qquad \text{(d-7)}$$

The expression is plotted in Figure D-7. Given the parabola of Figure D-7, partial differentiation yields:

$$\Delta |f_b|_{av} = \frac{df}{dt} \left(1 - \frac{2\tau}{P}\right) \Delta\tau . \qquad \text{(d-8)}$$

This equation is valid in the region given by

$$\left| \tau - \frac{P}{2} \right| \geq \frac{\Delta\tau}{2} \qquad \text{(d-9)}$$

remembering that $0 \leq \tau \leq P$. But to be certain of the ends, let

$$\frac{\Delta\tau}{2} \leq \tau \leq P - \frac{\Delta\tau}{2} . \qquad \text{(d-10)}$$
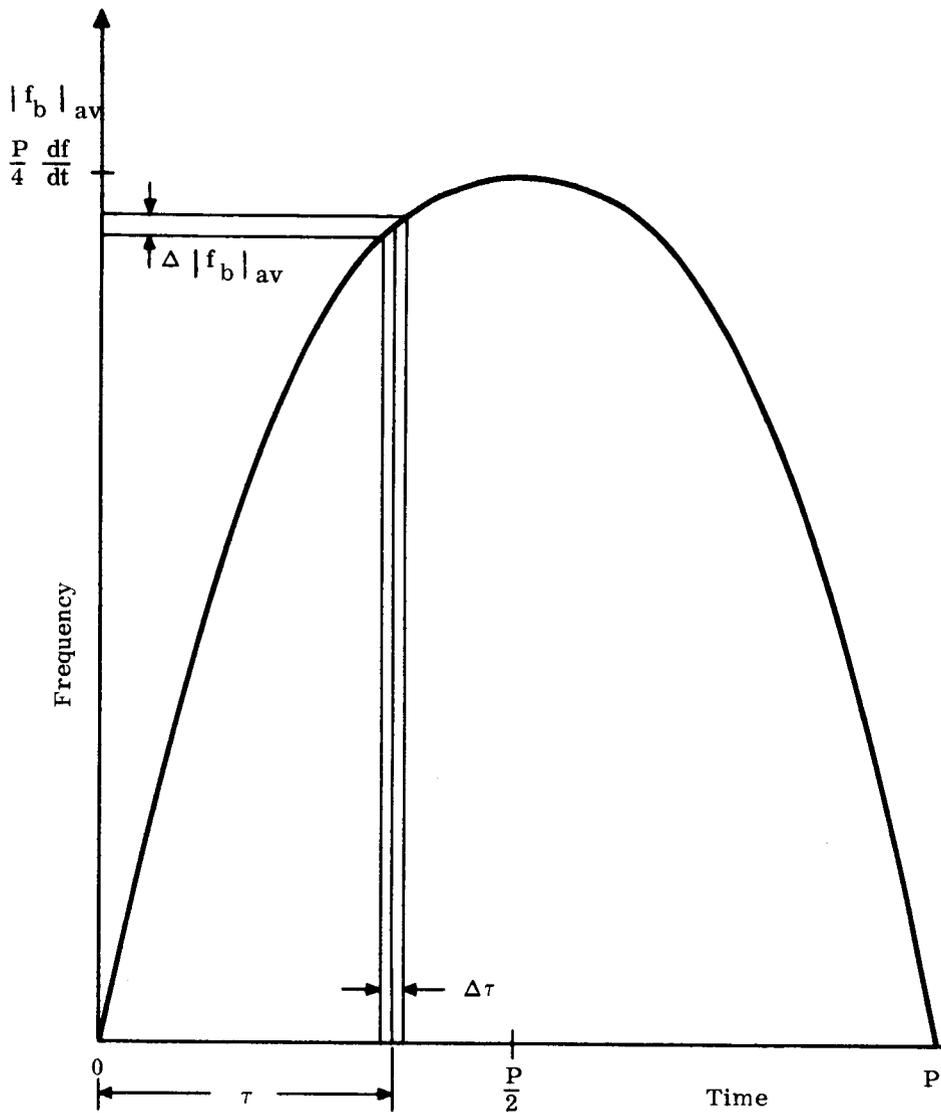
Figure D-7. $|f_b|$ avg.

If however $\left| \tau - \dfrac{P}{2} \right| \leq \dfrac{\Delta \tau}{2}$, then our sampling with $\Delta \tau$ takes $|f_b|_{av}$ to the peak of the parabola and part-way down the other side. Thus, $\Delta |f_b|_{av}$ is the peak value of the parabola minus the lowest end of the projection of the sample $\Delta \tau$ upon the curve.

$$\Delta |f_b|_{av} = \frac{df}{dt} \frac{P}{4} - \tau' \frac{df}{dt} \left[1 - \frac{\tau'}{P} \right] , \qquad\qquad (d\text{-}11)$$

$$\Delta |f_b|_{av} = \frac{df}{dt} \left[\frac{P}{4} - \tau' + \frac{\tau'^2}{P} \right] \qquad\qquad (d\text{-}12)$$

when

$$\left| \tau - \frac{P}{2} \right| \leq \frac{\Delta \tau}{2} \qquad\qquad (d\text{-}13)$$

and

$$\tau' \equiv \frac{P}{2} + \left| \frac{P}{2} - \tau \right| + \frac{\Delta\tau}{2} \quad . \tag{d-14}$$

If τ and $\Delta\tau$ are expressed as fractions of P, values of $\left| \Delta \left| f_b \right|_{av} \right|$ versus τ for various $\Delta\tau$ are given in Figures D-9 and D-10 in units of $P \frac{df}{dt}$.

$\Delta \left| f_b \right|_{av}$ is in units of $P \frac{df}{dt}$. Reconsidering the original signal:
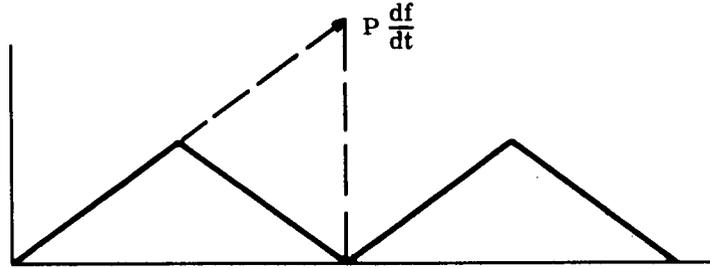
Figure D-8. $\Delta \left| f_b \right|$ avg.

$$\Delta \left| f_b \right|_{av} = 2 \left(\frac{P}{2} \frac{df}{dt} \right) \quad . \tag{d-15}$$

In other words, to get $\Delta \left| f_b \right|_{av}$ as a multiple of the total frequency excursion of the original signal, its numerical value, as determined by the equations on the previous pages, should be multiplied by two. Or

$$\Delta \left| f_b \right|_{av} = F(\tau) \left(P \frac{df}{dt} \right) = 2F(\tau) \left(\frac{P}{2} \frac{df}{dt} \right) \quad . \tag{d-16}$$

It is desirable to avoid using the center portion of the range ambiguity interval because $\Delta f_{b\,av}$ is small in that region. In order to resolve an increment of range, the change in $\Delta f_{b\,av}$ must at least equal the resolution of the counter that measures $f_{b\,av}$. If the center portion of the ambiguity interval is used, the total frequency excursion $P \frac{df}{dt}$ becomes much larger than would be necessary if that portion were avoided. As an example, if the range resolution is 0.05 P, Figure D-9 shows that the change in $\Delta f_{b\,av}$ at the center of the interval is only 0.000625 $P \frac{df}{dt}$ for an increment of range change. The total frequency sweep would then have to be $(1/6.25 \times 10^{-4})$ N or 1600N, where N is the counting resolution of the frequency counter. On the other hand, if the portion of the interval between 0.46 and 0.54 of the interval is avoided, the minimum $\Delta f_{b\,av}$ is 0.004 $p \frac{df}{dt}$, so that the frequency excursion is only 250N. Since the signal bandwidth is a function of $P \frac{df}{dt}$, there is a large saving in bandwidth by avoiding the use of the center of the ambiguity interval.
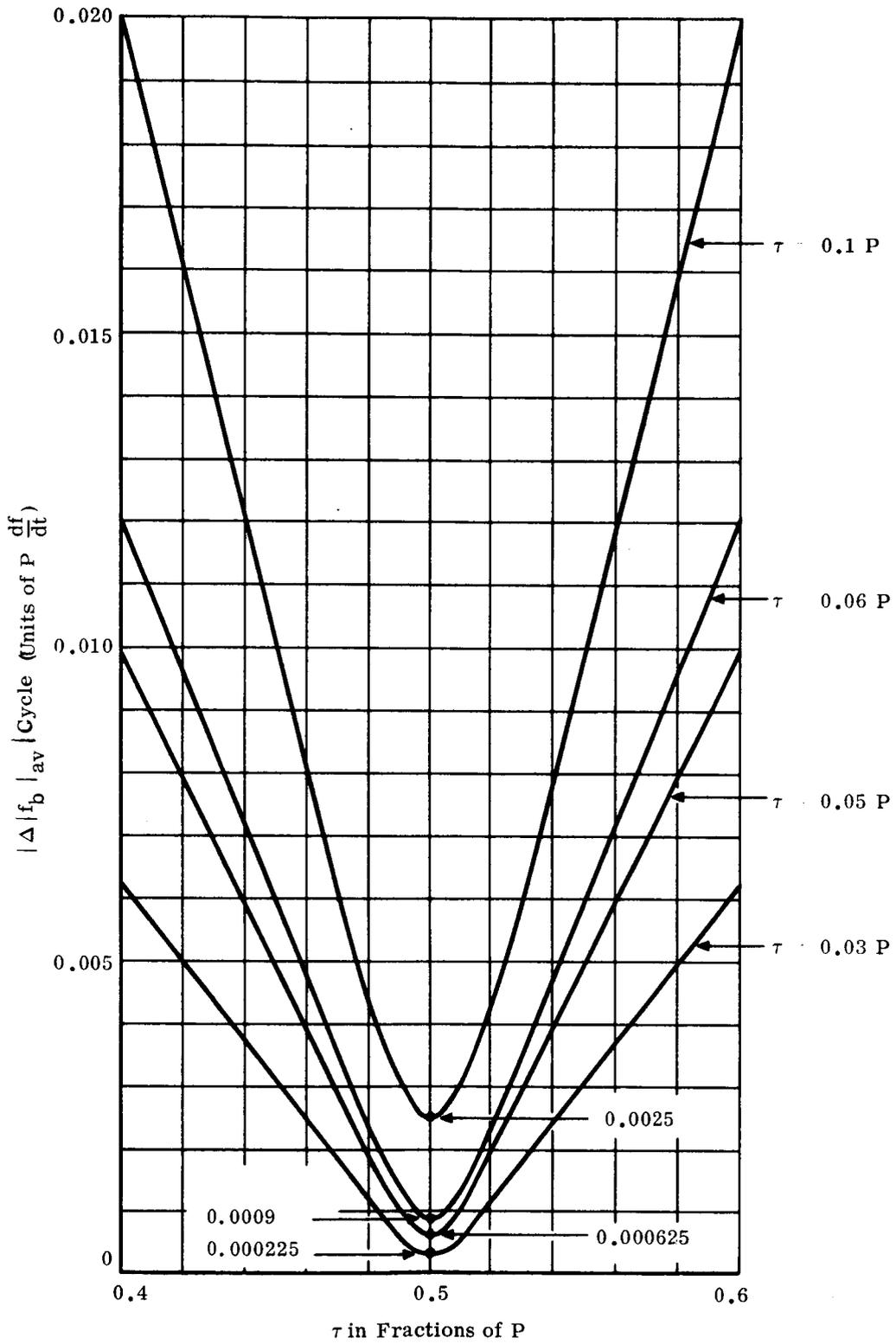
Figure D-9. τ in Fractions of P
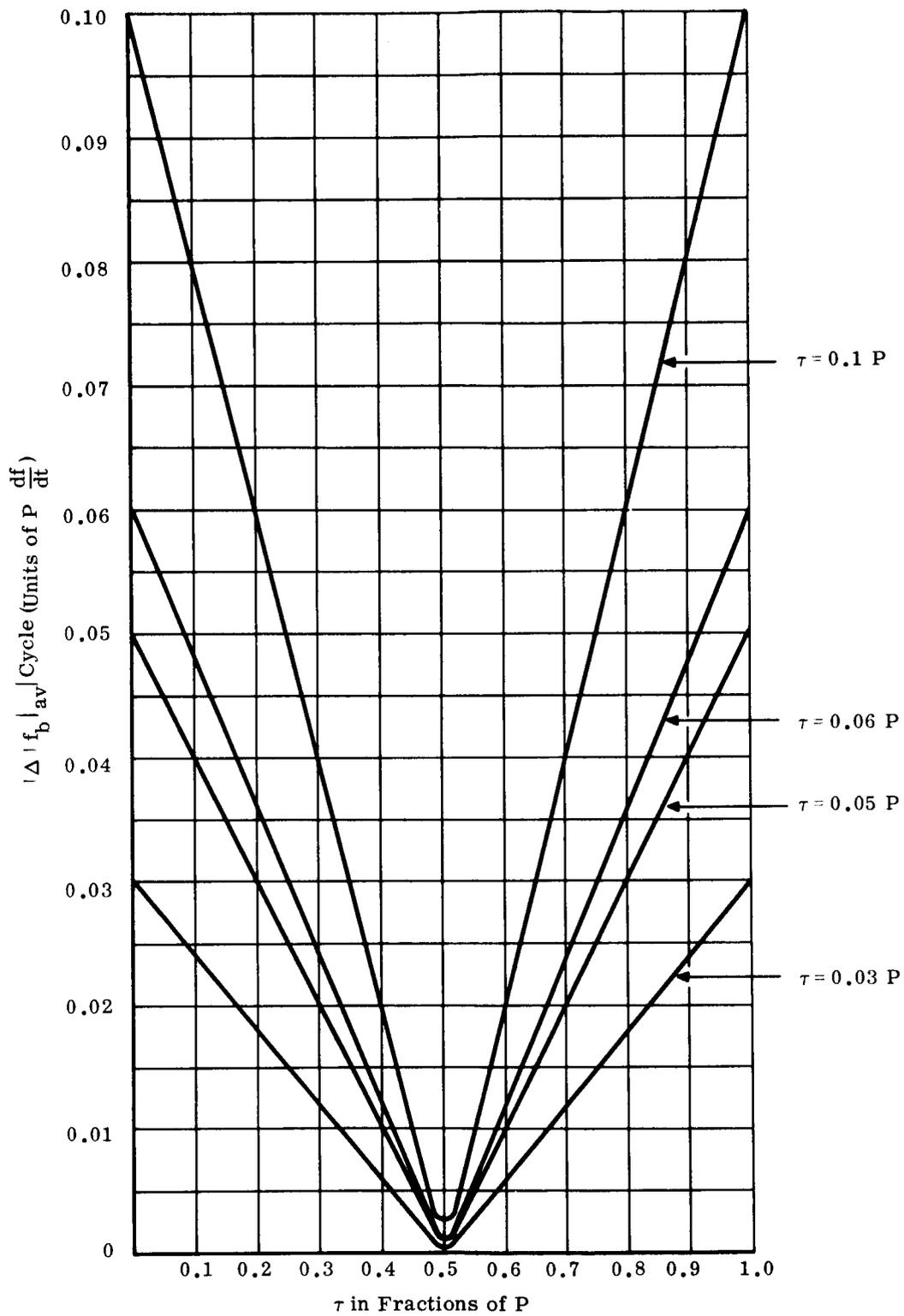
Figure D-10. τ in Fractions of P

The ends of the interval, where τ is small, should also be avoided for two reasons. One is that f_b becomes vanishingly small as τ approaches zero or P. But τ must be large enough to provide a sufficient number of cycles during the interval. The other reason for avoiding the ends of the interval is that τ must be large enough to accommodate doppler shift as explained previously.

It is obvious that the center and ends of the ambiguity interval cannot be avoided by ignoring range intervals, for the user can be at any range from the satellite. It is, therefore, necessary to avoid these portions by shifting the interval in space.

This can be done by changing P, the period of the sweep. The longest range ambiguity that must be resolved is approximately the radius of the earth. A synchronous satellite is approximately six earth radii from the earth's surface. The sweep period for resolving the largest ambiguity is thus one-seventh the propagation time from the satellite to the limb of the earth; that is, to where the most distant user is located.

The ambiguity interval may be shifted in range by one-fourth its length by changing the sweep period by one twenty-eighth of its value. For shorter range ambiguity intervals, the proportionate shift in period is smaller. Figure D-11 illustrates how the ambiguity interval is shifted with respect to a user's position.
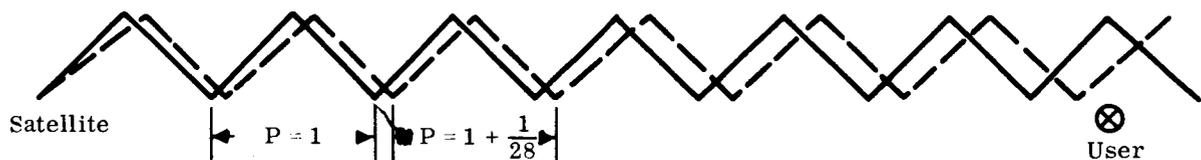
Figure D-11. Ambiguity Interval Shift

In the application of the method, there is no prior knowledge of the user's relative position in the ambiguity interval, hence, a measurement must be made to determine if he is in a favorable part of the interval. If he is, the measurement sequence proceeds to the next smaller interval. If not, the interval must be shifted appropriately to place him in a favorable part. The test must be made for each successive interval throughout the range measurement.

The measurement of range to the user requires that the sweep period be known accurately, for it is in effect the calibrated unit of measurement on the distance scale. The accuracy of the positions of the ambiguity intervals in space is determined by the timing accuracy of the frequency swept waveform.

To insure that the ambiguity interval limits are within the specified accuracy for the system, the frequency of the modulating waveform must be known to within the ratio of range resolution to the total range. Fortunately, the state-of-the-art in frequency measurement and control is such that the requirement is easily met. The maximum range is approximately 1.3×10^8 feet. An accuracy of one part in 0.65×10^{-8} would suffice for one foot resolution. The factor of two results because of the two-way propagation for active

ranging. Suitable oscillators are available as standard items. When the unfavorable parts of the ambiguity interval are avoided by changing the period of the triangular sweep, the required frequency sweep is the number of cycles that can be resolved by the counter, multiplied by the number of range resolution elements in the sweep period. From this simple relationship, we can derive an approximation to the signal bandwidth. Let

$$\text{ambiguity interval} = \frac{P}{2}$$

$$\text{range resolution interval} = \Delta\tau$$

$$\text{counting resolution} = N$$

(N refers to the counting resolution of the counter, not to the stability of the reference oscillator. Counters normally have a resolution of ±1 count regardless of the number of events counted), then the swept frequency is

$$f_s = N \frac{P}{2\Delta\tau} \text{ cycles .} \tag{d-17}$$

From Carson's rule for FM modulation, the bandwidth is the sum of the swept frequency plus twice the highest modulating frequency, f_n. For a triangular waveshape, we approximate the modulating frequency as three times the reciprocal of the triangular waveform period

$$f_n \overset{\sim}{=} \frac{3}{P} . \tag{d-18}$$

The transmitted signal bandwidth is:

$$f_{bw} \overset{\sim}{=} N \frac{P}{2\Delta\tau} + \frac{6}{P} . \tag{d-19}$$

The relationship of the terms in this expression are shown in Figure D-12. Approximating the highest frequency component of the triangular waveform by $f_n = 3/P$, implies that modulation sideband components higher than the third harmonic will not be passed by the receiver. The effect is to introduce a small frequency modulation component that averages out over a sweep cycle and also a small amount of amplitude modulation.

Minimum bandwidth may be determined by differentiating and setting to zero the expression for f_{bw}. Thus, minimum bandwidth occurs when

$$P^2 = \frac{120\Delta\tau}{N} . \tag{d-20}$$

Ambiguity must be resolved without excessive bandwidth. This may be done using the method described previously, keeping bandwidth constant and using a difference $\frac{df}{dt}$ for each step.

Another method is by the use of a second modulating waveform, superimposed on the transmitted waveform with a wave-length in space sufficient for the larger ambiguity interval. It may be sinusoidal in form. This technique has been called double modulated FM (Reference 13). The envelope of the sinusoidal modulation is detected and the transmitted and received phases compared.
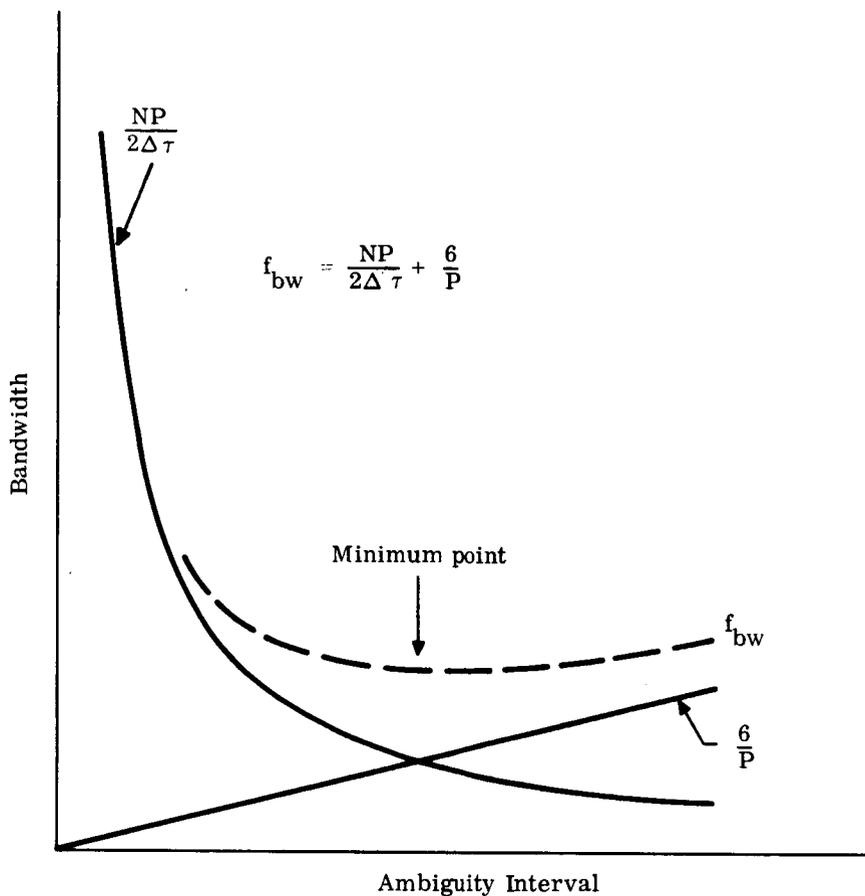
Figure D-12. Relationship of Ambiguity Interval and Bandwidth

The figure shows a graph with "Bandwidth" on the vertical axis and "Ambiguity Interval" on the horizontal axis.

$$f_{bw} = \frac{NP}{2\Delta\tau} + \frac{6}{P}$$

with the curve labeled $\frac{NP}{2\Delta\tau}$, the minimum point marked, and curves labeled f_{bw} and $\frac{6}{P}$.

C. IMPLEMENTATION CONSIDERATIONS

Figures D-13 and D-14 are block diagrams of ground station, satellite and user equipments for the FM-CW triangular modulation method. For illustration, we assume a system operating on VHF aeronautical frequencies and a desired range resolution of one thousand feet.

A highly stable oscillator, preferably a cesium beam standard or a high stability quartz oscillator, is suggested as the basic frequency reference. A cesium beam standard drifts approximately one part in 10^{-11} during the life of the cesium beam tube; while a high stability quartz oscillator has a stability of approximately five parts in 10^{-11} per day. When used in a system in which the user transponds, frequency stability of approximately one part in 10^{-7} is required. However, if the system is to be used in a passive direct ranging mode, as well—one in which the users receive signals and compare their modulation phase with an on-board reference, the timing of the signals must be maintained to high accuracy with reference to standard time. The relatively small additional cost for providing the passive mode capability strongly suggests that the high stability oscillator be furnished.

The output of the oscillator drives a clock pulse generator, thus, furnishing high accuracy timing for the modulation phase. A ramp generator develops a triangular waveform that sweeps between two voltage levels. The triangular waveform is applied to the modulator of the transmitter, so that the transmitted frequency is swept linearly between two values.

Figure D-13. Equipment Block Diagram

Satellite Transponder

User Transponder
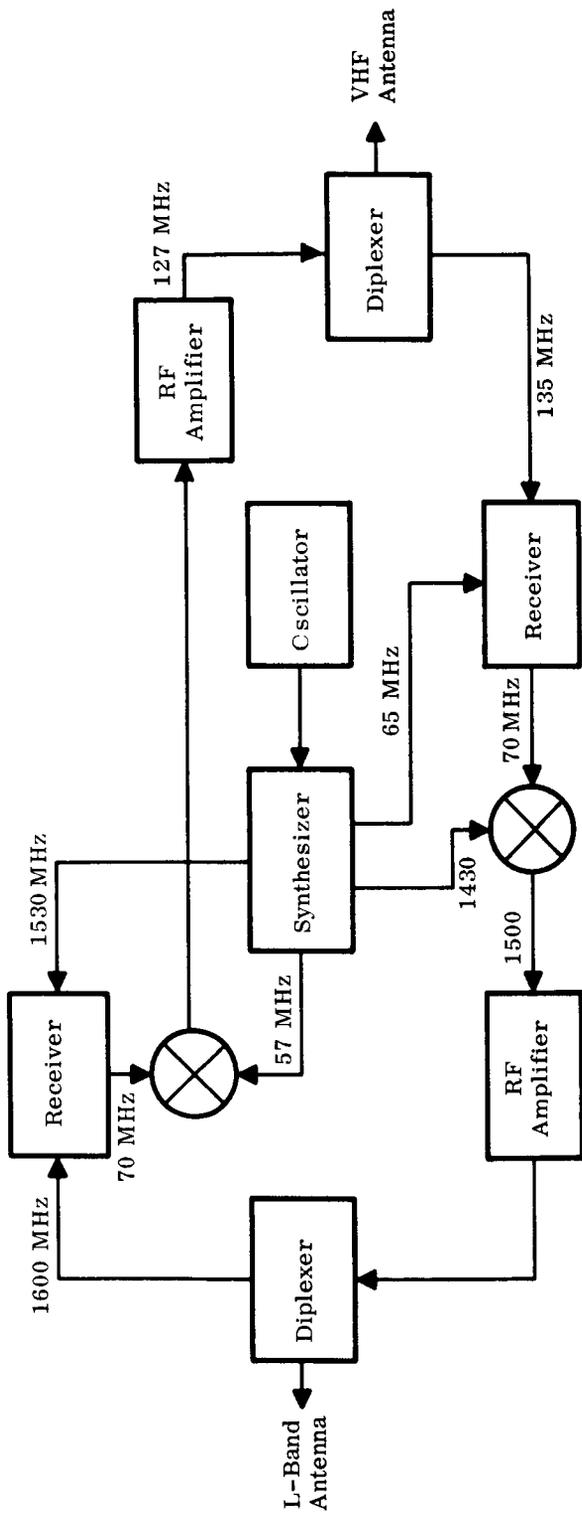
Figure D-14. Equipment Block Diagrams
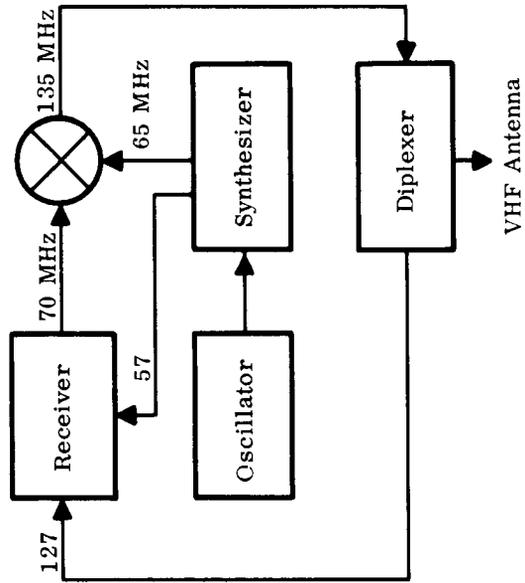
(Diagram does not include
address decoder and logic)

Frequency sweep, counting resolution and range resolution at minimum bandwidth are related by Equation (d-21).

$$P^2 = \frac{12\Delta\tau}{N} \qquad\qquad (d-21)$$

$\Delta\tau$ = 2 x 10^{-6} seconds for 1000 feet range resolution

N = 4 a value selected to allow for the ±1 count resolution of a normal counter plus possible false or missed counts due to noise. .

$$P = 2.45 \times 10^{-3} \qquad\qquad (d-22)$$

The sweep frequency can be determined from Equation (d-17).

$$f_s = N \frac{P}{2\Delta\tau} \qquad\qquad (d-23)$$

$$= 2450 \text{ Hz.}$$

Doppler shift in the VHF aircraft band will not exceed approximately 800 Hz and, therefore, no additional allowance must be made to accommodate the shift.

Radio frequency bandwidth for the highest range resolution is the sum of both terms in Equation (d-19).

$$f_{bw} = 2450 + \frac{6}{2.45 \times 10^{-3}} = 4900 \text{ Hz} \qquad\qquad (d-24)$$

If the range to the user is completely unknown, he can be at any range from the subsatellite point to the horizon; the initial triangular sweep period is 40 x 10^{-3} seconds. The ramp generator, Figure D-13, is adjusted to generate a voltage that will sweep the transmitter frequency 2450 Hz in $\frac{P}{2}$, or 20 x 10^{-3} seconds.

The signal will be transponded by the satellite and user equipments. Each of the transponders is designed to cancel the effects of drift in its local oscillator, by using the same frequency source for all frequency translation. Doppler shift cannot be cancelled out by such a technique.

After approximately one-half second, the signal is returned from the user, and heterodyned with the signal being transmitted to produce f_b.

The ambiguity interval shift logic circuit resolves the ambiguity, caused by the increasing and decreasing sweeps, by noting if the increasing received frequency follows the increasing transmitted frequency by less than or more than $\frac{P}{2}$. It then observes the waveform to determine if doppler shift is larger than $\tau \frac{df}{dt}$ and also observes the average beat frequency to determine that it is below its maximum possible value—in this case, $\frac{2450}{2}$ or 1225 Hz. If one of the conditions is not met, the logic circuit shifts the period of the sweep waveform appropriately so that the user's range is within the acceptable portion of the ambiguity interval.

When it has been established that the user is in the acceptable range, the Time Interval Pulse Generator opens the Gate and the beat frequency is counted over an integral number of frequency sweep intervals. The total count is divided by the time of the counting interval to provide a measure of range within the ambiguity limits.

Resolution depends on the location of the user within the ambiguity interval. As explained previously, resolution is highest when the received frequency is close to the transmitting frequency, and decreases linearly until the frequency difference approaches its maximum value. For the parameters selected, the resolution will be in the range of 0.6 percent to approximately 5 percent, depending on where the user is located in the interval. For 5 percent resolution, he is located to within 3200 x 0.05 = 160 nautical miles.

The sweep interval is then changed, but the magnitude of the frequency sweep is not. The sweep interval, P, is longer than the two-way radio propagation time for the range interval, or

$$P > \frac{320}{162} \times 10^{-3} = 1.98 \times 10^{-3} \text{ seconds.}$$

The measurement sequence, described above, is repeated for the new interval with the frequency sweep remaining 2450 Hz. Range resolution is better than 0.05 x 160 = 8.0 nautical miles.

A third and fourth repetition of the process will locate the user within the desired 1000-foot resolution.

Four steps were needed because there was no a priori knowledge of the user's range and because the user was assumed to be in the unfavorable portion of the ambiguity interval. In practice, it is unlikely that four steps would be necessary; two, and, in some cases, one may suffice for most measurements in an operational system, where the ranges are known approximately because they are measured frequently.

Numerous variations are possible in a system implementation. For example, the time for a range measurement can be reduced, if the waveform transmitted from the ground station contains a sequence of different sweep periods to insure that all the ambiguities are resolved, and that at least one sequence of pulses in each ambiguity interval is proper for each possible location within the interval.

The basic user equipment is simple—consisting of a hard limited transponder. Its cost compares favorably with that required for other modulation methods. It must receive and transmit simultaneously; therefore, a diplexer is required. When the receive and transmit frequencies can be separated by approximately 10 percent, and wide bandwidth antennas are readily designed for the application, a diplexer can be provided at reasonable cost. For aircraft at VHF frequencies, a diplexer may present a more serious problem, because it is difficult to design wide bandwidth antennas that are aerodynamically acceptable.

A discussion of the means that might be used for addressing users and for communicating, which might be an important adjunct to the ranging function, has not been included. The communication functions can be provided in essentially the same manner as in other modulation methods. Except for detection circuits, there need be no essential differences in the address recognition, communication and display equipments.

APPENDIX E

PSEUDO RANDOM CODE TECHNIQUES

A. INTRODUCTION AND CODE GENERATION

In the past 10 years, there has been considerable effort on pseudo noise (PN) code application to communications and to range measurements. This section of the report will draw from this work and analyze properties of PN codes to determine their applicability for the system use, addressed by this study. Perhaps one of the easiest ways to approach an explanation and analysis of PN codes is through the mechanics of their generation. The ease with which these codes can be generated is one of the major reasons for their wide application.

A pseudo noise sequence, or code, is generally generated from a binary shift register and logic circuit. The shift register is a series of digital flip flop circuits, that are arranged so as to shift a binary word through the register in step with an applied train of "clock" or "shift" pulses. The individual stages of the register, when triggered by a shift pulse, take the state of the previous stage in the previous bit time, as shown in Figure E-1. Thus, the shift register can move a sequence through itself in exact synchronism with a shift pulse train. To use a shift register to generate a code, it is necessary to use the output of the register to provide the input through some suitable logic. One of the simplest forms of this logic has been called by names, such as "half adder" or modulo two "adder." This logic is most precisely described by the "truth" table, shown in Figure E-2.

If the register of Figure E-1 were connected by half-adder logic, as shown in Figure E-3, the register, when driven by a clock pulse train, would generate a sequence 255 bits long, without repeating. Looking at the first line of the truth table of Figure E-2, note the output of a half adder is a zero state if its inputs are both zero. Thus, if the generator of Figure E-3 were set up with all zeros in the eight stages, this state would never change. Should the eight stages be preset to any other configuration of ones and zeros, however, as the clock or shift pulses drive the register, it would take on, successively, all possible states, except the all zeros state. The serial sequence seen at the output would therefore, be a repetitive sequence 255 bits long. Expressed mathematically, the sequence length is 2^N-1 bits long; where N is the number of stages in the shift register. Since 2^N is the maximum number of different binary states N stages can possibly have, and since the all-zero state must be excluded from the sequence, the value 2^N-1 is a maximal length sequence.

The logic arrangement, shown in Figure E-3, was picked specifically as an arrangement that will generate a maximal length code. The code generated is unique to the logic shown. Other logic or feedback, configurations may or may not generate maximal length codes. The particular arrangement shown was picked from a set of tables so that the sequence would be maximal length.
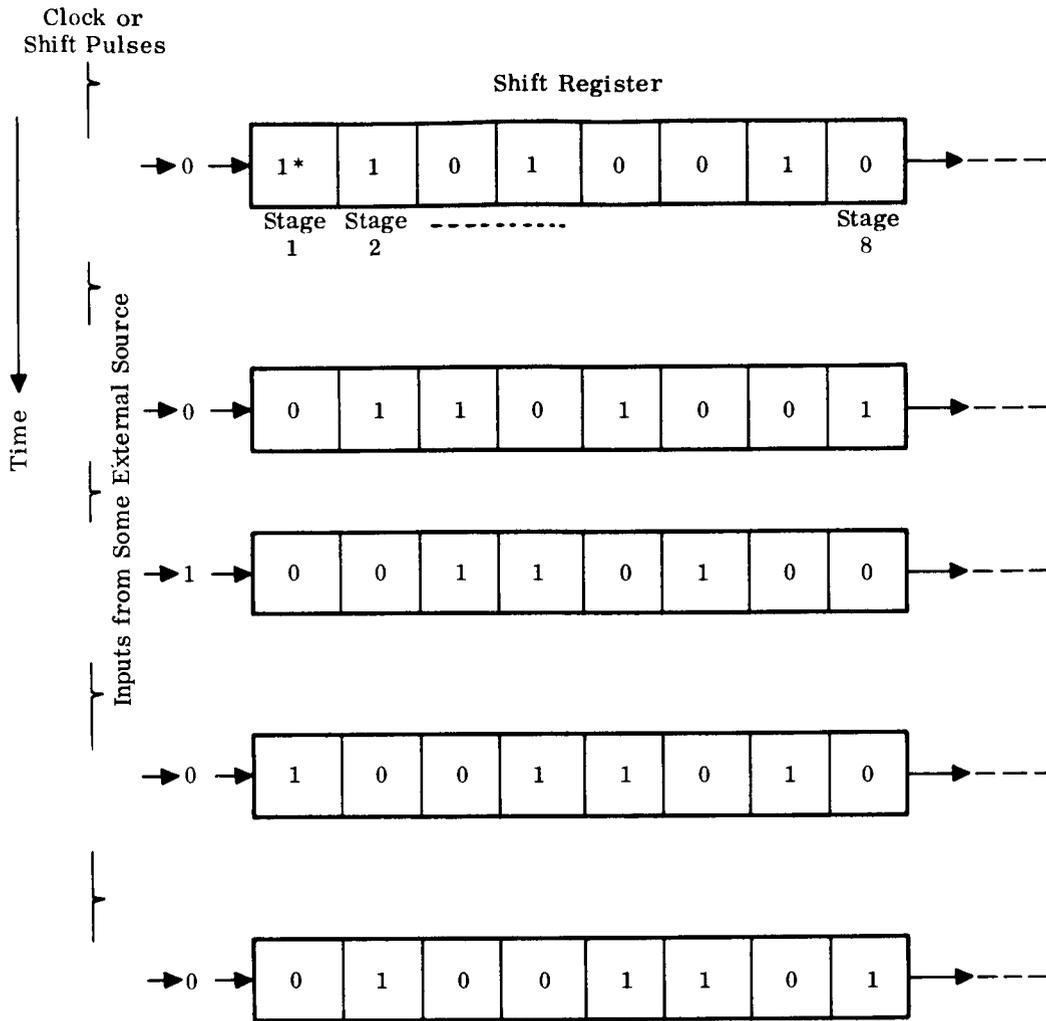
*Numerals indicate the binary state of individual stages.

Figure E-1. Basic Shift Register Action

Figure E-4 shows cyclic diagrams depicting maximal and non-maximal length codes. In the reference of these diagrams, the register, in generating a PN code, proceeds around the cycle of states the small circles, each representing a particular state of the stages of the generator. The codes generated are cyclic in nature. When a non-maximal length sequence is generated, there may also be other independent sequences that can be generated by the same logic. The sequence generated can, therefore, be dependent upon the state from which the register starts. As indicated in the non-maximal length cycle A, there may be one or more states that lead to a sequence, but which are never repeated once the sequence starts. By varying the length of the register involved, the feedback logic and, perhaps, in the case of non-maximal length codes, the initial state of the register, codes of most any length can be generated.

Input A

Input B ──────────▶⊕──────────▶ Output
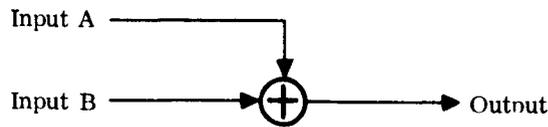
Truth Table

Inputs			
	A	B	Out
Binary State	0	0	0
	0	1	1
	1	0	1
	1	1	0

Figure E-2. Truth Table and Logic Symbol for Half-Adder

8 Stage Shift Register

Output
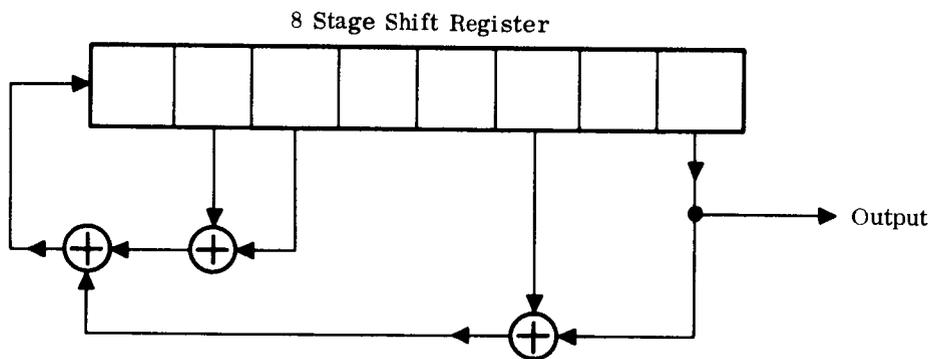
Figure E-3. Sequence Generator (Pseudo Noise)

B. CODE PROPERTIES

In order to examine the application possibilities of the codes, it is helpful to look at some of the properties of the PN codes generated, as described. A property of major interest for PN codes is the auto correlation function. Figure E-5 shows a plot of the autocorrelation function of a typical code. In this case, $\Delta\tau$ represents the time shift of the code with respect to an identical reference code. The autocorrelation function, of course, repeats once per cycle period of the code. From the point of maximum correlation, the function decreases linearly to complete loss of correlation in a one bit time space. Thus, the pseudo noise codes are uniquely definable every bit period of their cyclic length.
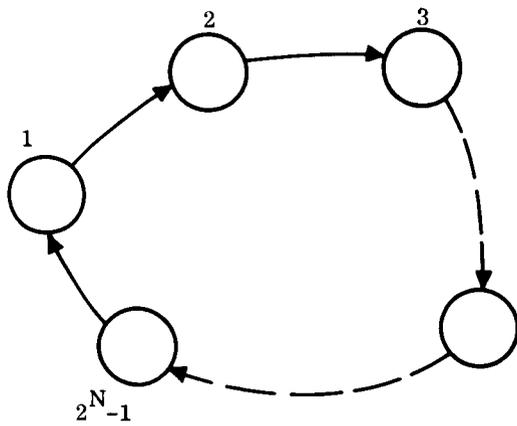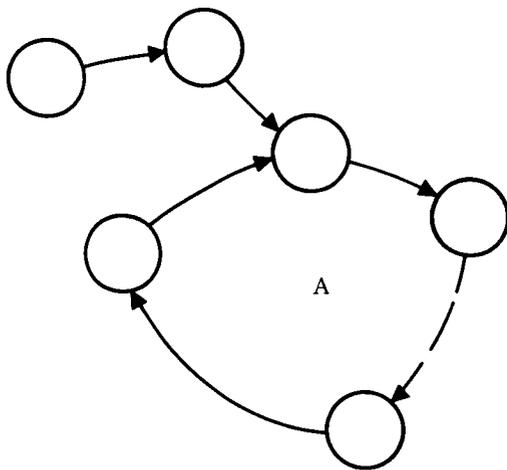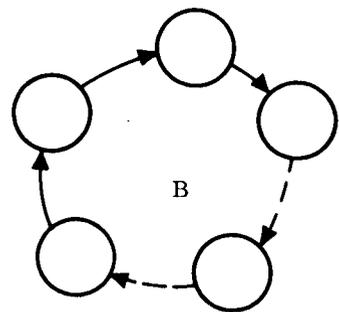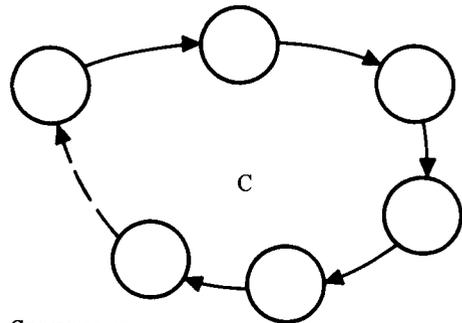
Maximal Length Sequence

Non-Maximal Length Sequences

Figure E-4. Maximal and Non-Maximal Length Sequences

Even the shortest possible code may cause untenable acquisition times; therefore, acquisition short cuts are desirable. A short cut analyzed in Para. H of this Appendix is to take advantage of the properties of a code generated by a partitioned generator. As an example, assume a code of length l_p is made up of a logical combination of three sub codes of lengths l_K, l_M, and l_N. The lengths of the sub codes will satisfy the relation:

$$l_p = (l_K) \times (l_M) \times (l_N) \ .$$

(e-1)

The "average length" of the sub codes will therefore be:

$$l_{AV} = 3\sqrt{l_p} \ .$$

(e-2)

If such a partition were physically realizable, the problem would then be to acquire three codes of length l_{AV}. Taking advantage of the cross-correlation peaks to trigger decision thresholds, the three codes can be searched individually, acquiring first one sub-code, then another and finally the third to achieve full acquisition. Thus, only 3X (l_{AV}) bit times need to be searched to acquire. In practice, it is often not possible to realize exactly $3\sqrt{l_p}$ as a code length; thus, the product of $l_K \cdot l_M \cdot l_N$ is an approximation to the ideal. In this case $l_K + l_M + l_N$ bit times must be searched to acquire. The value $3 \ 3\sqrt{l_p}$ is a lower bound on maximum acquisition time and is not always exactly realizable. The value is, however, close to practical results and a valuable estimation for a three part code, or more generally $J \ J\sqrt{l}$, where J equals the number of sub-code lengths used.

D. PSEUDO NOISE RECEIVER IMPLEMENTATION

Perhaps, the most common form of receiver for use with pseudo noise codes is shown in Figure E-6. This receiver multiplies the received PN signal with a locally generated PN reference.
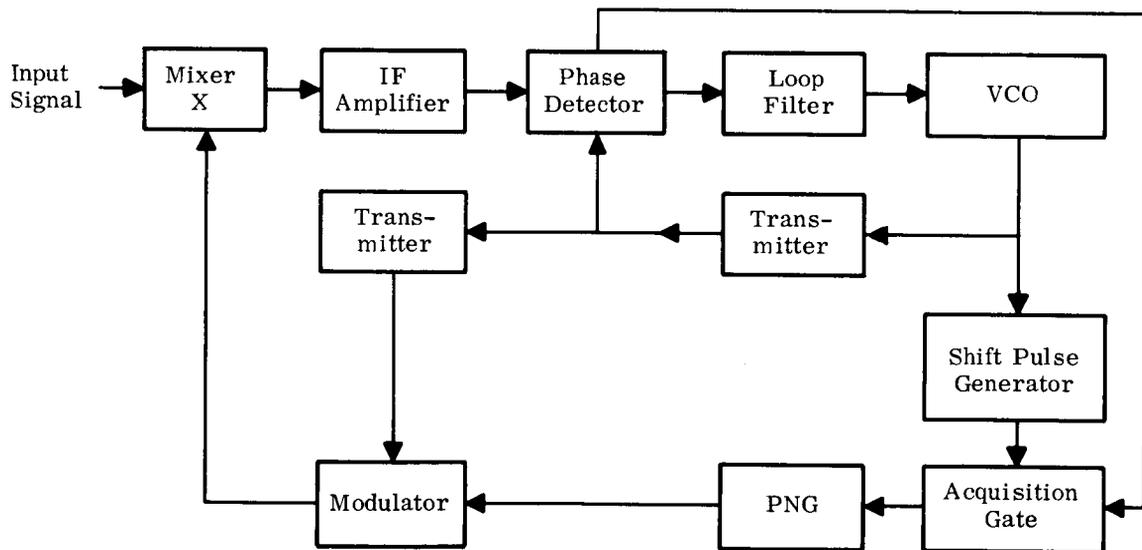
Figure E-6. Pseudo Noise Receiver Implementation
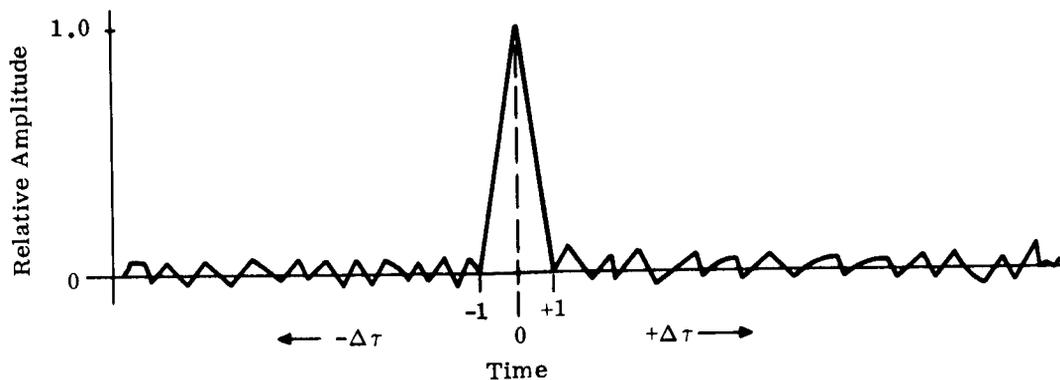
Figure E-5. Autocorrelation Function

A property of maximal length codes that serve a very useful generation function is the so-called shift and add property. This property states that if a maximal length code is shifted an arbitrary amount and modulo two summed with the unshifted code, the result will be the original code shifted to some point in time. This property of the codes makes it possible to generate nearly arbitrary time shifts of a maximal length code by multiple logical half addition. Especially when very long codes are used, this can considerably simplify generating large time displacement code references.

If a maximal length code of length 2^N-1 is modulo two added to a maximal length code of length 2^M-1, the result will be a sequence $2^{(M+N)}-1$ bits long. In general, logical combination can be used to generate a certain length sequence from two or more shorter sequences. This is sometimes referred to as a "partition generator" approach. If a sequence that has been generated from such a logical combination is cross correlated with one or more of the sequences that form its parts, the correlation function will exhibit a shape similar to that of the autocorrelation function, although not as high. The amplitude of the cross-correlation functions will be a function of how many parts of the total sequence are represented. At least, for sequences that have been made up of only a few parts, the cross-correlation peaks are easily discernable from the "self noise" that exists off correlation.

C. PSEUDO NOISE CODE ACQUISITION

It must be assumed that when a signal is first received at the receiver, a search through all bit-time periods in the code must be made in order to acquire the received signal. While this is a worst case, it is one that is as likely to occur as any other and therefore the assumption must be made. In order to make a "correlation" or "no correlation" decision in each bit time, enough time (t_{CD}) must be allowed for the decision level to be reached. Thus, this particular time increment t_{CD}, is dependent upon the bandwidth of the decision filters. For a sequence of l_p bits length, the acquisition time for the code alone is $\left[t_{CD} \times l_p \right]$. If, in addition to a search of the code, a search of frequency uncertainty must be made, the value of t_{CD} must be lengthened sufficiently to allow the decision filters to "look" at the entire frequency uncertainty. For example, if the total unknown frequency area were to be ten times larger than the decision bandwidth, then the acquisition time would be $10 \times t_{CD} \times l_p$. It is rather easy to see that the acquisition search time can get very large for a long code. In the particular application addressed by this report, it is, therefore, desirable to keep the code as short as possible, i.e., only long enough to resolve range ambiguities.

Assuming the two signals are correlated, the output of interest of the mixing process will be a spectral line at the frequency difference of the received PN signal and the local reference. This signal is then amplified and applied to a phase detector. In the phase detector, the correlated spectral line is compared with a multiple of the VCO to develop a phase lock signal for the VCO. The VCO is also used to generate shift pulses for the pseudo noise generator (PNG). This receiving approach assumes the transmitted PN signal bit rate and carrier are phase coherent. Generally, a higher multiple of the VCO than the phase detector reference is modulated to provide the first mixer reference. By means of removing or inserting shift pulses, the PNG output can be shifted in time to mechanize a systematic time search for correlation. Generally an output of the phase detector is used to sense correlation and stop the time search process when full correlation has been achieved. If it is necessary to search out frequency uncertainties, this can be done with a VCO sweep signal in conjunction with the shift pulse gating for time search. The receiver briefly described here is of course, a basic phase locked receiver and can be analyzed as such.

In the succeeding sections of this Appendix, considerable analysis has been done to predict the performance of PN codes to the application of this study. Based upon the results of this analysis, an example is presented to illustrate possible performance of a pseudo noise system. The receiver of Figure E-6 is assumed to derive its input signal from a standard VHF receiver IF. To maintain coherence, the LO of the VHF receiver will be synthesized from the phase locked receiver VCO output. The phase detector operates at 5 MHz. Since the phase locked receiver can be unlocked by noise, a signal-to-noise ratio in the loop of 10 db is assumed to be minimum allowable for consistent operation. While 6 db is generally considered to be loop threshold for unlock, the probability of unlock, at this signal-to-noise ratio is quite high. For the 10 db signal-to-noise ratio, the phase jitter on the 5 MHz VCO derived signal can be calculated from the phase error relation of Section V,

$$S/N = \frac{1}{2 \, (\phi_\epsilon)^2} \qquad \text{(e-3)}$$

$$\phi_\epsilon = \frac{1}{\sqrt{2 \times 10}} = \frac{0.5}{\sqrt{5}} = 0.224 \text{ rad} \qquad \text{(e-4)}$$

At 5 MHz, this corresponds to a time jitter of:

$$\delta_T = \frac{0.224 \times 1/5 \times 10^6}{2\pi} = 7.1 \text{ nanoseconds } (1\sigma). \qquad \text{(e-5)}$$

In practice, the receiver would be built with a 5 MHz oscillator, which would be divided to produce the desired bit rate of 8 KHz. The division process generally does not reduce the time jitter. Therefore, the 7.1 nanosecond figure is a good estimate of the locally generated rms code jitter. This number is, of course, much, much smaller than the other range errors that will occur at VHF and as such, can be ignored.

E.   MEASUREMENT ACCURACIES

A major source of range measurement error for the pseudo-noise receiver is the receiver's ability to locate the peak of the correlation function. As can be seen from the preceding discussion on phase lock, the post-acquisition time jitter will be very small; however, the acquisition accuracy must be treated separately.

The accuracy with which the correlation function can be acquired is a function of the shape of the correlation function of the pseudo noise signal. The shape of the correlation function is a triangle for unfiltered pseudo noise; however, this is not the case for a filtered spectrum. As shown at the end of this section, the equations for the autocorrelation function were derived for a spectrum filtered at the first zero of the $(\sin x/x)^2$ shape. This function was evaluated by the time share computer to obtain data on both the autocorrelation function $R(\tau)$ and its derivative $R'(\tau)$. The data is included with the programs used to generate the information. Figure E-7 shows the autocorrelation functions of the filtered and unfiltered spectra.

Two approaches to acquisition of a PN correlation function have been popular. One approach is to differentiate the output of a single correlator and use the result as a control function to position a servo controlled local generator. Another approach is to use two correlators with their reference inputs displaced in time relative to each other. This is done quite simply by tapping separate stages of the shift register for the reference signals. The outputs of the two correlators are then differenced to yield a control function such as shown in Figure E-8.

With either approach, a servo control loop can be mechanized to center the local generator once the code search has established some correlation between the local generator and the received signal. In both cases, amplitude sensing of the correlation function must be used to sense acquisition and turn loop control over to the VCO phase lock instead of correlation lock. In the case of the single correlator, the peak of the correlation function must be sensed. For the double correlators the null of the differenced output can be used. As can be seen in Figure E-7, the filtering process has rounded off the peak of the auto correlation function. This makes the determination of the peak amplitude difficult, since the slope becomes so small. Figure E-9 is based upon the slope of the autocorrelation function and is an estimate of the signal-to-noise ratio required to make an acquisition decision to a given range accuracy. These curves ignore the factors of frequency errors, amplitude errors, equipment drift and the like.

The coherent pseudo noise system can make very precise range measurements once phase lock is established. The accuracy of those measurements, however, is a function of the accuracy of the acquisition decision, based upon the correlation function.

F.   ACQUISITION TIME

A factor which has a major effect on actual acquisition time is the frequency errors that the system must tolerate. As an illustrative example, assume one kilohertz frequency uncertainty and one watt transmitted power. Using the second column of Table III-1, as a power budget, some calculations can be made.

For a phase locked loop signal-to-noise ratio of 10 db, the bandwidth factor would be: $42 - 27.5 = 14.5$ db or 28.2 Hz closed loop noise bandwidth. Assuming double correlators and 16,000-foot acquisition accuracy, a 20 db acquisition circuit signal-to-noise ratio is necessary or a 2.8 Hz acquisition decision filter (ω_D).
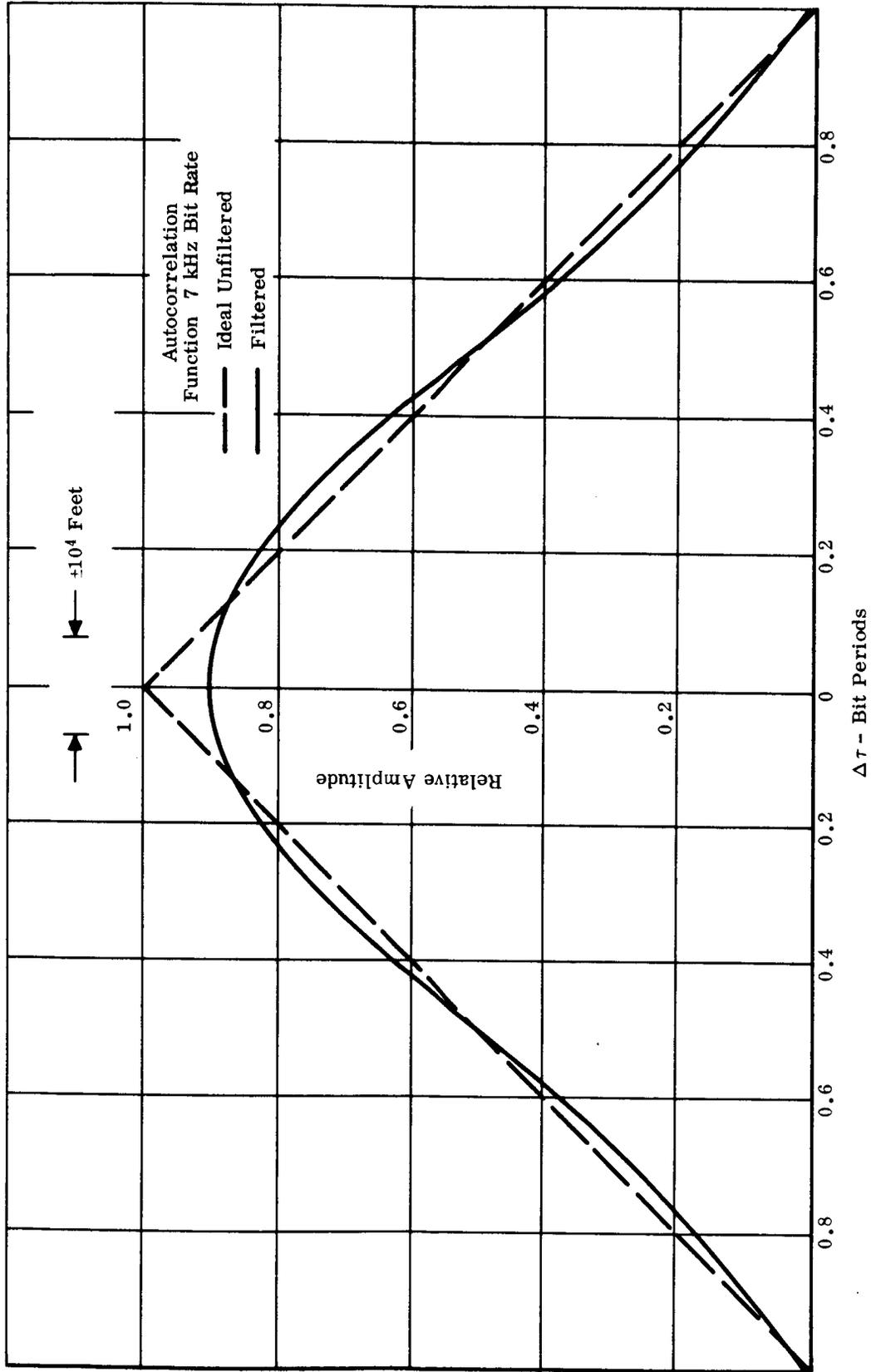
Figure E-7. Autocorrelation Functions—Filtered and Unfiltered Spectra
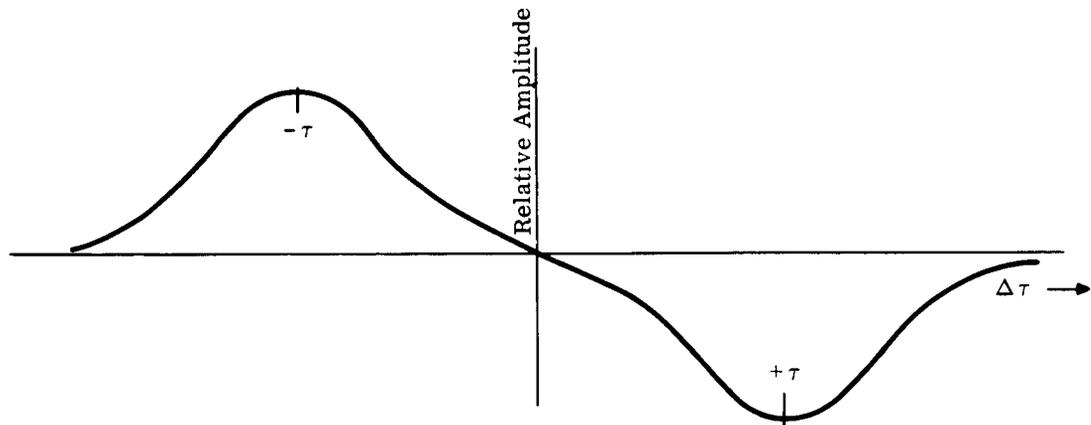
Figure E-8. Double Correlator Control Function

For initial correlation-no correlation decision, the output of the phase locked loop phase detector can be used. As such, the frequency error can be searched out in accordance with the results of Section V. Figure V-3 indicates a sweep rate of 100 Hz per second[2]. Thus, ten seconds will be used to search out the 1.0 kHz frequency uncertainty per bit of code. For the 24 bits indicated in the analysis, a worst case time of 240 seconds is necessary for establishment of correlation. For the 7 kHz bit rate being considered, the 16,000-foot range accuracy represents approximately 0.11 bit periods. Therefore, a search of nine time slots may be required to establish full acquisition. The double correlator will have resolved plus or minus limiting the search to one side of the control function of Figure E-8. For a $0.8/\omega_D$ time to decision per time slot, this will require 7.2/2 x 2.8 seconds or 0.406 second. Thus, a worst case acquisition would take 240.4 seconds.

This acquisition time is very much a function of the received signal level. The curves for search rate for a phase locked loop (Figure V-3) show:

$$SR = K \overline{B_n}^2 \qquad (e-6)$$

where

SR = sweep rate

B_n = closed loop noise bandwidth

K = constant depending upon loop type, generally $0.1 \leq K \leq 0.16$.

For a constant signal-to-noise ratio, the bandwidth is going to be directly related to transmitted power. Therefore, the search time due to frequency uncertainty will vary inversely with the square of transmitted power. Thus, 36 watts transmitted power could reduce the establishment of correlation to 40 seconds. The time search for range accuracy acquisition varies inversely with bandwidth. Therefore the time search of the correlation function would be 0.011 second. Of course, a direct time reduction can also be made by limiting the frequency uncertainty.

Acquisition accuracy for 7 kHz Bit-Rate
Pseudo-Noise ranging with the sequence
filtered at First Spectral Zero.
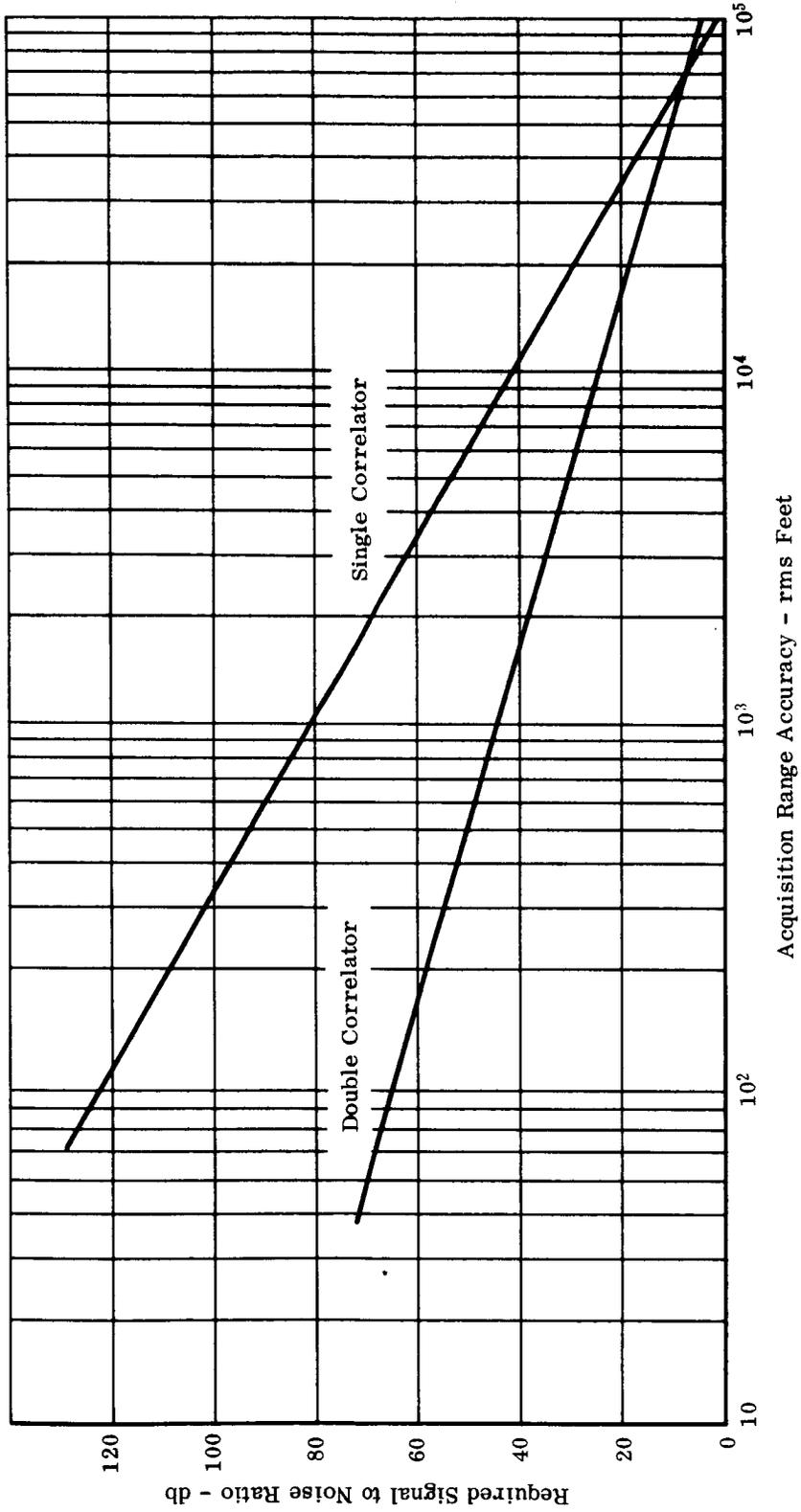
Figure E-9. Acquisition Accuracy

G. SUMMARY

In a practical sense, pseudo noise codes require a relatively sophisticated receiver and can require very long times for acquisition. Trade-offs of acquisition time for power can be made. However, if very short times are a goal, the power required can get very high.

The range accuracy is a direct function of the bit rate used for the code. For low bit rates, rather impractical signal-to-noise ratios for acquisition would be necessary. Time for signal-to-noise ratio trade-offs can be made in this area also. However, this will only hold as long as circuit noise after the decision filter remains small with respect to the noise out of the filter.

In general, PN codes appear to require too much time for acquisition and too costly a receiver for the applications envisioned in this study.

H. DETAILED ANALYSES

1. FREQUENCY BAND OCCUPANCY AND INFORMATION TRANSFER RATE

In a pulsed communication system with fixed sampling rate, the power spectrum of the signal waveform is shaped for bandwidth economy. A major consideration is minimizing inter-symbol interference and a minor consideration is reducing delay distortion. It is well-known that an ideal low-pass filter with a cutoff frequency at ω_v produces an impulse response of $\sin \omega_v t / \omega_v t$ type with uniformly spaced zeros at $t = k\pi/\omega_v$. (Where $k = \pm 1, 2, \ldots$.) When sampled at the proper times, no inter-sample interference occurs between such responses. But, because of the relatively large side lobes of this impulse response, errors in the sampling times give rise to an intolerable residual inter-symbol interference. Furthermore, relatively large delay variation associated with the sharp frequency cutoff of practical filters will time disperse the response and enhance the inter-symbol interference.

According to theoretical results of Nyquist, the oscillating nature of the pulse tail can be reduced and a more linear phase characteristic can be achieved by replacing this sharp cutoff by a more gradual cutoff characteristic having even symmetry about the original cutoff frequency ω_v. He has shown that such modification retains all the zero crossing points of the $\sin \omega_v t / \omega_v t$ type response and adds more zeros. In particular, we may consider the raised cosine type spectrum for a low-pass filter; i.e., for some real parameters, $u = |\omega| - \omega_v$ and $0 < \omega_x < \omega_v$, the even symmetric angular spectrum $A(|\omega|)$ can be specified in terms of u as A*(u), thus

$$A(|\omega|) = A^*(u) = \begin{cases} 1 & -\omega_v < u < -\omega_x \\ \frac{1}{2}(1 - \sin \pi u/2\omega_v) & |u| \leq \omega_x \\ 0 & u > \omega_x \end{cases} \qquad (e\text{-}7)$$

The resulting impulse response of $A(|\omega|)$ is given by Sunde (15) as

$$a(t) = \frac{\omega_v}{\pi} \frac{\sin \omega_v t}{\omega_v t} \frac{\cos \omega_x t}{\left[1 - (2\omega_x t/\pi)^2\right]} \cdot \qquad\qquad (e-8)$$

Figure E-10 shows the normalized response with ω_x/ω_v as a parameter, for $\omega_x/\omega_v = 1$, 3/4, and 1/2. With the reduction of ω_x/ω_v from unity to half, the maximum values of the side lobes of the response are slightly increased, but the uniformly spaced zeros of a(t) are still spaced $\pi/\omega_v = 1/2f_v$ apart.
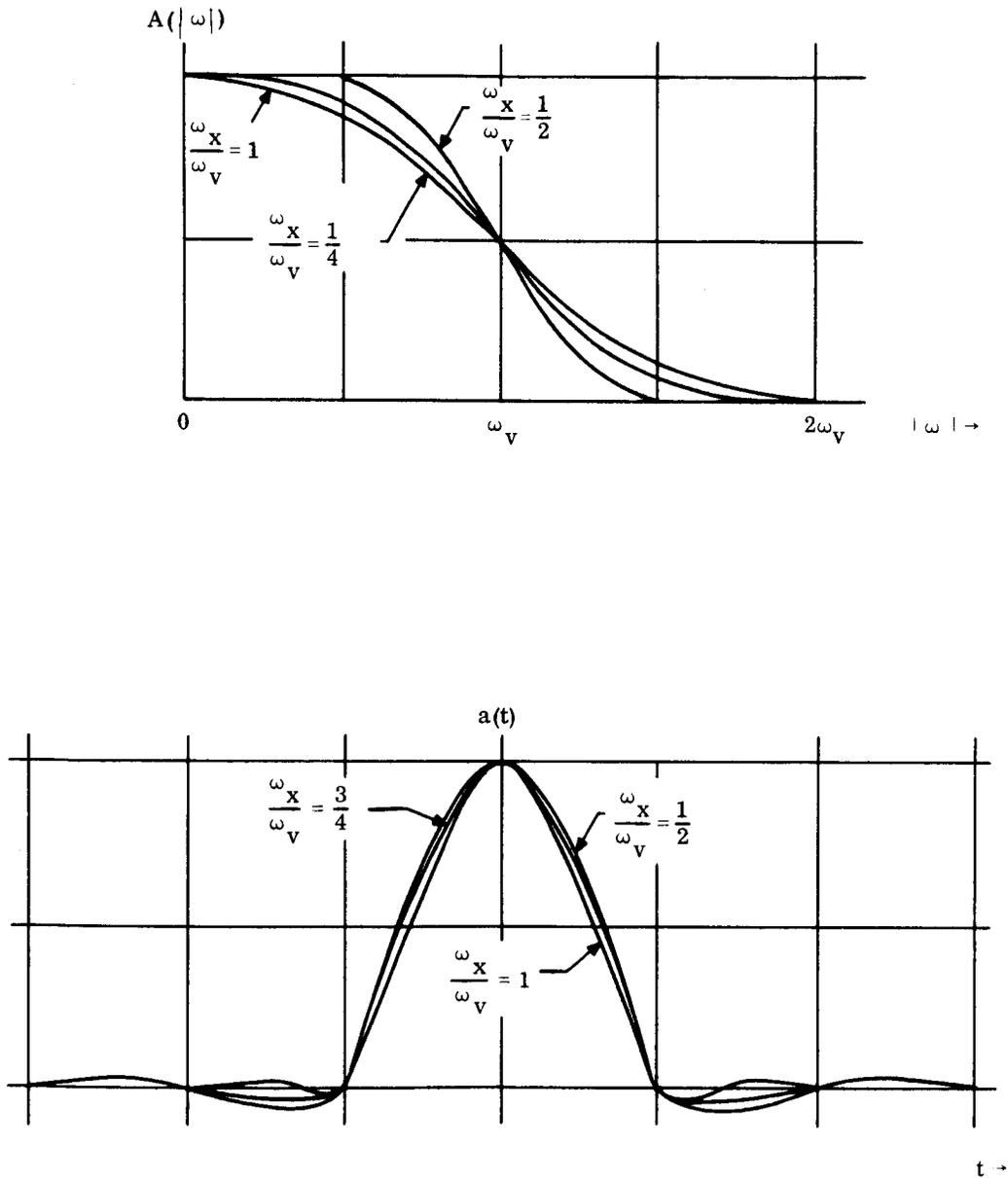
Figure E-10. Impulse Responses Obtainable with Various Sinusoidal Roll-Offs

For a signal transmitted through a time-varying medium, as long as the fading rate is relatively small in comparison with the rate of information transfer, the time variation in the medium can be neglected. The carrier phase can be determined at the receiver through a phase locked loop. In such a situation, it is highly desirable to use a product detector. The communication system is thus operated in a phase-coherent mode with vestigial sideband-suppressed carrier transmission to minimize the frequency band occupancy of the system. On a minimum bandwidth system, where only a single sideband is transmitted, the rf spectrum occupies a frequency band from the carrier frequency ω_c to an upper cutoff frequency ω_d. Since such transmission characteristics involve a sharp cutoff at the carrier frequency ω_c, it has all the drawbacks associated with the ideal low pass filter discussed before. On the other hand, a vestigial symmetric case requires a slightly greater frequency band occupancy, but it allows a gradual filter roll-off. By applying the Nyquist criterion used before, with the upper and lower sideband roll-off characteristic matched to yield an odd symmetry about the carrier frequency ω_c, the equivalent low-pass characteristics of the system will be preserved.

For some real parameters, let $v = |\omega| - \omega_c$ and $0 < \omega_y < \omega_d - \omega_c$. The bandpass angular frequency spectrum $B(|\omega|)$, which has even symmetry about zero frequency, is defined in terms of $B^*(v)$ as follows:

$$B(|\omega|) = B^*(v) = \begin{cases} 0 & \text{for } -\omega_v < v < -\omega_y \\[2mm] \frac{1}{2}\left(1 + \sin\frac{\pi v}{2\omega_y}\right) & |v| \leq \omega_y \\[2mm] 1 & \omega_y < v \leq \omega_d \\[2mm] 0 & v > \omega_d \end{cases} \qquad (e\text{-}9)$$

When the low band-pass analogy is applied, the resultant angular frequency spectrum $\tilde{c}(|\omega|)$ of the vestigial sideband system is defined as

$$\tilde{c}(|\omega|) = B(|\omega|)\left[A(|u + \omega_c|) + A(|u - \omega_c|)\right]. \qquad (e\text{-}10)$$

Recalling $v = |\omega| - \omega_c$, we can express $\tilde{c}(|\omega|)$ in terms of v as $c^*(v)$, then

$$\tilde{c}(|\omega|) = c^*(v) = \begin{cases} 0 & \text{for } -\omega_c < v < -\omega_y \\[2mm] \frac{1}{2}\left(1 + \sin\frac{\pi v}{2\omega_y}\right) & |v| \leq \omega_y \\[2mm] 1 & \omega_y < v < \omega_v - \omega_x \\[2mm] \frac{1}{2}\left[1 - \sin\frac{\pi(v - \omega_v)}{2\omega_x}\right] & \omega_v - \omega_x \leq v \leq \omega_v + \omega_x \\[2mm] 0 & v > \omega_v + \omega_x = \omega_d \end{cases} \qquad (e\text{-}11)$$

E-14

where the zero frequency of the low pass characteristic $A(|\omega|)$ is identified with the carrier frequency ω_c. The corresponding double sideband characteristic $\tilde{c}^*(|\omega|)$ is the band-pass analog of $B(|\omega|)$. It should equal the linear sum $c^*(v) + c^*(-v)$ which has even symmetry about $v = 0$; therefore

$$\bar{\bar{c}}(|\omega|) = \tilde{c}^*(|v|) = c^*(v) + c^*(-v) = \begin{cases} 0 & \text{for} & |v| > \omega_v + \omega_x \\[2mm] \dfrac{1}{2}\left[1 + \sin\dfrac{\pi(v + \omega_v)}{2\omega_x}\right] & -\omega_v - \omega_x \leq v \leq -\omega_v + \omega_x \\[2mm] 1 & -\omega_v + \omega_x \leq v \leq \omega_v - \omega_x \\[2mm] \dfrac{1}{2}\left[1 - \sin\dfrac{\pi(v - \omega_x)}{2\omega_x}\right] & \omega_v - \omega_x \leq v \leq \omega_v + \omega_x \end{cases} \qquad \text{(e-12)}$$

The condition that assumes that the maximum response of $\bar{\bar{c}}(|\omega|)$ remains equal to unity is given by

$$\omega_x + \omega_y \leq \omega_v = \omega_d - \omega_x \quad . \qquad \text{(e-13)}$$

For the case of utmost symmetry when $\omega_x = \omega_y$, the above relation is reduced into $\omega_x + \omega_y = 2\omega_x \leq \omega_v$ or $\omega_x \leq \omega_v/2$.

Then, in the expression of $\tilde{c}(|\omega|)$, the true $B(|\omega|)$ $A(|u + \omega_c|)$ has even symmetry with respect to $-(\omega_c + \omega_v/2)$ and the term $B(|\omega|)$ $A(|u - \omega_c|)$ has similar even symmetry with respect to $\omega_c + \omega_v/2$. If we also let $\omega_x = \omega_y = \omega_v/2$, and

$$w = |\omega| - \omega_c - \frac{\omega_v}{2} = |\omega| - \omega_c - \omega_x \quad , \qquad \text{(e-14)}$$

the band-pass characteristic will have even symmetry with respect to its band center frequency $w = 0$, and we may rewrite $\tilde{c}(|\omega|)$ in terms of w as $\hat{c}(|w|)$

$$\tilde{c}(|\omega|) = \hat{c}(|w|) = \begin{cases} \dfrac{1}{2}\left(1 - \sin\dfrac{\pi w}{2\omega_x}\right) & \text{for} \ |w| \leq \omega_x \\[3mm] 0 & |w| > \omega_x \quad . \end{cases} \qquad \text{(e-15)}$$

According to analysis by Sunde (15), such a roll-off characteristic is one of those closest to a linear phase characteristic as shown in Figure E-11. Therefore, it will only introduce negligible delay distortion.

All the spectra discussed above are shown in Figure E-12.
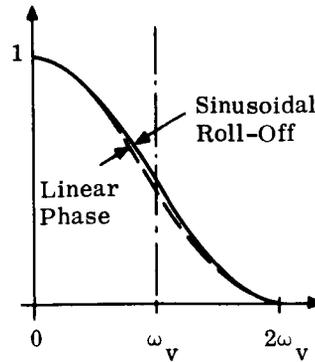
Figure E-11. Cross Approximation of the Sinusoidal Roll-Off Lowpass Filter
Characteristic to the Linear Phase Lowpass Filter Characteristic

2. SIGNAL-TO-NOISE RATIO AND FREQUENCY BAND OCCUPANCY FOR VARIOUS MODES OF
TRANSMISSION

If a modulating signal $v(t)$ modulates a sinusoidal carrier $A \cos \omega_c t$ by amplitude modulation, the
modulated signal $s(t)$ can be expressed as

$$s(t) = A \left[1 + v(t) \right] \cos \omega_c t \quad . \tag{e-16}$$

But if a balanced modulator is used to suppress the carrier, then the resulting double sideband signal
$s_2(t)$ is given as

$$s_2(t) = A v_2(t) \cos \omega_c t \quad . \tag{e-17}$$

Let $v_2(t) = \cos \omega_x t$, then,

$$s_2(t) = A \cos \omega_x t \cos \omega_c t = \frac{A}{2} \cos (\omega_c + \omega_x)t + \frac{A}{2} \cos (\omega_c - \omega_x)t \quad . \tag{e-18}$$

If a high pass filter with cutoff frequency ω_c is used to remove the lower sideband, then the single sideband
signal is given as

$$s_1(t) = A v_1(t) \cos \omega_c t \quad . \tag{e-19}$$

Suppose we let $v_1(t) = \sqrt{2} \cos \omega_x t$, then, after extracting the lower sideband, we obtain

$$s_1(t) = \frac{A}{\sqrt{2}} \cos (\omega_c - \omega_x)t \quad . \tag{e-20}$$

The peak power, $A^2/2$, of the double sideband signal $s_2(t)$ is reached when both sidebands reach their peak
power together. But the single side signal $s_1(t)$ also has a peak power of $A^2/2$. Thus, $s_1(t)$ and $s_2(t)$ have
identical transmitting peak power.
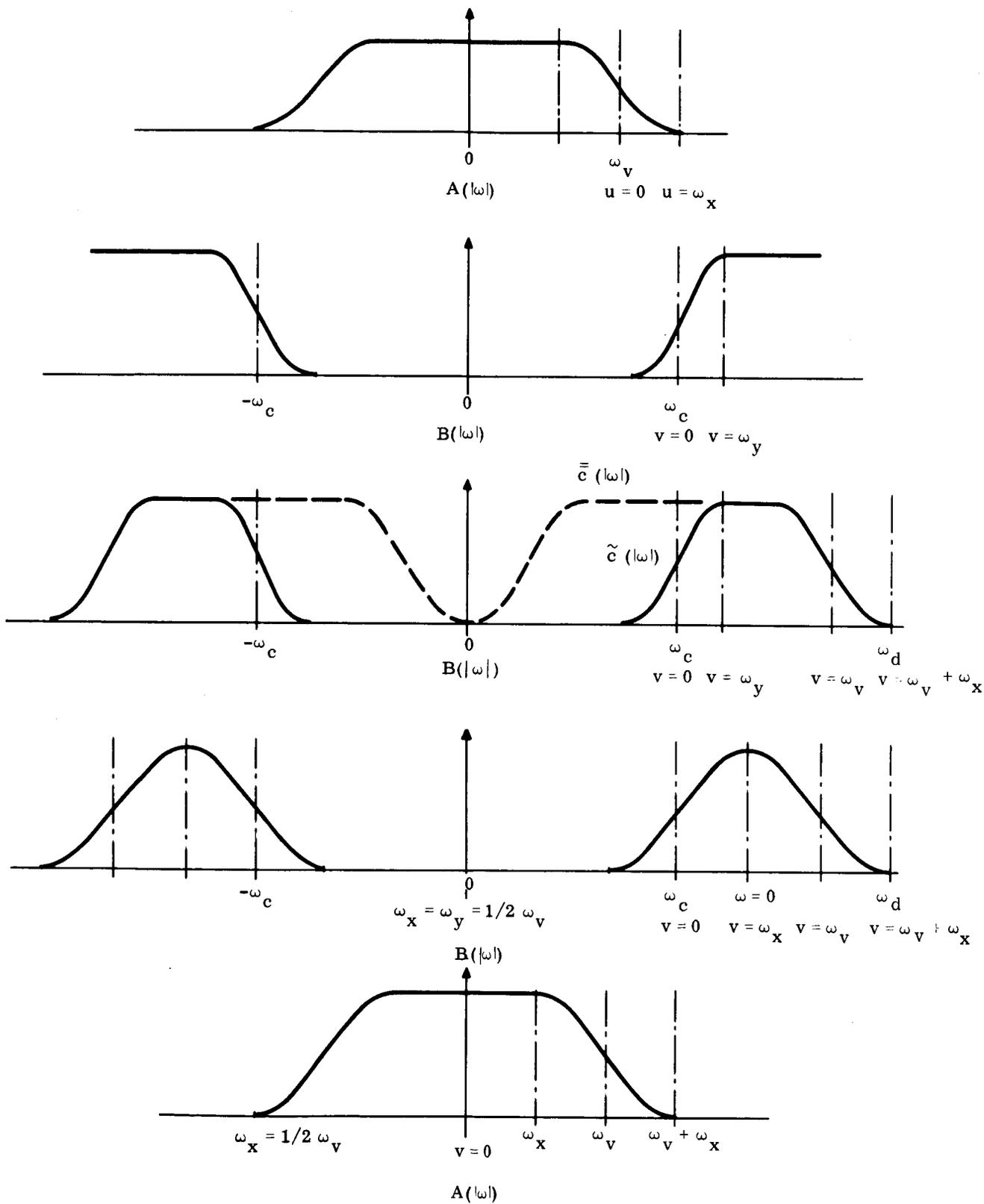
Figure E-12. Various Spectra

When a product detector with local oscillator voltage $K_m \cos \omega_c t$ is used to demodulate either the double or single sideband signal, the recovered signals $r_2(t)$ and $r_1(t)$ respectively, are given as

$$\tilde{r}_2(t) = K_m s_2(t) \cos \omega_c t = \frac{AK_m}{2} \cos \omega_x t + \frac{AK_m}{4} \cos (2\omega_c + \omega_x)t + \frac{AK_m}{4} \cos (2\omega_c - \omega_x)t \ . \qquad \text{(e-21)}$$

$$\tilde{r}_1(t) = K_m s_1(t) \cos \omega_c t = \frac{AK_m}{2\sqrt{2}} \cos \omega_x t + \frac{AK_m}{2\sqrt{2}} \cos (2\omega_c - \omega_x)t \ . \qquad \text{(e-22)}$$

After passing through an ideal low-pass filter with cutoff frequency ω_v, when $\omega_v > \omega_x$, the resultant baseband outputs $r_2(t)$ and $r_1(t)$, respectively, are given as

$$r_2(t) = \frac{AK_m}{2} \cos \omega_x t \ . \qquad \text{(e-23)}$$

$$r_1(t) = \frac{AK_m}{2\sqrt{2}} \cos \omega_x t \ , \qquad \text{(e-24)}$$

where the peak power of the recovered baseband signal for a double sideband system is exactly twice the corresponding peak power of a single sideband system.

If stationary white Gaussian noise $n(t)$ is associated with the transmitted signal for either $s_2(t)$ or $s_1(t)$ with power per unit bandwidth given as N_o, then $n(t)$ has a uniform power density of $N_o/4\pi$ per radian from $\omega = -\infty$ to ∞.

Since we are only interested in narrow band communication systems, we may replace the white noise $n(t)$ by a corresponding narrow band noise process expressed as

$$n(t) = n_c(t) \cos \omega_c t - n_s(t) \sin \omega_c t \ , \qquad \text{(e-25)}$$

where both $n_c(t)$ and $n_s(t)$ are stationary, white, Gaussian noise with identical power density per unit bandwidth as $n(t)$. Thus, the demodulated noise output from the product detector $\tilde{r}_n(t)$ is given as

$$\tilde{r}_n(t) = \frac{K_m n_c(t)}{2} + \frac{K_m n_c(t)}{2} \cos 2\omega_c t - \frac{K_m n_s(t)}{2} \sin 2\omega_c t \ . \qquad \text{(e-26)}$$

After passing through an ideal low pass filter with cutoff frequency ω_v, $\omega_v < 2\omega_c$, the baseband noise process $r_n(t)$ is given as

$$r_n(t) = \frac{K_m n_c(t)}{2} \ . \qquad \text{(e-27)}$$

If the receiver filter has a transfer function $H(\omega)$, the mean square output noise power σ_N^2 is given as

$$\sigma_N^2 = \frac{K_m^2}{4} E\left\{n_c^2(t)\right\} = \frac{K_m^2 N_o}{16\pi} \int_{-\infty}^{\infty} \left| H(|\omega|) \right|^2 d\omega = \frac{K_m^2 N_o}{8\pi} \int_{0}^{\infty} \left| H(|\omega|) \right|^2 d\omega \ . \qquad \text{(e-28)}$$

For the double sideband signal, the ideal bandpass filter $\left| H_2(|\omega|) \right|^2$ is specified by

$$\left| H_2(\omega) \right|^2 = \begin{cases} 1 & \text{for} & |\omega - \omega_c| \le \omega_v \\ 0 & & |\omega - \omega_c| > \omega_v \quad , \end{cases} \qquad \text{(e-29)}$$

and for the single sideband signal, $\left| H_1(|\omega|) \right|^2$ is specified by

$$\left| H_1(\omega) \right|^2 = \begin{cases} 0 & \text{for} & \omega - \omega_c < 0 \\ 1 & & 0 \le \omega - \omega_c < \omega_v \\ 0 & & \omega - \omega_c > \omega_v \quad . \end{cases} \qquad \text{(e-30)}$$

Therefore,

$$\sigma_{N_2}^2 = \frac{K_m^2 N_o f_v}{2} \quad \text{and} \quad \sigma_{N_1}^2 = \frac{K_m^2 N_o f_v}{4} \quad . \qquad \text{(e-31)}$$

The corresponding peak signal-to-mean square noise power ratios ρ_2 and ρ_1 for double and single sideband systems, respectively, are given by

$$\rho_2 = \frac{\dfrac{A^2 K_m^2}{4}}{\dfrac{K_m^2 N_o f_v}{2}} = \frac{A^2}{2 N_o f_v} \qquad \text{(e-32)}$$

$$\rho_1 = \frac{\dfrac{A^2 K_m^2}{8}}{\dfrac{K_m^2 N_o f_v}{4}} = \frac{A^2}{2 N_o f_v} \quad , \qquad \text{(e-33)}$$

where $f_v = \omega_v / 2\pi$. It is clear that $\rho_1 = \rho_2 = \rho$.

If we use a gradual roll-off filter instead of using an ideal bandpass filter, for the double sideband system, and also minimize inter-symbol interference, the previous analysis shows that the desired recovered signal spectrum is given by Equation (e-8). The corresponding bandpass characteristic for the case $2\omega_x = \omega_v$ is given by Equation (e-12) as $\overline{\overline{c}}_v(|\omega|)$. According to the matched filter theory for minimizing the interference caused by added Gaussian noise, the transmitting and receiving filter must have conjugate spectra so that

$$\left| H_2^*(\omega) \right| = \left| \sqrt{\overline{\overline{c}}_v(|\omega|)} \right| \quad . \qquad \text{(e-34)}$$

Then the power spectrum of the receiving filter is given by

$$\left| H_2^*(|\omega|) \right|^2 = \bar{\bar{c}}_v(|\omega|) \tag{e-35}$$

and

$$\int_{-\infty}^{\infty} \left| H_2^*(|\omega|) \right|^2 d\omega = 2 \int_0^{\infty} \bar{\bar{c}}_v \left| |\omega| \right| d\omega = 2 \int_0^{\infty} \left| H_2(\omega) \right|^2 d\omega = 2\omega_v \ . \tag{e-36}$$

Similarly, for the vestigial sideband case, when the desired bandpass characteristic is specified by Equation (e-15) as $\tilde{c}(|\omega|)$. Also according to matched filter theory, the receiving filter spectrum is

$$\left| H_1^*(|\omega|) \right| = \left| \sqrt{\tilde{c}(|\omega|)} \right| \ . \tag{e-37}$$

The power spectrum of the receiving filter is given by

$$\left| H_1^*(|\omega|) \right|^2 = \tilde{c}(|\omega|) \ , \tag{e-38}$$

and

$$\int_{-\infty}^{\infty} \left| H_1^*(|\omega|) \right|^2 d\omega = 2 \int_0^{\infty} c(|\omega|) d\omega = 2 \int_0^{\infty} \left| H_1(\omega) \right|^2 = \omega_v \ . \tag{e-39}$$

Therefore, the peak signal-to-mean-square-noise-power ratio for the double sideband gradual roll-off case and vestigial sideband case are given as ρ_2^* and ρ_1^*, respectively, and

$$\rho = \rho_2^* = \rho_1^* = \frac{A^2}{2N_o f_v} \ . \tag{e-40}$$

Therefore, the four different coherent systems, discussed, all have identical information transfer rates and identical peak signal-to-mean-square-noise power ratios, ρ. Their only difference is the frequency band occupancy.

The vestigial sideband system is chosen as an engineering compromise because of its smaller frequency band occupancy and lower sidelobe response. In the subsequent analysis, it is more convenient to determine signal-to-noise ratios on the basis of the double sideband system.

3. MULTIPLEXING SPEECH AND RANGING CHANNELS AND RANGE DETERMINATION

We next consider a particular communication system, in which two channels are multiplexed together—one channel for speech transmission and the other for ranging by a continuous waveform. The speech channel requires an information transfer rate of 7×10^3 pulses per second to avoid any sacrifice of its quality. We assume the ranging channel can handle the same information transfer rate theoretically. But in practice, its critical rate of information transfer may be much slower. Therefore, the total information transfer rate of the combined communication system is assigned a rate of 1.4×10^4 bits per second.

According to the previous analysis in this section, if a coherent system with vestigial sideband transmission is adopted, a frequency band occupancy of 1.4×10^4 Hertz is adequate. We will consider both time division multiplex, TDX, and frequency division multiplex, FDX, for multiplexing the two channels.

For coherent transmission with limited frequency band occupancy, the speech signals are most efficiently transmitted by bi-polar pulse amplitude modulation. Because such a system has a suppressed carrier, it need not transmit any pulses at all during pauses in speech. Thus, it achieves an important energy saving for satellite application.

Because the two channels are multiplexed together, the speech channel is always available except for the small portion of time set aside for user identification signals. The ranging channel contains a continuous sequentially modulated waveform, originated at the ground station. The user's equipment automatically transponds the ranging signal. Ranging and speech signals from the ground station are relayed from the satellite and demultiplexed in the user's receiver. The received ranging signal is multiplexed with the user's speech signal and shifted to a different frequency to be rebroadcast toward the satellite for relay to the ground station. Range determination is made at ground station. The total round trip delay τ_d is expressed in terms of discrete phase or number of bit intervals m_i, a positive interger, and θ the vernier phase is a positive real constant, $0 \le \theta \le 2\pi$, that is

$$\tau_d = \left(m_i + \frac{\theta}{2\pi} \right) d_r \quad , \tag{e-41}$$

where, $d_r = 1/f_v$ is the bit duration or interval, and $f_v = F_o/k_v$, where k_v is a positive integer and F_o is the frequency band occupancy of the combined communication system. The range accuracy is determined by the accuracy of the vernier phase, θ, which is determined by a phase locked loop.

The variance of range error σ_r^2 in the direction of a unit vector $\vec{u}$, that lies on the straight line joining the user and satellite, is

$$\sigma_r^2 = \sigma_\theta^2 \, d_r^2 \, \frac{c^2}{4\pi^2} \quad , \tag{e-42}$$

where σ_θ^2 is the variance of the vernier phase error θ_ϵ and c is the velocity of an electromagnetic wave. The variance σ_d^2 of the actual range error is the projection of σ_r^2 onto the user's horizon in the case of an aircraft, its artificial horizon. Let $\vec{h}$ be a unit vector in the direction determined by the intersection of the plane of incidence of the electro-magnetic wave and the user's artificial horizon. Then, the cosine of the angle Ψ between $\vec{u}$ and $\vec{h}$ is given by $|\vec{u} \cdot \vec{h}|$ and we have

$$\sigma_d^2 = |\vec{u} \cdot \vec{h}|^2 \sigma_\theta^2 \, d_r^2 \, \frac{c^2}{4\pi^2} \quad . \tag{e-43}$$

The signal-to-noise ratio is least at the edge of the coverage, and at the same time, the vector $\vec{u}$ almost lies on the user's artificial horizon; that is $|\vec{u} \cdot \vec{h}| \approx 1$, and $\sigma_d^2 \approx \sigma_r^2$, so that this case results in the maximum range error.

4. MODULATION BY BINARY SEQUENCE FOR RANGING AND ACQUISITION TIME

If the ranging system must cover a long unambiguous time interval T, the transmitted signal must equal or be larger than T. If such a signal s(t) is to be generated by a binary sequence v(t), modulating a sinusoidal carrier wave of angular frequency ω_c by amplitude modulation, the relation between s(t) and v(t) is given by

$$s(t) = A \left[1 + v(t)\right] \cos \omega_c t. \tag{e-44}$$

But if a balanced modulator is used to suppress the carrier, then

$$s(t) = Av(t) \cos \omega_c t. \tag{e-45}$$

Now if v(t) is a binary sequence having the values +1 and -1, clearly

$$s(t) = \begin{cases} +A \cos \omega_c t & \text{for } v(t) = +1 \\ -A \cos \omega_c t & v(t) = -1 \; . \end{cases} \tag{e-46}$$

The waveform of such a signal is called phase shift keying where the phase is shifted by π radians between the "mark" and "space" of the modulation signal.

Let $f_c = \omega_c/2\pi$, and let the binary sequence have a bit-rate of f_v bits per second, then, for the case of coherent transmission, the carrier frequency f_c must be an integral multiple of the bit-rate f_v. According to the previous analysis for the product detector, the frequency band occupancy is equal to the bit-rate of information transfer. Therefore, the bit duration of the binary sequence is given by

$$d_r = \frac{1}{f_v} \; . \tag{e-47}$$

Now if such system is time multiplexed into two channels, one of which is for ranging and the other for speech transmission, the length of the ranging binary sequence must be ℓ bits long such that

$$\ell \geq \frac{T}{2d} \; . \tag{e-48}$$

In the receiver, the component of greatest concern is the product detector. If the two inputs to the product detector are PSK signals of identical bit duration, as described before,

$$s_c(t) = A_c V_c(t) \cos (\omega_c t + \theta_c) \tag{e-49}$$

$$s_b(t) = A_b V_b(t) \cos (\omega_b t + \theta_b) \; , \tag{e-50}$$

then, the response of the product detector is

$$r_1(t) = \frac{A_c A_b K_m}{2} V_c(t) V_b(t) \cos\left[(\omega_c - \omega_b)t + (\theta_c - \theta_b)\right]$$

$$+ \frac{A_c A_b K_m}{2} V_c(t) V_b(t) \cos\left[(\omega_c + \omega_b)t + (\theta_c + \theta_b)\right] \quad , \qquad \text{(e-51)}$$

where K_m is the net gain through the product detector.

If one of the sidebands (say the upper one) is removed by a linear filter, and the resultant response $r(t)$ and the bit duration between the two input signals has been synchronized, (i.e., $\theta_c - \theta_b = 0$), then

$$r(t) = \frac{A_c A_b K_m}{2} V_c(t) V_b(t) \cos(\omega_c - \omega_b)t = A V_1(t) \cos(\omega_c - \omega_b)t \qquad \text{(e-52)}$$

where

$$A = \frac{A_c A_b K_m}{2} \qquad \text{and} \qquad V_1(t) = V_c(t) V_b(t) \quad . \qquad \text{(e-53)}$$

Since $V_c(t)$ and $V_b(t)$ are binary sequences, $V_1(t)$ must also be a binary sequence, which takes on the values +1 or −1, according to the following Karnaugh chart

		$V_b(t)$	
		+1	−1
$V_c(t)$	+1	+1	−1
	−1	−1	+1

It is interesting to note that, if a very narrow bandpass filter has its center frequency at angular frequency $\omega_c - \omega_b$, it will integrate $r(t)$. If both $V_c(t)$ and $V_b(t)$ have period T, then the sampled output of the narrow bandpass filter, after multiplication by $\cos(\omega_c - \omega_b)t$ in a product detector, is the correlator output for two input modulated sequences, as shown in Figure E-13.
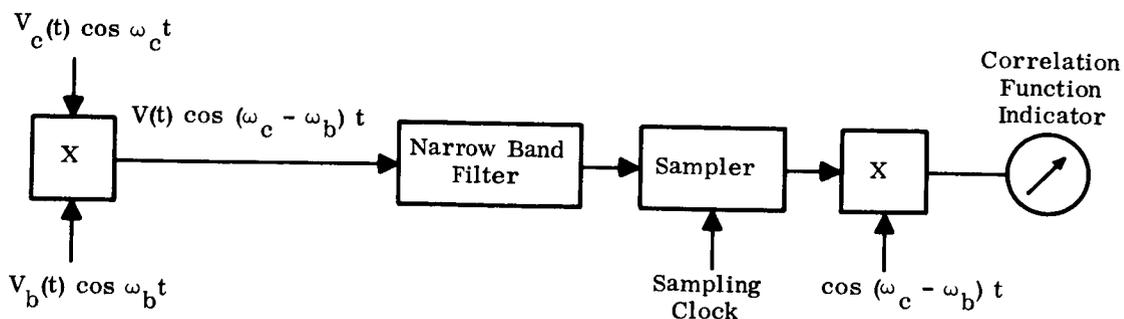

Figure E-13. The Correlator of Two Binary Sequence Modulated Waveforms

We now define acquisition time, $\hat{T}_{acq}$, as the time required for determining the discrete phase m_i or the exact number of bits that the received sequence has been shifted relative to the transmitted sequence. For any long sequence with a two level correlation function, such a determination can be made by selecting the peak output among all the outputs of a set of correlators whose reference sequences consist of every possible shifted version of the transmitted sequence. If an integrator takes k bit-durations to quench after sampling before it is ready for the next integration, and the total number of correlators in the receiver is h, then

$$\hat{T}_{acq} = \frac{\ell(\ell + k) d_r}{h} \quad , \tag{e-54}$$

is a good measure of the acquisition time. It is obvious $\hat{T}_{acq}$ is proportional to ℓ^2. The integer k is usually small compared with ℓ. Therefore, for a single receiver having a single correlator, $T_{acq} \approx \ell T$.

In order to reduce T_{acq}, the most obvious solution is to increase the number of correlators in the receiver. But for long unambiguous range, ℓ is large, and we would need hundreds of correlators before T_{acq} would be reduced to a practical value. An alternate way to achieve the same goal has been proposed by Golomb (16) and realized by Easterling (17). It has been successfully applied in spacecraft ranging systems. Since our problem is similar to the spacecraft ranging problem, a slightly modified version of this ranging system will be discussed in detail.

An apparent difficulty of binary correlator receivers is that simplicity can only be achieved with long acquisition time. This may be remedied, however, by adopting a composite sequence formed from many shorter sequences. For instance, if several sequences all have different periods, such that the common factor between any pair of periods is unity, then the common period of the composite sequence must be equal to the product of the periods of the component sequences. If a total of J component sequences is used to form a composite sequence in this way, the approximate period of component sequences is only the Jth- root of the period of the composite sequence. Then, according to the Chinese remainder theory (reference [18]) which states that

every system of linear congruence

$$x_1 = c_1 \bmod \ell_1$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$\cdot \tag{e-55}$$
$$x_j = c_j \bmod \ell_J$$

in which the moduli ℓ_j ($1 \leq j \leq J$) are pair wise relative prime, is solvable, and the modulus solution x is the product of all the moduli in the system, i.e. .

$$x = c \bmod \prod_{j=1}^{J} \ell_j = c \bmod \ell \quad . \tag{e-56}$$

E-24

Therefore, the ℓ discrete phases of the composite sequence is uniquely specified by a J-triple vector

$$c = (c_1, \ldots, c_J), \quad c_j \bmod \ell_j \quad \text{for } 1 \le j \le J,$$

(e-57)

the set of discrete-phases of all component sequences. Therefore, the approximate acquisition time of a ranging system, using such a composite sequence, is

$$T^J_{acq} = \frac{\left(\sum_{j=1}^{J} \ell_j\right)(\ell + k)\, d_r}{h} \approx \frac{J\, \ell^{1/J}(\ell + k)\, d_r}{h} \;.$$

(e-58)

But the ranging system, using composite sequences, requires a frequency band occupancy J times the original frequency band occupency, because now J component sequences must be transmitted in parallel.

5. BINARY SEQUENCES THAT ARE GENERATED BY BOOLEAN FUNCTION OF COMPONENT SEQUENCES AND THEIR CORRELATION FUNCTIONS

In a pseudo-random sequence of length ℓ (ℓ odd), there are $(\ell - 1)/2$ minus ones and $(\ell + 1)/2$ plus ones or vice versa. We assume here that the first case is true, and from now on, we use the notation "+" for "+1" and "-" for "-1." The relative frequencies with which the sequence assumes the individual value + or - are then given by

$$P(-) = \frac{(\ell - 1)}{2}$$

$$P(+) = \frac{(\ell + 1)}{2} \;.$$

(e-59)

Two sequences with relative prime periods, ℓ_1 and ℓ_2, respectively, $\left[\text{i.e., } (\ell_1, \ell_2) = 1\right]$, take on all possible discrete phases with respect to each other during the combined period $\ell = \ell_1 \ell_2$. Knowing the phase of either sequence does not give any information concerning the other sequence. If we state this condition differently, we consider the sampling of two sequences randomly. The occurrence of a particular digit in one sequence does not affect the probability of the corresponding digit in the other sequence. Thus, we are assured that the two sequences of mutually prime periods are mutually independent statistically. We may use this property to compute various in-phase and out-of-phase correlation functions between sequences. For example, we may choose the first sequence X: $\left\{+ + + - + - -\right\}$ and the second sequence Y: $\left\{+ + -\right\}$. The relative frequency of the value + or -, occurring in the X and Y sequence, are given as

$$X \begin{cases} p(-) = \dfrac{3}{7} \\[2mm] p(+) = \dfrac{4}{7} \end{cases}$$

(e-60)

$$Y \begin{cases} p(-) = \dfrac{1}{3} \\[2mm] p(+) = \dfrac{2}{3} \end{cases}$$

(e-61)

Then the joint relative frequencies of certain XY pairs occurring in the correlator are

$$p(-,-) = \frac{3}{7} \times \frac{1}{3} = \frac{3}{21}$$

$$p(-,+) = \frac{3}{7} \times \frac{2}{3} = \frac{6}{21}$$

$$p(+,-) = \frac{4}{7} \times \frac{1}{3} = \frac{4}{21} \qquad\qquad (e\text{-}62)$$

$$p(+,+) = \frac{4}{7} \times \frac{2}{3} = \frac{8}{21} \quad .$$

The above states of the auto-correlation function can be represented by the Karnaugh graph, where the resultant state is the product of the corresponding states in X and Y sequences as shown below at the left, and the corresponding joint relative frequencies are as show below at the right.

The relative frequency of bits between X and Y sequences, being in agreement, is the sum of the joint relative frequencies of all the states denoted + in the left square, and the relative frequency of bits in disagreement is the sum of the joint relative frequencies of all the states denoted by − in the left square. Therefore, the normalized correlation function $\rho_{xy}(j)$, according to the formula given in Equation (e-62), is as follows

$$\rho_{xy}(j) = p(+,+) + p(-,-) - p(-,+) - p(+,-) = \frac{1}{21} \qquad j \bmod \ell \quad . \qquad (e\text{-}63)$$

Such a Boolean sequence, formed by the Boolean function of two or more sequences, digit by digit, has the desired properties for ranging. Although it has multiple peaks in its auto-correlation function, the sub-peaks may be used for the alignment of component sequences. In particular, the most important properties of the composite sequence are: 1) the number of occurrences of + or − must be balanced; and 2) there should be a dominating sequence such that in the auto-correlation function, the sub-peak only occurs when the dominating component sequence is in phase. If only one non-dominating sequence is in phase or only several non-dominating sequences are in-phase at the same time, the correlation function must remain at small value.

Eastering gives such a Boolean function for three component sequences, X, Y and Z, when X is the dominating component sequence. The Boolean sequence U is defined by

$$Y = X (Y \& Z). \qquad (e\text{-}64)$$

where, all the component sequences are pseudo random sequences, having mutually prime periods ℓ_x, ℓ_y and ℓ_z. The period of U is $\ell = \ell_x \ell_y \ell_z$.

The "AND" ("OR") logic used to combine the sequences Y and Z and the product ("AND") logic used to combine X and (Y & X) are given by the following Karnaugh graphs.

$$
\begin{array}{c}
\text{Y} \\
\begin{array}{cc}
- & + \\
\end{array} \\
\text{Z} \quad
\begin{array}{c}
- \\
+ \\
\end{array}
\begin{array}{|c|c|}
\hline
- & - \\
\hline
- & + \\
\hline
\end{array} \\
\text{"OR"}
\end{array}
\qquad
\begin{array}{c}
\text{Y \& Z} \\
\begin{array}{cc}
- & + \\
\end{array} \\
\text{X} \quad
\begin{array}{c}
- \\
+ \\
\end{array}
\begin{array}{|c|c|}
\hline
+ & - \\
\hline
- & + \\
\hline
\end{array} \\
\text{"AND"}
\end{array}
$$
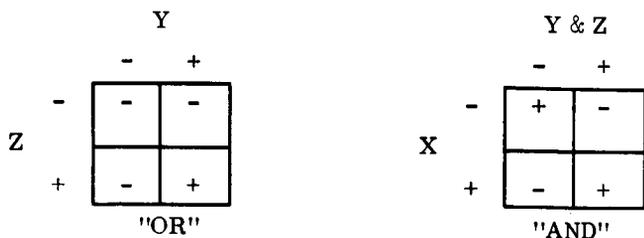
Every possible state of this Boolean sequence is given by

$$
\left.
\begin{array}{lcccccccc}
\text{Z} & - & + & + & - & - & + & + & - \\
\text{Y} & - & - & + & + & + & + & - & - \\
\hline
\text{(Y\&Z)} & - & - & + & - & - & + & - & - \\
\text{X} & - & - & - & - & + & + & + & + \\
\hline
\text{U} & + & + & - & + & - & + & - & - \\
\end{array}
\right\} \quad \text{(e-65)}
$$

The computations of the correlation function of binary Boolean sequences with pseudo random sequences as components are much simplified by assuming the component sequences to have an equal number of +'s and -'s—the longer the sequence period, the better the approximation. Under such an approximation, when computing the auto-correlation function of U and u (where the lower case letter is used to differentiate the out-of-phase version of the same sequence denoted by the capital letter), the following four different situations between corresponding values of U and u occur equally often:

(-, -) (-, +) (+, -) and (+, +) .

Since the particular Boolean function, specified U, is symmetric in Y and Z, we note, in computing the correlation of U and u, that the correlation with the Y component alone in phase must be identical with the correlation with the Z component alone in phase, and the correlation with X and Y components in phase must be identical with the correlation with X and Z components in phase. In computing the correlation, we list every possible state of U and u sequences on adjoining sides of a square; then in every element of the square matrix, if the column state is in agreement with the row state, + is entered, otherwise - is entered. The number of + and - entires in the square matrix must now be counted. The out-of-phase correlation is easily computed in this way, for each entry occurs equally often. When some components of the sequence are in phase, certain combinations of states in the square matrix become forbidden states. If we shade all of these forbidden states, the correlation can be found by counting the number of agreements and disagreements in the remaining unshaded squares of the square matrix. In Figure E-14, we have determined the auto-correlation function of the U sequence by this method. The results can be summarized as follows:

For no components in phase, $\rho_u = 0$

For Y component alone in phase, $\rho_u = 0$

$$U = X(Y \& Z)$$

			Z	-	+	+	-	-	+	+	-
			Y	-	-	+	+	+	+	-	-
			X	-	-	-	-	+	+	+	+
Z	Y	X	U	-	-	+	-	+	-	+	+
-	-	-	-	+	+	-	+	-	+	-	-
+	-	-	-	+	+	-	+	-	+	-	-
+	+	-	+	-	-	+	-	+	-	+	+
-	+	-	-	+	+	-	+	-	+	-	-
-	+	+	+	-	-	+	-	-	+	-	-
+	+	+	-	+	+	-	+	+	-	+	+
+	-	+	+	-	-	+	-	-	+	-	-
-	-	+	+	-	-	+	-	-	+	-	-

ρ_u = 0, No Component in Phase

ρ_u = 1, All Components in Phase

	Y	-	-	+	+	+	+	-	-
Y	U	-	-	+	-	+	-	+	+
-	-	+	+					-	-
-	-	+	+					-	-
+	+			+	-	+	-		
+	-			-	+	-	+		
+	+			+	-	+	-		
+	-			-	+	-	+		
-	+							+	+
-	+	-	-					+	+

ρ_u = 0, Y Component in Phase

	X	-	-	-	-	+	+	+	+
X	U	-	-	+	-	+	-	+	+
-	-	+	+	-	+				
-	-	+	+	-	+				
-	+	-	-	+	-				
-	-	+	+	-	+				
+	+					+	-	+	+
+	-					-	+	-	-
+	+					+	-	+	+
+	+					+	-	+	+

ρ_u = 1/4, X Component in Phase

		Z	-	+	+	-	-	+	+	-
		Y	-	-	+	+	+	+	-	-
Z	Y	U	-	-	+	-	+	-	+	+
-	-	-	+		-			+		-
+	-	-								
+	+	+	-		+			-		+
-	+	-								
-	+	+								
+	+	-	+		-			+		
+	-	+								
-	-	+	-		+			+		-

ρ_u = 0, Y & Z Components in Phase

		Y	-	-	+	+	+	+	-	-
		X	-	-	-	-	+	+	+	+
Y	X	U	-	-	+	-	+	-	+	+
-	-	-	+	+						
-	-	-	+	+						
+	-	+			+	-				
+	-	-			-	+				
+	+	+					+	-		
+	+	-					-	+		
-	+	+							+	+
-	+	+							+	+

ρ_u = 1/2, X & Y Components in Phase

Figure E-14. Auto-Correlation Function of U

For Z component alone in phase, $\rho_u = 0$

For Y and Z components in phase together, $\rho_u = 0$

For X component alone in phase, $\rho_u = 1/4$

For X and Y components in phase together, $\rho_u = 1/2$

For X and Z components in phase together, $\rho_u = 1/2$

For all X, Y and Z components in phase, $\rho_u = 1$.

The above method can also be applied to compute the cross-correlation function between a Boolean sequence and one of its component sequences; likewise the cross-correlation function between a Boolean function and some of its composite sub-sequences, formed by several components, can be computed. The cross-correlation function of the U sequence is computed by this method, as shown in Figure E-15. The results can be summarized as follows:

For no component in phase, $\rho_{UX} = 0$

For X component in phase, $\rho_{UX} = -1/2$

For no component in phase, $\rho_{UY} = 0$

For Y component in phase, $\rho_{UY} = 0$

For no component in phase, $\rho_{UZ} = 0$

For Z component in phase, $\rho_{UZ} = 0$

For X component in phase alone, $\rho_{U(XY)} = 0$

For X & Y component in phase together, $\rho_{U(XY)} = -1/2$

For X component in phase alone, $\rho_{U(XZ)} = 0$

For X & Z components in phase together, $\rho_{U(XZ)} = -1/2$

Based upon the known properties of the auto-correlation function and cross-correlation function of the Boolean sequence U, we can formulate the acquisition procedure as follows:

First step: cross-correlate U with X component alone

    X component not in phase $\rho_{UX} = 0$

Second step: stepping X component until

    X component in phase $\rho_{UX} = -1/2$

Third step: auto-correlate u with U, ($|\rho_u|$ decreases)

    X component in phase alone $\rho_u = 1/4$

Fourth step: stepping Y component until

    X and Y components in phase together $\rho_u = 1/2$

Fifth step: stepping Z component until

    X, Y and Z components all in phase $\rho_u = 1$

The block diagram of such a device, with a phase locked-loop to acquire the clock phase, is shown in Figure E-16. This particular Boolean function, which generates the U sequence, is limited to three sequences. This is sufficient in our application. But there exist other good Boolean functions, which can be used to generated Boolean sequences with larger numbers of components, if longer unambiguous time intervals are required.

$\rho_{ux} = 0$, X Component out of Phase

V = XY, No. Z Component

$\rho_{ux} = -1/2$, X Component in Phase

$$\left[X_{(Y \& Z)} \right] \left[XY \right]$$
$$= (Y \& Z)Y$$

$\rho_{uv} = 0$, X Component in Phase
Y Component out of Phase

$\rho_{uy} = 0$, Y Component out of Phase

$$\left[X_{(Y \& Z)} \right] \left[XY \right]$$
$$= (Y \& Z)Y$$

$\rho_{uv} = -1/2$, X Component in Phase
Y Component in Phase

$\rho_{uy} = 0$, Y Component in Phase

Figure E-15. Cross-Correlation of U and Its Components

From Receiver I-F

Phase Locked Loop

$X \left[Y \And Z \right] \cos w_c t$
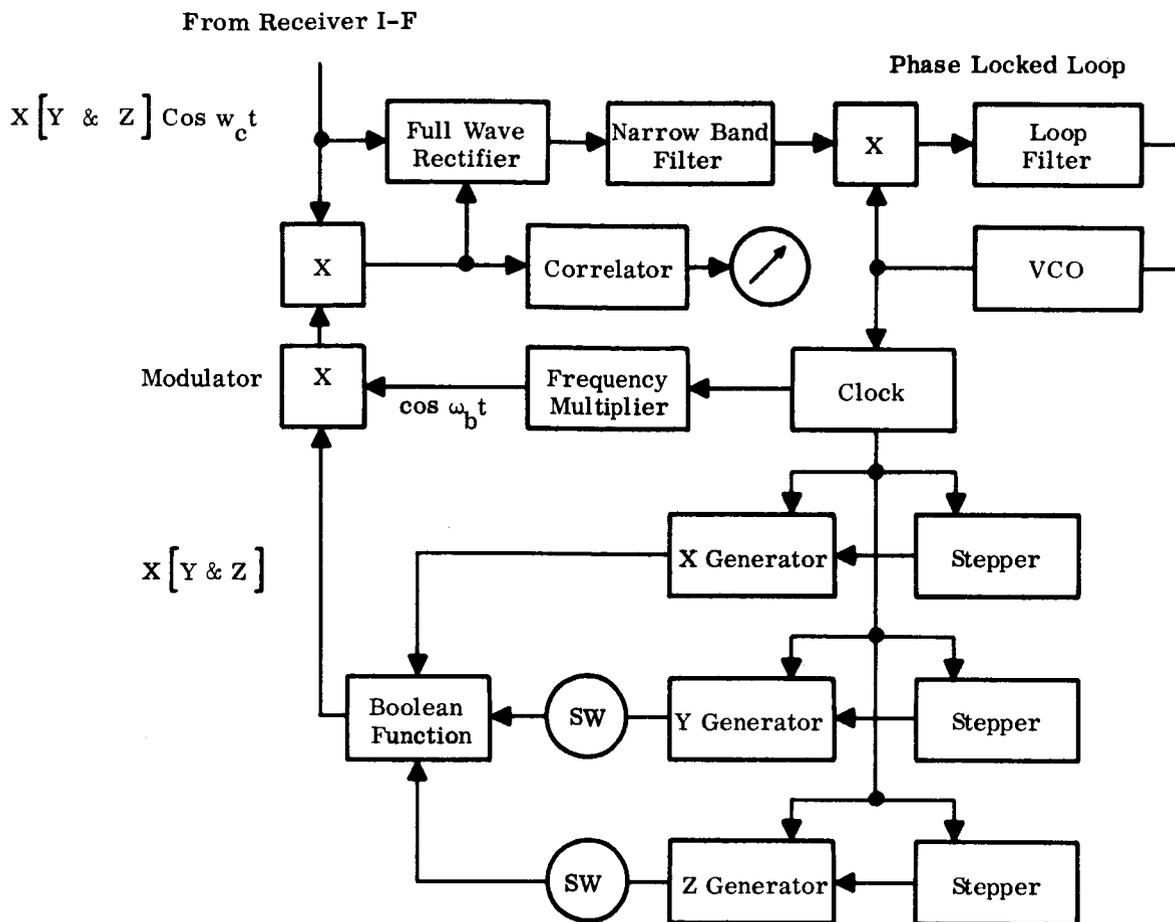
Figure E-16. Block Diagram of Correlation Receiver

6. PSEUDO-RANDOM SEQUENCES AND SHIFT REGISTER SEQUENCE GENERATORS

All the pseudo-random sequences of length $\ell = 4t - 1$ are known only when: (a) $4t - 1$ is a prime number; (b) $4t - 1 = 2^k - 1$; or (c) $4t - 1$ is equal to the product of two successive odd numbers, both of which are prime. These three types of pseudo-random sequence are called respectively: (a) Legendre sequences; (b) m-sequences; and (c) Jacobi sequences.

a. Legendre Sequences

A Legendre sequence (reference [20]) is also called a quadratic residue sequence, when $\ell = 4t - 1$ is a prime number. Then, a two-level sequence $\left\{ x_j \right\}$ of length ℓ can be constructed by putting

$$x_j = \begin{cases} 1 & \text{if j is a quadratic residue mod } \ell \\ -1 & \text{otherwise.} \end{cases} \tag{e-66}$$

When ℓ is an odd prime number, the integer j is a quadratic residue mod ℓ, if $y^2 = i \pmod{\ell}$ has a solution and j and ℓ are mutually prime, i.e., $(j, \ell) = 1$. Then x_i can also be defined by the Legendre symbol

$$x_j = \left(\frac{j}{\ell} \right) . \tag{e-67}$$

b.   m-Sequence (21)

The m-sequence stands for the maximal length, linear recurring sequence. Its construction is by recursive formula, which can be derived from the algebraic theory of such sequences (reference [21]). If an m-sequence of period $2^k - 1$ is desired, it can be based upon any irreducible k^{th} degree polynomial over a Golois Field of two elements. Since any field must contain the zero element and the unity element, the two element fields must be composed of these two elements. For any degree k, associated with the irreducible polynomial

$$f(x) = \sum_{j=0}^{n} \hat{C}_j \hat{X}_j = 0 , \tag{e-68}$$

there is the recursive relation

$$\hat{C}_o \hat{X}_m + \hat{C}_1 \hat{X}_{m-1} + \dots + \hat{C}_k \hat{X}_{m-k} = 0, \tag{e-69}$$

which defines a binary sequence $\{\hat{X}_j\}$ in elements 0 and 1 of length $2^k - 1$, with any initial k states $(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_k)$, with all $\hat{X}_i$ not all equal to zero for $1 \le j \le k$. If it is an irreducible polynomial of degree k, the coefficients $\hat{C}_o$ and $\hat{C}_k$ must be non-zero. Therefore, $\hat{C}_o = \hat{C}_k = 1$. Then the above recurrence relation defines $\hat{X}_{k+1}$ from assigned $\hat{X}_1, \hat{X}_2, \dots, \hat{X}_k$. The formula for defining the logic to produce a successor is given as

$$\hat{X}_{k+1} = \hat{X}_1 + \sum \hat{X}_{k-j} \tag{e-70}$$

$$1 \le j \le k - 1$$

$$\hat{C}_j \ne 0 .$$

The most simple irreducible polynomial is a trinomial, for which the summation in the above formula consists of only a single term. Unfortunately, not all values of k have an irreducible trinomial. Watson (22) lists a primitive (i.e., period $2^k - 1$) irreducible polynomial for all k, $k \le 100$.

Now we need an m-sequence $\{X_j\}$ with binary values +1 and -1. Such a sequence is defined as

$$X_j = (-1)^{\hat{X}_j} . \tag{e-71}$$

The corresponding formula, for defining the logic to produce the successor, is given by

$$X_{k+1} = X_1 \prod X_{k-j} \tag{e-72}$$

$$1 \le j \le k - 1$$

$$\hat{C}_j \ne 0$$

and will generate a periodic sequence of length ℓ, if initial states $X_1, \dots, X_k$ are not all +1.

c. Jacobi Sequence (23)

If ℓ_i is an odd prime for i = 1, 2, and $\ell = \ell_1 \ell_2$, then the Jacobi symbol (i/ℓ) is related to the Legendre symbols (i/ℓ_1) and (i/ℓ_2) as follows

$$\left(\frac{i}{\ell}\right) = \left(\frac{i}{\ell_1}\right)\left(\frac{i}{\ell_2}\right) \qquad \text{where } (i,\ell) = 1. \tag{e-73}$$

Now let $\ell_2 = \ell_1 + 2$, then the Jacobi sequence of length ℓ is defined as

$$X_i = \begin{cases} (i/\ell) & \text{if } (i,\ell) = 1 \\ 1 & i = 0 \bmod \ell_2 \\ -1 & \text{otherwise.} \end{cases} \tag{e-74}$$

Such a Jacobi sequence is also called a twin prime sequence.

All the pseudo-random sequences can be generated by shift registers. For a shift register of k consecutive 2-state memory cells, regulated by a common clock, at each clock pulse, the state (+ or -) of each memory cell is shifted into the next cell on the line. A shift register is converted into a periodic sequence generator, through the addition of a suitably connected feedback loop. The output state of the feedback loop is determined by the kth-state in the memory cell of the shift register before the shifting. The feedback loop can be divided into direct and logical portions. The successor state is the product of the output of the direct and logical portions. The output of the direct portion is the product of the state of the kth cell and some odd numbers of states of particular chosen cells, among the 1st to the $(k-1)^{\text{th}}$ cells. The output of the logic portion is the Boolean function of the states of the 1st to $(\ell-1)^{\text{th}}$ cells inclusive. Such a shift register is sketched in Figure E-17.

For the generation of an m-sequence 2^k-1 bits long, a k-cell shift register sequence generator is adequate. For an irreducible trinomial, the direct portion is the product of the 1st and another state. For certain k, such as k = 8, 12, 13, 14, 16 and 19, there is no irreducible trinomial. For k ≤ 20, the output of the direct portion is the product of 4 states. For the m-sequence generator, the Boolean function of the indirect portion can be omitted, or alternatively provides an affirmative "-" output, when all first k-1 cells are in the "+" state, to assure that the generator be self-starting, even when initially loaded with all + states.

A possible logic for the direct portion of the k-cell m-sequence generator is listed in Table E-1, for k ≤ 20.

For generating Legendre or Jacobi sequences, the length of this shift register must be at least equal to the longest run of identical states in the sequence, if there are no identical words of length equal to the longest run within the period. Otherwise, the least length is that, which assures no identical words within the period. Baumert has listed the Boolean functions in the logical portion for generating such sequences by shift register generators, for lengths up to 107 bits, while the logic in the direct portion of such a generator is fixed as $X_1 X_2$, in the book cited as reference (19).

$$X_j = \begin{cases} +1 \\ -1 \end{cases}$$
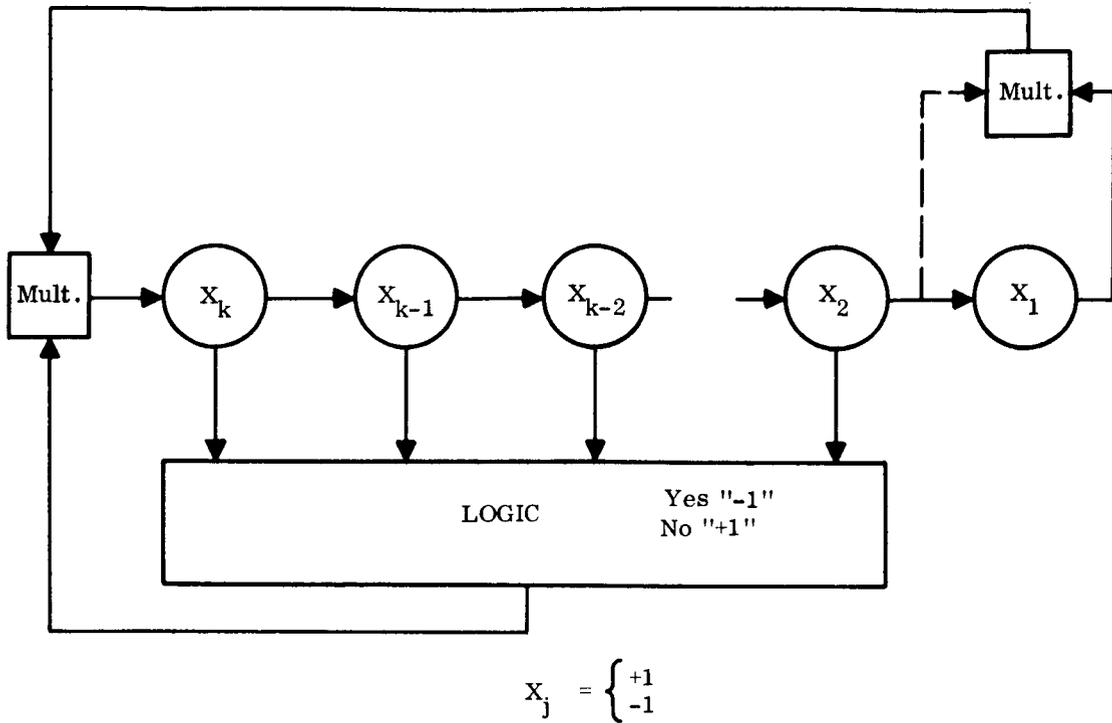
Figure E-17. Shift Register Sequence Generator

Table E-1

Logic of Direct Portion

k	direct portion	k	direct portion	k	direct portion	k	direct portion
1	X_1	6	X_1X_2	11	X_1X_3	16	$X_1X_3X_4X_6$
2	X_1X_2	7	X_1X_2	12	$X_1X_2X_3X_{11}$	17	X_1X_4
3	X_1X_2	8	$X_1X_5X_6X_7$	13	$X_1X_2X_3X_{13}$	18	X_1X_8
4	X_1X_2	9	X_1X_5	14	$X_1X_2X_3X_{13}$	19	$X_1X_2X_3X_6$
5	X_1X_3	10	X_1X_4	15	X_1X_2	20	X_1X_4

7. DETERMINE THE PERIODS OF COMPONENT SEQUENCES IN THE BOOLEAN SEQUENCE AND THE ACTUAL IMPROVEMENT IN ACQUISITION TIME

We assume an ambiguous propagation time interval $T = 58 \times 10^{-3}$ seconds, which is longer than the two-way range ambiguity interval for a synchronous satellite ranging to users on the earth. First, we determine the bit duration d_r of the ranging sequence, by letting $k_v = 2$, then the bit duration for the TDX and FDX ranging sequence is given respectively as

$$(d_r)_T = \frac{2}{1.4 \times 10^{-4}} = 1.43 \times 10^{-4} \text{ seconds.} \tag{e-75}$$

$$(d_r)_F = \frac{2}{7 \times 10^{-4}} = 2.86 \times 10^{-4} \text{ seconds.} \tag{e-75}$$

E-34

Then the length ℓ of period of the ranging sequence, expressed in bits, is given respectively as

$$\ell_T = \frac{58 \times 10^{-3}}{1.43 \times 10^{-4}} \geq 406 \quad , \tag{e-77}$$

$$\ell_F = \frac{58 \times 10^{-3}}{2.86 \times 10^{-4}} \geq 203 \quad . \tag{e-78}$$

For the component Boolean sequence, $U = X(Y \& Z)$, we choose X with period $\ell_x = 2$ for both TDX and FDX systems. Let $\ell_y = 11$ and $\ell_z = 19$ for TDX system, then

$$\ell_T = \ell_x \cdot \ell_y \cdot \ell_z = 2 \times 11 \times 19 = 418 \geq 406. \tag{e-79}$$

But $\ell_y = 7$ and $\ell_z = 15$ for FDX system. Then

$$\ell_F = \ell_x \cdot \ell_y \cdot \ell_z = 2 \times 7 \times 15 = 210 \geq 203 \quad . \tag{e-80}$$

For determining the acquisition time, replace the approximation value $J\sqrt{\ell}$ by the exact value

$$S = \sum_{j=1}^{J} \ell_j \quad , \quad \text{then we have} \tag{e-81}$$

$$S_T = (\ell_x + \ell_y + \ell_z)_T = 2 + 11 + 19 = 32 \tag{e-82}$$

$$S_F = (\ell_x + \ell_y + \ell_z)_F = 2 + 7 + 15 = 24 \quad . \tag{e-82}$$

8. AUTO-CORRELATION FUNCTION SHAPE

For the case of Pseudo Noise that has been filtered, the auto-correlation function $R_o(\tau)$, will not be a triangle, but instead will be "rounded off." To calculate the shape of this function, the pseudo noise signal's power spectrum $S_i(\omega)$ is considered to be filtered, so as to only have energy out to the first zero of the spectrum. To calculate the $R_o(\tau)$, the auto-correlation function $R_i(\tau)$ of the unfiltered $S_i(\omega)$ is convolved with $h(\tau)$, which is the impulse response of an ideal filter, which cuts $S_i(\omega)$ off at the first zero.

With $R_o(\tau)$ defined, the function is differentiated to form $R_o'(\tau)$, so that the slope of $R_o(\tau)$ can be calculated.

The input signal, $f_i(t)$, to an ideal low-pass filter, $H(\omega)$, is a pulse defined by:

$$f_i(t) = \begin{cases} 1 & \text{for } |t| < 1/2 \\ 0 & \text{elsewhere.} \end{cases} \qquad \text{(e-84)}$$

Thus, the input signal can be represented by
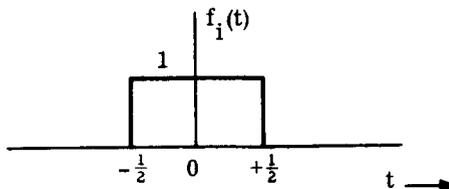
where

$$H(\omega) = \begin{cases} 1 & \text{for } |\omega| \leq 2\pi \\ 0 & \text{elsewhere.} \end{cases} \qquad \text{(e-85)}$$
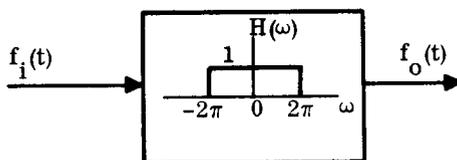
The frequency spectrum of the input signal, $F_i(\omega)$, is thus given by:

$$F_i(\omega) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-j\omega t} \, dt = \frac{\sin \frac{\omega}{2}}{\frac{\omega}{2}} \, . \qquad \text{(e-86)}$$

Now the Power Spectrum of the input signal, $S_i(\omega)$, is defined by

$$S_i(\omega) = F_i(\omega) \, F_i^*(\omega) = \frac{\sin^2 \left(\frac{\omega}{2}\right)}{\left(\frac{\omega}{2}\right)^2} \, , \qquad \text{(e-87)}$$

where * denotes complex conjugate. The auto-correlation function is then determined to be:

$$R_i(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_i(\omega) \, e^{+j\tau\omega} \, d\omega \, . \qquad \text{(e-88)}$$

Or

$$R_i(\tau) = \begin{cases} \tau + 1 & \text{for } -1 \leq \tau \leq 0 \\ -\tau + 1 & \text{for } 0 \leq \tau \leq 1 \\ 0 & \text{elsewhere} \end{cases} \qquad \text{(e-89)}$$
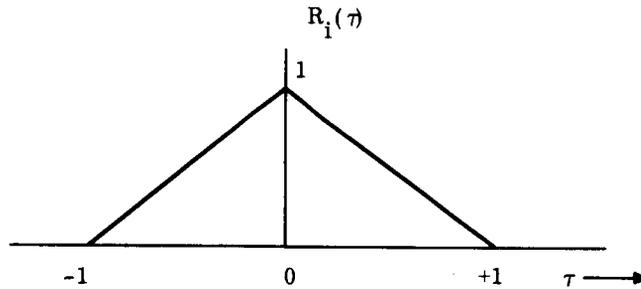
The auto-correlation function of the output signal, $R_o(\tau)$, is given by

$$R_o(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) \frac{\sin^2\left(\frac{\omega}{2}\right)}{\left(\frac{\omega}{2}\right)^2} e^{j\tau\omega} d\omega \quad . \qquad \text{(e-90)}$$

Or alternately

$$R_o(\tau) = R_i(\tau) \circledast h(\tau) \quad , \qquad \text{(e-91)}$$

where $\circledast$ denotes convolution and $h(\tau)$ is the inverse Fourier Transform of $H(\omega)$. That is:

$$h(\tau) = \frac{1}{2\pi} \int_{-2\pi}^{2\pi} e^{j\tau\omega} d\omega = \frac{\sin 2\pi\,\tau}{\pi\,\tau} \quad . \qquad \text{(e-92)}$$

Using Equation (e-91)

$$R_o(\tau) = \int_{-\infty}^{\infty} R_i(\tau - \lambda) h(\lambda) d\lambda \qquad \text{(e-93)}$$

or

$$R_o(\tau) = \int_{\tau}^{\tau+1} \frac{\sin 2\pi\lambda}{\pi\lambda} (\tau - \lambda + 1) d\lambda + \int_{\tau-1}^{\tau} \frac{\sin 2\pi\lambda}{\pi\lambda} (-\tau + \lambda + 1) d\lambda \quad . \qquad \text{(e-94)}$$

Or

$$R_o(\tau) = \frac{\cos 2\pi\lambda}{2\pi^2}\bigg|_{\tau}^{\tau+1} - \frac{\cos 2\pi\lambda}{2\pi^2}\bigg|_{\tau-1}^{\tau}$$

$$+ (1 - \tau) \int_{\tau-1}^{\tau} \frac{\sin 2\pi\lambda}{\pi\lambda} \, d\lambda + (1 + \tau) \int_{\tau}^{\tau+1} \frac{\sin 2\pi\lambda}{\pi\lambda} \, d\lambda \quad .$$

(e-95)

Now let

$$2\pi\lambda = y \quad , \tag{e-96}$$

so that

$$d\lambda = \frac{dy}{2\pi} \quad , \tag{e-97}$$

then

$$R_o(\tau) = \frac{2(1 - \tau)}{2\pi} \int_{2\pi(\tau-1)}^{2\pi\tau} \frac{\sin y}{y} \, dy + \frac{2(1 + \tau)}{2\pi} \int_{2\pi\tau}^{2\pi(\tau+1)} \frac{\sin y}{y} \, dy \quad . \tag{e-98}$$

Thus

$$R_o(\tau) = \frac{(1 - \tau)}{\pi} \left\{ S_i(2\pi\tau) - S_i\left[2\pi(\tau - 1)\right] \right\}$$

$$+ \frac{(1 + \tau)}{\pi} \left\{ S_i\left[2\pi(\tau + 1)\right] - S_i(2\pi\tau) \right\} \quad . \tag{e-99}$$

Where $S_i(x)$ is the sine-integral of X defined by

$$S_i(x) = \int_o^x \frac{\sin y}{y} \, dy \quad . \tag{e-100}$$

Now the derivative of the output auto-correlation function, $R_o'(\tau)$, is

$$R_o'(\tau) = \frac{d R_o(\tau)}{d\tau} = \frac{(1 - \tau)}{\pi} \left[\frac{\sin 2\pi\tau}{2\pi\tau} - \frac{\sin 2\pi(\tau - 1)}{2\pi(\tau - 1)}\right] + \frac{1 + \tau}{\pi} \left[\frac{\sin 2\pi(\tau + 1)}{2\pi(\tau + 1)}\right.$$

$$\left. - \frac{\sin 2\pi\tau}{2\pi\tau}\right] + \frac{1}{\pi}\left\{S_i\left[2\pi(\tau+1)\right] - 2S_i(2\pi\tau) + S_i\left[2\pi(\tau - 1)\right]\right\} \quad . \tag{e-101}$$

LIST

```
100 PRINT "TAU","R(TAU)","SI INDEX"
110 LET N7=2
120 GO SUB 910
130 LET P=3.14159265
140 LET H1=0
150 LET H2=0
160 LET H3=0
170 LET K=1
180 FOR T=2*K*1E-9 TO K*1E-8 STEP K*1E-9
190 LET S1=0
200 LET S2=0
210 LET S3=0
220 LET R1=0
230 LET H1=0
240 LET H2=0
250 LET H3=0
260 LET R2=0
270 LET R3=0
280 LET A=T-1
290 LET E1=SIN(2*P*A)
300 LET E2=SIN(2*P*(A+1))
310 LET E3=SIN(2*P*(A+2))
320 FOR N=0 TO 100
330 LET R2=2*N+1
340 IF N=0 THEN 360
350 GO TO 380
360 LET R3=1
370 GO TO 400
380 LET R3=R3*(2*N)
390 LET R3=R3*(2*N+1)
400 LET R4=R2*R3
410 LET C=A
420 IF H1=1 THEN 490
430 GOSUB 870
440 LET S1=S1+D1
450 IF ABS(D1)<(1E-4)*ABS(S1) THEN 470
460 GO TO 490
470 LET H1=1
480 LET A1=N
490 LET C=A+1
500 IF H2=1 THEN 570
510 GOSUB 870
520 LET S2=S2+D1
530 IF ABS(D1)<1E-20 THEN 550
540 GO TO 570
550 LET H2=1
560 LET A2=N
570 LET C=A+2
580 IF H3=1 THEN670
```

This function, when evaluated for $\tau = 0$, gives

$$R_o'(\tau)\Big|_{\tau=0} = R_o'(o) = \frac{1}{\pi}\left\{\int_o^{2\pi} \frac{\sin y}{y} \, dy + \int_o^{-2\pi} \frac{\sin y}{y} \, dy\right\} \qquad \text{(e-102)}$$

or

$$\boxed{R_o'(o) = 0} \qquad \cdot \qquad \text{(e-103)}$$

Tables of $S_i(x)$ are available in the literature. Using the series expansion for $S_i(x)$ the following time share computer programs were run to calculate both $R_o(\tau)$ and $R_o'(\tau)$.

```
590  GOSUB 870
600  LET  S3=S3+D1
610  IF  ABS(D1)<1E-15 THEN  650
620  IF  N<10 THEN  750
630  IF  ABS(D1)<(1E-4)*ABS(S3) THEN  650
640  GO  TO  750
650  LET  H3=1
660  LET  A3=N
670  IF  H1=1 THEN  690
680  GO  TO  750
690  IF  H2=1 THEN  710
700  GO  TO  750
710  IF  H3=1 THEN  730
720  GO  TO  750
730  LET  N=100
740  LET  N=100
750  NEXT  N
770  LET  J1=((1-T)/P)*(S2-S1)
780  LET  J2=((1+T)/P)*(S3-S2)
790  LET  R0=J1+J2
800  PRINT  T,R0,A1;A2;A3
810  GOSUB 910
820  NEXT  T
830  LET  K=10*K
840  IF  K=1E9 THEN  860
850  GO  TO  180
860  GO  TO  1010
870  LET  B=2*P*C
880  LET  R1=((-1)†N)*(B†(2*N+1))
890  LET  D1=R1/R4
900  RETURN
910  LET  N7=N7+1
920  IF  N7=51 THEN  940
930  RETURN
940  LET  N7=2
950  FOR  N8=0 TO  13
960  PRINT
970  NEXT  N8
980  PRINT  "TAU","R(TAU)","SI  INDEX"
990  PRINT
1000 RETURN
1010 END
```

```
100 PRINT "TAU","R'(TAU)","SI INDEX"
110 LET N7=2
120 GO SUB 910
130 LET P=3.14159265
140 LET H1=0
150 LET H2=0
160 LET H3=0
170 LET K=1
180 FOR T=2*K*1E-9 TO K*1E-8 STEP K*1E-9
190 LET S1=0
200 LET S2=0
210 LET S3=0
220 LET R1=0
230 LET H1=0
240 LET H2=0
250 LET H3=0
260 LET R2=0
270 LET R3=0
280 LET A=T-1
290 LET E1=SIN(2*P*A)
300 LET E2=SIN(2*P*(A+1))
310 LET E3=SIN(2*P*(A+2))
320 FOR N=0 TO 100
330 LET R2=2*N+1
340 IF N=0 THEN 360
350 GO TO 380
360 LET R3=1
370 GO TO 400
380 LET R3=R3*(2*N)
390 LET R3=R3*(2*N+1)
400 LET R4=R2*R3
410 LET C=A
420 IF H1=1 THEN 490
430 GOSUB 870
440 LET S1=S1+D1
450 IF ABS(D1)<(1E-4)*ABS(S1) THEN 470
460 GO TO 490
470 LET H1=1
480 LET A1=N
490 LET C=A+1
500 IF H2=1 THEN 570
510 GOSUB 870
520 LET S2=S2+D1
530 IF ABS(D1)<1E-20 THEN 550
540 GO TO 570
550 LET H2=1
560 LET A2=N
570 LET C=A+2
580 IF H3=1 THEN670
```

```
590 GOSUB 870
600 LET S3=S3+D1
610 IF ABS(D1)<1E-15 THEN 650
620 IF N<10 THEN 750
630 IF ABS(D1)<(1E-4)*ABS(S3) THEN 650
640 GO TO 750
650 LET H3=1
660 LET A3=N
670 IF H1=1 THEN 690
680 GO TO 750
690 IF H2=1 THEN 710
700 GO TO 750
710 IF H3=1 THEN 730
720 GO TO 750
730 LET N=100
740 LET N=100
750 NEXT N
760 LET J0=(1/P)*(S1+S3-(2*S2))
770 LET J1=((1-T)/P)*(E2/T-E1/(T-1))
780 LET J2=((1+T)/P)*(E3/(T+1)-E2/T)
790 LET R0=J0+J1+J2
800 PRINT T,R0,A1;A2;A3
810 GOSUB 910
820 NEXT T
830 LET K=10*K
840 IF K=1E9 THEN 860
850 GO TO 180
860 GO TO 1010
870 LET B=2*P*C
880 LET R1=((-1)↑N)*(B↑(2*N+1))
890 LET D1=R1/R4
900 RETURN
910 LET N7=N7+1
920 IF N7=51 THEN 940
930 RETURN
940 LET N7=2
950 FOR N8=0 TO 13
960 PRINT
970 NEXT N8
980 PRINT "TAU","R'(TAU)","SI INDEX"
990 PRINT
1000 RETURN
1010 END
```

R(T)   16:00   CH2 TH 05/11/67

TAU	R[TAU]	SI	INDEX	
2.00000 E-9	.902826	10	1	10
3.00000 E-9	.902826	10	1	10
4.00000 E-9	.902826	10	1	10
5.00000 E-9	.902826	10	1	10
6.00000 E-9	.902826	10	1	10
7.00000 E-9	.902826	10	1	10
8.00000 E-9	.902826	10	1	10
9.00000 E-9	.902826	10	1	10
1.00000 E-8	.902826	10	1	10
2.00000 E-8	.902826	10	1	10
3.00000 E-8	.902826	10	1	10
4.00000 E-8	.902826	10	1	10
5.00000 E-8	.902826	10	1	10
6.00000 E-8	.902826	10	1	10
7.00000 E-8	.902826	10	1	10
8.00000 E-8	.902826	10	1	10
9.00000 E-8	.902826	10	1	10
1.00000 E-7	.902826	10	2	10
2.00000 E-7	.902826	10	2	10
3.00000 E-7	.902826	10	2	10
4.00000 E-7	.902826	10	2	10
5.00000 E-7	.902826	10	2	10
6.00000 E-7	.902826	10	2	10
7.00000 E-7	.902826	10	2	10
8.00000 E-7	.902826	10	2	10
9.00000 E-7	.902826	10	2	10
.000001	.902826	10	2	10
.000002	.902826	10	2	10
.000003	.902826	10	2	10
.000004	.902826	10	2	10
.000005	.902826	10	2	10
.000006	.902826	10	2	10
.000007	.902826	10	2	10
.000008	.902826	10	2	10
.000009	.902826	10	2	10
.00001	.902826	10	2	10
.00002	.902826	10	2	10
.00003	.902826	10	2	10
.00004	.902826	10	2	10
.00005	.902826	10	2	10
.00006	.902826	10	3	10
.00007	.902826	10	3	10
.00008	.902826	10	3	10
.00009	.902826	10	3	10
.0001	.902826	10	3	10
.0002	.902826	10	3	10
.0003	.902825	10	3	10
.0004	.902825	10	3	10

TAU	R[TAU]	SI	INDEX	
.0005	.902825	10	3	10
.0006	.902825	10	3	10
.0007	.902825	10	3	10
.0008	.902824	10	3	10
.0009	.902824	10	3	10
.001	.902824	10	4	10
.002	.902818	10	4	10
.003	.902808	10	4	10
.004	.902794	10	4	10
.005	.902776	10	4	10
.006	.902754	10	5	10
.007	.902728	10	5	10
.008	.902698	10	5	10
.009	.902664	10	5	10
.01	.902626	10	5	10
.02	.902026	10	6	10
.03	.901028	10	6	10
.04	.899631	10	7	10
.05	.897835	10	7	11
.06	.895647	10	8	11
.07	.893068	10	8	11
.08	.890096	9	8	11
.09	.886742	9	9	11
.1	.883006	9	9	11
.2	.82574	8	11	12
.3	.73719	8	13	13
.4	.626603	7	14	13
.5	.504898	6	16	14
.6	.383063	5	17	15
.7	.270769	4	18	16
.8	.175162	4	19	17
.9	.100304	3	20	17
1.	4.71249 E-2	1	21	18

TIME: 19 SECS.

RUN
WAIT.

TAU		R'[TAU]		SI	INDEX	
2.00000	E-9	-1.67638	E-8	10	1	10
3.00000	E-9	-1.30385	E-8	10	1	10
4.00000	E-9	-1.86265	E-8	10	1	10
5.00000	E-9	-2.60770	E-8	10	1	10
6.00000	E-9	-3.91155	E-8	10	1	10
7.00000	E-9	-5.96046	E-8	10	1	10
8.00000	E-9	-3.72529	E-8	10	1	10
9.00000	E-9	-4.84288	E-8	10	1	10
1.00000	E-8	-4.84288	E-8	10	1	10
2.00000	E-8	-9.12696	E-8	10	1	10
3.00000	E-8	-1.28523	E-7	10	1	10
4.00000	E-8	-1.75089	E-7	10	1	10
5.00000	E-8	-2.06754	E-7	10	1	10
6.00000	E-8	-2.51457	E-7	10	1	10
7.00000	E-8	-2.92435	E-7	10	1	10
8.00000	E-8	-3.29688	E-7	10	1	10
9.00000	E-8	-3.65078	E-7	10	1	10
1.00000	E-7	-4.00469	E-7	10	2	10
2.00000	E-7	-8.06525	E-7	10	2	10
3.00000	E-7	-1.22935	E-6	10	2	10
4.00000	E-7	-1.60933	E-6	10	2	10
5.00000	E-7	-2.00793	E-6	10	2	10
6.00000	E-7	-2.41026	E-6	10	2	10
7.00000	E-7	-2.81446	E-6	10	2	10
8.00000	E-7	-3.21679	E-6	10	2	10
9.00000	E-7	-3.62098	E-6	10	2	10
.000001		-4.00469	E-6	10	2	10
.000002		-8.00192	E-6	10	2	10
.000003		-1.20215	E-5	10	2	10
.000004		-1.60094	E-5	10	2	10
.000005		-2.00085	E-5	10	2	10
.000006		-2.39909	E-5	10	2	10
.000007		-2.79862	E-5	10	2	10
.000008		-3.20170	E-5	10	2	10
.000009		-3.60142	E-5	10	2	10
.00001		-4.00096	E-5	10	2	10
.00002		-8.00099	E-5	10	2	10
.00003		-1.19997	E-4	10	2	10
.00004		-1.60001	E-4	10	2	10
.00005		-2.00009	E-4	10	2	10
.00006		-2.40002	E-4	10	3	10
.00007		-2.79998	E-4	10	3	10
.00008		-3.20015	E-4	10	3	10
.00009		-3.60001	E-4	10	3	10
.0001		-4.00009	E-4	10	3	10
.0002		-8.00004	E-4	10	3	10
.0003		-1.19999	E-3	10	3	10
.0004		-1.59999	E-3	10	3	10

TAU	R'[TAU]		SI	INDEX	
.0005	-1.99998	E-3	10	3	10
.0006	-2.39997	E-3	10	3	10
.0007	-2.79997	E-3	10	3	10
.0008	-3.19997	E-3	10	3	10
.0009	-3.59997	E-3	10	3	10
.001	-3.99995	E-3	10	4	10
.002	-7.99984	E-3	10	4	10
.003	-1.19996	E-2	10	4	10
.004	-1.59993	E-2	10	4	10
.005	-1.99988	E-2	10	4	10
.006	-2.39981	E-2	10	5	10
.007	-2.79971	E-2	10	5	10
.008	-3.19958	E-2	10	5	10
.009	-3.59941	E-2	10	5	10
.01	-.039992		10	5	10
.02	-7.99394	E-2	10	6	10
.03	-.119798		10	6	10
.04	-.159522		10	7	10
.05	-.199073		10	7	11
.06	-.238398		10	8	11
.07	-.277459		10	8	11
.08	-.316208		9	8	11
.09	-.35461		9	9	11
.1	-.392618		9	9	11
.2	-.742347		8	11	12
.3	-1.01295		8	13	13
.4	-1.18027		7	14	13
.5	-1.23538		6	16	14
.6	-1.18495		5	17	15
.7	-1.04897		4	18	16
.8	-.856327		4	19	17
.9	-.639256		3	20	17
1.	-.427849		1	21	18

TIME: 21 SECS.

APPENDIX F

SQUARE LAW DETECTION AND POST DETECTION FILTERING OF ASYMMETRICAL SIGNALS

Let $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$ be the input and output random signals of a square law device.
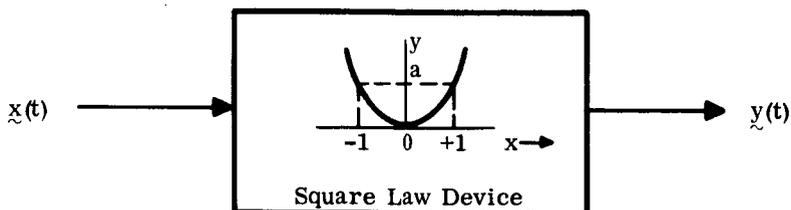
Figure F-1. Square Law Device

Where

$$\underset{\sim}{y}(t) = a \; \underset{\sim}{x}^2(t) \tag{f-1}$$

and "a" is a constant.

If $f_{\underset{\sim}{x}}(x)$ denotes the probability density function (p.d.f.) of the random input, then the p.d.f. of the random output, $f_{\underset{\sim}{y}}(y)$, is given by:

$$f_{\underset{\sim}{y}}(y) = \frac{1}{2\sqrt{ay}} \left[f_{\underset{\sim}{x}}\left(\sqrt{\frac{y}{a}}\right) + f_{\underset{\sim}{x}}\left(-\sqrt{\frac{y}{a}}\right) \right] U(y) . \tag{f-2}$$

Where

$$U(y) = \begin{cases} 1 & \text{for } y \geq 0 \\ \\ 0 & \text{elsewhere} \end{cases} \tag{f-3}$$

Also, if $F_{\underset{\sim}{x}}(x)$ denotes the probability that the random variable $\underset{\sim}{x}$ is less than or equal to the value x, then the probability that the random output y is less than or equal to the value y, $F_{\underset{\sim}{y}}(y)$, is given by:

$$F_{\underset{\sim}{y}}(y) = \left[2 \, F_{\underset{\sim}{x}}\left(\sqrt{\frac{y}{a}}\right) - 1 \right] U(y). \tag{f-4}$$

Now let $\underset{\sim}{x}(t)$ be a stochastic process, stationary in the strict sense (stationary of order two would suffice), with $f_{\underset{\sim}{x}}(x)$ and $f_{\underset{\sim}{x_1},\underset{\sim}{x_2}}(x_1, x_2, \tau)$, the first and second order probability density functions of x(t). Where it is understood that

$$\left.\begin{array}{l} \underset{\sim}{x_1} = \underset{\sim}{x}_1(t_1) \\ \underset{\sim}{x_2} = \underset{\sim}{x}_2(t_2) \end{array}\right\} \tag{f-5}$$

and

$$t_2 - t_1 = \tau .$$

The expected value of the output, $E\left[\underset{\sim}{y}(t)\right]$ is then given by:

$$E\left[\underset{\sim}{y}(t)\right] = \int_{-\infty}^{\infty} a\, x^2\, f_{\underset{\sim}{x}}(x)\, dx \qquad \text{(f-6)}$$

and the autocorrelation of the output, $R_{\underset{\sim}{yy}}(\tau)$, by:

$$R_{\underset{\sim}{yy}}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a^4\, x_1^2\, x_2^2\, f_{\underset{\sim}{x_1},\underset{\sim}{x_2}}(x_1, x_2, \tau)\, dx_1,\, dx_2 . \qquad \text{(f-7)}$$

For the case where $a = 1$,

$$\underset{\sim}{y} = \underset{\sim}{x}^2 \qquad \text{(f-8)}$$

and where $\underset{\sim}{x}$ is a stationary normal process, with zero mean and autocorrelation function $R_{\underset{\sim}{x}}(\tau)$, the p.d.f. becomes:

$$f_{\underset{\sim}{y}}(y) = \frac{1}{\sqrt{2\pi R_{\underset{\sim}{x}}(0)y}}\, e^{\frac{-y}{2R_{\underset{\sim}{x}}(0)}}\, U(y) \qquad \text{(f-9)}$$

and

$$R_{\underset{\sim}{yy}}(\tau) = R_{\underset{\sim}{y}}(\tau) = R_x^2(0) + 2R_x^2(\tau) \qquad \text{(f-10)}$$

and

$$E\left\{\underset{\sim}{y}^2(t)\right\} = R_{\underset{\sim}{y}}(0) = 3R_x^2(0) . \qquad \text{(f-11)}$$

The variance of $\underset{\sim}{y}(t)$, σ_y^2, is then

$$\sigma_y^2 = E\left\{\underset{\sim}{y}^2(t)\right\} - E^2\left\{y(t)\right\} = 2R_x^2(0) \qquad \text{(f-12)}$$

Now, if $S_{\underset{\sim}{x}}(\omega)$ is the power spectrum of $\underset{\sim}{x}(t)$, then the power spectrum of the output, $S_{\underset{\sim}{y}}(\omega)$, is given by:

$$S_{\underset{\sim}{y}}(\omega) = 2\pi R_x^2(0)\, \delta(\omega) + \frac{1}{\pi}\, S_{\underset{\sim}{x}}(\omega) \circledast S_{\underset{\sim}{x}}(\omega) \qquad \text{(f-13)}$$

where

$\circledast$ denotes convolution.

For the case where the process $\underset{\sim}{x}(t)$ is <u>ideal bandpass</u>, that is:

$$S_{\underset{\sim}{x}}(\omega) = \begin{cases} A & \text{for } \omega_2 \leq |\omega| \leq \omega_3 \\ 0 & \text{elsewhere} \end{cases} \qquad \text{(f-14)}$$

F-2
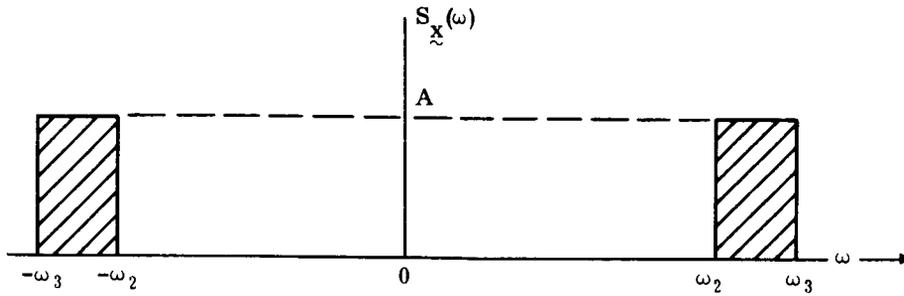
Figure F-2. Function $S_{\underset{\sim}{X}}(\omega)$

Then

$$R_{\underset{\sim}{X}}(0) = \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\underset{\sim}{X}}(\omega) \, e^{+j\tau\omega} \, d\omega \right\} \Bigg|_{\tau = 0}$$

(f-15)

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\underset{\sim}{X}}(\omega) \, d\omega = \frac{A}{\pi}(\omega_2 - \omega_1)$$

and

$$I(\omega) = S_{\underset{\sim}{X}}(\omega) * S_{\underset{\sim}{X}}(\omega) = \int_{-\infty}^{\infty} S_{\underset{\sim}{X}}(\lambda) \, S_{\underset{\sim}{X}}(\omega - \lambda) \, d\lambda$$

(f-16)

where

$S_{\underset{\sim}{X}}(\lambda - \omega)$ is shown below in Figure F-3.
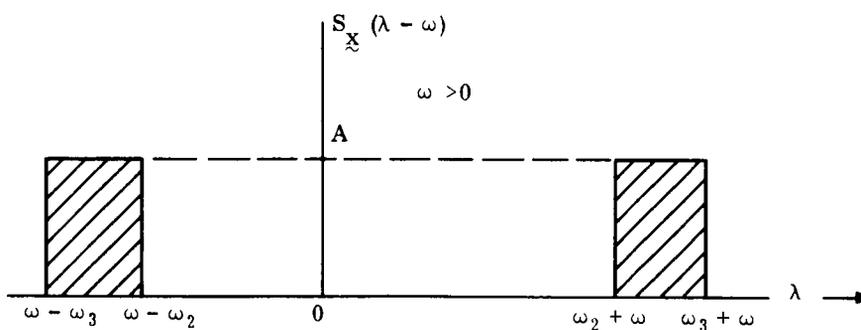
Figure F-3. Function $S_{\underset{\sim}{X}}(\lambda - \omega)$

Now $S_{\underset{\sim}{x}}(\omega - \lambda)$ is just $S_{\underset{\sim}{x}}(\lambda - \omega)$ reversed about zero, as shown in Figure F-4.
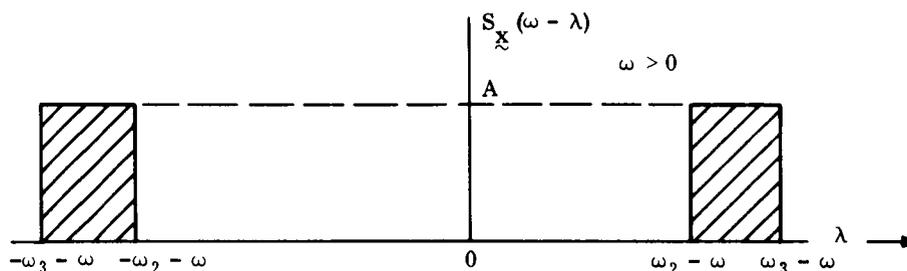
Figure F-4. Function $S_{\underset{\sim}{x}}(\omega - \lambda)$

Thus for

$$\underline{\omega \geq 0} \qquad\qquad\qquad\qquad\text{(f-17)}$$

for

$$0 \leq \omega \leq \omega_3 - \omega_2$$

$$I(\omega) = 2 \int_{\omega_2}^{\omega_3 - \omega} A^2\, d\lambda = 2\,A^2 (\omega_3 - \omega - \omega_2) \qquad\qquad\text{(f-18)}$$

for

$$\omega_3 - \omega_2 \leq \omega \leq 2\omega_2$$

$$I(\omega) = 0 \qquad\qquad\qquad\qquad\text{(f-19)}$$

for

$$2\omega_2 \leq \omega \leq \omega_2 + \omega_3$$

$$I(\omega) = \int_{\omega_2 - \omega}^{-\omega_2} A^2\, d\lambda = A^2(-\omega_2 - \omega_2 + \omega) = A^2(\omega - 2\omega_2) \qquad\text{(f-20)}$$

for

$$\omega_2 + \omega_3 \leq \omega \leq 2\omega_3$$

$$I(\omega) = \int_{-\omega_3}^{\omega_3 - \omega} A^2\, d\lambda = A^2(\omega_3 - \omega + \omega_3) = A^2(2\omega_3 - \omega) \qquad\text{(f-21)}$$

and, for

$$\omega > 2\omega_3$$

$$I(\omega) = 0 \qquad\qquad\qquad\qquad\text{(f-22)}$$

F-4

Because of symmetry

$$I(-\omega) = I(\omega),$$ (f-23)

so that summarizing

$$I(\omega) = S_{\underset{\sim}{x}}(\omega) * S_{\underset{\sim}{x}}(\omega) = \begin{cases} 2A^2\left[(\omega_3 - \omega_2) - |\omega|\right] & \text{for } 0 \le |\omega| \le \omega_3 - \omega_2 \\ 0 & \text{for } \omega_3 - \omega_2 \le |\omega| \le 2\omega_2 \\ A^2(|\omega| - 2\omega_2) & \text{for } 2\omega_2 \le |\omega| \le \omega_2 + \omega_3 \\ A^2(2\omega_3 - |\omega|) & \text{for } \omega_2 + \omega_3 \le |\omega| \le 2\omega_3 \\ 0 & \text{for } |\omega| > 2\omega_3 \end{cases}$$ (f-24)

$I(\omega)$ is shown below in Figure F-5.
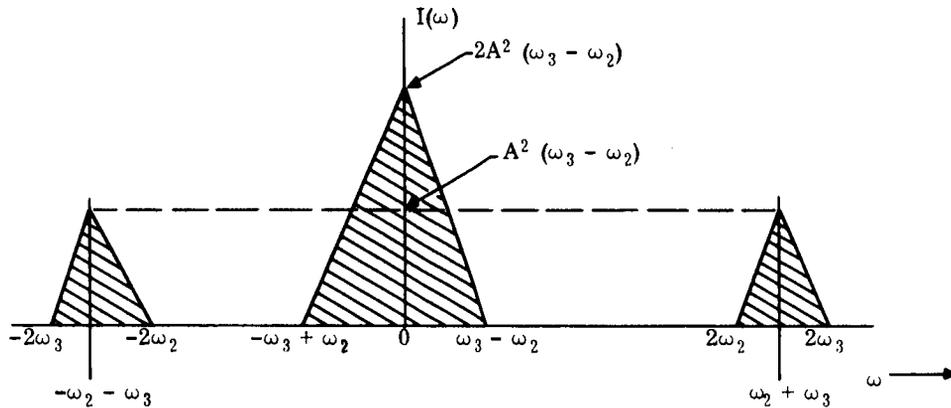

Figure F-5. Function $I(\omega)$

Substituting Equations (f-15) and (f-24) into Equation (f-13) yields

$$S_{\underset{\sim}{y}}(\omega) = \frac{2A^2}{\pi}(\omega_3 - \omega_2)^2 \delta(\omega) + \frac{1}{\pi} I(\omega)$$ (f-25)

Therefore, a sketch of the output power spectrum, $S_{\underset{\sim}{y}}(\omega)$, is as shown below:
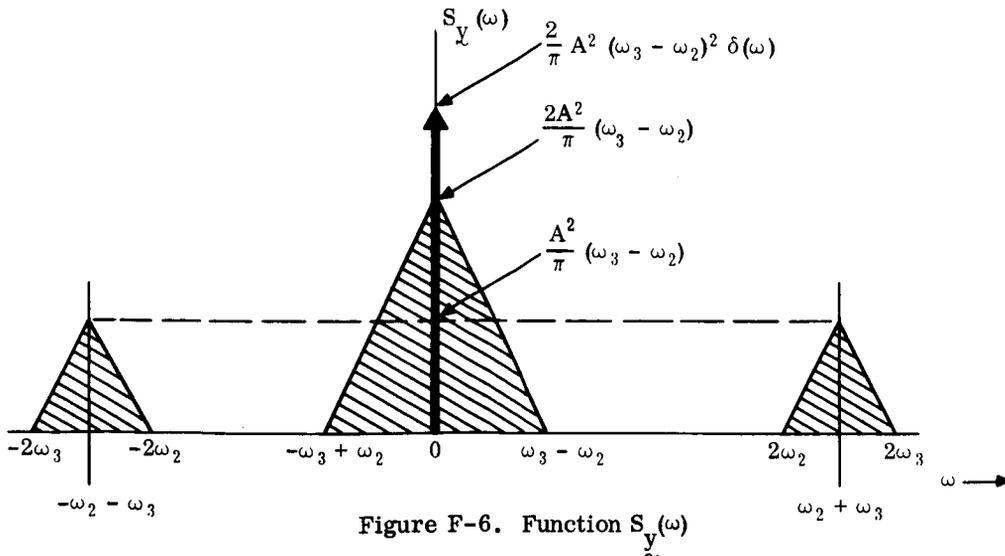

Figure F-6. Function $S_{\underset{\sim}{y}}(\omega)$

For the special case, where the input random process $x(t)$ is comprised of three random components corresponding to a carrier, $x_c(t)$, a tone, $x_1(t)$, and noise $n(t)$, then

$$x(t) = x_c(t) + x_1(t) + n(t) \tag{f-26}$$

where

$$x_c(t) = A_c \cos(\omega_c t + \theta_c) \tag{f-27}$$

$$x_1(t) = A_1 \cos(\omega_1 t + \theta_1). \tag{f-28}$$

θ_c and θ_1 are independent and random variables over the interval $\left[0 \text{ to } 2\pi\right]$ and where $n(t)$ is an ideal band-pass Gaussian random process, with zero mean. From Equation (f-8), the output of the square law device is given by:

$$y(t) = x_c^2 + x_1^2(t) + n^2(t) + 2x_c(t)x_1(t) + 2x_c(t)n(t) + 2x_1(t)n(t). \tag{f-29}$$

The autocorrelation of the output, $R_y(\tau)$, is thus given by:

$$
\begin{aligned}
R_y(\tau) &= E\left[y(t)\,y(t+\tau)\right] \\
&= E\left\{\left[x_c^2(t) + x_1^2(t) + n^2(t) + 2x_c(t)x_1(t)\right.\right. \\
&\quad \left. + 2x_c(t)n(t) + 2x_1(t)n(t)\right]\left[x_c^2(t+\tau)\right. \\
&\quad + x_1^2(t+\tau) + n^2(t+\tau) + 2x_c(t+\tau)x_1(t+\tau) \\
&\quad \left.\left. + 2x_c(t+\tau)n(t+\tau) + 2x_1(t+\tau)n(t+\tau)\right]\right\}.
\end{aligned}
\tag{f-30}
$$

Because θ_c, θ_1, and $n(t)$ have been assumed independent and because $n(t)$ is stationary with zero mean, expected values involving $n(t)$ or $n(t+\tau)$ to the first power drop out. Also, since the expected value of the product of two independent random variables is equal to the product of the expected value of each, then one may write:

$$
\begin{aligned}
R_y(\tau) &= R_{x_c^2,x_c^2}(\tau) + R_{x_1^2,x_1^2}(\tau) + R_{n^2,n^2}(\tau) + 4R_{x_c x_1,\ x_c x_1}(\tau) \\
&\quad + 4R_{x_c n, x_c n}(\tau) + 4R_{x_1 n, x_1 n}(\tau) + 2R_{x_c^2,x_1^2}(\tau) \\
&\quad + 2R_{x_c^2,n^2}(\tau) + 2R_{x_1^2,n^2}(\tau).
\end{aligned}
\tag{f-31}
$$

The power spectrum of the output of the square law device, $S_y(\omega)$, is then determined from

$$S_y(\omega) = \int_{-\infty}^{\infty} R_y(\tau)\, e^{-j\omega\tau}\, d\tau \tag{f-32}$$

or from Equation (f-31)

$$S_{\underset{\sim}{y}}(\omega) = S_{\underset{\sim}{x}_c^2, \underset{\sim}{x}_c^2}(\omega) + S_{x_1^2, x_1^2}(\omega) + S_{\underset{\sim}{n}^2, \underset{\sim}{n}^2}(\omega) + 4S_{\underset{\sim}{x}_c x_1, \underset{\sim}{x}_c x_1}(\omega)$$

$$+ 4S_{x_c n, \underset{\sim}{x}_c \underset{\sim}{n}}(\omega) + 4S_{x_1 \underset{\sim}{n}, x_1 \underset{\sim}{n}}(\omega) + 2S_{\underset{\sim}{x}_c^2, x_1^2}(\omega)$$

$$+ 2S_{\underset{\sim}{x}_c^2, \underset{\sim}{n}^2}(\omega) + 2S_{x_1^2, \underset{\sim}{n}^2}(\omega) \; . \qquad \qquad \left.\rule{0pt}{90pt}\right\} \text{(f-33)}$$

The following Fourier Transform pairs are easily determined.

$$R_{x_c, x_c^2}(\tau) = \frac{A_c^4}{4}\left(1 + \frac{\cos 2\omega_c \tau}{2}\right)$$

$$S_{x_c^2, x_c^2}(\omega) = \frac{\pi A_c^4}{2}\left[\delta(\omega) + \frac{\delta(\omega - 2\omega_c) + \delta(\omega + 2\omega_c)}{4}\right] \; . \qquad \left.\rule{0pt}{70pt}\right\} \text{(f-34)}$$

And

$$R_{x_1^2, x_1^2}(\tau) = \frac{A_1^4}{4}\left(1 + \frac{\cos 2\omega_1 \tau}{2}\right)$$

$$\qquad \qquad \text{(f-35)}$$

$$S_{x_1^2, x_1^2}(\omega) = \frac{\pi A_1^4}{2}\left[\delta(\omega) + \frac{\delta(\omega - 2\omega_1) + \delta(\omega + 2\omega_1)}{4}\right]$$

$R_{n^2, n^2}(\tau)$ and $S_{n^2, n^2}(\omega)$ can be obtained by substitution from Equation (f-10), and (f-24), respectively. Now

$$R_{x_c x_1, x_c x_1}(\tau) = E\left\{A_c \cos(\omega_c t + \underset{\sim}{\theta}_c) \cdot A_1 \cos(\omega_1 t + \underset{\sim}{\theta}_1) \cdot A_c \cos\left[\omega_c(t + \tau) + \underset{\sim}{\theta}_c\right]\right.$$

$$\left. \cdot A_1 \cos\left[\omega_1(t + \tau) + \underset{\sim}{\theta}_1\right]\right\} \qquad \text{(f-36)}$$

where again $\underset{\sim}{\theta}_1$ and $\underset{\sim}{\theta}_c$ are independent stationary random processes, uniformly distributed between $\left[0 \text{ and } 2\pi\right]$. So that

$$R_{x_c x_1, x_c x_1}(\tau) = \frac{A_1^2 A_c^2}{8}\left[\cos(\omega_c + \omega_1)\tau + \cos(\omega_c - \omega_1)\tau\right] \qquad \text{(f-37)}$$

with

$$S_{x_c x_1, x_c x_1}(\omega) = \frac{\pi A_1^2 A_c^2}{8}\left[\delta(\omega - \omega_c - \omega_1) + \delta(\omega + \omega_c + \omega_1)\right.$$

$$\left. + \delta(\omega - \omega_c + \omega_1) + \delta(\omega + \omega_c - \omega_1)\right] \qquad \left.\rule{0pt}{60pt}\right\} \text{(f-38)}$$

Also,

$$R_{x_c n, x_c n}(\tau) = E\left\{x_c(t)\, x_c(t + \tau)\, n(t)\, n(t + \tau)\right\}$$

$$= E\left\{x_c(t)\, x_c(t + \tau)\right\} \cdot E\left\{n(t)\, n(t + \tau)\right\} \qquad \left.\right\} \text{(f-39)}$$

$$= R_{x_c, x_c}(\tau) \cdot R_{n, n}(\tau)$$

which leads to

$$S_{x_c n, x_c n}(\omega) = \frac{1}{2\pi}\, S_{x_c, x_c}(\omega) * S_{n, n}(\omega) . \qquad \text{(f-40)}$$

However, since

$$S_{x_c, x_c}(\omega) = \frac{\pi A_c^{\,2}}{2}\left[\delta(\omega - \omega_c) + \delta(\omega + \omega_c)\right] \qquad \text{(f-41)}$$

one obtains

$$S_{x_c n, x_c n}(\omega) = \frac{A_c^{\,2}}{4}\left[S_n(\omega - \omega_c) + S_n(\omega + \omega_c)\right] \qquad \text{(f-42)}$$

and similarly

$$S_{x_1 n, x_1 n}(\omega) = \frac{A_1^{\,2}}{4}\left[S_n(\omega - \omega_1) + S_n(\omega + \omega_1)\right] . \qquad \text{(f-43)}$$

Now

$$R_{x_c^{\,2}, x_1^{\,2}}(\tau) = E\left[x_c^{\,2}(t)\right]\, E\left[x_1^{\,2}(t + \tau)\right]$$

$$= E\left[A_c^{\,2}\cos(\omega_c t + \theta_c)\right] \cdot E\left[A_1^{\,2}\cos\left\{\omega_1(t + \tau) + \theta_1\right\}\right] \qquad \left.\right\} \text{(f-44)}$$

$$= \frac{A_1^{\,2} A_c^{\,2}}{4}$$

and

$$S_{x_c^{\,2}, x_1^{\,2}}(\omega) = \frac{\pi A_1^{\,2} A_c^{\,2}}{2}\, \delta(\omega) . \qquad \text{(f-45)}$$

Similarly

$$R_{x_c^{\,2}, n^2}(\tau) = E\left[x_c^{\,2}(t)\right] \cdot E\left[n^2(t + \tau)\right] = \frac{A_c^{\,2}}{2}\, \sigma_n^{\,2} \qquad \text{(f-46)}$$

where $\sigma_n^{\,2}$ is the noise power in.

F-8

This gives

$$S_{x_c^2, n^2}(\omega) = \pi A_c^2 \sigma_n^2 \delta(\omega) \tag{f-47}$$

and

$$S_{x_1^2, n^2}(\omega) = \pi A_1^2 \sigma_n^2 \delta(\omega) \tag{f-48}$$

Substituting these results into Equation (f-33), yields:

$$
\begin{aligned}
S_y(\omega) = & \left[\frac{\pi}{2}(A_c^4 + A_1^4) + \pi A_1^2 A_c^2 + 2\pi\sigma_n^2(A_c^2 + A_1^2) \right] \delta(\omega) \\
& + \frac{\pi}{8} A_c^4 \left[\delta(\omega - 2\omega_c) + \delta(\omega + 2\omega_c) \right] + \frac{\pi}{8} A_1^4 \left[\delta(\omega - 2\omega_1) + \delta(\omega + 2\omega_1) \right] \\
& + S_{n^2, n^2}(\omega) + \frac{\pi A_1^2 A_c^2}{2} \left[\delta(\omega - \omega_c - \omega_1) + \delta(\omega + \omega_c + \omega_1) + \delta(\omega - \omega_c + \omega_1) \right. \\
& \left. + \delta(\omega + \omega_c - \omega_1) \right] \\
& + A_c^2 \left[S_n(\omega - \omega_c) + S_n(\omega + \omega_c) \right] + A_1^2 \left[S_n(\omega - \omega_1) + S_n(\omega + \omega_1) \right] .
\end{aligned}
\tag{f-49}
$$

Only those terms containing noise alone remain to be explicitly defined. These spectra are shown in Figure F-7. Now the noise power spectrum $S_n(\omega)$ is ideal bandpass of bandwidth $2\pi B$ radians per second, centered at $\omega = \pm(\omega_1 + \omega_c)/2$. Thus,
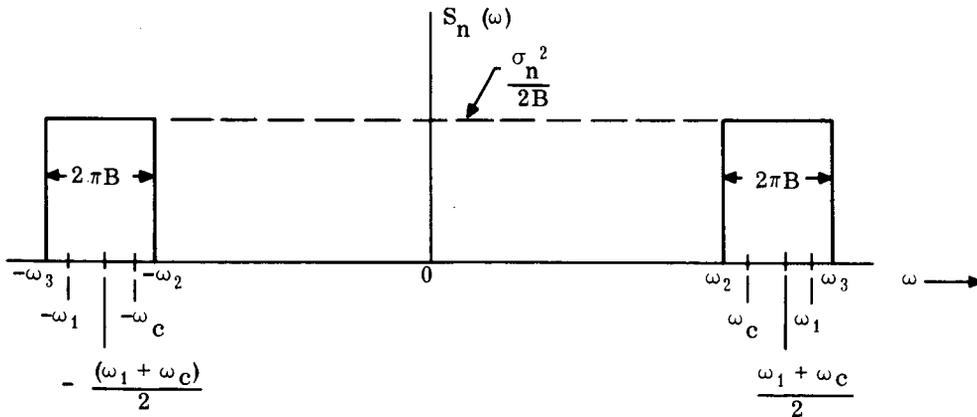

Figure F-7. Function $S_n(\omega)$

where

$$\omega_2 = \frac{\omega_1 + \omega_c}{2} - \pi B$$

$$\omega_3 = \frac{\omega_1 + \omega_c}{2} + \pi B \tag{f-50}$$

$$A = \frac{\sigma_n^2}{2B}$$

Substituting these relations into Equation (f-25) the following sketches result:
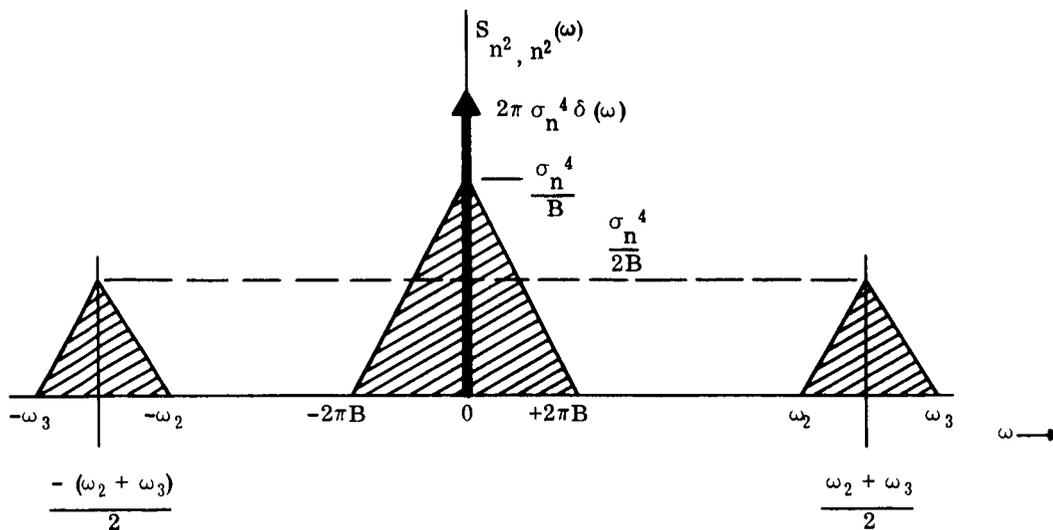

Figure F-8. Function $S_{n^2, n^2}(\omega)$
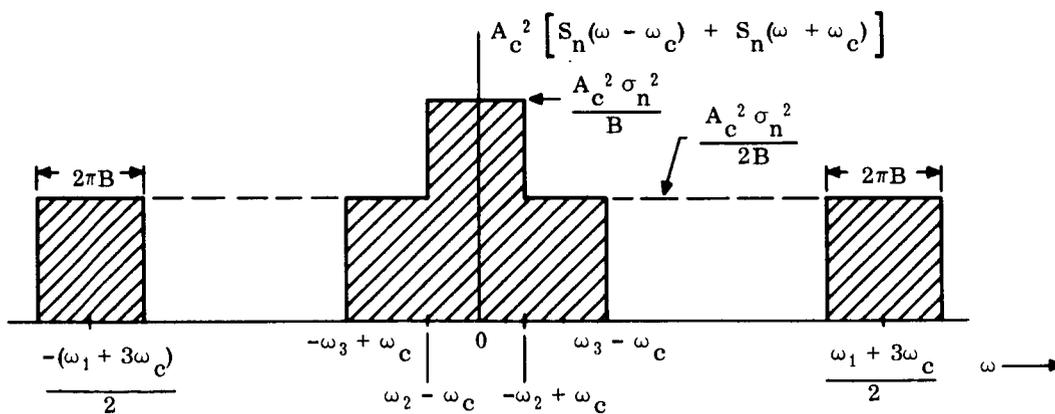
Figure F-9. Function $A_c^2 \left[S_n(\omega - \omega_c) + S_n(\omega + \omega_c) \right]$
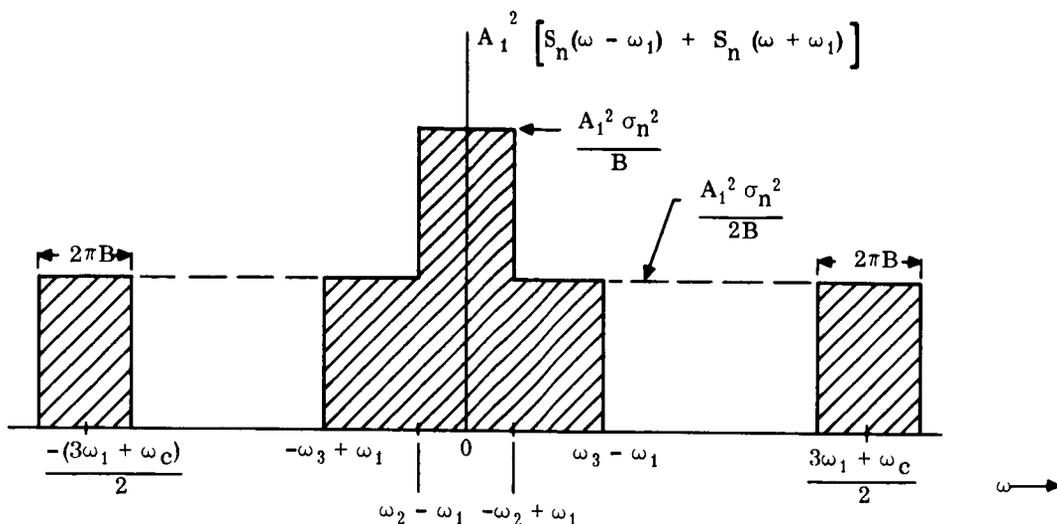
Figure F-10. Function $A_1^2 \left[S_n(\omega - \omega_1) + S_n(\omega + \omega_1) \right]$

Now let us suppose that the output of the squaring device is passed through an ideal low-pass zonal filter, so that only those terms in the vicinity of zero frequency remain. Denoting the output of the zonal filter by $\underset{\sim}{Z}(t)$, the power spectrum of $\underset{\sim}{Z}(t)$, $S_{\underset{\sim}{Z}}(\omega)$ will be given by:

$$
\begin{aligned}
S_{\underset{\sim}{Z}}(\omega) = \pi & \left[\frac{(A_c^4 + A_1^4)}{2} + A_1^2 A_c^2 + 2\sigma_n^2 (A_c^2 + A_1^2) \right] \delta(\omega) \\
& + \frac{\pi A_1^2 A_c^2}{2} \left[\delta(\omega - \omega_c + \omega_1) + \delta(\omega + \omega_c - \omega_1) \right] \\
& + \hat{S}_{n^2, n^2}(\omega) + A_c^2 \left[\hat{S}_n(\omega - \omega_c) + \hat{S}_n(\omega + \omega_c) \right] \\
& + A_1^2 \left[\hat{S}_n(\omega - \omega_1) + \hat{S}_n(\omega + \omega_1) \right]
\end{aligned}
$$

(f-51)

where the $\wedge$ denotes the selection of that part lying within the low-pass filter.

Let the low-pass filter have cutoff frequency ω_r, where

$$\omega_r = 2\pi B .$$

(f-52)

Thus, the power spectrum out of the low-pass filter, $S_{\underset{\sim}{Z}}(\omega)$, is as shown below:
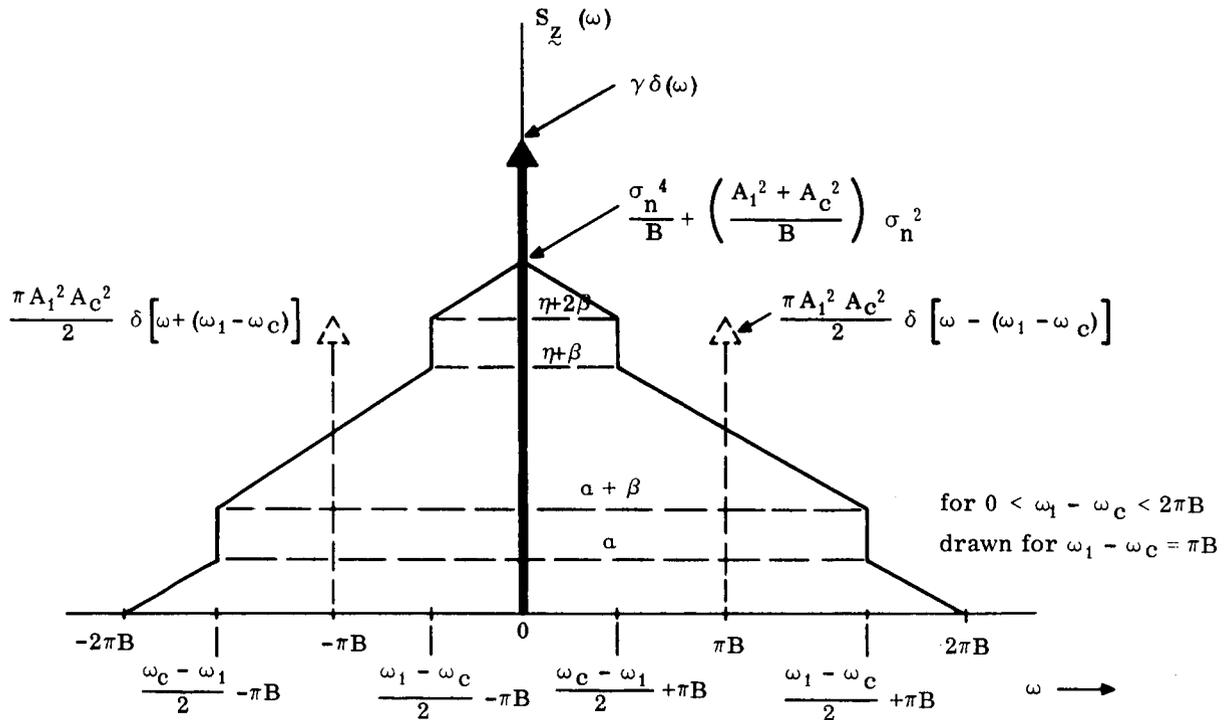
Figure F-11. Power Spectrum at Low-pass Filter Output

Where

$$\gamma = \pi \left[\frac{A_1^4 + A_c^4}{2} + A_1^2 A_c^2 + 2\sigma_n^2 (A_c^2 + A_1^2) + 2\sigma_n^4 \right] \qquad \text{(f-53)}$$

$$\alpha = \frac{\sigma_n^4}{2B} \left(1 - \frac{\omega_1 - \omega_c}{2\pi B} \right) \qquad \text{(f-54)}$$

$$\beta = \left(\frac{A_1^2 + A_c^2}{2B} \right) \sigma_n^2 \qquad \text{(f-55)}$$

and, where

$$\eta = \frac{\sigma_n^4}{2B} \left(1 + \frac{\omega_1 - \omega_c}{2\pi B} \right) . \qquad \text{(f-56)}$$

The impulses located at $\pm(\omega_1 - \omega_c)$ correspond to the carrier-"tone" cross term and shall be referred to simply as "tone." Thus, the tone has frequency

$$\omega = \omega_1 - \omega_c. \tag{f-57}$$

It is now of interest to determine the tone-to-noise ratio at the output of an ideal bandpass filter of width $2\pi B_1$, centered at $\omega = \omega_1 - \omega_c$, into which the signal $\underset{\sim}{z}(t)$ is fed. Let the output of this filter be denoted by $\underset{\sim}{V_1}(t)$ as shown below:
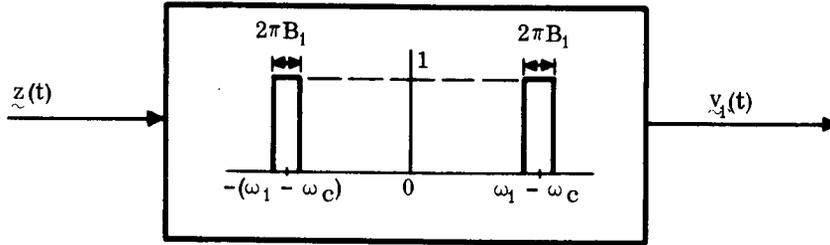

Figure F-12. Ideal Bandpass

The power spectrum of $\underset{\sim}{V_1}(t)$, $S_{\underset{\sim}{V_1}}(\omega)$, is then given by:

$$S_{\underset{\sim}{V_1}}(\omega) = \begin{cases} S_{\underset{\sim}{z}}(\omega) & \text{for } \omega_1 - \omega_c - \pi B_1 \leq |\omega| \leq \omega_1 - \omega_c + \pi B_1 \\ 0 & \text{elsewhere} \end{cases}. \tag{f-58}$$

Thus the noise power out of this filter, P_{N_1}, is given by

$$P_{N_1} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\underset{\sim}{V_1}}(\omega)\, d\omega \tag{f-59}$$

or by symmetry

$$P_{N_1} = 2\left[\frac{1}{2\pi} \int_{\omega_1 - \omega_c - \pi B_1}^{\omega_1 - \omega_c + \pi B_1} S_{\underset{\sim}{z}}(\omega)\, d\omega \right]. \tag{f-60}$$

From Figure F-11, one finds

$$S_{\underset{\sim}{z}}(\omega) = \begin{cases} \dfrac{\sigma_n^4}{B}\left(1 - \dfrac{\omega}{2\pi B}\right) & \text{for } \dfrac{\omega_1 - \omega_c}{2} + \pi B < \omega \leq 2\pi B & \text{Region I} \\[4mm] \dfrac{\sigma_n^4}{B}\left(1 - \dfrac{\omega}{2\pi B}\right) + \left(\dfrac{A_1^2 + A_c^2}{2B}\right)\sigma_n^2 & \text{for } \dfrac{\omega_c - \omega_1}{2} + \pi B < \omega < \dfrac{\omega_1 - \omega_c}{2} + \pi B & \text{Region II} \\[4mm] \dfrac{\sigma_n^4}{B}\left(1 - \dfrac{\omega}{2\pi B}\right) + \dfrac{\left(A_1^2 + A_c^2\right)}{B}\sigma_n^2 & \text{for } 0 \leq \omega < \dfrac{\omega_c - \omega_1}{2} + \pi B & \text{Region III} \\[4mm] 0 & \text{for } \omega \geq 2\pi B & \text{Region IV} \end{cases} \tag{f-61}$$

F-13

For the ideal bandpass falling completely within any of these regions, Equation (f-60) yields:

$$P_{N_1} = \begin{cases} \dfrac{\sigma_n^4 B_1}{\pi B}\left[2\pi - \dfrac{(\omega_1 - \omega_c)}{B}\right] & \text{Region I} \\[3ex] \dfrac{\sigma_n^4 B_1}{\pi B}\left[2\pi - \dfrac{(\omega_1 - \omega_c)}{B}\right] + \dfrac{B_1}{B}\sigma_n^2 (A_1^2 + A_c^2) & \text{Region II} \\[3ex] \dfrac{\sigma_n^4 B_1}{\pi B}\left[2\pi - \dfrac{(\omega_1 - \omega_c)}{B}\right] + \dfrac{2B_1}{B}\sigma_n^2 (A_1^2 + A_c^2) & \text{Region III} \\[3ex] 0 & \text{Region IV} \end{cases} \qquad \text{(f-62)}$$

From Equation (f-51), the tone output power P_{x_c,x_1}, is given by:

$$P_{x_c,x_1} = \frac{A_1^2 A_c^2}{2}. \qquad \text{(f-63)}$$

Thus, the tone-to-noise power output, Γ_0, is determined to be

$$\Gamma_0 = \frac{A_1^2 A_c^2}{2 P_{N_1}} \qquad \text{(f-64)}$$

with P_{N_1} given by Equation (f-62).

The carrier power input, P_{X_c}, is given by

$$P_{X_c} = \frac{A_c^2}{2}. \qquad \text{(f-65)}$$

The noise power input, P_n, is found from Figure F-7 to be

$$P_n = \sigma_n^2. \qquad \text{(f-66)}$$

So that the input, carrier-to-noise power ratio, Γ_i, is given by

$$\Gamma_i = \frac{A_c^2}{2\sigma_n^2}. \qquad \text{(f-67)}$$

Now if one denotes by, Γ_1, the ratio of these two ratios, then

$$\Gamma_1 = \frac{\Gamma_0}{\Gamma_i} = \left[\frac{A_1^2 A_c^2}{2 P_{N_1}} \middle/ \frac{A_c^2}{2\sigma_n^2}\right] = \frac{A_1^2 \sigma_n^2}{P_{N_1}}. \qquad \text{(f-68)}$$

Where again, P_{N_1} is given by Equation (f-62). It is observed that the σ_n^2 will be divided by σ_n^4, so that with

$$A_1^2 = \mu_1 A_c^2 \tag{f-69}$$

one may write

$$\Gamma_1 = \frac{\Gamma_o}{\Gamma_i} = \frac{\mu_1 \left[\dfrac{A_c^2}{2\sigma_n^2}\right]}{\left[\dfrac{P_{N_1}}{2\sigma_n^4}\right]} \quad . \tag{f-70}$$

So that, the final result has been expressed in terms of the input carrier-to-noise power ratio and the input carrier-to-"tone" power ratio.

Figure F-13 shows the power spectrum $S_Z(\omega)$ for the special case where:

Special Case

$$\left.\begin{array}{l} B \;\; = \;\; 25 \text{ kHz} \\[2mm] f_1 \;\; - \;\; f_c \;\; = \;\; 18.4 \text{ kHz} \quad . \end{array}\right\} \tag{f-71}$$

For this case, it is observed that the tone is in Region II. Now, the bandpass filter will fall completely within this region provided

$$B_1 \;\; < \;\; 7.0 \text{ kHz} \tag{f-72}$$

So that, with this restriction, one obtains from Equation (f-62),

$$P_{N_1} = \frac{\sigma_n^4 B_1}{\pi B}\left[2\pi - \frac{\omega_1 - \omega_c}{B}\right] + \frac{B_1}{B}\sigma_n^2(A_1^2 + A_c^2) \quad . \tag{f-73}$$

Thus

$$\Gamma_1 = \frac{\mu_1\left(\dfrac{A_c^2}{2\sigma_n^2}\right)}{\dfrac{B_1}{B}\left[1 - \dfrac{f_1 - f_c}{B} + (1 + \mu_1)\left(\dfrac{A_c^2}{2\sigma_n^2}\right)\right]} \tag{f-74}$$

for Region II.

Letting

$$B_1 = 90 \text{ Hz} \tag{f-75}$$
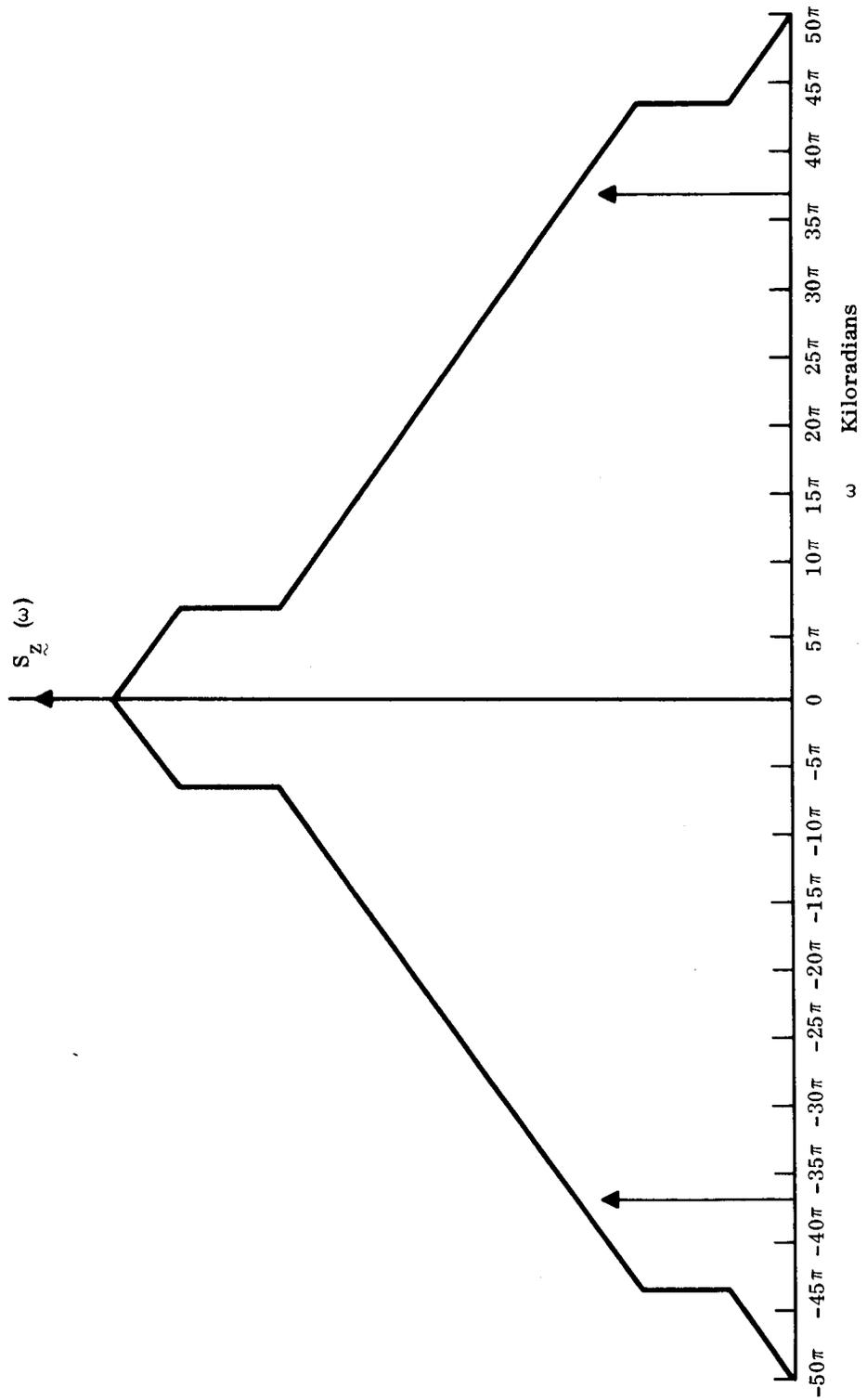
Figure F-13. Power Spectrum—Special Case

and substituting from Equation (f-71) gives

$$\Gamma_1 = \frac{\mu_1 \left(\dfrac{A_c^2}{2\sigma_n^2}\right) \times 10^3}{0.95 + 3.6 (1 + \mu_1) \left(\dfrac{A_c^2}{2\sigma_n^2}\right)} \; .$$

(f-76)

Now for

$$10 \log_{10} \left[\frac{A_c^2}{2\sigma_n^2}\right] = 1 \text{ db},$$

(f-77)

$$\frac{A_c^2}{2\sigma_n^2} \approx 1.26$$

(f-78)

and, with

$$\mu_1 = \frac{1}{6}$$

(f-79)

$$\Gamma_1 \approx 33.64.$$

(f-80)

APPENDIX G

MATCHED FILTERS AND AMBIGUITY FUNCTIONS

A filter matched to a signal $s(t)$, has an impulse response $h(t)$ given by

$$h(t) = s*(\Delta - t). \tag{g-1}$$

Where * means complex conjugate and Δ is an arbitrary time delay to make $h(t)$ realizable, i.e., $h(t)$ must be zero for negative time because; $h(t)$ represents the response of the filter at time t to an impulse applied at time zero. A response for negative time means that the filter responds before the impulse is applied.

For real signals, $s*(\Delta - t) = s(\Delta - t)$. Therefore $h(t) = s(\Delta - t)$, showing that the matched filter impulse response is $s(t)$, reversed in time and delayed by an arbitrary amount Δ.

The response $x(t)$ of a filter, with an impulse response $h(t)$, to a signal $s_1(t)$ is given formally by the following convolution integral

$$x(t) = \int_{-\infty}^{\infty} s_1(\tau) \, h(t - \tau) \, d\tau . \tag{g-2}$$

Let the signal $s_1(t)$ be the signal $s(t)$, offset in frequency by ν Hz. Further, let $h(t)$ be the impulse response of a filter matched to $s(t)$. Ignoring the time delay factor Δ,

$$h(t) = s*(-t) .$$

Now

$$s(t) = \mu(t)e^{j2\pi f_o t}$$

therefore,

$$s_1(t) = \mu(t)e^{j2\pi \nu t} e^{j2\pi f_o t}$$

and

$$h(t) = s*(-t) = \mu*(-t) \, e^{j2\pi f_o t} .$$

If above expressions for $s_1(t)$ and $h(t)$ are substituted into the convolution integral, the following is obtained:

$$x(t) = e^{j2\pi f_o t} \int_{-\infty}^{\infty} \mu(\tau) \, \mu*(\tau - t) \, e^{j2\pi \nu \tau} \, d\tau . \tag{g-3}$$

The integral (g-3) gives the complex envelope of the filter response. Suppressing the exponential outside the integral, because we are interested in the response envelope, and arbitrarily interchanging t and τ, gives

$$x_1(\tau, \nu) = \int_{-\infty}^{\infty} \mu(t) \mu^*(t - \tau) e^{j2\pi\nu t} dt \qquad (g-4)$$

which is nearly the quantity defined as the ambiguity function of the signal $s_p(t)$. The ambiguity function is usually taken as the normalized function $x(\tau, \nu) = (1/2E)x_1(\tau, \nu)$, where E is the energy contained in the physical signal $s_p(t)$.

$x(\tau, \nu)$ is important in the study of the waveforms for ranging, because it displays not only the filter output as a function of time τ, but also shows how the filter output behaves as a function of frequency offset ν.

Because the matched filter output is usually detected by a square law or similar device, performance of the matched filter is conveniently given in terms of the square magnitude of the ambiguity function = $|X(\tau, \nu)|^2$. When this function is plotted above the τ, ν plane, a surface called the ambiguity surface results.

A consideration of some simple, single pulses illustrates these ideas. The physical signal of a carrier, amplitude modulated by a rectangular pulse, is

$$s_p(t) = A \cos \omega_0 t \qquad 0 < t < T_0$$
$$= 0 \qquad \text{otherwise} . \qquad (g-5)$$

The pulse energy is $E = 1/2 \, A^2 T_0$. The complex envelope in this case is real and is

$$\mu(t) = A \qquad 0 < t < T_0$$
$$= 0 \qquad \text{otherwise.} \qquad (g-6)$$

The ambiguity function

$$X(\tau, \nu) = \frac{1}{2E} \int_{-\infty}^{\infty} \mu(t) \mu^*(t - \tau) e^{j2\pi\nu\tau} dt \qquad (g-7)$$

becomes

$$X(\tau, \nu) = \frac{1}{A^2 T_0} \int_{\tau}^{T_1} (A)(A)e^{j2\pi\nu\tau} dt \qquad 0 \leq \tau \leq T_0$$

$$= \frac{1}{T_0} \left(\frac{1}{j2\pi\nu}\right) \left[e^{j2\pi\nu T_0} - e^{j2\pi\nu\tau}\right] .$$

Evaluating $X(\tau, \nu)$ for $-T < \tau\ 0$ shows the ambiguity surface. The squared magnitude of $X(\tau,\nu)$ is

$$|X(\tau, \nu)|^2 = X^*(\tau, \nu)$$

$$= \left(\frac{|\tau| - T_o}{T_o}\right)^2 \left(\frac{\sin \pi\nu(\tau - T_o)}{\pi\nu(\tau - T_o)}\right)^2 \quad - T_o \leq \tau \leq T_o . \tag{g-8}$$

Figure G-1 is a sketch of this equation. The sketch illustrates much of what is known from experience. For example, a reasonable approximation of the idealized pulse and matched filter described above are provided by nearly rectangular radar pulses and the IF amplifier-filter used in their reception. In this case, the IF output pulse width is nearly twice that of the input, in agreement with Figure G-1 for $\nu = 0$. Also detuning the pulse from the center frequency of the IF amplifier, so as to place the pulse carrier at the edge of the passband, results in an IF output with appreciable amplitude confined to the regions near the leading and trailing edges of the pulse. This effect is shown in Figure G-1 by taking a cut through the surface parallel to the τ axis, through $\nu = 1/T_o$; the resulting curve peaks at $\tau = \pm T_o/2$.

$$s_p(t) = A \cos \omega_o t \quad \text{for } 0 < t < T_o$$

$$s_p(t) = 0 \quad \text{for } 0 > t > T_o$$
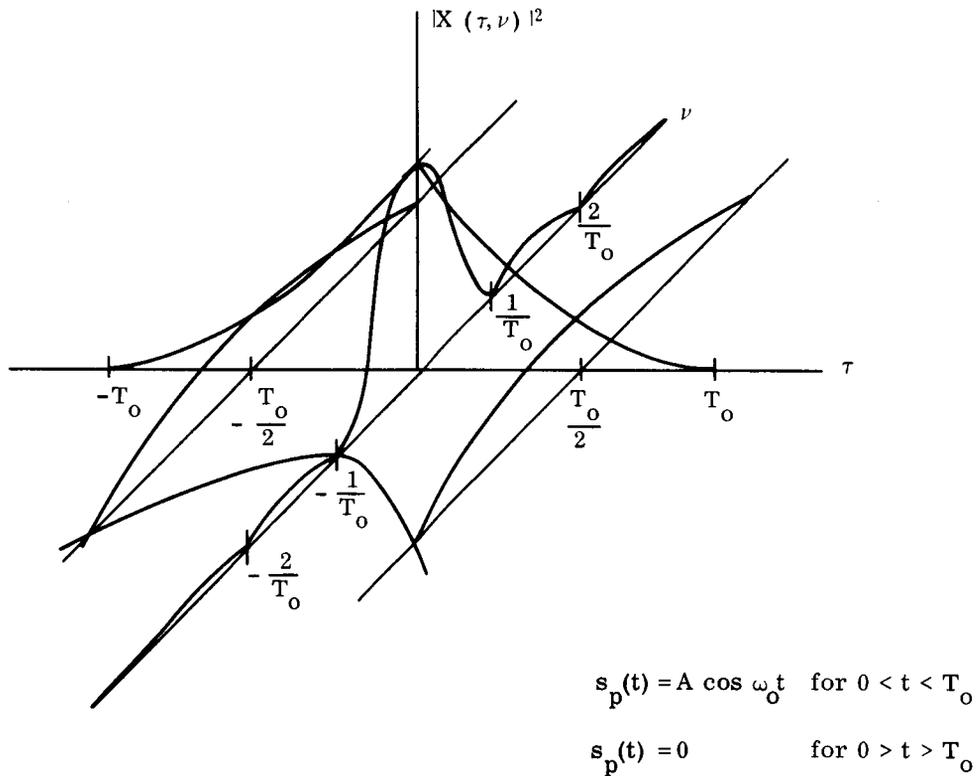
Figure G-1. Plot of Ambiguity Surface for Single Rectangular Pulse

To illustrate the effect of pulse shape on the ambiguity surface, consider a carrier, amplitude modulator with a Gaussian pulse:

$$s_p(t) = Ae^{-at^2} \cos \omega_0 t \qquad \text{(g-9)}$$

in this case

$$\mu(t) = Ae^{-at^2} \qquad \text{(g-10)}$$

and the energy of the pulse is

$$E = \frac{A^2}{2} \sqrt{\frac{\pi}{2a}} \ . \qquad \text{(g-11)}$$

The ambiguity function is, from (g-7) and (g-10)

$$X(\tau,\nu) = \frac{1}{2E} \int_{-\infty}^{\infty} \left(Ae^{-at^2}\right)\left(Ae^{-a(t-\tau)^2}\right) e^{j2\pi\nu t} dt \qquad \text{(g-12)}$$

$$= \frac{A^2}{2E} e^{-a\tau^2/2} \int_{-\infty}^{\infty} e^{-2a(t-\tau/2)^2} e^{j2\pi\nu t} dt \ .$$

The integral can be evaluated from the known Fourier transform of a Gaussian pulse, giving

$$X(\tau,\nu) = \frac{A^2}{2E} \sqrt{\frac{\pi}{2a}} e^{-a\tau^2/2 - \pi^2\nu^2/2a} e^{j\pi\nu\tau} \qquad \text{(g-13)}$$

which, using (g-11) to eliminate E, results in

$$|X(\tau,\nu)|^2 = e^{-(a\tau^2 + \pi^2\nu^2/a)}. \qquad \text{(g-14)}$$

Figure G-2 shows a sketch of the intersection of this surface with two principle planes. It drops off from 1 at $\nu = \tau = 0$ according to a Gaussian curve.

Another single pulse of considerable interest is that in which the frequency is swept linearly during the pulse. This type of pulse is often called a Chirp pulse. It is convenient to take the pulse as having a Gaussian envelope, as the previously considered pulse had. The complex waveform is:

$$s(t) = Ae^{-(a-jb)t^2} e^{j\omega_0 t} \qquad \text{(g-15)}$$

and the complex envelope is

$$\mu(t) = Ae^{-(a-jb)t^2} = Ae^{-at^2} e^{jbt^2} \ . \qquad \text{(g-16)}$$

It is the factor b in the exponent that generates the frequency sweep. The instantaneous radian frequency is given by $d\left[\phi(t)\right]/dt$, where, in this case, $\phi(t) = bt^2$. Therefore

$$\omega(t) = \frac{d\phi(t)}{dt} = 2bt \ \text{rad/sec}$$

and

$$f(t) = \frac{\omega(t)}{2\pi} = \frac{1}{\pi} \, bt \text{ cycles/sec}$$

which shows that, assuming $b > 0$, the instantaneous frequency increases linearly with time. The ambiguity function

$$X(\tau, \nu) = \frac{1}{2E} \int_{-\infty}^{\infty} \mu(t)\, \mu^*(t - \tau)\, e^{j2\pi\nu\tau}\, dt$$

can be manipulated into the form

$$X(\tau, \nu) = \frac{A^2}{2E} \, e^{-(a/2 + jb)\tau^2} \int_{-\infty}^{\infty} e^{-2a(t - \tau/2)^2} e^{j2\pi\left(\frac{b\tau}{\pi} + \nu\right)t}\, dt$$

where, with a change of frequency variable, the integral can be evaluated from known Fourier transforms. This gives

$$X(\tau, \nu) = e^{-a\tau^2/2 \, - \, (\pi^2/2a)(b\tau/\pi + \nu)^2} \, e^{-j(2b\tau^2 + \pi\tau\nu)} \, . \tag{g-17}$$

This represents the complex envelope at the matched filter output.


$$s_p(t) = Ae^{-at^2} \cos \omega_o t$$
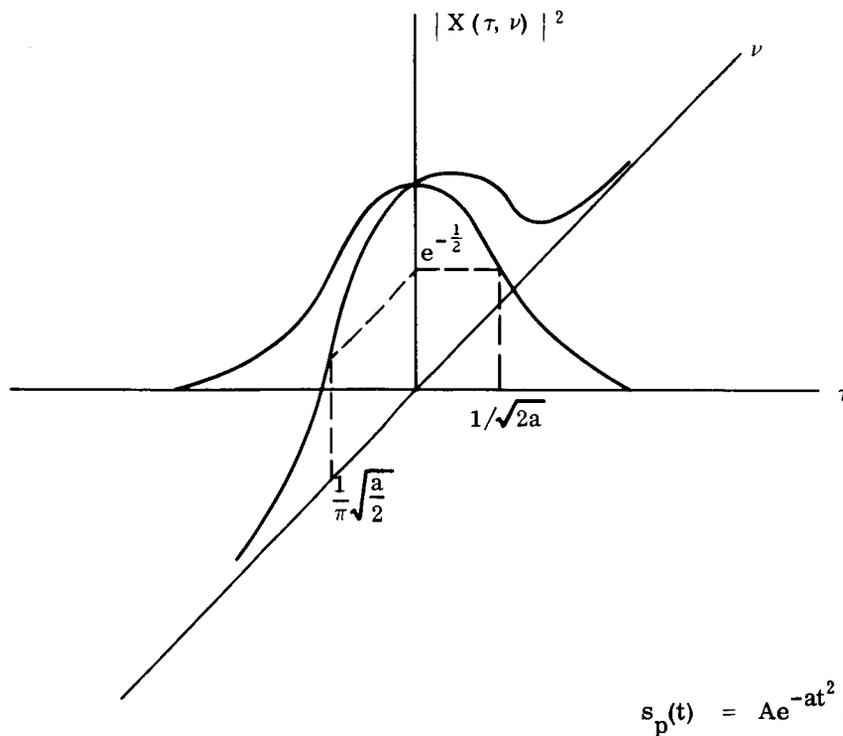
Figure G-2. Plot of Ambiguity Surface for Single Gaussian Pulse

The ambiguity surface is given by:

$$|X(\tau, \nu)|^2 = e^{-(1/a)\left[(a^2 + b^2)\tau^2 + (2\pi b\nu\tau) + (\pi^2\nu^2)\right]}. \tag{g-18}$$

The exponent in (g-18) can be recognized as the equation of an ellipse, whose axes are rotated with respect to the τ, ν coordinates.

Figure G-3 is a sketch of the ambiguity surface for this Chirp, or swept frequency pulse.
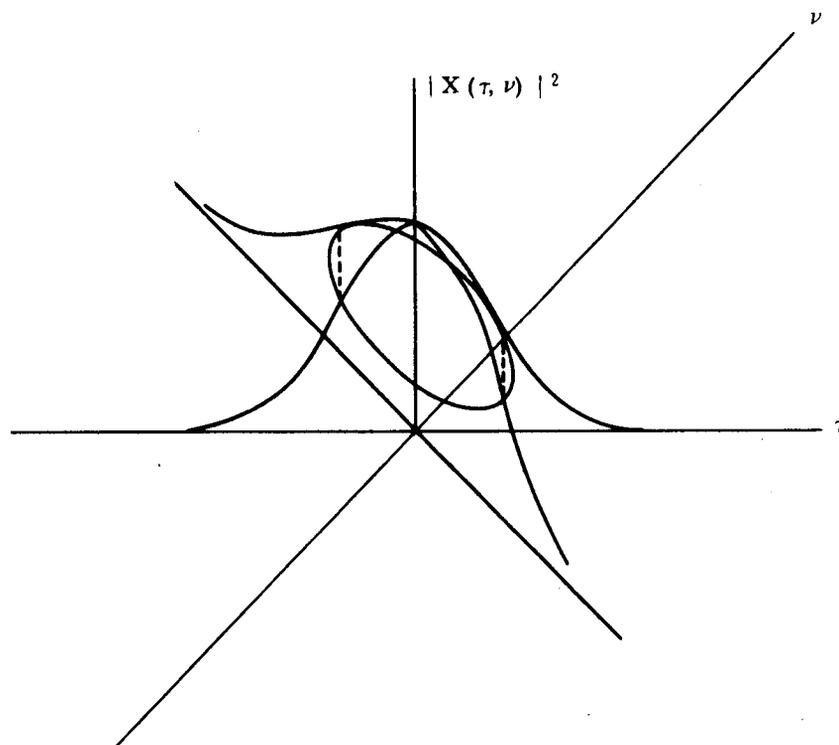

Figure G-3. Sketch of the Ambiguity Surface of a Swept Frequency Pulse

The sketch and Equation (g-18) show several significant features. The first term in the exponent, which contains b^2, shows that sweeping the frequency narrows the response in the τ direction for a fixed pulse duration, as controlled by a. This shows that the ranging accuracy is better than that of an unswept frequency pulse, as would be expected, because sweeping the frequency results in a larger β_0 in Equation (b-7) of Appendix B. The second factor, however, contains $\nu\tau$ term, which causes the rotation of the ellipse axis. Equation (b-6) shows that $\alpha > 0$ for the waveform and consequently the range accuracy is reduced. This is a result which also shows that frequency shifts and time shifts are coupled: a frequency shift causes a proportional time shift. The last factor shows that, at $\tau = 0$, the response drops off with frequency shift ν at

the same rate as an unswept-frequency pulse of the same time duration. The implication of the coupling between τ and ν is that unless we know one of the parameters exactly, there is an error introduced in the measurement of the other parameter. In the normal ranging situation, neither τ nor ν is known precisely and if range is to be measured, ν must be controlled so as to make the range uncertainty due to frequency shifts small.

This background in ambiguity functions and surfaces provides a basis for interpreting and applying some work by Rihaczek.[4,5]

Rihaczek considers envelope-recurrent pulse trains. Although it is not difficult to visualize pulse trains excluded from this class, many pulse trains of practical interest are included. Envelope-recurrent pulse trains are those in which each of the component pulses has the same real envelope and the same bandwidth. The spacing between pulses, the phase structure within each pulse, and the center frequency of each pulse may be varied to achieve desired ambiguity functions. Figure G-4 shows a sketch of the envelope-recurrent pulse train. The envelope shown is arbitrary. N is the number of pulses in the train. Pulse staggering is restricted to $|\Delta T_n| \leq T/2$ to preserve pulse train characteristics.
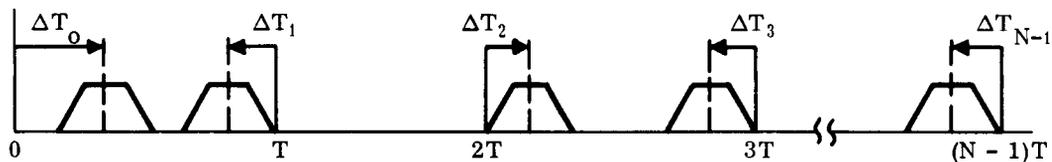

Figure G-4. Envelope-Recurrent Pulse Train with Staggered Interpulse
Spacing (Average Value of $\Delta T_n = 0$)

With these restrictions the following cases were considered.

A. UNIFORM PULSE TRAIN

For this case $\Delta T_n = 0$ and each of the component pulses has the same phase structure, bandwidth and carrier frequency.

The absolute value of the ambiguity function for such a pulse train is:

$$\left| X(\tau, \nu) \right| = \frac{1}{N} \sum_{m=-(N-1)}^{N-1} \left| X_p(\tau - mT, \nu) \right| \cdot \left| \sum_{n=0}^{N-1-m} e^{j2\pi\nu nT} \right| .$$
(g-19)

The ambiguity surface is

$$\left| X(\tau, \nu) \right|^2 = \frac{1}{N^2} \sum_{m=-(N-1)}^{N-1} \left| X_p(\tau - mT, \nu) \right|^2 \cdot \frac{\sin^2 \pi\nu(N - |m|)T}{\sin^2 \pi\nu T} \; . \qquad \text{(g-20)}$$

$X_p(\tau - mT, \nu)$ is the ambiguity function for the component pulses of the train delayed by the factor mT from $\tau = 0$. Since the pulses are subject only to the restrictions mentioned previously, the pulse shape, and therefore the ambiguity function, can take on a wide variety of forms. For example, the component pulses might be the rectangular, Gaussian or swept frequency pulses, discussed previously, or others.

The $\sin^2 \pi\nu(n - |m|)T/\sin^2\pi\nu T$ factor in Equation (g-20) samples the component pulse ambiguity surface in the frequency shift direction. Thus, Equation (g-20) shows that the ambiguity surface of the uniform pulse train is generated by repeating the ambiguity surface of the component pulses at intervals of T along the τ axis, scaling each surface in amplitude by $(1 - m/N)^2$, and sampling each surface in the frequency shift direction. The number of time intervals T along the τ axis from $\tau = 0$ can be interpreted as m.

Figure G-5a, the gross ambiguity surface of the uniform pulse train, illustrates the repetition of the component pulse surface along the τ axis.

Figure G-5b, a cut through the surface at $\tau = 0$, illustrates the fine structure in the ν direction, introduced by the sampling function.

Figure G-5c, shows the more detailed sketch of the pulse train ambiguity surface.

These figures illustrate two important features of pulse trains. First, signal energy can be reduced in large regions of the frequency shift axis, in which energy exists for the single pulse case. Secondly, this is accomplished without reducing the range resolving power of the central surface. That is, the width of the central surface in the τ direction is the same as for a single pulse. Unfortunately, the uniform pulse train ambiguity surface has subsidiary peaks along the τ axis which, although separated from the central surface, can cause false responses.

The reduction of energy in certain frequency bands has important implications for frequency division multiplexing. As shown in Figure G-5b, the width of the regions containing large amounts of energy is about 1/NT Hz at the half power points and 2/NT at the base. This means that a filter matched to a uniform pulse train, identical to that for which the diagrams are drawn except displaced in carrier frequency by ν Hz, $1/NT < |\nu| < 1/T - 1/NT$, will have only a small output given by the level between spikes, in response to the pulse train. Consequently two systems using identical, (except for a small difference in carrier frequency), uniform pulse trains can operate in substantially the same frequency band. This can be extended to more than two systems. The limit is ultimately reached when the responses between spikes of the several pulse trains is comparable to the energy in the spikes. Frequency shifts due to oscillator instabilities or doppler frequency shifts put a lower limit on the width in the ν direction of the usable central surface. The sensitivity of the filter response to frequency offset is proportional to N. Saturation of receivers and power amplifiers may be another limitation on the use this special form of frequency division.
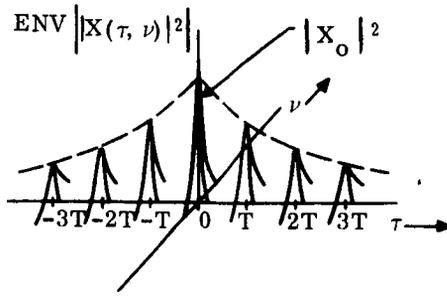
G-5a. Gross Structure of the Ambiguity Surface
of the Uniform Pulse Train
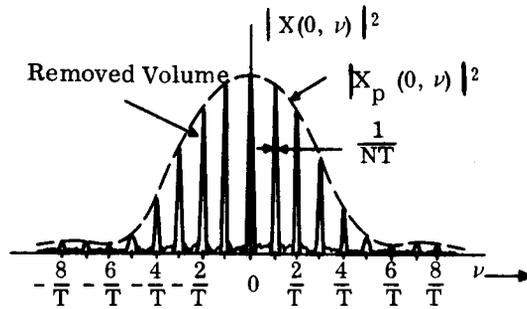
G-5b. Fine Structure of the Ambiguity Surface
along the Frequency Shift Axis
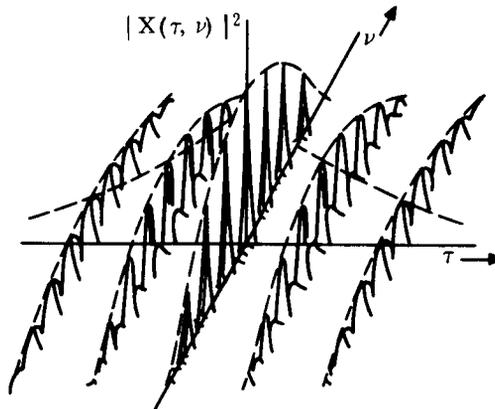
G-5c. Ambiguity Surface of the Uniform
Pulse Train

Figure G-5. Illustrations of the Ambiguity Surfaces of Uniform
Pulse Trains (from Reference[4])

B. STAGGERED REPETITION INTERVAL PULSE TRAINS

This pulse train consists of N identical pulses, but, in contrast to the uniform pulse train, the pulses are more or less randomly shifted from the positions they would occupy in the uniform train case. This shifting implies that, when simple pulses are used, the pulse train is restricted to a lower duty cycle than the uniform pulse train, so as to avoid pulse overlap. When the component pulses are more complex, avoiding overlap is not necessary if the system is linear, because the pulses can be resolved after compression. Amplitude limiters or other nonlinearities would cause signal suppression and distortion.

The expression for the ambiguity function for staggered repetition interval trains, given in reference [4], is complicated and will not be given here. Figure G-6 shows sketches of the ambiguity surface for this case. It is apparent that secondary responses in both the τ and ν directions have been suppressed by the pulse staggering. Because of the normalizing factor 2E in the definition of the ambiguity function, the value of $|X(\tau, \nu)|^2$ for the waveform is always unity at $\tau = \nu = 0$, even though, as in Figure G-6, the pulse train contains more energy than the single pulse.

It is possible to suppress the responses along the τ axis by $1/N^2$, compared to the central peak of unity normalized amplitude. Similar reductions are possible along the ν axis.
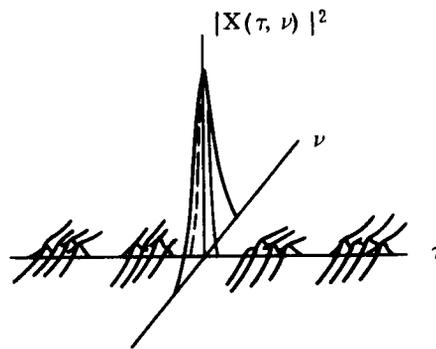
G-6a. Gross Ambiguity Surface Showing the Suppression of Responses Along the τ Axis
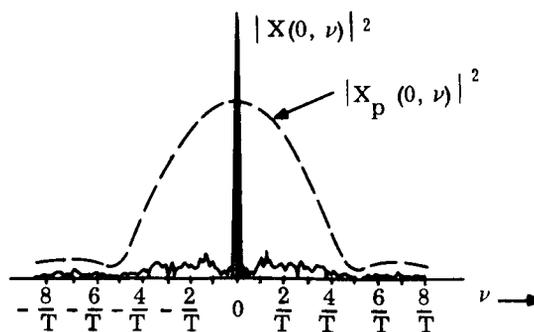
G-6b. Smearing of the Doppler Ambiguities by Staggering of the Repetition Interval (Note Decrease of Surface Width in ν Direction)

Figure G-6. Illustrations of Ambiguity Surfaces for Staggered Repetition Interval Pulse Trains (from Reference [4])

C. CODED-PULSE PULSE TRAINS

This type of pulse train again consists of N pulses. Each pulse has the same real envelope and the same bandwidth. However, the pulses have high dimensionality* and are chosen so as to have mutually small cross-correlation. This might, for example, be accomplished by phase coding within the pulse.

When pulses in such a pulse train are repeated periodically, the secondary peaks of the ambiguity surface in the τ direction are reduced, possibly to a level of $1/N^2$, compared to the central peak, as a result of the low cross-correlation between pulses. In the ν direction, however, the uniform repetition rate causes secondary peaks. These secondary peaks can be reduced by staggering the repetition interval.
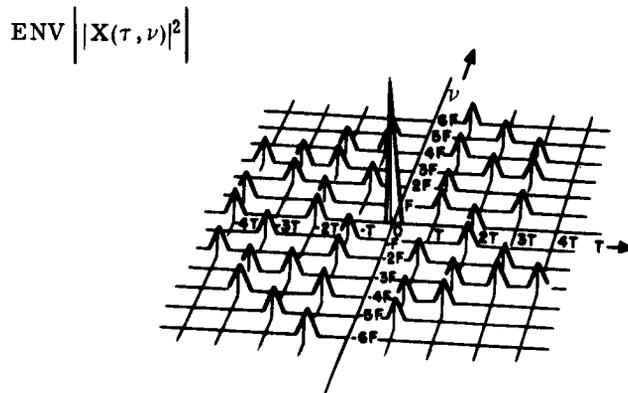

Figure G-7a. Gross Structure of the Ambiguity Surface
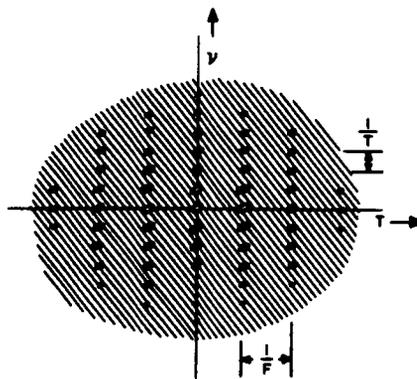
Figure G-7b. Fine Structure of the Central Ambiguity Surface of the Pulse Train

Figure G-7. Sketches of Ambiguity Surface of Pulse Train Consisting of Uniformly Spaced, Pseudo Randomly Frequency Shifted Pulses. (From Reference[4] .)

---

* See Appendix H for a discussion of dimensionality.

D. MULTIPLE CARRIER PULSE TRAINS

The individual pulses again have the same envelope and bandwidth but different carrier frequencies. It is assumed that the pulses in the train do not sensibly overlap in either frequency or trim and are separated in frequency by integer multiples of a frequency differences F.

Figure G-7 shows the structure of the ambiguity surface for the case of pseudo-random shifting of the carrier frequency of uniformly spaced pulses.

Figure G-7a illustrates the ultimate reduction of the gross structure of the secondary surfaces possible with a properly selected pulse train. To achieve such an ambiguity surface, it is required that the pulse train be so constructed that if it is moved in the time-frequency plane, component pulses will not coincide with the original positions of the pulses in more than one place at a time for any movement.

Figure G-7b shows the fine structure of the central surface. The amplitude of the secondary surfaces can be reduced by staggering the interpulse spacing.

An examination of Figures G-5, G-6, and G-7 shows that waveform selection can be used to move the volume of the ambiguity surface around in the τ, ν plane. But, decreasing the volume in one region increases it in another. This is an illustration of a general theorem that states that the volume under an ambiguity surface is equal to unity, i.e.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |X(\tau, \nu)|^2 \, d\tau d\nu = 1 \tag{g-21}$$

Figure G-5, for the uniform pulse train, illustrates this effect quite clearly. Repeating the pulse causes a large decrease in the volume of the central surface, compared to a single pulse, due largely to a decrease in width in the ν direction. The volume decrease results in the removed volume being shifted to many secondary, somewhat isolated in this case, responses spread over the τ, ν plane.

If the time-bandwidth product of a waveform is large, the central spike contains relatively little volume, as is illustrated by the preceding figures. The volume, therefore, is largely contained in secondary responses surrounding the central spike. These secondary responses can be made uniformly low by the methods discussed. When this is done, it is conceptually and analytically convenient to have a general approximate description of the ambiguity surface. The so-called thumb tack surface, shown in Figure G-8, provides such a description. Note that the volume in the pedestal, representing the secondary response, is unity. This representation is especially useful when considering filter responses to several waveforms, identical to the one to which it is matched, except shifted in time and/or frequency. It shows that the average interference from a displaced waveform is about $1/\beta_0 t_0$ below the response of the matched waveform, assuming both contain the same energy.

Cross ambiguity functions X_{12} (τ, ν) describe, in general, a filter response to waveforms other than that to which it is matched. The definition is similar to Equation (g-4), if both waveforms are normalized to unit energy, i.e.,

$$X_{12} (\tau, \nu) = \int_{-\infty}^{\infty} \mu_1 (t) \, \mu_2{}^* (t - \tau) \, e^{j2\pi\nu t} \, dt \qquad\qquad (g\text{-}22)$$

It has been shown (Reference [7]) that the volume under the cross-ambiguity surface $|X_{12}$ $(\tau, \nu)|^2$ is also unity This means that two large $\beta_0 \, t_0$, equal (unit) energy waveforms, having low crosscorrelation functions for all frequency shifts, have a cross-ambiguity surfaces, $|X_{12}$ $(\tau, \nu) |^2$, representable by the pedestal only of Figure G-8. Therefore, in general, the interference between equal energy waveforms is about $1/\beta_0 \, t_0$ below the matched waveform response.

The "thumb tack" representation assumes small average value of the $t\phi'$ (t) product, that is, α. When α for the pulse train is not small, it is not possible to realize a thumb tack surface. It is possible, however, to construct pulse trains with small α from pulses, not identical, which individually have large α.
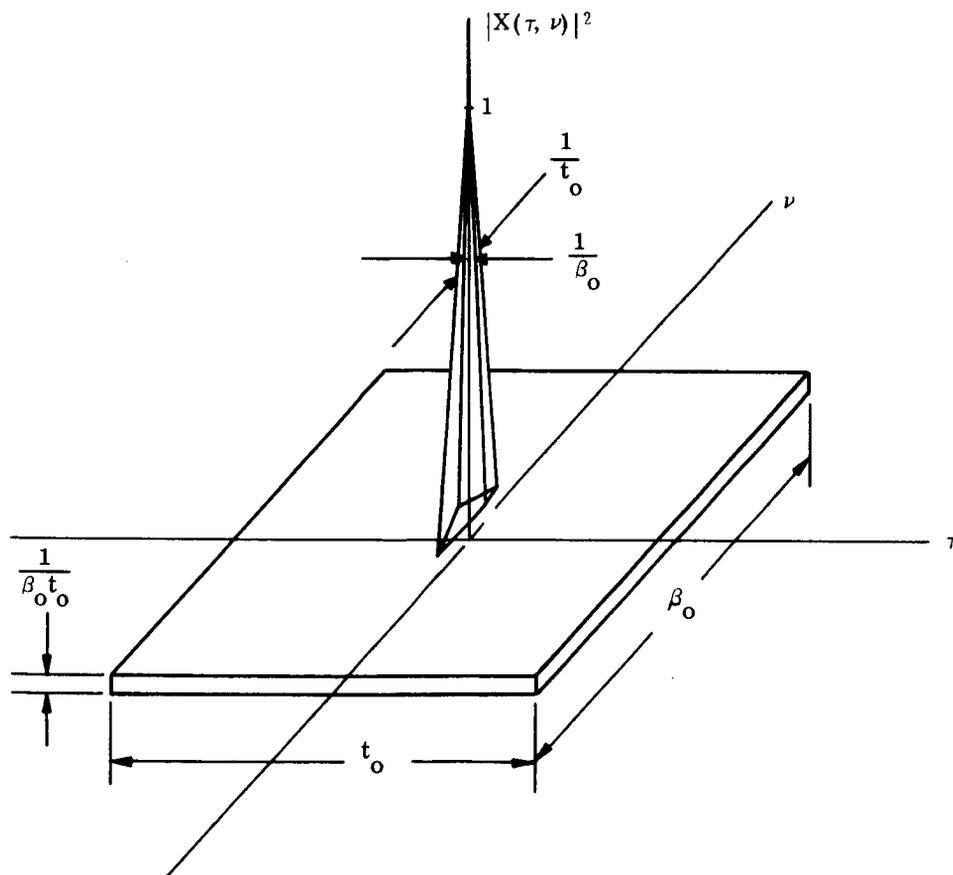

Figure G-8. Thumb Tack Ambiguity Surface

To illustrate the use of ambiguity diagrams, consider the uniform pulse train. It might be thought that the relatively clear intervals along the τ axis in Figure G-5 could be used effectively in time division multiplexing of pulse trains. This is true, in some cases, but for ranging from satellites, it is not attractive because, since a response can come from anywhere on approximately half the earth surface, the ambiguous range delay is about 43 msec. This means the ambiguity surface of the pulse train can have no large secondary peaks within 43 msec. of the central peak. Therefore, T must be greater than 43 msec. and $1/T < 23$ Hz. To gain appreciable energy increase over a single pulse, $N > 10$. The half width at the base of the central surface would then be only about 2.3 Hz. With such a narrow width, oscillator instability and doppler frequency shifts would cause severe fading. Automatic frequency control can be used to reduce fading; but, the long processing time makes the technique unattractive.

It is important to note that the half width of this central spike in ν is the reciprocal of the pulse train duration, NT. This is generally true for coherently processed pulse trains, for which the phase factor α is small.

Figure G-6 shows that repetition interval staggering reduces secondary peaks in both the τ and ν directions. This effect can be used to remove the restrictions on T (i.e., T > 43 msec.). The restriction on the train duration, resulting from frequency instability remains, but is not important because T, and therefore the train duration, can be made small. Note that the secondary surfaces can be smeared out, but not eliminated. The smeared surfaces introduce interference but, being at a low level, it is analogous to noise.

APPENDIX H

RELATIONSHIP BETWEEN SYSTEM PARAMETERS

Some of the relationships between system parameters are derived in this appendix. Because the concept of dimensionality of waveforms and waveform space is useful in deriving these relationships, this section will start with a discussion of this concept. As with the use of bandwidth, time duration and phase parameters discussed previously, the results obtained using the concept of dimensionality will be approximately correct. But, the accuracy of the approximation will depend on the specific problem considered.

Probably, the most straight-forward way to illustrate the dimensionality of a waveform space is through the use of the sampling theorem.

Effectively this theorem states that any baseband waveform with frequency components, confined to the interval $|f| < W_w$ Hz, can be described by amplitude samples of the waveform, separated by $1/2W_w$ seconds, interpolated by appropriate sync functions. More specifically, if $s(t)$ is a baseband waveform then

$$s(t) = \sum_{i=-\infty}^{\infty} a_i \ \text{sinc}\,(2W_w t - i) \tag{h-1}$$

where

$$a_i = s\left(\frac{i}{2W_w}\right) \quad , \quad \text{sinc } x = \frac{\sin \pi x}{\pi x}$$

The a_i are the amplitudes of the waveform at $t = i/2W_w$, where i is an integer. From a filter viewpoint, the theorem means that any wave form, with frequency components confined to $|f| < W_s$, can be generated by applying appropriate amplitude samples spaced $1/2W_s$ seconds to an ideal low pass filter with cutoff frequency W_s Hz. If the samples are spaced closer than $1/2W_s$ second, they either cannot be picked independently, or if they are independent, the filter will introduce correlation between samples, so that the output will be a waveform capable of being defined by samples, separated by $1/2W_s$ seconds.

In view of the above, if a waveform is confined to a time interval of T_w second and the frequency interval $|f| < W_w$, equation (h-1) can be written

$$s(t) = \sum_{i=1}^{2W_w 1_w} a_i \ \text{sinc } (2W_w t - i) \tag{h-2}$$

This approximation is useful for $2W_w T_w$ larger than 2 or 3 and becomes better as $2W_w T_w$ increases. It is possible theoretically to confine a waveform in both time and frequency, but for engineering purposes the

portion of the signal outside the time and frequency intervals can be made small enough, especially for large $2W_wT_w$, so that the approximation is useful.

One interpretation of equation (h-2) is that, if a waveform is confined to frequency and time intervals of $|f| < W_w$ and T_w, it can be defined by $2W_wT_w$ or less parameters, which are the amplitudes of samples spaced $1/2W_w$ second. Therefore, the dimensionality of the waveform is defined as

$$D_w = 2W_wT_w \qquad \text{(h-3)}$$

Similarly, if bandwidth of an ideal low pass channel is $|f| < W_s$ and the duration of a signal in that channel is limited to T_s, the dimensionality of the waveform space is

$$D_s = 2W_sT_s \qquad \text{(h-4)}$$

A similar result is obtained in the case of bandpass waveforms. If $s(t)$ is such a waveform, confined to a bandwidth, W_w, centered at f_c, i.e., its positive frequency components are restricted to the interval

$$f_c - \frac{W_w}{2} < f < f_c + \frac{W_w}{2} \;,$$

it can be written as

$$s(t) = s_i(t) \cos \omega_c t - s_q(t) \sin \omega_c t \qquad \text{(h-5)}$$

where $\omega_c = 2\pi f_c$

$s_i(t)$ and $s_q(t)$ are the amplitudes of the in-phase and quadrature components of the waveform and can be independent. They are baseband waveforms, confined to the frequency interval $|f| < W_{w/2}$. From the previous discussion of baseband waveforms, it is apparent that if $s(t)$ is T_w seconds long, both $s_i(t)$ and $s_q(t)$ can require, at most, T_wW_w parameters in their description. Since they can be independent and each can require T_wW_w parameters, then $s(t)$ can require, at most, $2T_wW_w$ parameters.

Therefore, if either baseband or passband waveforms are restricted to a bandwidth W_w and a duration T_w, they can contain no more than $2T_wW_w$ independent parameters, taken here as amplitude samples. The waveform is said to have $2T_wW_w$ dimensions, corresponding to the maximum number of independent amplitude samples. Such a waveform will fit into a waveform space, if $T_w \leq T_s$ and $W_w < W_s$, which requires the $D_w \leq D_s$. Given the dimensionality of a waveform space, some method of defining useful waveforms fitting the space is needed. The following is a discussion of one method that appears appropriate for application to pulse train ranging and gives a useful relationship between waveform duration, bandwidth and the number of waveforms available in the waveform space.

With D_s, the dimensionality of the waveform space, it has been shown that communication using this space is done with the least power, if the information is coded into M symbol waveforms, $M = D_s + 1$, which correspond to the vertices of a regular polyhedron, with M vertices, centered at the origin of the space. The crosscorrelation coefficient of these waveforms is $-1/D_s = -1/M - 1$, corresponding geometrically, to the projection of a vector from the origin to a vertex on a similar vector, drawn to any other vertex. Figure H-1 illustrates the case where $D_s = 2$, $M = 3$. The regular polyhedron is an equilateral
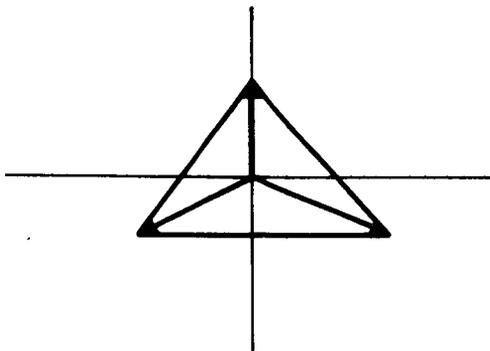
Figure H-1. Waveform Space ($D_s = 2$, $M = 3$)

triangle. The three vectors represent the waveforms. For baseband waveforms, the projections of a vector on the coordinates represents the amplitude of two samples, a_1 and a_2, from which the waveform can be found by applying equation (h-1).

For bandpass waveforms, the projections and the coordinates represent the amplitude of simultaneous samples of $s_i(t)$ and $s_q(t)$. Figure H-1, in this case, represents three equal amplitude, constant frequency pulses, differing from each other only by 120 degrees in carrier phase.

In summary, this method of specifying waveforms gives, in a space of D_s dimension, M waveforms

$$M = D_s + 1 = 2W_s T_s + 1 \tag{h-6}$$

with mutual crosscorrelation coefficients

$$\gamma = -\frac{1}{D_s} = -\frac{1}{(M-1)} \tag{h-7}$$

The energy content is the same for all M waveforms. Equation (h-7) shows that the waveforms become nearly orthogonal as D_s becomes large, i.e. $\gamma \to 0$ as D_s increases. It should be noted that, although the crosscorrelation coefficients are small, the autocorrelation and crosscorrelation functions of the waveforms are not necessarily small for all time shifts. However, the correlation functions can be varied by rotation of reference axes, and it will be assumed that usefully low values of the correlation functions can be obtained by a suitable rotation.

Because of the role played by the phase information, detection of these waveforms must be done coherently, such as by a matched filter. But one important implementation of pulse train ranging utilizes simple, amplitude modulated pulses which are non-coherently detected, e.g. by envelope or square law envelope detectors. Also, it is convenient to make these pulses of equal amplitude. In this case, since phase information is not utilized by the detector and amplitude variations are not allowed, about the only variable available is the time of occurrence of the pulse. If the simple pulses, at different time positions, are to have small crosscorrelation coefficients, they must, in the bandpass case, be separated by at least $1/W_s$ seconds. This gives $T_s W_s$ positions for the pulses in a time interval of T_s seconds. This will be taken as the

number of orthogonal waveforms available in a bandpass space of D_s dimensions, when simple envelope detected pulses are used.

The methods discussed, for defining waveforms and relating these to the waveform space, permit the derivation of the relationship between bandwidth, time and address coding.

Let W_c be the assigned channel bandwidth and W_u be the portion of W_c usable for transmissions. If ν_{st} is the maximum plus and minus deviation of a transmission center frequency from the channel center frequency, the restriction on the transmission bandwidth is

$$W_{ut} = W_c - 2|\nu_{st}| \text{ Hz} \qquad \text{(h-8)}$$

ν_{st} is the sum of frequency shifts due to doppler and transmitter oscillator instability. Consequently the transmission is confined to W_{ut} to avoid transmission outside the assigned channel.

In addition to transmitted center frequency shifts, receiver instabilities may further restrict the usable bandwidth. Considering the receiver instabilities only and letting ν_{sr} be the frequency shift of the receiver center frequency, the restriction on the usable bandwidth is

$$W_{ur} = W_c - 2|\nu_{sr}| \qquad \text{(h-9)}$$

Considering both possible frequency shifts, gives the maximum usable bandwidth as

$$W_s = W_c - 2\left(|\nu_{st}| + |\nu_{sr}|\right) \text{ Hz} \qquad \text{(h-10)}$$

It is often possible to make the contribution of the receiver frequency shift to bandwidth reduction small by receiver AFC; thereby, making ν_{dt} the major consideration. W_s is the bandwidth limitation of the waveform space.

The time dimension of the waveform space, T_s, will usually be restricted by addressing rate requirements or equipment complexity. The time duration of coherently processed waveforms T_w will similarly be restricted by equipment complexity or possibly by frequency errors. The error frequency for coherent processing is the difference between the waveform center frequency, when applied to the matched filter, and the filter center frequency. Temperature changes, for example, may change the filter center frequency. If ν_f is the shift in filter frequency, (it may be positive or negative), an error frequency may be defined as

$$\nu_e = \nu_{st} - (\nu_{sr} + \nu_f) \qquad \text{(h-11)}$$

As discussed in Appendix G, the matched filter response can be significantly decreased by frequency offsets (errors). The exact relation depends upon the waveform and can be worked out from the definition of the ambiguity function. Usually, it is desirable to limit $|\nu_e| < 1/t_o$, (where t_o is the RMS waveform duration described in Appendix B), giving a bound on the coherent processing time of

$$t_o < \frac{1}{|\nu_e|} \qquad \text{(h-12)}$$

T_w, which corresponds to conventional definition of waveform duration, is proportional to t_o. But, the definition of t_o does not give a general proportionality constant. The constant can be found for any specific waveform

from the equations of Appendix B. Consequently, ν_e will place an upper bound on T_w. But, the exact value has to be worked out for a specific waveform.

To apply the concept of dimensionality and the method discussed for waveform coding to pulse trains consider the time dimension of the waveform space divided into N intervals, T seconds long, as shown in Figure H-2. Each T second interval contains a pulse, T_w seconds long.
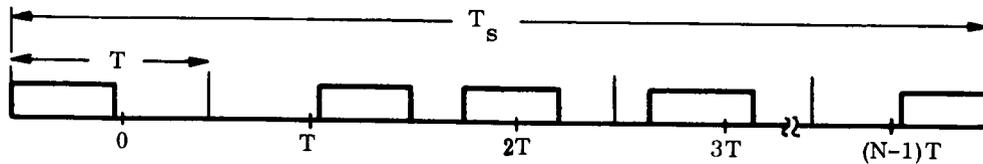

Figure H-2. Time Dimension of Waveform Space

The pulses, in general, are staggered about the midpoints of the T second intervals and have the same real envelope and bandwidth, but not necessarily, the same center frequency and phase function.

Let the bandwidth of the pulse train space, which is equivalent to the waveform space discussed before, be W_s and the time duration T_s, giving a dimensionality of $D_S = 2T_S W_S$, as before. The pulse waveform dimensionality is $D_w = 2T_w W_w$, which must be less than or equal to the pulse space dimensionality, $D_p = 2TW_s$. From the definitions, $T_w \leq T \leq T_s$, $W_w \leq W_s$ and $D_w \leq D_p \leq D_s$. If U is the number of users to be addressed, that number of distinguishable waveforms must be specified in the D_s space. This can be done by considering the pulse train as a N digit code word with one digit for each pulse space. Each digit can take on $D_p + 1 = M_p$ different values, which are represented by the number of nearly orthogonal waveforms (symbols) available in D_p space. If D_w is less than D , frequency and/or time shifted versions of a pulse waveform are taken as a distinct waveform, provided the frequency shift is greater than $1/T_w$ or that the time shift is greater than $1/W_w$. It is assumed that frequency and time shifts are integral multiples of $1/T_w$ and $1/W_w$, and the allowed shifts do not take the waveform out of the pulse space. These assumptions permit M_p waveforms in D_p space, even if D_w is smaller than D_p.

In a code word of N digits, each of which can be any one of M_p symbols, the number of different code words is

$$
\begin{aligned}
U &= (M_p)^N \\
&= \left[2TW_s + 1 \right]^N \\
&= \left[\frac{2T_s W_s}{N} + 1 \right]^N
\end{aligned}
$$

(h-13)

This equation relates the number of addresses, pulse train space dimensions and the number of pulses in the train.

When N = 1, $U = 2T_sW_s + 1$. This is the case when the pulse train is coherently processed as a whole. It means-assuming $2T_sW_s \gg 1$ (which is usually true), the dimensionality of the pulse train space is equal to the number of users. When the number of user addresses is large, this may require processing times longer than permitted by economic or frequency stability requirements.

Increasing N and coherently processing only the pulse waveform, decreases the coherent processing time requirements and increases the number of addresses available.

APPENDIX J

LIST OF SYMBOLS

LIST OF SYMBOLS

SECTION IV. Multiple Pulse Train Ranging

E	–	energy of a pulse train
N_o	–	single sided noise power density
T_S	–	period (duration) of signal
W_S	–	waveform space bandwidth
D_S	–	dimensionality of waveform space
D_P	–	dimensionality of pulse space
D_W	–	dimensionality of pulse waveform
M_P	–	number of waveforms in pulse space
N	–	number of pulses in a pulse train
m	–	number of available waveforms
ν_d	–	one-way doppler frequency shift
$\dot{R}$	–	range rate
c	–	propagation velocity
f_c	–	carrier frequency
t_o	–	time duration of pulse train
L	–	number of simultaneous user responses
q	–	number of addresses in use at a given time
σ_τ	–	range delay measurement error
β_o	–	bandwidth of pulse train
U	–	number of user addresses required
t_{op}	–	coherent processing time of matched filter
T	–	period (duration) of pulse

SECTION V. Side Tone Modulation

T_P	–	mean-time-to-unlock
B_n	–	two-sided noise bandwidth of loop
S/N	–	signal-to-noise ratio

SECTION V. (CONT.)   Side Tone Modulation

ϕ_ϵ		phase jitter
$(S/N)_o$	–	output signal-to-noise ratio
$(S/N)_i$	–	input signal-to-noise ratio
B_{CL}	–	noise bandwidth of carrier loop
B_{TL}	–	noise bandwidth of tone loop
T_A	–	acquisition time
T_C	–	carrier loop acquisition time
T_T	–	tone loop acquisition time
SR	–	sweep rate
Δf	–	total frequency uncertainty
M	–	no. of messages—limit of mixing messages
T_M	–	message length in time
T_D	–	data gathering time
T_G	–	time guard band between messages
η	–	data gathering technique efficiency
f_d	–	doppler frequency shift
C/N	–	carrier-to-noise ratio

APPENDIX A.   A Multiple Access Approach

t_o	–	reference (origin) time
A	–	address word
R	–	range word
D	–	data word

APPENDIX B.   Waveform Descriptions and Parameters

$s_p(t)$	–	a physical signal (function)
$\mu(t)$	–	complex envelope
β_o	–	bandwidth of pulse train
t_o	–	time duration of pulse train

LIST OF SYMBOLS

APPENDIX B. (CONT.) Waveform Descriptions and Parameters

α — phase structure of pulse train

σ_τ — range delay measurement error

$M(f)$ — Fourier transform of $\mu(t)$

$\phi(t)$ — phase function of $\mu(t)$

E — energy of ranging waveform

N_o — white, Gaussian noise power density

APPENDIX C. FM Pulse Compression

B_b — total bandwidth of transmitted signal

B_f — bandwidth of individual filter

R — compression ratio

t_t — width of trigger pulse

t_p — width of transmitted pulse

t_D — time required for detection of received signal

$(S/N)_v$ — voltage signal-to-noise ratio

V_s — signal voltage

V_n — noise voltage

$(S/N)_p$ — output power signal-to-noise ratio (single pulse rcvr)

T_D — total detection time

M — no. of messages—limit of mixing messages

T_M — message length in time

T_A — acquisition time

T_G — time guard band between messages

T_P — round trip propagation time

η — data gathering technique efficiency

F_T — time figure of merit for growth capacity

APPENDIX D.   FM-CW Triangular Modulation

f_c —   carrier frequency

S_t —   rms carrier threshold

k_1 —   rms noise voltage per unit bandwidth

Bf_D —   bandwidth of an FM signal

P_t —   signal power at threshold

C/N —   carrier-to-noise ratio

P —   sweep period

τ —   time delay between transmitted and received waveforms (round trip)

f_b —   beat, or difference, frequency between transmitted and received waveforms

f_1 —   frequency at start of sweep

f_2 —   frequency at end of sweep

δf —   frequency shift due to doppler effect

$\Delta_n P$ —   measurement resolution in units of P, the sweep period

N —   counting resolution

$\Delta \tau$ —   range resolution interval

f_s —   sweep frequency

$P/2$ —   ambiguity interval

f_n —   highest modulating frequency

f_{bw} —   transmitted signal bandwidth

APPENDIX E.   Pseudo Random Code Technique

N —   number of stages in shift register

$\Delta \tau$ —   time shift of code with respect to reference code

M —   number of stages in shift register $(M \neq N)$

t_{CD} —   time for correlation decision

l_P —   number of bits in a code sequence

l_K —   number of bits in a code sub-sequence

l_M —   number of bits in a code sub-sequence

APPENDIX E (CONT.) Pseudo Random Code Technique

l_N	–	number of bits in a code sub-sequence
l_{AV}	–	average bit length of sub-sequences
S/N	–	signal-to-noise ratio
ϕ_ϵ	–	phase error
δT	–	time jitter
$R(\tau)$	–	auto-correlation function
$R'(\tau)$	–	derivative of auto-correlation function
SR	–	sweep rate
B_n	–	closed loop noise bandwidth
ω_D	–	bandwidth of decision filter
K	–	constant
J	–	number of sub-codes in pseudo code
ω_v	–	cut-off frequency of low pass filter
ω_c	–	carrier frequency
$A(\omega)$	–	a symmetric angular frequency spectrum
ω_d	–	upper cut-off frequency—single sideband r.f. spectrum
$B(|\omega|)$	–	bandpass angular frequency spectrum
$c(|\omega|)$	–	resultant angular frequency spectrum
$v(t)$	–	modulating signal
$s(t)$	–	modulated signal (carrier + modulation)
$s_1(t)$	–	single sideband signal
$s_2(t)$	–	double sideband (suppressed carrier) signal
A	–	carrier amplitude
$v_1(t)$	–	modulating signal for single sideband case
$v_2(t)$	–	modulating signal for double sideband case
K_m	–	local oscillator signal amplitude
$r_1(t)$	–	recovered single sideband signal
$r_2(t)$	–	recovered double sideband signal

APPENDIX E (CONT.) Pseudo Random Code Technique

$n(t)$	–	white Gaussian noise associated with r.f. signal
N_o	–	noise power (per unit bandwidth)
$r_n(t)$	–	baseband noise
$H(\omega)$	–	transfer function of receiver filter
σ_n^2	–	mean square output noise power
f_v	–	cut-off frequency of receiver low pass filter
ρ	–	peak signal-to-mean square noise power ratio
τ_d	–	round trip delay time
m_i	–	number of bit intervals
d_r	–	bit duration or interval
σ_r^2	–	variance of range error
ϕ_ϵ	–	vernier phase error
σ_θ^2	–	variance of vernier phase error
c	–	propagation velocity
Tacq	–	acquisition time
h	–	total number of correlators in receiver
$p(\ )$	–	relative frequency of binary sequence ( )
$\rho_{xy}^{(j)}$	–	normalized correlation function
S_T	–	number of bits to be searched in time multiplex case
S_F	–	number of bits to be searched in frequency multiplex case
$R_o(\tau)$	–	auto-correlation function
$S_i(\omega)$	–	power spectrum, pseudo noise input signal
$h(t)$	–	impulse response of an ideal filter
$F_i(\omega)$	–	frequency spectrum of input pseudo noise signal

APPENDIX F.   Square Law Detection and Post Detection Filtering of Asymmetrical Signals

$\underset{\sim}{x}(t)$ –   input random signal

$\underset{\sim}{y}(t)$ –   output random signal

$f_{\underset{\sim}{x}}(x)$ –   probability density function of the random input

$f_{\underset{\sim}{y}}(Y)$ –   probability density function of the random output

$E\left[\underset{\sim}{y}(t)\right]$ –   expected value of the output

$R_{\underset{\sim}{y}}(\tau)$ –   autocorrelation of the output

$R_{\underset{\sim}{x}}(\tau)$ –   autocorrelation of the input

σ_y^2 –   variance of $\underset{\sim}{y}(t)$

$S_{\underset{\sim}{x}}(\omega)$ –   power spectrum of $\underset{\sim}{x}(t)$

$S_{\underset{\sim}{y}}(\omega)$ –   power spectrum of output

$\underset{\sim}{x}_c(t)$ –   input carrier

$\underset{\sim}{x}_1(t)$ –   input tone

$\underset{\sim}{n}(t)$ –   input noise

$S_n(\omega)$ –   noise power spectrum

$2\pi B$ –   bandwidth of ideal bandpass filter (rad/sec)

$\underset{\sim}{z}(t)$ –   output of zonal filter

$S_{\underset{\sim}{z}}(\omega)$ –   power spectrum of $\underset{\sim}{z}(t)$

ω_v –   cut-off frequency of low-pass filter

P_N –   noise power out of filter

Γ_o –   output tone-to-noise power ratio

Γ_i –   input carrier-to-noise power ratio

Γ_1 –   ratio of Γ_o to Γ_i

APPENDIX G.   Matched Filters and Ambiguity Functions

$s(t)$ –   signal function

$h(t)$ –   impulse response function

Δ –   arbitrary time delay

APPENDIX G (CONT.)   Matched Filters and Ambiguity Functions

ν — offset frequency

τ — time shift

E — energy contained in a pulse

T_0 — period of pulse train

N — number of pulses in a train

β_0 — bandwidth of pulse train

APPENDIX H.   Relationship Between System Parameters

W_W — bandwidth of signal

$s(t)$ — baseband waveform

W_S — cutoff frequency of low-pass filter

T_S — period (duration) of signal

T_W — period of waveform

D_W — dimensionality of waveform

D_S — dimensionality of waveform space

f_c — center frequency

M — integer—no. of symbol waveforms

W_C — assigned channel bandwidth

W_u — portion of W_c useable for transmissions

ν_{st} — maximum total deviation of transmission center frequency

ν_{sr} — frequency shift of receiver center frequency

ν_f — shift in filter frequency

ν_e — error frequency

t_0 — time duration of rms waveform

U — no. of users to be addressed

D_p — pulse space dimensionality

N — no. of digits in code word (no. of pulse spaces)

REFERENCES

[1] Woodward, P.M.; "Probability and Information Theory with Application to Radar," McGraw Hill Book Company, New York.

[2] Univac; "Autoscan, An Automatic Satellite/Computer Aid to Navigation," NASA 706, 1 February 1964.

3 Eckler, A.R.; "The Construction of Missile Guidance Codes Resistant to Random Interference," BSTJ, 39, pp. 973-994; July 1960.

4 Rihaczek, A.W.; "Radar Resolution Properties of Pulse Trains," PROC. IEEE, pp. 153-164; February 1964.

5 Rihaczek, A.W.; "Radar Signal Design for Target Resolution," PROC. IEEE, pp. 116-128; February 1965.

6 Helstrom, C.W.; "Statistical Theory of Signal Detection," New York, Pergamon Press; 1960.

7 Stutt, C.A.; "A Note on Invariant Relations for Ambiguity Functions," IRE Trans. on Information Theory, pp. 164-167; December 1959.

8 Pettit, R.H.; "Pulse Sequences with Good Autocorrelation Properties," Microwave Journal, pp. 63-67; February 1967.

9 Chesler, D.; "Performance of a Multiple Address RADA System," IEEE Trans. on Communication Tech., pp. 369-375; August 1966.

10 Urkowitz, H.; "Energy Detection of Unknown Deterministic Signals," Proc. IEEE, pp. 523-531; April 1967.

11 Viterbi, A.M.; "Phase Locked Dynamics in the Presence of Noise by Fokker-Planck Techniques," Proceedings of IEEE; December 1963.

12 Smith, B.M.; "A Semi-Empirical Approach to the PLL Threshold," IEEE Transactions on Aerospace and Electronic Systems; July 1966.

13 Skolnik, M.I.; "Introduction to Radar Systems," McGraw-Hill Book Company, Inc., 1962.

14 Nichols & Rauch; "Radio Telemetry," Second Edition, John Wiley and Sons, Inc., New York, 1956.

15 Sunde, E.D.; "Theoretical Fundamentals of Pulse Transmission-I," B.S.T.J. 33, May 1954, pp. 579-660.

16 Golomb, S.W., "Deep Space Range Measurement," Research Summary No. 36-1, JPL-Pasadena, California, 15 February 1960, pp. 39-42.

17 Easterling, M., "Long Range Precision Ranging System," Tech. Report No. 32-80, JPL-Pasadena, California, 10 July 1961, pp. 1-7.

18 LeVeque, W.J., "Elementary Theory of Numbers," Theorem 3-12, p. 52, Addison-Wesley, Reading, Massachusetts, 1962.

19 Golomb, S.W., Ed., "Digital Communication with Space Applications," Chapter 5, pp. 65-105, Prentice Hall, Englewood, New Jersey, 1964.

20 Plotkin, M., "Binary Sequence with Specified Minimum Distance," Trans. IRE IT-6, 4 September 1954, pp. 38-49.

21 Zierler, N., "Linear Recurring Sequence," J. Soc. Ind. & App. Math. 7 (1959) 31-48.

22 Watson, E.J., "Primitive Polynomials (Mod. 2), "Mathematics of Computation 16, 368-69 (1962).

23 Brauer, A., "On a New Class of Hadamard Determinante," Mathmatische Zeitschrift 58, 219-25 (1953).