DOT/TSC-RA-78-06

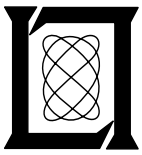
Project Report
ATC-86

Alternate Waveforms for a Low-Cost Civil GPS Receiver

H.F. Vandevenne
E.J. Kelly

14 April 1980

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS



Prepared for the Transportation Systems Center,
Cambridge, MA

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/TSC-RA-78-06		

4. Title and Subtitle	5. Report Date
Alternate Waveforms for a Low-Cost Civil GPS Receiver	14 April 1980
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Herman F. Vandevenne and Edward J. Kelly	MIT LIN ATC-86

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Massachusetts Institute of Technology Lincoln Laboratory P.O. Box 73 Lexington, Massachusetts 02173	11. Contract or Grant No. DOT/FA-79-WAI-091
	13. Type of Report and Period Covered Project Report
12. Sponsoring Agency Name and Address Transportation Systems Center Department of Transportation Cambridge, Massachusetts	14. Sponsoring Agency Code

15. Supplementary Notes

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology.

16. Abstract

This report examines the technical feasibility of alternate waveforms to perform the GPS functions and to result in less complex receivers than is possible with the GPS C/A waveform.

The approach taken to accomplish this objective is (a) to identify, after a quick broad overview, a few promising waveforms, (b) to complete the architectural synthesis of a GPS system based on the new waveforms, (c) to analyze the performance of these systems in providing ranging and communication capability and (d) to assess the impact on receiver hardware and software.

One conclusion is that the class of pulsed waveform, when combined with emerging matched filter technology, permits the use of a greatly simplified receiver design. A major part of the report consists of description of a pulse-based GPS design. Although very specific, the design was intended as a vehicle for performance analysis only and many refinements and alternatives within the same class are possible.

Other classes of waveforms such as tone ranging and chirped FM waveforms, although adequate, do not seem superior to the present PN encoded CW waveform for GPS.

Finally, if the "alternate" waveform is simply the present C/A waveform but with substantially increased power level, it is shown that with the present designs the performance will be greatly increased or, alternatively, performance margin can be traded for a simplified receiver implementation.

17. Key Words	18. Distribution Statement
Global Positioning System Navigation Receiver Ranging Waveforms Tone Ranging Clock Stability CCD Matched Filter Application	Document is available to the public through the National Technical Information Service, Springfield, Virginia, 22151

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	
Unclassified	Unclassified	126	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ABSTRACT

This report examines the technical feasibility of alternate waveforms to perform the GPS functions and to result in less complex receivers than is possible with the GPS C/A waveform.

The approach taken to accomplish this objective is (a) to identify, after a quick broad overview, a few promising waveforms, (b) to complete the architectural synthesis of a GPS system based on the new waveforms, (c) to analyze the performance of these systems in providing ranging and communication capability and (d) to assess the impact on receiver hardware and software.

One conclusion is that the class of pulsed waveform, when combined with emerging matched filter technology, permits the use of a greatly simplified receiver design. A major part of the report consists of description of a pulse-based GPS design. Although very specific, the design was intended as a vehicle for performance analysis only and many refinements and alternatives within the same class are possible.

Other classes of waveforms such as tone ranging and chirped FM waveforms, although adequate, do not seem superior to the present PN encoded CW waveform for GPS.

Finally, if the "alternate" waveform is simply the present C/A waveform but with substantially increased power level, it is shown that with the present designs the performance will be greatly increase or, alternatively, performance margin can be traded for a simplified receiver implementation.

CONTENTS

1.	INTRODUCTION	1
	1.1 Background	1
	1.2 Scope of Study	1
	1.3 Methodology	4
	1.4 Further Outline of Report	7
2.	REVIEW OF WAVEFORM CHOICES	8
	2.1 Choice of Waveforms for Ranging	8
	2.2 Basic Types of Ranging Systems	8
	2.3 Satellite Configurations, Frequency Choices, Channel Sharing	9
	2.4 TDMA Waveform Choices	10
	2.4.1 Constant Carrier Single Pulse	11
	2.4.2 Waveforms with Thumbtack Ambiguity Function	12
	2.4.3 Waveforms with Sheared Ridge Ambiguity Function	13
	2.4.4 Class of Harmonic Signals	13
3.	USE OF PRESENT C/A WAVEFORM WITH MORE POWER	15
	3.1 Context for Study on Increased Power	15
	3.2 Summary Statements on Impact of More Power on GPS Receiver	16
	3.2.1 Code Synchronization	16
	3.2.2 Possible Use of Shorter Codes	17
	3.2.3 Use of Matched Filter for Code Synchronization	17
	3.2.4 Increased Ranging Accuracy	17
	3.2.5 Increased Performance of Carrier PLL	17
	3.2.6 Bit Error Rates	18
	3.2.7 Shorter Dwell-Time per Satellite	18
	3.2.8 Possibility of Using Less Costly VCO	20

CONTENTS (CONTINUED)

4. TDMA PULSED GPS WAVEFORM 21

4.1 Waveform Description 21

4.1.1 Power Level 21

4.1.2 Signal Format 22

4.1.3 Waveform 25

4.1.4 Ranging and Synchronization 25

4.1.5 Data Modulation 27

4.2 Receiver Design 31

4.2.1 Receiver Hardware Design 31

4.2.2 Receiver Packaging and Sizing 36

4.3 Conclusions on the TDMA Pulsed Waveform Design 38

5. TDMA TONE RANGING GPS WAVEFORM 41

5.1 Tone Ranging System Description 43

5.2 Analysis Defining Main System Parameters 50

5.2.1 Signal Format 50

5.2.1.1 Determination of Time Interval for Radiating Pure Carrier 50

5.2.1.2 Determination of Time Interval T for Radiating Tones 51

5.2.1.3 Choosing Time Interval for Transmitting Data 52

5.2.2 Choosing Number of Tones and Tone Frequencies 53

5.3 Discussion on Tone Ranging Performance 61

6. PRFDMA WAVEFORM 63

6.1 Ranging 63

6.2 Data Demodulation 66

6.3 Satellite Identification 66

6.4 Clock Stability and Computer Load Implications 66

7. IMPLICATIONS OF TDMA FORMAT 68

CONTENTS (CONTINUED)

8. HANDOVER FROM ALTERNATE WAVEFORM MODE TO P MODE 72

9. CLOCK STABILITY REQUIREMENTS IN TDMA SYSTEMS 74

APPENDIX A

ANALYSIS OF PERFORMANCE OF GPS C/A WAVEFORM WITH MORE POWER 89

A.1 Code Synchronization 89

A.1.1 Sequential Code Synchronization 89

A.1.2 Possible Use of a Shorter Code 93

A.1.3 Use of Matched Filter 98

A.1.4 Improved Code Tracking 98

A.2 Performance of Carrier Tracking Loop 99

A.3 Data Demodulation 99

A.4 Satellite Scheduling 100

APPENDIX B

DETAILS OF TDMA PULSED GPS RECEIVER 102

APPENDIX C

POWER DISTRIBUTION IN TONES FOR GPS TONE RANGING SYSTEM 110

REFERENCES 114

List of Illustrations

3-1. MINIMUM PHASE ERROR AND OPTIMAL LOOP BANDWIDTH AS A FUNCTION OF C/N_0 19

4-1. PULSED SIGNAL FORMAT 24

4-2. PSE AND AVERAGE MESSAGE RUNLENGTH WITHOUT ERROR VERSUS SNR FOR 16-ARY PPM 30

4-3. OVERALL BLOCK DIAGRAM 32

4-4. RF/IF PORTION 33

4-5. SPECIAL PURPOSE LOGIC 34

4-6. PACKAGING OF TDMA PULSED GPS RECEIVER 37

5-1. TONE RANGING GPS SIGNAL FORMAT 48

5-2. BLOCK DIAGRAM OF GPS TONE RANGING RECEIVER. 49
5-3. SUPPRESSION FACTOR AS FUNCTION OF FM MODULATION INDEX. 60
6-1. SKETCH OF PRFDMA WAVEFORM 64

9-1. DEFINITION OF X 78

A-1. TEST TIME ON ONE CODE PHASE FOR INITIAL CODE ACQUISITION FOR MAXIMUM FREQUENCY UNCERTAINTY OF ± 4 kHz. 96
A-2. TEST TIME ON ONE CODE PHASE FOR INITIAL CODE ACQUISITION FOR MAXIMUM FREQUENCY UNCERTAINTY. 97
B-1. PROCESSING WAVEFORMS. 103
B-2. SYNC PROCESSING FLOW CHART 104

B-3. SPECIAL PURPOSE LOGIC CARD D-1 105

B-4. SPECIAL PURPOSE LOGIC CARD D-2 106

B-5. SPECIAL PURPOSE LOGIC CARD D-3 107

B-6. DATA DEMODULATOR COUNTERSTATES 108

B-7. PULSED GPS RECEIVER: CASE REMOVED 109

List of Tables

4-1. POWER BUDGET 23

4-2. DATA DEMODULATION SUMMARY 29

4-3. SUMMARY OF FEATURES AND PERFORMANCE 40

5-1. GPS TONE RANGING POWER BUDGET 45

5-2. BASIC TONES, f_n, AND TONES ACTUALLY MODULATING CARRIER: f'_n 46

5-3. ERROR BUDGET OF GPS TONE RANGING SYSTEM 47

5-4. RANGE ACCURACY AND HIGHEST TONE 56

5-5. CHOOSING NUMBER OF TONES IN TERMS OF PROBABILITY OF FALSELY RESOLVING AMBIGUITY 58

5-6. SUMMARY OF FEATURES AND PERFORMANCE 62

A-1. HIGH POWER GPS BUDGET 90

SUMMARY

1. Introduction

Lincoln Laboratory was tasked, in an interagency agreement sponsored by the FAA through DOT/TSC, to examine alternate waveforms to the presently planned Clear Acquisition (C/A) GPS waveform. The purpose was to review, select and propose waveforms which would perform the GPS ranging and data carrying functions for navigation, using the GPS satellites, but would result in greatly simplified (and therefore less costly) civil user receivers. Quantitative performance criteria were not specified, but in a broad comparison, performance of the alternate waveform, in terms of navigation output, should at least equal the performance achievable by a sequential receiver using the GPS C/A waveform.

Performance criteria to be included were ranging performance, data handling capability, time-to-first fix, acquisition and re-acquisition capability, and sensitivity to multipath.

The general approach specified was to perform a quick broadbrush review of waveforms, including the present C/A waveform with higher power level, and select a very few promising waveforms for a more substantial design effort. The more detailed designs then would allow credible numerical performance characteristics to be obtained.

Complexity of civil user equipment was to be the major criterion for recommending any changes, in particular alternate or additional waveforms, to GPS because user equipment cost, which reflects the complexity, was deemed to represent the largest component of the total cost of a widely accepted and implemented system.

The principal results of the task outlined above were reported in briefings to the sponsor and a set of four Working Papers. They are summarized in this report.

2. Outline of Principal Results and Conclusions

Following a short review of the principal classes of waveforms, based on their properties to determine range and range rate when optimally processed, a few classes were identified as interesting: the burst pulsed high peak power,

low duty cycle, waveforms combined with some method of bandspreading; the pulsed waveform with a regular PRF; the multitone waveform; the present C/A waveform but with more power. For each class a detailed design was proposed making choices for a number of additional parameters such as data modulation, features for synchronization, and more. While there was often a good reason for specific choices, the designs and their characteristics are considered only representative of what can be obtained using the waveform. More refined designs may result in different and more pertinent choices of parameters.

Early in the study certain "desirable" signal structure features were identified which were adopted in the designs where appropriate. One such feature was the channel sharing method based on TDMA (Time Division Multiple Access). The pulse-burst GPS waveform design and the multitone GPS waveform both use TDMA.

The advantages of TDMA, irrespective of the specific waveform, are rather compelling for a low cost receiver and are discussed in some detail in Section 7 of this report. One such TDMA format would allow each pair of antipodal satellites to transmit in one of twelve timeslots, the total cycle of twelve slots with fixed assignments repeating itself indefinitely. Each timeslot can be further subdivided into signal segments performing specific functions and into a "guardtime" to guarantee that signals transmitted in consecutive timeslots by satellites at large range differences to a user can not overlap. In the pulse-burst design a first segment consists of a burst of pulses for synchronization and ranging. A second segment of pulses transmitted at a lower duty cycle serves the purpose of transmitting the data. In the multitone design the first segment is reserved for transmission of pure carrier for phase-locking, a short second segment is reserved for transmission of tones for ranging for which phase locking can occur very fast since the carrier frequency offset is known from the first segment, and in a third segment DPSK modulated carrier is radiated for data transmission. Briefly stated the principal conclusions on the pulse-burst design (Section 4 in this report) are as follows.

If a high peak power (2kw), low duty cycle pulse burst waveform, having about one-third the average power and the same information content as the current C/A waveform, were added to or would replace the C/A signal on the GPS satellites, it would be possible to produce a high performance navigation receiver of considerably less complexity than the present single channel sequential receiver. The specific PN-encoded pulse burst waveform design and the receiver design take maximum advantage of the features of the emerging CCD (Charge Coupled Device) matched filter technology. The receiver, as compared to a receiver using the C/A signal, exhibits comparable or better accuracy, greater update rate (of the order of one per second) which can be made adaptive depending on user platform maneuver condition, and lower time to first fix (of the order of 20 seconds). This receiver would exhibit less sensitivity to multipath and to signal dropouts because it would continuously track all satellites in view and could essentially instantly re-acquire a lost signal.

These advantages of an additional waveform must be weighted against the cost and difficulty of producing high power pulses on board the satellites. A brief survey has indicated that suitable space qualified power generation hardware is or soon will be available. Also interference and interoperability, even if the new waveform had to replace the C/A waveform, do not seem to present unsurmountable problems. No detailed analysis however was done on these issues.

The advantages of an additional waveform consisting of multitones were not as clear-cut. The specific design (Section 5) allowed demonstration of the capability of adequate performance with consideralby less peak power requirements (8 dB) than the pulse burst system described above. The system would be sensitive to interference however. The core of the receiver hardware would consist of a bank of PLL (Phase Locked Loops) operating in a narrow band around a set of known but changing frequencies. Unless great strides are made in developing cheap PLL's in the half megahertz range, no great cost savings are expected.

A third class of waveforms which was considered (Section 6) consists of an unending PN-encoded pulse train, with a PRF (Pulse Repetition Frequency) specific for each antipodal satellite pair. The advantage of this waveform over the TDMA pulsed GPS design lies in the fact that a lower duty cycle satellite transmitter is needed and the user clock stability requirements can be relaxed (from one to 6 parts in 10^7 per second). As was the case for the TDMA designs, the use of signals from more than the minimal four satellites is possible at little additional computational cost while eliminating problems associated with satellite selection algorithms and problems associated with change-over between selected configurations. The specific PRFDMA (Pulse Repetition Frequency Division Multiple Access) design requires, during a signal acquisition, additional computation for pulse sorting which was absent in the other designs. And finally, the class of the present CW PN encoded GPS waveform with a slight parameter variations was considered. The waveform is ill-suited to a sequential receiver for the data collection function. To resolve this difficulty satisfactorily one may resort to a separate data collecting channel and an additional signal, or dedicate time intervals to collecting data while suspending the tracking function. In general, more power yields better performance of the delay locked loops and phase locked loops resulting in faster (re-) synchronization and possibly higher ranging accuracy. The higher power level makes matched filter applications attractive. The CCD correlator, around which the burst-pulse GPS waveform was designed, is less well suited for the C/A waveform because of the long code (1023 chips) and the need to make it programmable, although technology advances may change this situation.

In this study we also investigated several important issues transcending specific waveform choices, such as handover capability from an alternate waveform to the P-mode and the clock stability requirements for several range update rates. The former was found not to present any serious difficulties. The clock stability requirement for the TDMA pulse (burst) design and the multi-tone design, apart from bias and frequency offset, was characterized by (the square root of) the Allan variance of one part in 10^7 per second, to keep the clock error contribution to the ranging error below 13 feet. For the PRFDMA design the clock stability requirements could be relaxed to about 6 parts in 10^7 for the same contribution to the ranging error.

1. INTRODUCTION

1.1 Background

The basic reason for studying the possible use of GPS by civil aviation is that the system will almost certainly become fully operational and the potential for taking advantage of this navigational resource by non-military users will also exist. The system is being developed and will be maintained by DOD, under present plans; hence a navigational aid with outstanding capabilities in accuracy and coverage will be in place for use by the military. This system can also be used by the civilian sector, as OMEGA, LORAN and TRANSIT have been used.

A difficult issue to address is the establishment of cost-performance goals for airborne navigational equipment. It is not unreasonable for a user to expect a new generation of equipment to provide better performance at a lower cost than the one it replaces, particularly if the major development costs of the system (including some classes of user equipment) are borne by the Defense Department. Performance, however, is a multidimensional quantity which includes not only position determination accuracy and update rate, but coverage, reliability and the efficiency with which the user's interface with the equipment tells him what he wants to know regarding where he is and how he is to get where he wishes to go. Without performance criteria, new systems concepts tend to be compared with existing systems in artificial ways, such as how much better the new one will be with respect to isolated performance measures. What one wants to know, instead, is how much better off the user will be, in a comprehensive and general sense, if he adopts the new system. This viewpoint does not lend itself to numerical precision, but we intend to be guided by it in assessing the various alternatives and their impact on user equipment within the scope of this study.

1.2 Scope of the Study

The present study is focused on user equipment suitable for civil aircraft, and using GPS signals. We are primarily concerned with receiver simplicity, reliability and operational characteristics. No attempt will be made to predict actual dollar costs of equipment alternatives, since predictions of

this kind require in each case a detailed design and production cost breakdown, together with a practiced judgment regarding market realities which we do not pretend to possess. However, we are very much aware that cost to the user in dollars is the real issue, and we intend to address that issue as closely as possible by using engineering judgment in place of the techniques of production cost forecasting. We also believe that this is a reasonable approach to the problem at its present state, since the system alternatives to be considered are sufficiently varied as to cause the resulting receiver/processors to differ fundamentally from one another. At a later stage, when design choices of a much more specific kind are to be made, within the context of a chosen signal and data format, then hard dollar costing efforts would be appropriate. But, in this study we rely heavily on complexity as an indicator of cost, along with a general knowledge of the cost trends of the major component classes which are bound to be represented in the user equipment.

There is need for a study of equipment alternatives for civil aviation users even in the context of the presently planned waveforms and signal parameters. The military themselves envision an entire spectrum of user equipments for these signals. The questions of the relative advantages of multichannel versus sequential receivers and the possible use of the wideband P channel versus exclusive use of the relatively narrowband C/A channel are still open, in the context of minimum-cost receivers meeting civil performance goals. More detailed questions, such as loop parameter optimization and processing algorithms for position determination and satellite selection also merit study.

However, the problem is much broader than this, since it includes the general option of another signal for civil users. No detailed ground rules have been stated regarding such a signal, beyond a general willingness to provide a modest amount of space, weight and power, aboard the satellites, for this purpose. Obviously, such an add-on package must not interfere with the original equipment in any way. Still another option is the modification of one or both of the GPS signals to facilitate the design of low-cost receivers for civil use.

It is, therefore, natural to divide the study into two parts: one devoted to exploring the potential of parameter variations within the framework of the present GPS signal/data structure, and the other devoted to an evaluation of the possibilities of other generic classes of waveform and data structure. The problems of carrier frequency and average power are common to both parts of the study, while consideration of short-term and long-term duty cycles for the satellite transmitter also arise in the second part.

The signal requirements for a one-way ranging system such as GPS are basically a time-tagged ranging waveform and a means of conveying data which allow the user to compute the satellite's position. The time tagging consists of the provision of recognizable points in the waveform together with an inclusion in the data of the times of transmission of each of these points. The present GPS waveforms (P and C/A) are CW, using biphase PN coding for ranging, and synchronous, biphase DPSK modulation for data transmission. The major alternative ranging waveforms, especially pulsed waveforms and CW tone ranging waveforms will be considered. Each ranging waveform can be combined with any of several feasible data modulation techniques to complete a waveform definition. In addition, the data structure can be varied. This is an important consideration since the relevant data remains unchanged for long periods in satellite memory and can be transferred to the user for acquisition of precise and approximate information on the ephemerides of all the satellites, clock anomalies, ionospheric propagation models, and so on.

Besides these alternatives for the signals from individual satellites, there are important alternatives for sharing the available channel among the 24 satellites. Presently, GPS uses code division (CDMA), which combines well with the PN ranging waveform. The other principle alternatives, TDMA and FDMA, will be considered for each of the waveform alternatives.

Thus, the signal structures to be studied can be broken down in a number of ways:

GPS modification or new, additional signal
Carrier frequency and power levels
Ranging waveform used
Data modulation technique employed
Structure of the transmitted data
Nature of the multiple-access technique used.

Each of the many alternatives resulting from this multidimensional breakdown will be examined from the point of view of the implications for the receiver in terms of complexity, as discussed above.

1.3 Methodology

It would be desirable to establish the potential of each of the signal options described above for yielding satisfactory performance using a simple receiver. As mentioned earlier, this performance is not a single number, but a set of qualities, many not amenable to precise numerical measure. Some performance features which can be assessed numerically are:

Accuracy of a single position fix
Update rate of effectively independent fixes
Information data rate
Time to a first fix (TTFF) from start with varying degrees of *a priori* information.

Even these qualities are described by averages and probability distributions because they depend upon many variables, such as satellite geometry and user latitude. Other qualities must be described parametrically, such as:

Probability of loss of ranging data and time to re-acquisition as a function of user dynamics and orientation.

A final class of performance features is qualitative, but no less important, in nature. These include:

> The initialization requirements of the system, i.e., initial loading of current GPS system data (if this is required) and the ease with which this is accomplished.
> The nature of the output information displayed to the user and the ease with which he can obtain navigational guidance from his equipment.

Fortunately, these last two items are largely independent of the signal waveform and, although somewhat dependent on data structure, can to a large extent be discussed as features of any user system based on GPS, as compared to those based on VOR/DME, Loran-C, etc.

Returning to the problem of complexity/performance goals, these are of direct relevance in the case of two of the performance qualities which can be most reliably determined for a system concept, namely positional accuracy and update rate. These requirements vary greatly between the air traffic control extremes of trans-oceanic flight and final terminal approach. Correspondingly, the cost-effectiveness of a navigational system depends more on which of these ATC regimes it serves satisfactorily than whether its positional accuracy is 100 feet or 1000 feet. To relate the two requires the missing ingredient of performance goals.

If every signal option were evaluated in terms of all the performance measures discussed above, a large matrix of numerical and qualitative data would be obtained. Each element of that matrix would be the result of optimizing a receiver design, balancing the various conflicting performance requirements, for that particular signal option. Besides being a task of impractical magnitude, the results would be only as good as the optimization performed, and the subjective choices made in such an optimization.

Instead of such a thorough evaluation we divided the waveforms up into major categories based on the prime function of the GPS waveform, namely, ranging and by the exercise of judgment rather than detailed evaluation rejected those which did not seem competitive. This process led to three major classes of waveforms, namely the pulsed TDMA waveform, the tone ranging TDMA waveform and the present CW class. Rather than consider all possible variations and refinements in each class, a complete system design was performed, based on a large set of specific choices, for each major class. The resultant candidate is then used as a vehicle for performance analysis and in its broad conclusions representing the whole class. If an actual system design choice were to be made one could further investigate the variations in each class for optimization. Such work was considered beyond the scope of this study.

Some issues were found to transcend the choice of waveform modulation. For example, it appears that multiple access by time division is inherently desirable, from the point of view of receiver costs, because one receiver/processor channel can then perform ranging on all visible satellites in turn. A less obvious result is that with ranging data available from all usable satellites, an elaborate satellite selection algorithm can be dispensed with, and superior navigation obtained with a nominal increase in the complexity of the position computation algorithm. Some general conclusions such as this can be made without detailed system designs, but often the basic truths of a subject as complex as this one are perceived only after working out many detailed examples.

Another example may be the role played by multipath. It is hard to conclude much from general statements on multipath, but an assessment of its effects on a few candidate signal/receiver designs may permit some generalizations. It was decided not to pursue the effects of multipath in our detailed designs.

The main results of this study are the selection of a few promising alternate waveforms for which a GPS waveform and receiver design was carried through the initial stages to verify performance claims. It emerges that the class of pulsed TDMA waveforms is promising enough to warrant further development and, perhaps, detailed further work by the civilian agencies in order to exploit the potential of NAVSTAR GPS.

Most of the material presented in this work, and portions not reported here, has been published in the form of Lincoln Laboratory working papers.

1.4 Further Outline of the Report

In Section 2 we discuss the winnowing of the options based mainly on the capability to perform ranging, which is still the most important function to be performed. Out of this discussion emerge signal design concepts qualifying for further evaluation. They are:

a) Section 3, the basic GPS C/A signal structure, but enhanced in power level by a significant amount.

b) Section 4, a TDMA pulsed waveform system

c) Section 5, A TDMA tone ranging system.

d) Section 6, a PRFDMA pulsed waveform system.

In Sections 7,8 and 9 we discuss issues transcending specific waveform choices such as the operational implications of TDMA, handover capability and the role of clock stability in a low-cost receiver.

2. REVIEW OF WAVEFORM CHOICES

2.1 Choice of Waveforms for Ranging

In this section we review quickly the several ranging methods and provide some rationale for having limited our study to a few classes of waveforms for each of which we will design a prototype GPS system.

For purposes of navigation the user is in an unknown position (within bounds) and tries to establish his position relative to several sources at known locations (with the sources often providing the data from which their location can be determined). One method of obtaining relative location is through ranging or range differencing, and then establishing the user position by multilateration.

Some navigation systems, such as INS, do not rely on external signals, but are capable of integrating displacement and, given correct initial conditions, provide user position. We do not consider such systems, not even as ranging aids, but concentrate instead on methods of ranging based on externally transmitted and received signals.

2.2 Basic Types of Ranging Systems

There are many possible system approaches to ranging. First there is the interrogate-reply type, where the user interrogates several sources (e.g. satellites) and deduces two-way range from the delay in the reply. The most rudimentary form of interrogate-reply is the radar system based on an interrogation with RF pulses. If an object reflects a portion of the energy in the pulses, the delay with which the echo is received is proportional to the two-way range. Another form is the cooperative interrogate-reply method such as is used in an Air Traffic Control Beacon system or an IFF system. Several such satellite based navigation systems were studied at Lincoln Laboratory (Ref. 1). One further distinguishes such systems into simplex and duplex systems depending on whether interrogate-reply happens sequentially or simultaneously (using frequency or code orthogonality for example).

The second type of ranging system is one in which the several sources transmit time-marked signals from which the user deduces one-way range. If the source location is unknown or continuously changing as is the case for satellites, the signals must carry sufficient information to make it known. The data must also allow the time markings to be expressed in source-clock time. The latter allows the user to deduce the time of transmission of the signals. The ranging task is complicated by the fact that the transmitted time is expressed in source-clock time and the time of reception in receiver-clock time. If however, one considers the several source-clocks to be synchronized (maybe after adding corrections obtained from the transmitted data) the only error derives from the user clock. Assuming for a moment that all range-readings are taken simultaneously then the user clock error constitutes just one more unknown to be determined together with three position coordinates. A set of four ranges allows the user to obtain a solution. Non-simultaneous readings will introduce additional errors due to user clock imperfections. The GPS C/A and P modes are ranging systems of this second type and this is the only type considered further.

2.3 Satellite Configurations, Frequency Choices, Channel Sharing

The choice of sources leads to such diverse ranging systems as Omega, Loran, Transit and now NAVSTAR GPS. It is beyond the scope of this study to consider new satellite configurations or even the addition of, for example, a geo-stationary satellite to the presently projected GPS configuration. The same holds for frequency choice. No new frequency choices were considered. However, since two frequencies (L_1 = 1575 MHz, L_2 = 1227 MHz) have been allocated, either one or both could be considered for possible waveform change.

Since ranging to several sources is necessary, the problem of channel sharing arises. The C/A signals of GPS are shared via CDMA with a separation as expressed by the peak cross-correlation of the PN codes of 21.6 dB for maximum doppler. Within the class of CDMA waveforms variations on codelength or code generation procedures have been explored (Ref. 2,3) in the context of a less costly GPS receiver (called Spartan). PN encoding in a CDMA setup fulfills simultaneously the role of providing processing gain for the matched filter, and signal separation.

FDMA is generally undesirable if each signal already requires substantial bandwidth for other reasons. If one user receiver must receive all signals, FDMA does not seem amenable to a simple receiver and is not considered further.

TDMA, provided the signal (re-)acquisition problem is kept simple, is the solution resulting in the optimal use of bandwidth and power. In the TDMA systems proposed in this study each satellite in turn is assigned a fixed timeslot. Guardtime of about 18 msec to avoid signal overlap is provided corresponding to the maximum possible range difference (5600 km) of two satellites to a user. TDMA is ideally suited to a one-channel sequential receiver. Section VII in this report highlights its benefits. Obviously during each time-slot a wide variety of waveforms may be transmitted.

We have considered one additional mode of channel sharing, called PRFDMA, where each of the several signals consists of an infinite train of very short pulses with different PRF's (Section 6).

2.4 TDMA Waveform Choices

A choice of waveform is based on a set of criteria, in this case ranging and communication performance, not only in terms of accuracy and bit error rates but also in terms of easy signal (re-)acquisition and synchronization. And finally there is simplicity of implementation of the receiver.

The requirements for data communication are mild; only 50 bps must be reliably transmitted, and even so the same 1500 bit message is repeated about 120 times. Such redundancy could be exploited to an extent to coast through signal fades.

The requirement for ranging accuracy is the most demanding. Ranging is a function performed by radars. All waveforms of interest have been analyzed and synthesized in connection with radar applications. Waveforms are classified according to the nature of their ambiguity function (Ref. 4,5). The ambiguity function is the squared envelope of the matched filter response to the received signal, often represented as a surface with independent variables the mismatch in delay and doppler between the signal and a local replica. Desirable features of the surface are a sharply defined central peak for zero mismatch in delay

and doppler, if range and rate is to be measured, or a sharp ridge extending in the doppler dimension, if only range is to be measured, and very low even sidelobes. At the same time, the behaviour of the surface in the immediate vicinity of the peak or ridge should have, or allow easy generation of, an indication of the direction in which the peak lies, for the purpose of corrective action. We discuss waveforms with these desirable ambiguity function properties in mind.

2.4.1 Constant Carrier Single Pulse

The most basic ranging waveform is the constant carrier single pulse with basic characteristic a time-bandwidth product of unity. The ambiguity function is concentrated in a unity-size cell in delay-doppler space with a negligible portion of the pulse energy in the sidelobes. Sidelobes are determined by the exact pulse shape. Ranging accuracy relates to how well the peak of the ambiguity function can be located. For an ideal square pulse the ambiguity function in the delay dimension is triangular* and would in principle result in perfect accuracy. For a pulse of practical interest, the accuracy is often expressed by a lower bound on the standard deviation of the ranging error

$$\sigma_\tau \geq \frac{1}{2\pi\beta\sqrt{\frac{2E}{N_o}}}$$

where β expresses RMS signal bandwidth. In this expression β is the second central moment of $|S(f)|^2$ where $S(\cdot)$ is the Fourier transform of the modulating signal; N_o is one-sided spectral density of white noise; E is the signal energy. The value of σ_τ approaches more closely the asymptotic lower bound for higher E/N_o.

The above formula is useful for sensitivity analysis. It shows that increasing bandwidth and E/N_o increases accuracy. One way of increasing E would be to chose a longer pulse. However, it can be shown that longer pulses result in decreased bandwidth β, the two effects canceling each other at least in the delay dimension. (Accuracy in the doppler direction, which is not considered here, improves with longer pulses).

*See, for example, p. 486, 488 in Skolnik, "Introduction to Radar Systems".

If it is true, within limits, that a short pulse and a long pulse can give equal ranging performance, one may look at a regular train of short pulses which contains the same energy as the single big pulse and which are coherently integrated. The result is a "bed-of-spikes" ambiguity function with improved delay and doppler resolution capability, but with the possibility of incorrectly relating to the wrong spike. The central spike (which has dimension $\frac{1}{BT}$ as do all spikes) reflects the good properties of a single narrow spike in the delay dimension and those of a fictitious long pulse (as long as the train) in the doppler dimension. Such a train is a first example of waveforms where some form of modulation (here amplitude modulation) results in a large time-bandwidth product.

Other classes of so-called pulse-compression signals with large TB-products have been developed. They comprise the waveform classes with thumbtack, sheared ridge and bed-of-spikes ambiguity functions.

2.4.2 Waveforms with Thumbtack Ambiguity Function*

In this class the large time-bandwidth product TB, resulting in a resolution cell size of 1/TB, is generated by some non-linear phase function which can be removed by a matched filter receiver. Generally Doppler mismatch disturbs this match and therefore the effect of signal compression. In the class of waveforms generating thumbtack ambiguity function there is generally a large potential for interference from high Doppler-delay sidelobes.

Consider first the single-pulse waveform and focus attention on the phase modulation. Whether continuously varying phase variation is considered (resulting, for example, in the subclass of frequency-shift codes), or constant-phase steps (subclasses are phase-shift codes with further division into poly phase-shift codes, or its simplest version, the phase-reversal codes) the differences lie more in the realm of implementation than in the potential for ranging performance as their ambiguity functions are all of the same type, as long as there are no regularities in, for example, for the phase-reversal class, the sequence of phase-reversals. It is easy to see, for example for the phase-reversal codes, that a high time-bandwidth product, TB, means a long code (T) and a high chipping rate (B).

*For illustration, see, for example, Reference 4, p. 133.

As a second case consider a single pulse but amplitude modulated. The amplitude may vary continuously, or in steps. As long as there is no regularity in the way the amplitude varies a thumbtack ambiguity function results. Even if the extreme amplitude variation of on-off is used, but the resulting pulses are irregularly spaced, the irregular time spacing (sometimes called staggering) will avoid doppler ambiguities becoming pronounced as they were for a regular pulse train. Instead of irregular spacing, the pulses could be pseudorandomly coded in phase or frequency with a similar result.

Generally the thumbtack function is the "ideal" ambiguity function for ranging even if the sidelobe levels are high, although less pronounced than is the case in, for example, bed-of-spikes type ambiguity functions.

2.4.3 Waveforms with Sheared Ridge Ambiguity Function*

This is the class principally of linear FM waveforms or stepped-frequency approximations to linear FM, and coherent pulse trains with linear frequency shifting from pulse to pulse. The basic drawback of this class of waveforms is that no independent measurement capability exists in range and doppler. While a frequency sweep is easy to implement, one needs an aid to resolve the range-doppler ambiguity. One such solution is to have an independent doppler reading (for example, from a pulse with no FM modulation) while another solution is to have two FM signals each with different slope. In our study this class was not further considered since such "fixes" would result in more complicated designs than in the case with the other classes discussed here.

2.4.4 Class of Harmonic Signals

This is the general class including harmonic tones and specially constructed square waveforms such as BINOR (Ref. 6). The latter replaces the sinusoids of the tones by geometrically related square waves as far as the generation of the baseband waveform is concerned. They may differ in the choice of RF modulation where BINOR could use DSB with suppressed or reduced carrier, while tones use phase or frequency modulation. The square waves approach is known not to be superior to PN encoding in performance or ease of implementation and was therefore not studied further.

*For illustration, see, for example, Reference 4, p. 237.

The tone ranging approach is interesting because of the properties of the ambiguity function in the delay dimension. Accuracy is determined by the choice of the highest frequency. Ambiguity (repetition of the highest peak where tones are cophasal) is determined by the lowest harmonic tone. In between peaks the ambiguity function, around the highest peak, varies monotonically. This slope is an indication of the direction in which the delay has to be changed in order to achieve the peak. This is in strong contrast with PN-encoding where, outside of the region of alignment of about two chips, no indication exists as to whether to advance or retard the local code replica for better code coincidence. Therefore, for long codelengths and low signal levels, as is the case for the present GPS signal, where matched filtering is impractical at present, one has potentially a serious code acquisition and synchronization problem. This is even more serious for the sequential receiver where the several codes are continuously to be re-synchronized.

The potential of tone ranging to allow fast synchronization was promising enough to warrant a complete waveform design effort (Section 5).

3. USE OF PRESENT C/A WAVEFORM WITH MORE POWER

In this section we summarize the results of a study on the effects on receiver complexity if the present C/A waveform would be retained but the available power level would be increased by a substantial amount (in the range of 5 to 15 dB).

3.1 Context for the Study on Increased Power

A generic GPS receiver is considered and its several functions analyzed. In the analysis the following requirements have been assumed:

The presently proposed C/A GPS waveform must be used.

The receiver must be sequential.

The receiver must have simultaneous ranging and data collecting capability.

The output must be 3D position and velocity with accuracies comparable to those presently advertised for GPS.

The emphasis on the use of more power is not necessarily to provide more accurate ranging but simplified circuitry, but neither will a highly degraded ranging capability be acceptable.

The envisioned user class requires that the receiver be capable of coping with moderate dynamics (1g). Higher dynamics are allowed to reduce performance but not to make the receiver totally fail.

The noise consists of ambient noise plus unintentional radar interference in a 1 MHz band at the L_1 frequency.

The GPS waveform dictates certain receiver functions. This study addresses the performance of these functions when the power level is increased.

Specifically:

The code synchronization function is analyzed. Of concern is the time it takes to synchronize. The sync function provides the ranging capability and allows carrier despreading.

The carrier acquisition and tracking function is analyzed. Again acquisition time is of concern as is the sensitivity of the phase tracking loop to user dynamics. The principal outputs are doppler reading and data demodulation capability.

The data demodulation function is analyzed in terms of error rates as a function of signal-to-noise ratio.

The sequencing function is analyzed. The "dwell-time" per satellite determines user track update rate and to a certain extent, complexity of tracking algorithm. Other related issues are the relationship of bandwidth of the Phase Locked Loop (PLL) to dwell time, and the effect of dwell time on allowable oscillator quality.

3.2 Summary Statements on Impact of More Power on GPS Receiver

In this section we summarize the conclusions of our study. The details of the analysis are in Appendix A.

3.2.1 Code Synchronization

An increase of power by an amount of 10 dB would reduce the synchronization detection time on a given code phase to as low as one msec. That time is particularly suited to the C/A code since it corresponds to a full code length of 1023 chips. The corresponding doppler which can be tolerated is ± 500 Hz. The above calculation is valid for a detection threshold of at least 10 dB over the 1 kHz bandwidth. In normal operation of a sequential receiver with dwell time on each satellite of the order of one second and a relatively cheap Voltage Controlled Oscillator (VCO) of 10^{-7} Hz/Hz, it is expected that the received carrier frequency at the end of a three-second period will still be within the bandwidth, for a user with modest (1g) dynamics. It is also expected that the code phase can be predicted to within 1 or 2 chips. Re-acquisition of the code then will only require at most a few code phase tests and at worst one or two frequency steps. Total re-acquisition time will then require only a few msec.

The situation is entirely different at cold start when no a priori knowledge is available on doppler frequency or code phase. For the maximum doppler uncertainty of 4342 Hz from satellite motion and 1 kHz from user motion the maximum synchronization time could be reduced to 20 seconds for a 10 dB power increase above the present GPS level.

3.2.2 Possible Use of Shorter Codes

If there were any pressing reason to hold down code-acquisition time, or if still more power than an increase of 10 dB were available, one might consider the use of a shorter code and/or incoherent integration to great advantage. But the use of shorter codes will narrow the margin between code cross correlation sidelobes and the code auto correlation function peak. That margin determines multiple code access capability. The -23.9 dB margin (no doppler present) for the present C/A code shrinks to -17.5 dB for a shorter 127 bit Gold code. This margin could become even less due to different doppler affecting the different codes.

3.2.3 Use of Matched Filter for Code Synchronization

The use of shorter codes could bring new technology to bear on the code synchronization problem such as the use of matched filters in the form of charge coupled devices (CCD). The CCD matched filter obviates the need for code-phase search and the delay-locked loop, and realizes the full correlation processing gain. At present, however, the applications of the programmable CCD correlator in advanced engineering developments, are still limited to codelengths of up to 256 chips.

3.2.4 Increased Ranging Accuracy

The code tracking accuracy in terms of the chip splitting capability increases roughly as the square root of the signal to noise. For example, a 10 dB increase in power roughly results in an increased chip splitting capability of a factor of 3.

3.2.5 Increased Performance of Carrier PLL

A trade-off study was performed on increase in performance of the carrier PLL as a function of increased power. The presently considered PLL works with an adjustable bandwidth to allow rapid acquisition on the one hand and sufficiently

accurate phase tracking on the other hand. The degree to which bandwidth can be narrowed depends to a large extent on the dynamics of the user. Sudden accelerations translate into accelerations of phase. The transient phase behavior and the steady-state loop errors then depend on the order of the PLL and in particular on the filter in the PLL. The result of increased SNR is that the phase error due to dynamical effects dominates the noise-induced error more and more. The optimal loop bandwidth becomes larger with increased SNR. Figure 3.1 illustrates, for signal-to-noise ratios at the input of the PLL of 25 to 45 dB, what the optimal choice of bandwidth B_L would be when the phase error components are weighted (weight 2 for steady-state error resulting from acceleration versus weight 1 for standard deviation of the noise induced phase error). The analysis applies to the second order PLL. Larger optimal bandwidths here mean for example that the receiver becomes less sensitive to carrier frequency offsets (either from motion or from clock errors) or, for a given offset, can phase-lock in a shorter time. This is important for the sequential receiver where one switches to a different signal at a rate of maybe once, and possibly many times, a second. Initial acquisition will require, in a cold-start mode, fewer frequency bands to be searched.

3.2.6 Bit Error Rates

An increase in signal-to-noise ratio contributes directly and indirectly to lowering the bit error rate. First it increases the detection energy available after a given integration time. And secondly, the use of a wider bandwidth in the PLL gives substantial protection against cycle slip in a dynamic user environment, thus lowering the occurrence of bit errors.

3.2.7 Shorter Dwell-Time per Satellite

In a sequential receiver higher SNR means a quicker code and carrier acquisition. This feature allows shorter dwell time per satellite to be more efficient in terms of collecting ranging information (but not necessarily for data collection). Shorter dwell time will result in a higher ranging and position update rate. For a given requirement on the accuracy of the smoothed user position a higher update rate allows individual range measurements to be less accurate. This may allow a simpler circuit for delay lock loop tracking

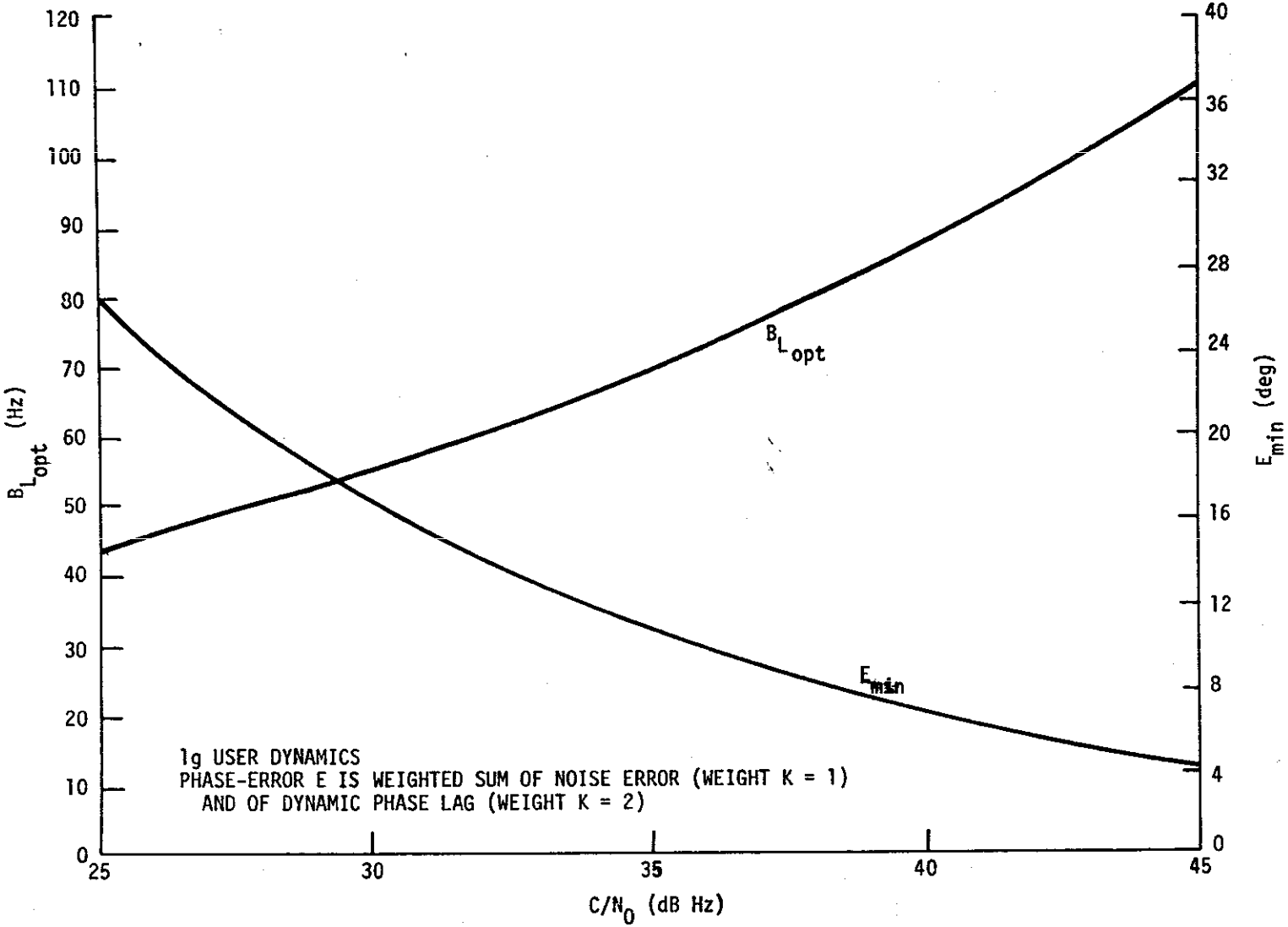


Fig. 3-1. Minimum phase error and optimal loop bandwidth as a function of C/N_o

to be sufficient. In the military GPS receiver implementations, the delay lock discrimination curve is usually obtained from the output of two or more correlator channels. One more correlator channel is used to perform the function of stripping the code from the signal and to provide the input to the Costas loop for carrier tracking and data demodulation. Simpler Delay Locked Loop (DLL) error generating methods, such as the use of tau-dither methods, may eliminate the use of several correlator channels.

3.2.8 Possibility of Using Less Costly VCO

In continuation of the argument made under point 3.2.7, it can be stated that if shorter dwell times can be implemented, the time between "visits" to the same satellite are shorter. Less frequency drift, will then accumulate with respect to the carrier frequency of a particular satellite while other satellite signals are collected. This allows the use of a less stable and consequently cheaper oscillator. For the GA application our designs have aimed at 1 g user dynamics, a 10^{-7} Hz/Hz oscillator and position update rates from 1 to 4 seconds (dwell times .25 to 1 sec.).

These eight points summarize our findings on the effect of a GPS C/A waveform with more power.

4. A TDMA PULSED GPS WAVEFORM

The advantages of TDMA waveforms as a class are rather compelling for a low-cost GPS receiver, and they are discussed in some detail in Section 7. In this section we introduce a particular TDMA example, namely a pulsed system. The fundamental tradeoff between the extremes of pulsed and CW waveforms lies between the cost and difficulty of generating high power pulses on board the satellites and the advantages of simplicity and performance gained for the receiver. In this report we emphasize the advantages for the receiver, which, in our opinion, are sufficiently great to balance the possible difficulties associated with high transmitter power. A brief survey has indicated to us that suitable, space-qualified power generation hardware is or soon will be available, and the interoperability of such a pulsed transmitter with the other functions of the GPS satellite hardware does not seem to present insurmountable difficulties, but this last point has not been analyzed in detail in this study.

Our main intent was to complete the design of a pulsed GPS waveform and derive performance from it which is considered representative for the class of pulsed systems. Although an attempt was made to select optimum choices many refinements are possible.

4.1 Waveform Description

In the system described here, the satellites transmit part of their message using short bursts of high energy pulses. The message portion is preceded by a set of identical pulses from which ranging information is derived. The L-band carrier is spread by a PN code (BPSK modulation) at a 4 MHz chipping rate. The data is PP (Pulse Position) modulated on the spread carrier. The message content is chosen to be the same as for the presently proposed GPS system.

4.1.1 Power Level

The first consideration is the needed power level. The signal carrier power to noise power density ratio (C/N_o) at the receiver front end is practically determined by the required detection energy (corresponding to chosen values for probabilities of detection and false alarm) and the frequency offset one must cope with (which fixes integration time).

For a chosen value of .97 probability of detection and 10^{-6} probability of false alarm, based on coherent integration, the required signal energy to noise power density ratio E_o/N_o is 14 dB. To simplify the particular TDMA pulsed design under consideration, we wanted to operate at a fixed receiver frequency, independently of the doppler present, or the clock error. To maintain an E_o/N_o of 14 dB under the worst frequency offset of 5.5 kHz with the above probabilities the integration time was chosen to be 64 μsec and the power ratio C/N_o to be 58 dBHz. The E_o/N_o for no frequency error is then 16 dB. For a receiver noise figure of 5 dB the required received power is -141 dBw.

Choosing a receiver antenna gain of 0 dB and a satellite antenna gain of 11.5 dB, a path loss of 184.5 dB at the longest user-satellite distance at 5^o elevation, and an atmospheric loss of 1 dB, the required transmitted peak power would be 2 kw. Table 4-1 summarizes the power budget.

4.1.2 Signal Format

We chose a strict TDMA format, allowing only the two satellites in antipodal positions to transmit simultaneously. The total timeframe, of as yet unspecified length, will therefore be divided into 12 slots. Since a time measurement will be obtained from at least four satellites per frame, transmitting in designated slots, the position update rate is determined by the frame length. Update rates should be higher than once per several seconds for adequate tracking performance. An argument can be made that higher update rates allow good tracking through maneuvers or allow the tracking filters to be substantially simpler. Most importantly a high update rate or short frame time avoids the possibility of the short-term clock drift accumulating any substantial timing error. The issue of clock stability in relation to these receivers is the subject of Section 9. We therefore opt for short frames.

In the design we chose a one half second frame length. In a strict TDMA system the timeslot per satellite pair is then 41.6 msec. To avoid overlap between the transmissions of different satellites (closest at zenith and farthest at lowest elevation) a "minimum" guard time of 17 msec is needed. We expanded this to 19.134 msec after other time parameters were chosen. The useful transmission time is divided into preamble and data transmission as shown in Fig. 4-1.

TABLE 4-1. POWER BUDGET

Transmitted Peak Power	+ 33 dBw (2 kw)
Satellite Antenna Gain	+ 11.5 dB
Path Loss on frequency L_1 on Longest Path (25824 km)	- 184.5 dB
Atmospheric Absorption	-1 dB
Receiver Antenna Gain	0 dB
Received Peak Power	141 below 1w or -141 dBw
Receiver Noise Figure of 5 dB	199 dB below 1w per Hz or -199 dB-joule
Ratio of Received Peak Power to Noise Power Density	58 dBHz
Per Pulse of 64 μsec	-42 dB sec
Ratio of Received Signal Energy to Noise Power Density	14 to 16 dB (up to 5.5 kHz frequency offset allowed).

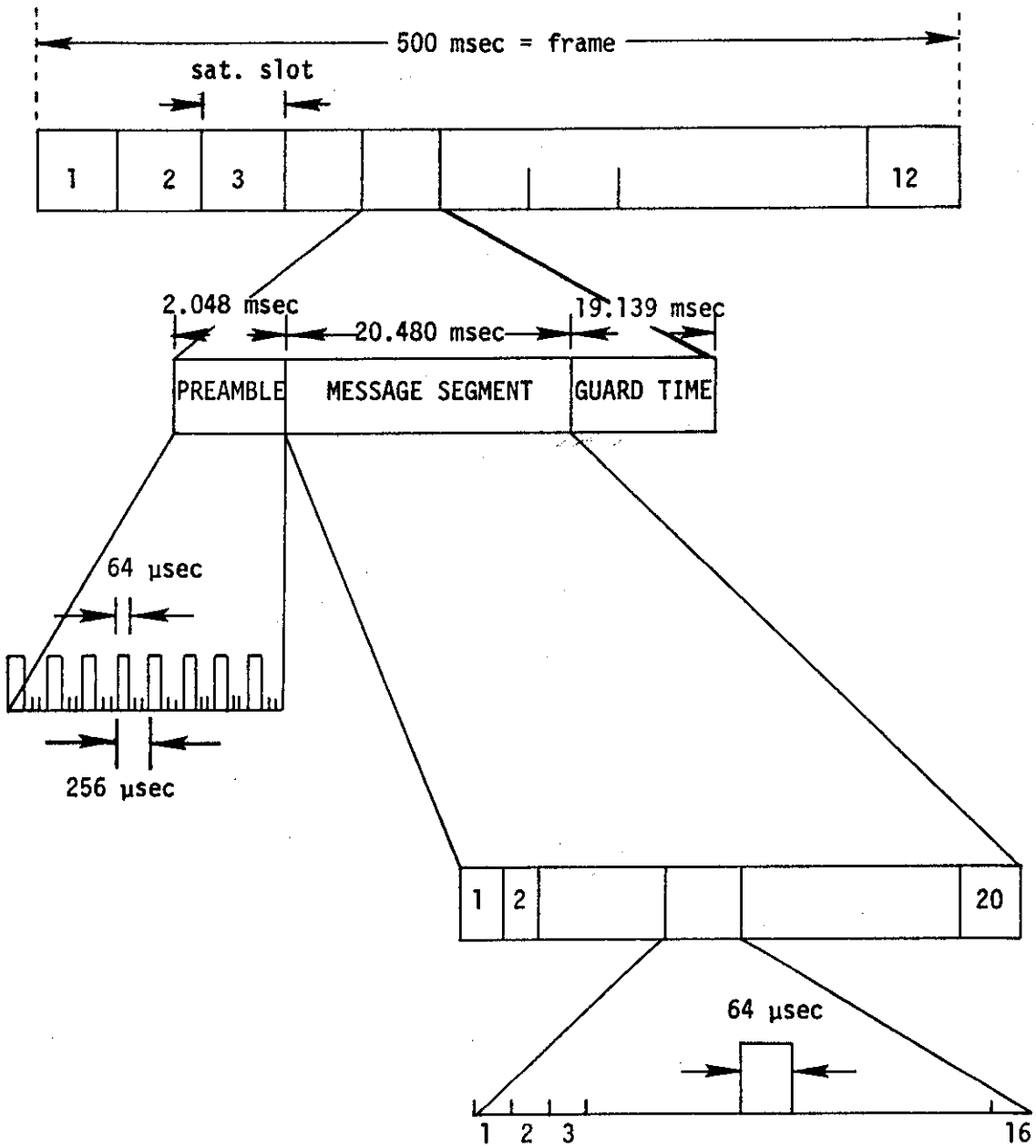


Fig. 4-1. Pulsed signal format.

The preamble consists of eight pulses, allowing synchronization with the PN code and accurate timing (ranging). The message portion consists of twenty pulses, each pulse conveying four data bits (see Section 4.1.5 on modulation).

4.1.3 Waveform

All pulses are 64 μsec long and identical. The L_1 carrier is BPSK modulated by a code 256 chips long, at a 4 MHz chipping rate. The higher chipping rate was chosen over the 1 MHz rate for the present C/A CPS signal in order to reduce the receiver's sensitivity to multipath by a factor of four. For a 4 MHz rate the altitude above which no multipath effects from a flat surface are felt reduces to about 2,000 ft. for a minimum elevation angle of the satellite of 5^{o}.

Since the signals of the 12 satellite pairs are strictly time multiplexed, a common single code can be used.

4.1.4 Ranging and Synchronization

As shown before, the preamble consists of 8 pulses. We consider here the issue of detection, synchronization and ranging.

Table 4-1 showed the SNR per pulse at the output of the matched filter to be 14 dB with 2 dB margin. This establishes a probability of detection (incoherent phase) of $P_D = 97\%$ and a probability of false alarm of 10^{-6}.

The 2 dB margin is used against the drop in SNR due to doppler mismatch. It can be easily shown that for a worst doppler of ± 5500 Hz the power falloff of the autocorrelation function is 1.8 dB for a 64 μsec integration time.

The range accuracy in terms of chip-splitting capability is

$$\sigma_\tau = \Delta \frac{1}{2\sqrt{E/N_o}} \leq 0.1\Delta$$

where σ_τ is the standard deviation of the TOA error from receiver noise only.

The sampling noise, at 8 MHz sampling rate, adds

$$\sigma_s = \Delta \frac{1}{2\sqrt{12}} .$$

The basic accuracy per pulse, for Δ = 250 nsec is then

$$\sigma_{R_1} = \sqrt{\sigma_\tau^2 + \sigma_s^2} \simeq 44 \text{ nsec.}$$

However, ranging will be based on the timing of 8 pulses resulting in

$$\sigma_R = \frac{1}{\sqrt{8}} \sigma_{R_1} \simeq 15.5 \text{ nsec or ft} ,$$

which is the basic ranging accuracy. Obviously other error sources add to this error. One is the errors n satellite position which is implicitly transmitted in the data and which can be calculated by the user. These " ephemeris errors" are estimated to be smaller than 5 feet. The main reason for keeping these errors so small is that the identical data are used in the P-mode with basic ranging accuracies claimed better than 5 ft and all other errors must therefore be kept commensurately small. Errors from tropospheric delays and satellite clocks are of the order of a few feet. Errors from multipath can be a substantial fraction of a chip. For this reason the higher chipping rate of 4 MHz was chosen. For this reason the altitude floor above which a user would not be affected by multipath is reduced from 8000 ft to about 2200 ft. The principal potential error contribution is from uncorrected ionospheric delay. Since a simple sequential GPS receiver probably does not use the dual frequency correction method one must rely on a model. Several such models have been proposed. It was estimated that with the use of a model statistically 50% of the error remains. One redeeming feature is that the error is strongly correlated in time, thus acting more like a temporary bias rather than a random error. The resulting user position will be slightly in error but in a consistent way.

The model likely to be used is similar to the General Dynamics "Semi-Static Crossline Ratio Technique." The technique is based on two tables stored in the user's processor and which correspond to the two seasonal extremes. Each table gives vertical group delay for 5 degree increments around the magnetic equator. For every 5 degrees increment in latitude away from the equator it provides the ratio of the vertical delay as related to the corresponding delay at the magnetic equator. Satellite data provide three coefficients for seasonal interpolation. The delay at any angle different from vertical can then be calculated.

4.1.5 Data Modulation

With a waveform based upon a single coded pulse, it is natural to think of pulse-position modulation (PPM) for data transmission because of its economy of power. In other words, a given bit error rate (BER) can be achieved with less signal-to-noise ratio than higher duty cycle waveforms. The message block is at most 22.5 msec in length, since the preamble requires 2.048 msec and the minimum dead time is 17.09 msec. This message block is roughly 350 pulse lengths in duration, and the only problem is to divide this interval into symbols, each of which conveys K bits by placing each pulse in one of $M = 2^K$ positions. Then $MN \simeq 350$, where N is the number of symbols per message block.

Pulse position modulation is a special case of M-ary signaling with orthogonal waveforms. In the incoherent case (appropriate here) the probability of a symbol error (PSE) is given by the expression

$$PSE(M) = \frac{1}{M} e^{\frac{-a^2}{2}} \sum_{m=2}^{M} \binom{M}{m} (-1)^m e^{\frac{a^2}{2m}} ,$$

where M is the number of pulse positions and

$$\frac{a^2}{2} = \frac{E_p}{N_o} = \frac{\text{energy per pulse}}{\text{noise power density}} .$$

A symbol error causes a number of bit errors and the probability that an arbitrary bit is in error, is given by

$$BER = \frac{1}{2} \frac{M}{M-1} PSE .$$

Allowing a 2 dB margin, our signal-to-noise ratio is 14 dB, and for this value of E_p/N_o the BER varies with M as shown in the following table:

K	M	BER
2	4	3.48×10^{-6}
3	8	6.86×10^{-6}
4	16	1.33×10^{-5} .

The larger the value of M, the lower the duty cycle of the waveform, but it is desirable to keep the BER on the order of 10^{-5} or lower, hence M=16 is the logical choice. Actually, since bit errors occur in groups with PPM, the BER as defined above is somewhat misleading. A better criterion would be expressed in terms of mean time between bit errors, i.e., the expected length of a run of correct bits. With M=16 and E_p/N_o = 14 dB, the symbol error probability is PSE = 2.49×10^{-5}, and it is easy to see that this yields the same error-free run length as an equivalent code with independent bit errors and a BER of 6.23×10^{-6}. Hence four bits per symbol is an adequate choice.

Since a symbol is now 1.024 msec long, the message block can hold no more than 21 symbols, and a reasonable value is 20 symbols, which makes the message block 20.48 msec long and leaves 19.14 msec for the guard time. The design then provides 80 bits per frame, which can be broken into 5 bits for satellite ID and 75 bits for data, for an overall data rate of 150 bps. Table 4-2. summarizes the data demodulation aspect of the pulsed design. Figure 4-2 illustrates for the chosen 16-ary PP modulation how the symbol error rates depend on SNR. The system design provides 14 dB SNR with a 2 dB margin. Even if the margin is exhausted the symbol error performance is quite satisfactory.

TABLE 4-2. DATA DEMODULATION SUMMARY

Data Message Length:	1500 bits
Data Rate:	80 bits/frame of .5 sec with
-	5 bits for ID and
-	75 bits of data
At 14 dB SNR	
- probability of symbol error:	$2.49\ 10^{-5}$
- Average length of error-free run:	$4.02\ 10^{4}$ symbols $= 1.61\ 10^{5}$ bits = 107 messages = 17.8 minutes
Bit error rate for independent errors:	$6.23\ 10^{-6}$ per sec

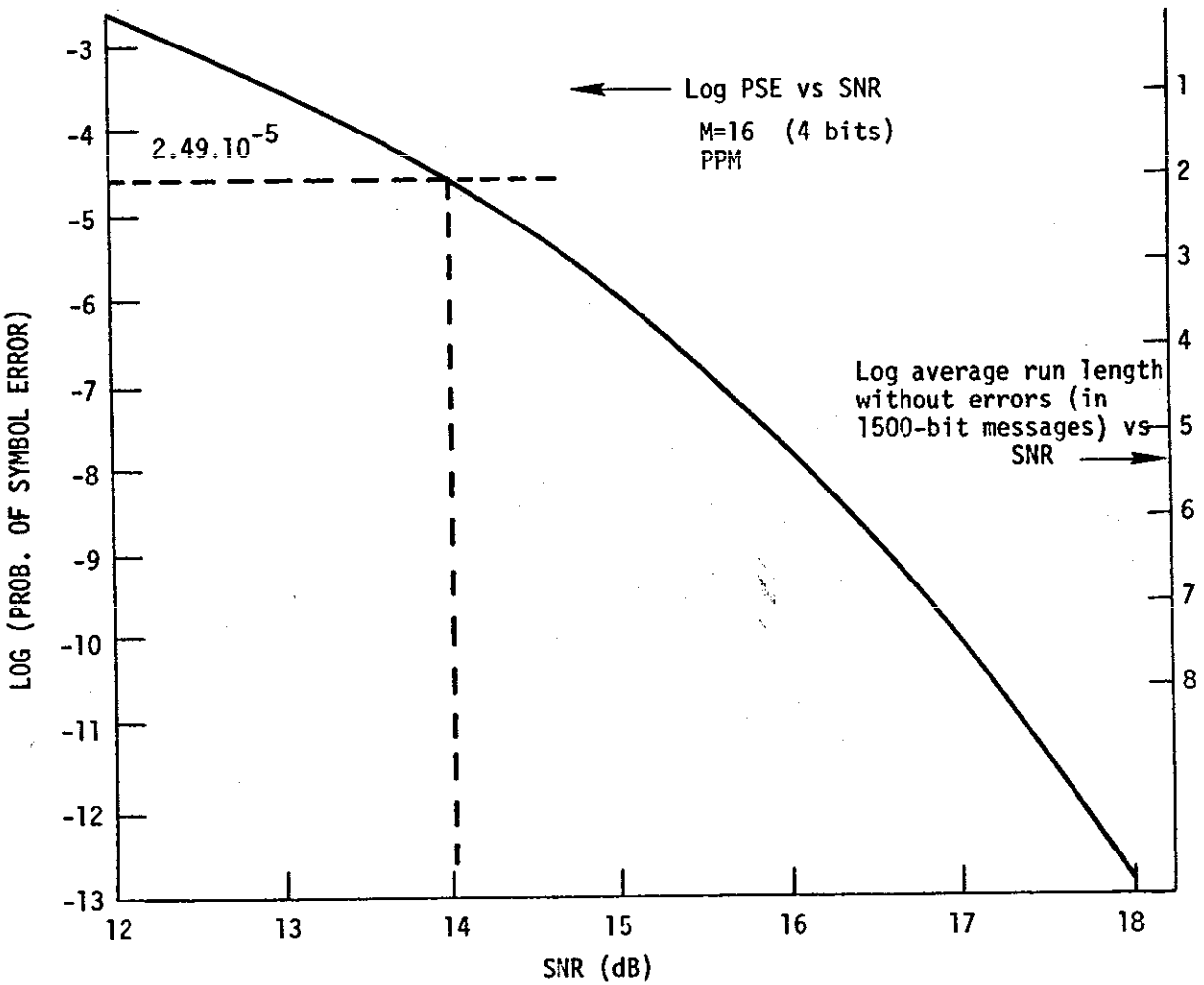


Fig. 4-2. PSE and average message runlength without error versus SNR for 16-ary PPM.

4.2 Receiver Design

The following section summarizes the results of a detailed look at the implementation aspects such as choice of hardware, size and to a limited extent cost for the TDMA waveform using PN encoded pulses.

The hardware implementation sketched out is only one of many approaches and is not intended as the model of future hardware but rather as a representation of one possibility for purposes of sizing (and to an extent costing as well).

4.2.1 Receiver Hardware Design

The overall receiver block diagram is shown in Fig. 4-3. The receiving system consists of three sections; the IF/Demodulator section, the Special-Purpose Logic, and the Microcomputer. The IF/Demodulator filters and amplifies incoming RF signals, heterodynes the desired signal to baseband, demodulates the PN coding in a correlator-pair which provides integration over the pulse duration, and converts the resultant correlator output into a digital output sample. The correlator output is also fed to the sync detection circuitry.

The special-purpose logic provides three functions; sync detection, time-of-arrival determination, and data demodulation. This information is fed to the microcomputer.

The microcomputer performs the data verification, selection of range samples to be used, and determination of position. As noted in the introduction, this study did not consider the computational aspects of position determination; that work remains to be done at a later date.

A block diagram of the IF/Demodulation circuitry prior to the correlators is shown in Fig. 4-4. The design approach is straightforward. The most difficult item is probably the relatively narrow-band filter used on the front end to reject interfering signals.

The baseband correlator, combining, and analog-to-digital conversion circuitry is essentially the same as that used in the development of the JRSVC radio*. The direct applicability of that work would reduce the effort required to configure the GPS receiver.

A block diagram of the special-purpose logic is shown in Fig. 4-5.

*Work presently underway at Lincoln Laboratory.

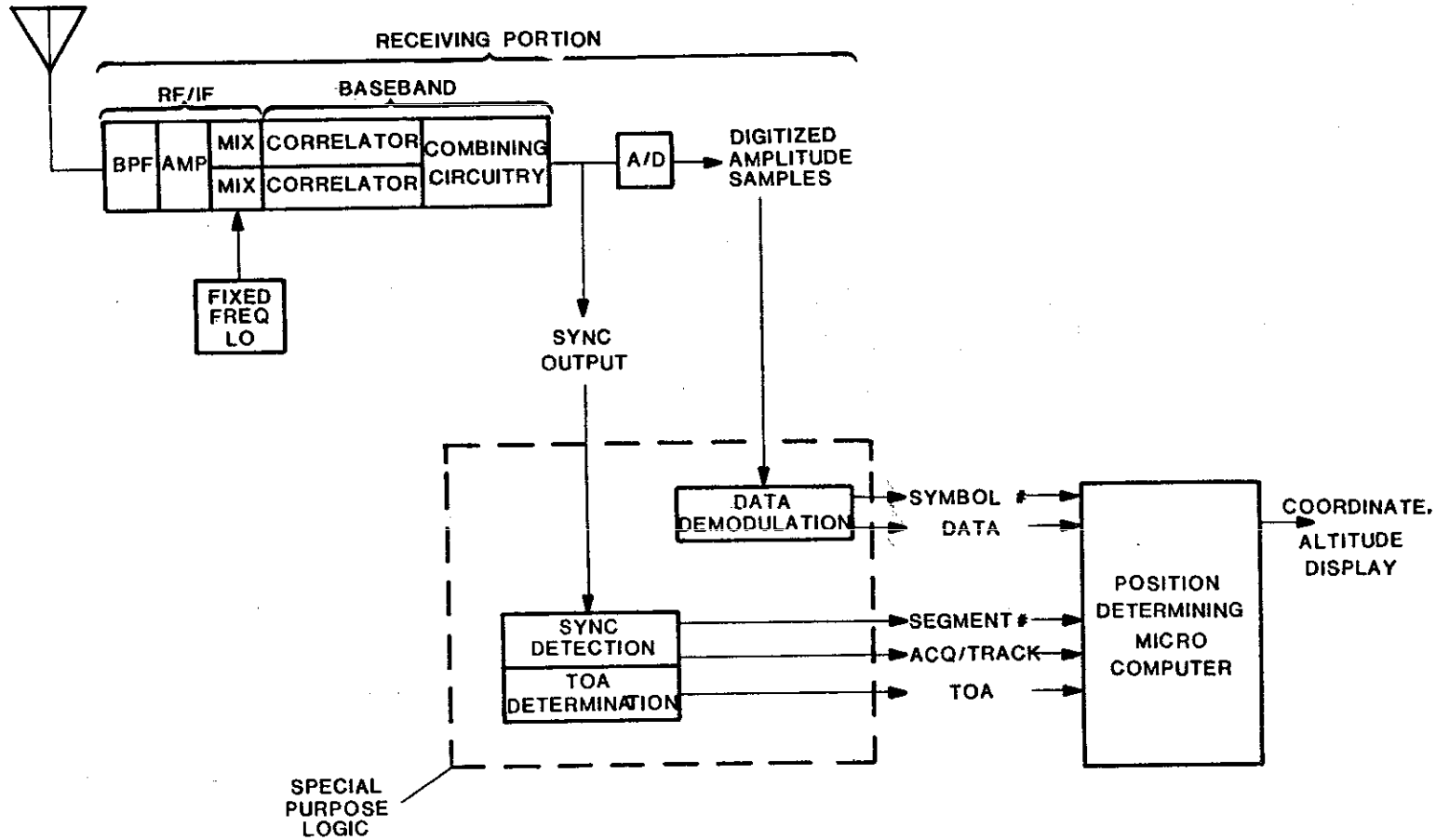


Fig. 4-3. Overall block diagram.

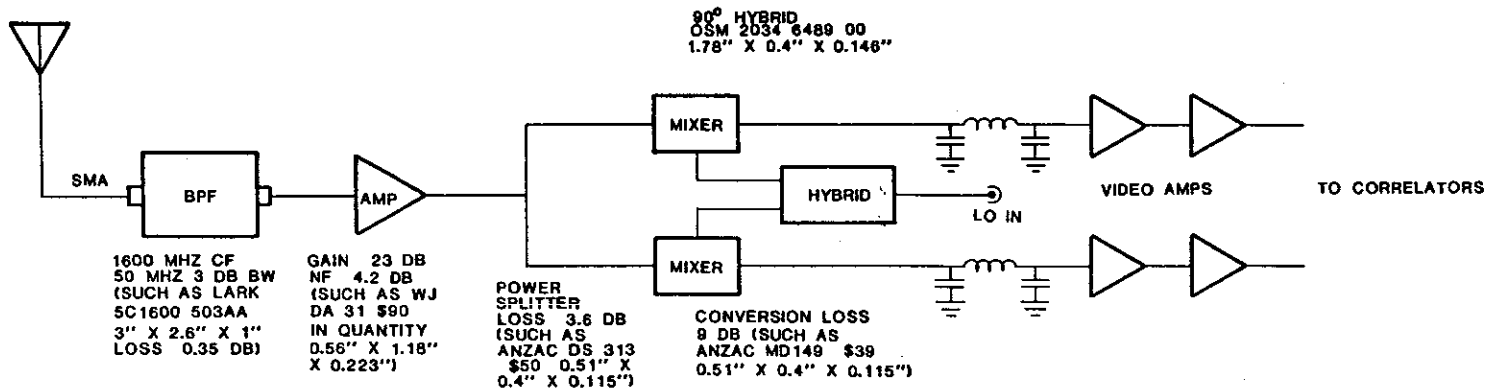


Fig. 4-4. RF/IF portion.

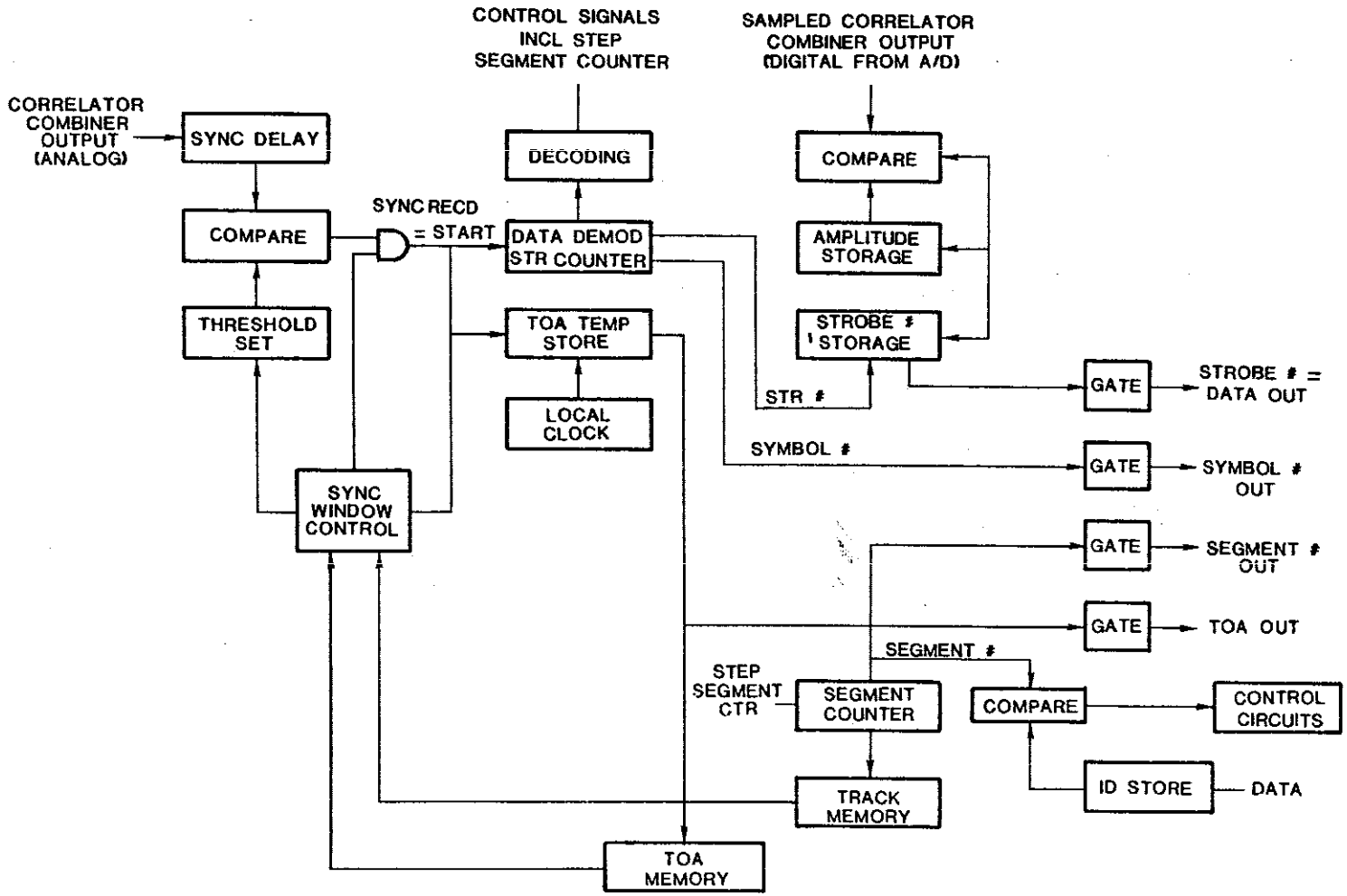


Fig. 4-5. Special purpose logic.

The analog output of the correlator combining circuitry is fed into the sync delay circuit, where it is thresholded and shifted into a shift register. As signals propagate through the shift register, a point in time will occur when sync pulses appear at all the taps of the register simultaneously, and this is the instant at which receipt of sync is recognized. The summed output of the register taps is compared to one of two possible values, and an output declared when the threshold is exceeded. The reason for the use of two threshold values is to allow signals to be acquired only when all eight sync pulses are detected, but to remain in track with less than eight detections. This reduces the probability that the signal will fall from track merely because one sync pulse was missed.

The detected sync starts the data demodulation counter which controls the demodulation of the data, and also loads the time-of-arrival register which stores the time at which the sync arrived. The data demodulation counter causes strobe signals to be generated which sample the A/D output at the sixteen possible data correlation times and determine the largest, or most probable, position of the data pulse. This position, which yields the data, is stored and subsequently read out to the computer.*

During initial acquisition the first sync signal received starts data demodulation and the ID signal recovered from the data is loaded into the segment counter to initialize the receiver with respect to the transmit pattern. IDs of subsequent signals should then agree with the segment counter reading at the time they are received; if not, a misalignment is present, and acquisition is reinitiated. Refinements of the alignment verification can readily be incorporated if desired.

Data is fed out from the special-purpose logic to the microcomputer at the end of each data symbol time. The microcomputer performs data manipulation and verification, as well as processing of time-of-arrival information to determine position.

*Detailed diagrams on sync processing logic resulting in logic cards labeled numbers 1 to 3 are included as Appendix B.

4.2.2 Receiver Packaging and Sizing

In order to arrive at some estimate of receiver complexity and size it was necessary to generate a first pass at the design and then estimate what it implies in terms of a receiver package. The receiver front end and special-purpose logic were considered in some detail, and drawings generated are included in the Appendix. Other portions were estimated based on similar work on the JRSVC radio or other projects at Lincoln Laboratory.

The receiver packaging as presently conceived is shown in Fig. 4-6. The "receiving" portion of the package, containing circuitry prior to the microcomputer is approximately 4.75 x 10 x 3 = 142.5 cubic inches. In the overall package as shown provision was made for three microcomputer boards. Depending on the speed of computation required, and the amount of memory required, this may not be sufficient. If not, the package would obviously grow beyond the 4.75 x 10 x 4.25 size (203 $in.^3$) shown in Fig. 4-6.

A rule of thumb for weight of equipment of this type is 25 cubic inches per pound, which would lead to a receiver weight estimate of approximately 6 pounds.

In summary, this initial broad-brush look at the design of a receiver for use with the proposed Pulsed GPS system indicates that adoption of a system of this type could lead to implementation of navigation receivers for the civilian market at moderate prices.

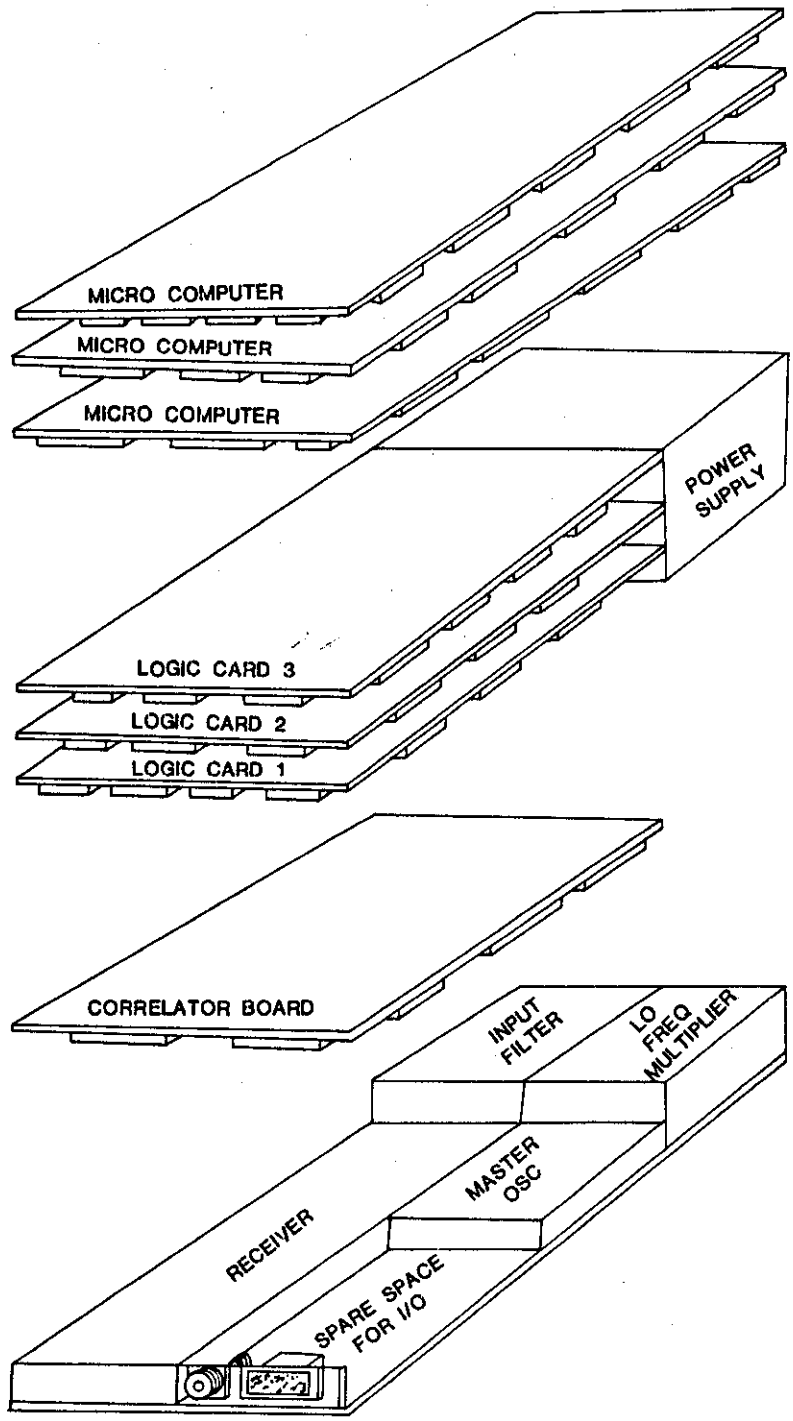


Fig. 4-6. Packaging of TDMA pulsed GPS receiver.

4.3 Conclusions on TDMA Pulsed Waveform Design

We have examined, out of the TDMA waveforms, the class of pulsed signals and composed a signal structure which, together with the emerging CCD technology, has led us to a design resulting in a simple receiver. The main advantages for the hardware are the absence of any DLL or PLL. Also the pulses are short enough to be insensitive to several kHz of frequency offset either due to relative user-satellite motion (Doppler) or oscillator instability.

Certain features such as the 4 MHz signal bandwidth, which is beneficial in obtaining high ranging accuracy, were chosen more for protection against multipath. The ceiling of approximately 8600 ft. above which there is no marked effect from ground related multipath for the GPS C/A signal is reduced to 2150 ft. So while we increased the bandwidth by a factor of four, it is still less than half that of that allotted to the GPS P-mode on the same frequency. At the same time, an average transmitted power of about one third that of the C/A mode is needed to provide very satisfactory performance for ranging and data transmission. The data rate provided in the design was three times the present C/A GPS signal data rate mainly to reduce the time-to-first-fix (TTFF).

The PN-code chosen was only 255 chips long compared to 1023 for the C/A GPS signal. The latter has interference from other GPS satellite signals to deal with (cross correlation sidelobes) while in the candidate design there is only one signal at the time. The number of chips in a pulse, whose length was determined by required energy per pulse for detection and the desire to be insensitive to the maximum occurring doppler, was then more a question of accurately determining TOA of the pulse in noise. The above limitation due to doppler and a low 14 dB SNR per pulse has been improved upon in the design by using a pulse burst (here 8) to synchronize for data demodulation and obtaining an improved measurement of pseudorange.

Finally, pseudorange readings on all visible satellites are obtained within .5 seconds. This provides the capability of varying user tracker update rates, depending on the computational capabilities provided. For example, a one second update rate may be desirable during maneuvers, while a

four second update interval may be quite acceptable in normal flight. This capability in turn may reduce the complexity of the user tracker. The .5 second pseudorange update rate allows several such measurements to be made per user tracker update, providing the capability of extrapolating (or interpolating) the measurements of the several satellites to a common time. This feature compensates for the loss of pseudorange-rate measurement available to the C/A GPS receiver with a phase-locked loop. The interpolation process is discussed in detail in Section 9.

TABLE 4-3. SUMMARY OF FEATURES AND PERFORMANCE OF TDMA PULSED GPS DESIGN

Power Levels	
Transmitted Average Power	7.2 w
Transmitted Peak Power	2 kw
Transmission Format	8 pulse burst over 2.048 msec. (25% short term duty cycle), 20 pulses spread over 20.48 msec.(6.25% short term duty cycle). Deadtime of 477 msec. (.35% overall duty cycle).
Power at Receiving Antenna	-141.25 dBw
Received C/N_o (5 dB noise figure)	56 dB-Hz
E/N_o per pulse	14 dB
Signal Design Features	
Bandwidth at baseband	4 MHz
Ranging Capability Provided by	sync on PN encoded pulse burst
Data Carrying Method	16-ary PP-Modulated pulses
Data Rate	150 bps
Data Repetition Period for user	10 sec
Receiver Features	
Ranging	CCD matched filter and special sync logic
Carrier Tracking Loop	Absent
Data Demodulation	PPM demodulation logic
Start-up Circuitry	Absent
Cold Start Capability	Superior to capability of C/A waveform
Range Update Rate	2 Hz simultaneously for all visible satellites
Measurement Type	Range
Sensitivity to Multipath	None above 2200 ft originating from ground reflections

5. A TDMA TONE RANGING GPS WAVEFORM

As pointed out in Section 2, Tone Ranging, based on the properties of the autocorrelation function, has the potential to be a very good ranging waveform if sufficient range ambiguity resolution is present. The latter is provided by using several tones. The tones can be chosen to be clustered to maximize bandwidth utilization. Tone ranging is vulnerable to jamming. However, in this study, jamming resistance or security were not considered.

Any alternative to the present continuous PN encoded waveform must provide ranging and data carrying capabilities. In this section the ranging capability is provided by a set of tones frequency modulating a carrier. Data are efficiently transmitted in separate time intervals, BPSK modulating the same carrier.

Tone ranging is an established technique (Ref. 6). A well-known system is GRARR (Goddard Range and Range Rate System) built in the mid-sixties. The system used a set of harmonic sidetones for tracking spacecraft in near-earth orbits and a combination of tones and PN coded signals for cislunar and translunar distances. Studies on tone ranging have also been performed in support of the AEROSAT program (Ref. 7).

Several things have been learned from these studies. We incorporate in our design some features without further motivation. The first one is the use of clustered tones instead of the fundamental tones needed for range ambiguity resolution. The second feature is the method for carrying data on a separate subcarrier or on the basic carrier itself in a time multiplexed mode. A third one is the choice of modulation (PM or FM). The constant signal envelope, in that case, allows the output amplifiers to be operating at their saturated level resulting in maximum output and efficiency. Although exponential modulation is used at transmission, no complementary demodulation can be used. The reason is that for the typical wide bandwidth of such modulations the received signal to noise levels (below 0 dB) do not allow regular FM or PM demodulation to be effective. Instead, a direct filtering of the sidetones, either at IF or even better (3 dB), at baseband, is used.

The chief difference between the GPS tone ranging problem and previous tone ranging systems is that not one but twenty-four radiating sources are involved. Each satellite must use basically the same set of fundamental tones. If all satellites use the same carrier frequency L_1 as well, a serious interference problem will exist. Indeed the FM generated RF spectrum consisting of the tones, harmonics and intertone harmonics is very wide, the width being determined by the modulation index. The only way continuous signals would not interfere would be when the corresponding satellites are not simultaneously visible to the user or the difference in doppler is sufficient to separate the spectral lines for the receiver.

A solution to this problem is to use a TDMA waveform where only two satellites (in antipodal positions) transmit simultaneously. Each pair of satellites then transmits in one of twelve assigned timeslots. The twelve timeslots make up the basic cycle time which will determine the position update rate for the user.

Another solution would be to have each pair of satellites transmit on a different carrier (FDMA). Since the useful spectrum consists only of a few lines it may be possible to "pack" the spectrum with 12 different carriers and their associated tones. Care must be taken that each carrier with its tones is allowed ± 5 kHz deviation for doppler.

In this study we do not pursue further the FDMA option since it does not seem amenable to making the receiver cheap. On the other hand, TDMA has some basic advantages as discussed in Section 7.

The design of a TDMA Tone Ranging prototype GPS system is optimized for the following general parameters:

The carrier frequency is L_1 = 1575 MHz.

User dynamics are generally limited to one g. One g dynamics will be assumed when optimizing system parameters such as, for example, loop bandwidth.

Ranging and data collecting capability must be present virtually simultaneously, i.e., both functions are exercised during the dwell-time on a single same satellite.

Basic data rates and basic ranging accuracies are the same as for the C/A GPS mode.

Track update rates must be higher than once per ten seconds but lower than once per second. Indeed, from practical experience with tracking civilian aircraft, update intervals of over ten seconds seem excessive and intervals smaller than one second do not greatly contribute to better performance. We will aim generally at a 2 to 4 second update interval.

In the remainder of this section we give a synoptic description of the prototype GPS tone ranging system, followed by a rationalization of some important parameter choices. Although it is concluded that a very adequate GPS tone ranging system could be designed, it is not concluded that it promises great simplicity for the receiver unless real breakthroughs occur in reducing the cost of phase locked loops.

5.1 Tone Ranging System Description

In this section we summarize the salient features of the GPS tone ranging system design: the power budget, the chosen tones, the expected ranging performance, the particular choice of waveform.

Table 5-1 shows the power budget for the GPS tone ranging system. The average power roughly equals that of the GPS C/A mode. The TDMA format with twelve time slots per frame is used, and a peak radiated power twelve times that of the C/A waveform is postulated. This is 5 dB less than the pulse system peak power.

Table 5-2 gives the clustered ranging tones. Table 5-3 summarizes the expected range error budget. The estimated ranging error, due to noise only, derived from the tones is 28 ft. The resolvable range ambiguity is 39.6 nmi. The total ranging error depends largely on that error and atmospheric delays and multipath. The latter is very situation-dependent and has not been studied further. To obtain a crude estimate of user position error, the ranging error must be multiplied by the GDOP. In practice a factor larger than GDOP should be applied since the range measurements are non-simultaneous. The tracker (for example, an adaptive Kalman-type filter) may on the other hand reduce the resulting error estimate again by a factor typically of two or three, depending on the user motion and on the validity of the tracker equations modeling of the real world.

The principal part of the design is the selection of the particular waveform. It seemed necessary to adopt a waveform at least similar to the one shown in Fig. 5-1 where, at the beginning of each slot, the pure carrier is transmitted, followed by a time interval where tones are transmitted, then followed by an interval of data transmission. The remainder of the timeslot is a guard time interval protecting against overlapping transmissions of satellites at different ranges. The advantage of a short "pure carrier" interval is that a variable bandwidth PLL can quickly acquire and lock onto the carrier. From the doppler on the carrier the precise frequencies of the several tones can then be predicted so that lock-on for the tone PLL at baseband can occur very fast. The length of the time interval for tones is then mostly determined by whether the tones are acquired in series or in parallel. For the design outlined here the probability of bit-error is less than $1.1\ 10^{-6}$.

Figure 5-2 shows a block diagram of a GPS tone ranging receiver the parameters of which are discussed in Section 5-2 and Appendix C.

TABLE 5-1. GPS TONE RANGING POWER BUDGET

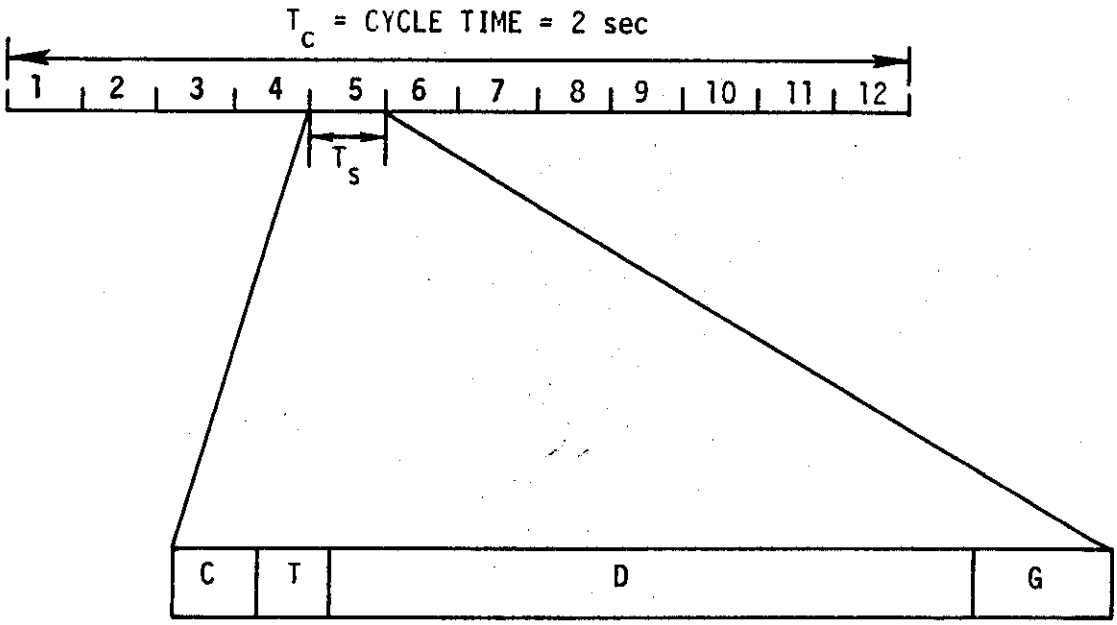
Radiated Power (25 w average; 12 x 25 w in TDMA) 300 w peak	25.0 dBw
Satellite Antenna Gain	11.5 dB
Path Loss at L_1 = 1575 MHz; distance 28524 km.	-184.5 dB
Atmospheric Loss	-1.0 dB
Receiver Losses (incl. polarization loss)	-2.0 dB
Receiver Antenna Gain	0.0 dB
Received Power	-151.0 dBw
Thermal Noise (kT)	-204.0 dB/Hz
Noise Figure	5.0 dB
	-199.0 dBw/Hz
Received C/N_o	48.0 dBHz
Margin	3.0 dB
Design Value for C/N_o	45.0 dBHz

TABLE 5-2. BASIC TONES, f_n, AND TONES FOR ACTUALLY MODULATING THE CARRIER: f_n'

	f_n	f_n'
n = 1	511.3 kHz	511.3 kHz
2	102.26 kHz	409.04 kHz
3	20.452 kHz	490.85 kHz
4	4.0904 kHz	507.21 kHz

TABLE 5-3. ERROR BUDGET OF GPS TONE RANGING SYSTEM

Basic Ranging Error σ_R		28 ft
Satellite Ephemeris		5 ft
Atmospheric Delay		5-50 ft
	RSS	29-58 ft
Multipath		(variable)



C: 20 msec, CARRIER ONLY

T: 20 msec, CARRIER FM MODULATED BY TONES

D: 109.7 msec, CARRIER DPSK MODULATED BY DATA

G: 17 msec, GUARD TIME

Fig.5-1. Tone ranging GPS signal format.

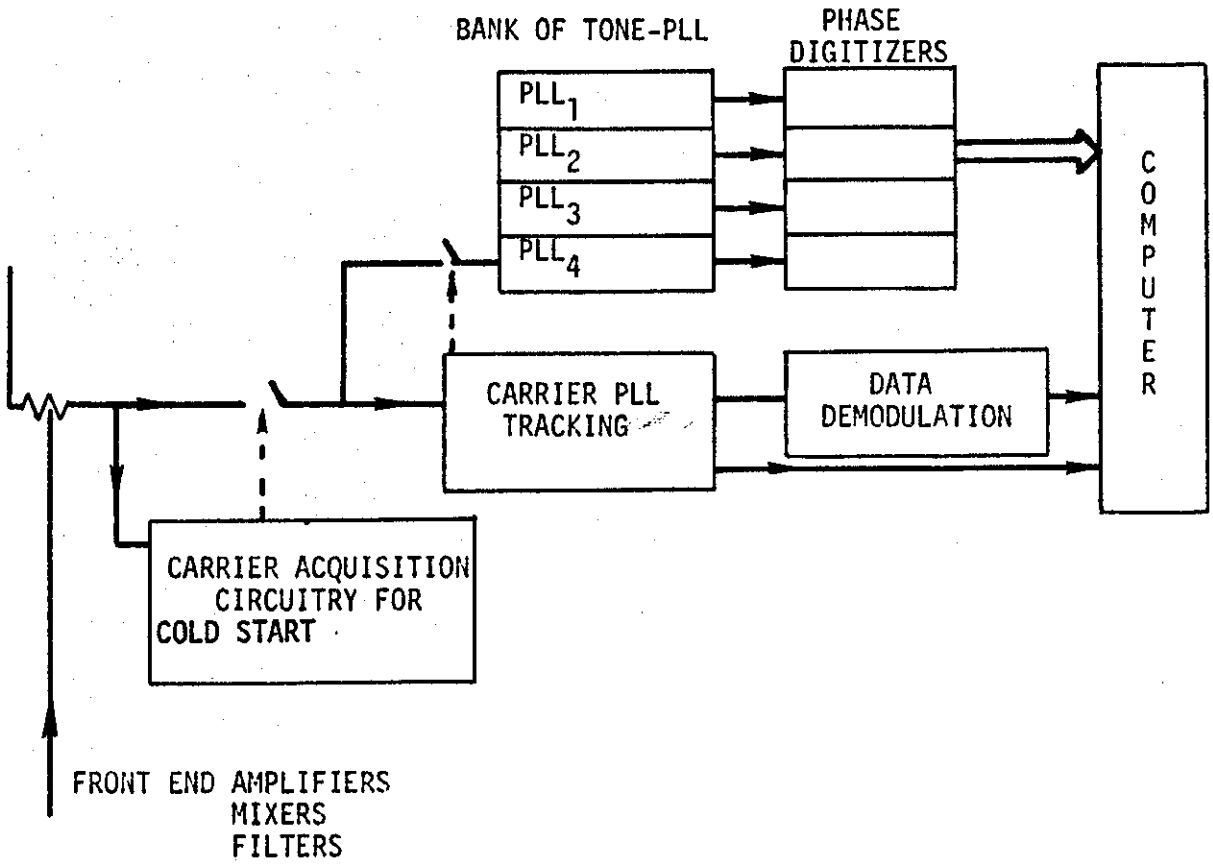


Fig. 5-2. Block diagram of GPS tone ranging receiver.

5.2 Analysis Defining the Main System Parameters

The main aspects of the system design are (1) the choice of the waveform, i.e., the determination of the several time intervals shown in Fig. 5.1; (2) choice of the tones and (3) choice of RF power in the tones or equivalently the choice of FM modulation indices for the tones.

5.2.1 Signal Format

A signal format well suited to the multitone ranging problem consists of a period of pure carrier, followed by a period of carrier FM modulated by the several sidetones, followed by a long period of carrier DPSK modulated by binary data. (See Fig. 5.1). The length of the various time intervals was chosen to allow:

acquisition from absolute cold start, i.e., when no a priori data on satellite whereabouts are available to the user;

re-acquisition within the time interval of pure carrier after every eleven timeslots of silence, coping with frequency drift caused by relative user-satellite motion and oscillator instabilities;

acquisition of the several tones and settling of the tone PLL's and read-out of the individual phases;

conveying a sufficient data rate.

This section discusses further the choice of basic cycle time T_c and the length of time blocks C, T, D and G of Fig. 5.1.

The cycle time T_c determines the rate at which the user position is updated by a range measurement from the same satellite. We want T_c to be on the order of two to four seconds. Also, T_c determines T_s, the timeslot for each satellite since $T_c = 12\ T_s$. The guardtime G is easily determined to be 17 msec, the time difference between the range to a satellite at 5^o elevation and at the zenith.

5.2.1.1 Determination of the Time Interval for Radiating Pure Carrier

The time interval C, during which pure carrier is radiated with a received C/N_o of 45 dBHz, must be long enough to allow phase-lock to occur. The frequency excursion which the PLL has to cope with depends on the length of time the phase tracking for the particular satellite was interrupted, the assumed oscillator drifts and the assumed user dynamics.

The PLL used will initially work with a large bandwidth B_L. The pull-in frequency threshold Δf_p, for a high gain second order loop is given by

$$\Delta f_p = \frac{1}{\pi} \sqrt{0.7\ W_m\ K_n} \qquad \text{Hz}, \tag{1}$$

for gain $K_n > \frac{Wm}{0.4}$, W_m being the loop eigenfrequency.

The pull-in time is then roughly

$$T_p = 4.2\ \frac{\Delta f_p^{\,2}}{B_L^{\,3}} \qquad \text{sec.} \tag{2}$$

For a short enough cycle time such as T_c = 2 sec, for user accelerations not more than a few g and for an oscillator with stability 10^{-7} Hz/Hz, only a few hundred Hz of frequency offset will have accumulated.

For B_L = 500 Hz, the pull-in range for a critically damped loop is roughly ± 400 Hz, in which case pull-in occurs within 5.5 msec. If the frequency offset is less than 225 Hz, lock-on occurs in about 1 msec.

The dynamic lag error for such large bandwidth is negligible and the statistical phase error, E, equals about 8^{o} for C/N_o of 45 dBHz. This phase-angle is important because cos(E) will scale the power available in the PLL of the individual tones at baseband. For this reason a variable bandwidth PLL could be used, with small bandwidth being used once lock-on has occurred. Choosing time interval C to be 20 msec for a cycle time T_c of 2 seconds therefore is adequate.

5.2.1.2 Determination of Time Interval T for Radiating Tones

In tone acquisition, two issues are of concern. The duration of the tones should be long enough for the individual PLL's to settle down. The second issue is the final phase accuracy with which the phase is being tracked.

The settling time again depends on the initial phase uncertainty and on the noise bandwidth of the tone-PLL.

Since we have a fixed set of tones, and the carrier on which the tones are modulated is phase-locked before the tones come on, the frequency offset on the tones at baseband is $\frac{fm}{fc} f_D$, where f_m denotes the tone frequency at baseband, fc the carrier frequency and f_D the doppler offset of the carrier. For the highest tone and maximum doppler f_D of 5 kHz, this frequency offset equals 1.6 Hz. In principle one could read f_D and correct the VCO of the tone PLL. However, the noise bandwidth of the tone PLL will be large enough to take care of 1.6 Hz initial frequency offset.

For a tone PLL bandwidth B_L of 70 Hz (tailored to a tone-to-noise-power density ratio of 35 dBHz), the phase error is about 10^o for 1 g dynamics. The PLL will then lock in about 7 msec for a 32 Hz offset. However, as argued above, the offset is usually much smaller.

In order to accumulate enough signal for a lock-indicator to work, one needs roughly $\frac{1}{B_{L,tone}}$ seconds or another 14 msec. However, the lock-on and signal integration happen simultaneously, consequently we allow for the time interval T a total of 20 msec.

5.2.1.3 Choosing Time Interval for Transmitting Data

A quick calculation shows that the data-interval will have to be considerably longer than the tone-interval. Also, to keep error rates down, high C/N_o is desirable for a given bit rate. All this leads us to have a separate data interval D, and to DPSK modulate the carrier directly during that interval. Data demodulation circuitry consists, for example, of a Costas loop with data detection. For C/N_o = 45 dBHz and about 11 dB SNR (for bit error rate 10^{-5} using DPSK data modulation), a data rate of 2.5 kbits per second could be supported.

If T_c is chosen to be 2 sec, the number of bits , R, which can be transmitted during the interval , D, where

$$D = \frac{T_c}{12} - G - T - C = 109.78 \text{ msec}$$

is 137 bits. This is sufficient since it exceeds the 50 bits per second for the present GPS design.

5.2.2 Choosing the Number of Tones and the Tone Frequencies

The number of tones and the choice of frequencies is based on the assumed initial uncertainty in the user's position (lowest tone), the required ranging precision (highest tone) and the capability of the tone phase lock loops to track the tone with a given phase error under assumptions of moderate user dynamics and of available input signal to noise power for the tone loops. The tones actually modulating the carrier are the sums and the differences of the basic tones on which the range calculation is based, and are "clustered" in the same part of the RF spectrum. FM or PM modulation of the carrier by the ranging tones seems generally preferable to AM from the point of view of the transmitter. The RF spectrum in that case is generally wide, many times the baseband of the modulating tones. In a tone ranging system no attempt is made to "demodulate" the signal and collect the wideband energy efficiently. Instead the RF spectrum is heterodyned to baseband and a narrowband filter (in the form of a PLL) is applied to the fundamental spectral line of each tone. The input SNR at the tone-PLL is a function of how the tone-modulating power is distributed among carrier and the tones. The distribution is determined by the modulation indices. Equality of indices for each of the clustered tones is appropriate in our design. The index then is chosen to maximize the power in the fundamental spectral line, with as little power as possible in intermodulation spectral lines and harmonics. At the same time the ratio of power in the fundamental line versus the total power is optimized.

In the following paragraphs some rationale is given for choosing the frequencies, the optimum number (four) of tones and the optimum index of 70%.

Tone ranging is based on the basic principle that delay in propagation of a CW signal is equivalent to phase shift, with an ambiguity equal to the signal period (assuming phase can be measured in the interval $[0,2\pi]$).

If the a priori uncertainty on the range (due to imprecise data on satellite position and user position) is M km, then the lowest frequency tone must have a frequency no higher than $f_L = \frac{c}{M}$, where c is the velocity of light.

By tracking the tone, the phase is determined. Let $\sin(\omega_L t_s)$ be the tone as transmitted by the satellite (forgetting the carrier modulation and demodulation process). Let $\sin(\omega_L t_R + \phi)$ be the received tone, with phase ϕ. Symbols t_s and t_R are time-readings according to the satellite clock and the receiver clock respectively. The unambiguous range portion is

$$\frac{\lambda_L}{2\pi}[\omega_L(t_R-t_s) + \phi] = c(t_R-t_s) + \frac{\lambda_L}{2\pi} \cdot \phi \ . \tag{3}$$

The portion $c(t_R-t_s)$ will be provided by the solution of four ranges to four satellites resulting in three dimensional user position and clock bias. For simplicity we assume the satellite times of the four satellites are perfectly synchronized and equal to true time. The term $c(t_R-t_s)$ represents the user clock bias. From formula (3) we see how the range error is related to the phase error and the wavelength:

$$\sigma_R = \sigma_\phi \cdot \frac{\lambda}{2\pi} \ . \tag{4}$$

In principle, if the phase error were arbitrarily small the range error would be arbitrarily small. However, σ_ϕ is limited by the signal-to-noise ratio in the PLL and the presence of dynamics. Therefore higher frequency tones must be chosen to resolve the basic uncertainty area. Finally the highest frequency tone provides the basic ranging accuracy in accordance with the same formula (4).

We continue our analysis for a clustered set of tones, since we have no intention of using these basic frequencies $(f_1 \ldots, f_N)$, but want to use clustered tones $(f'_1, \ldots. f'_N)$ instead obtained, for example, as follows:

$$\begin{aligned} f'_1 &= f_1 \\ f'_2 &= f_1-f_2 \\ f'_3 &= f_1-f_3 \\ &\vdots \\ f'_N &= f_1-f_N \quad , \end{aligned}$$

or as follows:

$$f_n' = \begin{cases} f_1 - f_n & \text{for n even} \\ f_1 + f_n & \text{for n odd.} \end{cases}$$

For tone ranging with clustered tones, the finest range-reading will be improved by a factor $\frac{1}{\sqrt{N}}$ or

$$\sigma_R \simeq \frac{1}{\sqrt{N}} \sigma_\phi \cdot \frac{\overline{\lambda}}{2\pi} \qquad (5)$$

for N clustered high frequency tones with average wavelength $\overline{\lambda}$, when each tone is tracked with a phase error having standard deviation σ_ϕ.

On the other hand, for ambiguity resolution, if the tones f_i' are tracked with error typified by σ_{ϕ_i} the corresponding error on f_i is $\sqrt{2}\ \sigma_{\phi_i}$, or worse by a factor of $\sqrt{2}$.

Table 5.4 shows values for σ_R according to formula (5) assuming a standard deviation of phase error of 10 degrees, which is obtainable with a PLL with input SNR of 35 dBHz and subject to 1 g dynamics. The SNR figure corresponds to a modulation index of 70% and is justified in Appendix B on Power Distribution. For σ_R equal to 30 ft (the most interesting design value) the highest frequency, which is also roughly the average cluster frequency, is then from 533 to 410 kHz, depending on the number of tones used.

The number of tones is determined by the desired probability of correctly resolving all range ambiguities from lowest to highest reconstituted tones, or equivalently by the probability that any of the N-1 ambiguities was falsely resolved. The latter probability, with a threshold of 10^{-6}, was used.

If each tone f_n' can be tracked with $\sigma_\phi = 10^o$ then the phase error on the reconstituted tone f_n is $\sqrt{2}\ \sigma_\phi \simeq 14^o$ with corresponding ranging error

$$\sigma_{R,n} = \sqrt{2}\ \sigma_\phi\ \frac{\lambda_n}{2\pi} \ .$$

TABLE 5-4. RANGE ACCURACY AND HIGHEST TONE

σ_R (ft)	N=3 f_1(Hz)	N=4 f_1(Hz)	N=5 f_1(Hz)
10	1.6 MHz	1.4 MHz	1.2 MHz
30	533 kHz	463 kHz	410 kHz
100	160 kHz	140 kHz	120 kHz
300	53.3 kHz	46.3 kHz	41 kHz

If $\frac{\lambda_n}{\lambda_{n-1}} = k_n$ then the probability that

$$\sigma_{R,n} > \lambda_{n-1}$$

resulting in incorrect ambiguity resolution is

$$P_{F_n} = 2 \text{ erf}\left(\frac{2\pi}{\sqrt{2}\,\sigma_\phi k_n}\right)$$

For N tones, there are N-1 ambiguity resolving measurements; and the overall probability of false ambiguity resolution is

$$P_F = 1 - (1-P_{F_N})\,(1-P_{F_{N-1}})\,\ldots(1-P_{F_2})$$

For equal ratio $k_n = k_{n-1} \ldots = \alpha$, one has $P_{F_N} = P_{F_{N-1}} = \ldots$, and then

$$P_F = 1-(1-P_{F_N})^{N-1}$$

or

$$P_{F_N} \simeq \frac{P_F}{N-1} \text{ for small } P_F \text{ .}$$

Table 5.5 shows how P_F and P_{F_N} vary as a function of N and the assumed σ_ϕ for f´ of 10°. Clearly 3 tones are too few, but 4 tones are satisfactory. Although 5 tones may seem better than 4, it constitutes an unnecessary additional complexity (one more PLL in a parallel setup). Therefore we chose the number of tones to be equal to 4.

Reverting to our previous table we see that $f_i = 463$ kHz is then the indicated highest frequency. The lowest frequency f_N, as pointed out, is determined by initial range uncertainty. If almanac data are available on all satellites, the initial range uncertainty is dominated by the user position uncertainty, say, 55 km. Then $f_N = 6$ kHz. Having determined the highest and lowest frequency and knowing the number of tones we determine the ratio α:

$$\frac{f_1}{f_4} = \alpha^3 = \frac{463}{6} \quad \text{or } \alpha = 4.26 \text{ .}$$

TABLE 5-5. CHOOSING THE NUMBER OF TONES IN TERMS OF PROBABILITY OF FALSELY RESOLVING AMBIGUITY

N	P_{F_N}	P_F
2	0.8	0.8
3	1.1×10^{-2}	2.2×10^{-2}
4	4×10^{-8}	1.2×10^{-7}
5	8×10^{-16}	3×10^{-15}

The tones therefore are

$$f_1 = 463 \text{ kHz}, \; f_2 = 108.69 \text{ kHz}, \; f_3 = 25.51 \text{ kHz}, \; f_4 = 6 \text{ kHz}.$$

However, we now adjust this set of frequencies using some engineering judgment. We adjust f_1 upward to 511.3 kHz which is one tenth of the basic GPS oscillator frequency. Since all tones must have definite phase relationships (for example, all phases are zero at t = 0) it is easier to have simple tone frequency ratios. For example, $\alpha = 5$ gives $f_2 = 102.26$, $f_3 = 20.452$, $f_4 = 4.0904$ kHz with finest accuracy 27.3 ft. (8.2 m) and resolvable initial uncertainty of 39.6 nmi (73.4 km) and probability of falsely resolving ambiguity of 1.1×10^{-6}.

The tones actually modulating the carrier are then, for example:

$$f_1' = 511.3 \text{ kHz}$$

$$f_2' = 409.04 \text{ kHz}$$

$$f_3' = 490.85 \text{ kHz}$$

$$f_4' = 507.21 \text{ kHz} \; .$$

Finally we discuss the choice of modulation index. In Fig. 5-3 it is shown that the suppression factor, which is defined as the ratio of power in the tone spectral line to the total carrier power, has a maximum for the FM modulation index Δ equal to .7. The SNR at the input of the tone PLL, for $C/N_o = 45$ dBHz, is then about 38 dBHz. The SNR available for the tone PLL actually used in our design was 35 dBHz, allowing for a 3 dB margin.

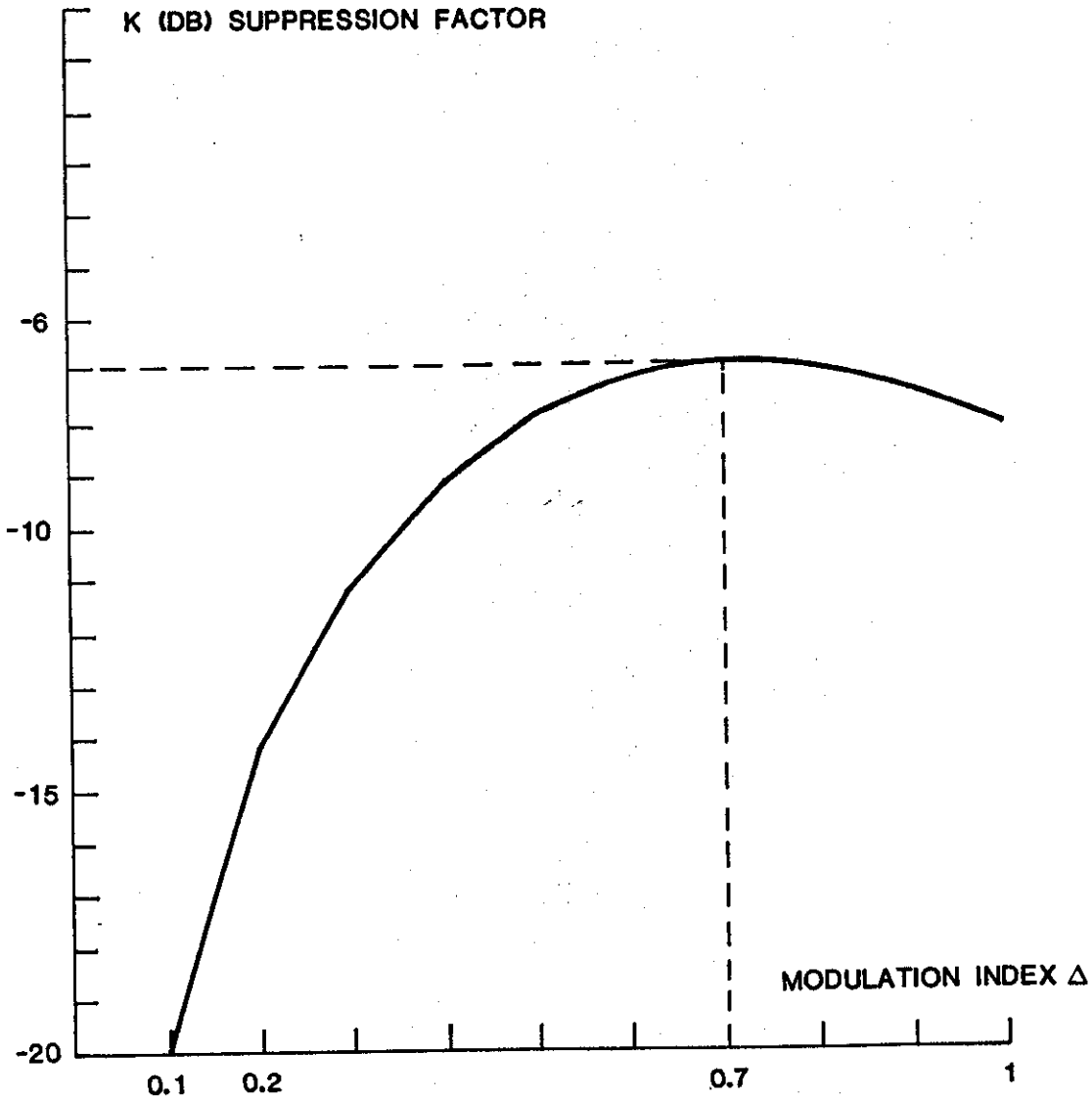


Fig. 5-3. Suppression factor as function of FM modulation index.

5.3 Discussion of Tone Ranging Performance

The main conclusion on tone-ranging as an alternate waveform is that one could indeed design an efficient tone-ranging system, as we have demonstrated.

Multitone systems without PN encoding are susceptible to multipath effects. It is possible however, that for a clustered tone system, multipath effects could be averaged out. This issue requires further study.

A multitone system without PN encoding could be easily jammed, since all that is necessary to disrupt it is to reproduce one or several tones.

The multitone system designed here, with periods of pure carrier being radiated, with some additional acquisition circuits, could "cold-start" more easily then the present C/A GPS system. A cold-start condition corresponds to total lack of a priori information on the satellite positions. The only knowledge assumed is a crude estimate of user position (error less than 40 nmi). Appendix B provides some details on a proposed method for cold-starting the multitone ranging system.

Tone ranging was found to be practical only in a TDMA context. Its ranging and data carrying capability can be favorably compared to those properties of the GPS C/A system. As an alternate waveform to the GPS C/A waveform it is generally more complex and is inferior to the pulsed system of Section 4, but its peak power requirements are lower.

Finally, we present Table 5-6 with a summary of the main features and performance characteristics of Tone-Ranging as presented by our chosen design.

The final criterion, cost, is not included. Based on our conceptual design of a multitone ranging system, there is little indication that it would cost less than a sequential GPS set due to the large number of phase-locked loops.

TABLE 5-6. SUMMARY OF FEATURES AND PERFORMANCE OF TONE RANGING SYSTEM

Power Levels	
Transmitted Average Power	25W
Transmitted Peak Power	25 dBW
Transmission	TDMA format 150 msec burst per 2 sec.
Power at Receiving Antenna	-151 dBW
Received C/N_o (5 dB noise figure)	48 dBHz
Signal Design Features	
Ranging Capability Provided by	Set of 4 tones
Data Carrying Method	DPSK of Pure Carrier
Data Rate	Burst of 137 bits in 109 msec or 66 bps avg.
Data Repetition Period for User	24 sec
Receiver Features	
Ranging Loop	Bank of 4 PLL's
Carrier Tracking Loop	Costas PLL
Data Demodulation	Coherent
Start-Up Circuitry	Yes
Cold Start Capability	Easy
Range Update Rate	0.5 Hz
Measurement Type	$\dot{R},R$
Sensitivity to Multipath	Sensitive

6. PRFDMA WAVEFORM

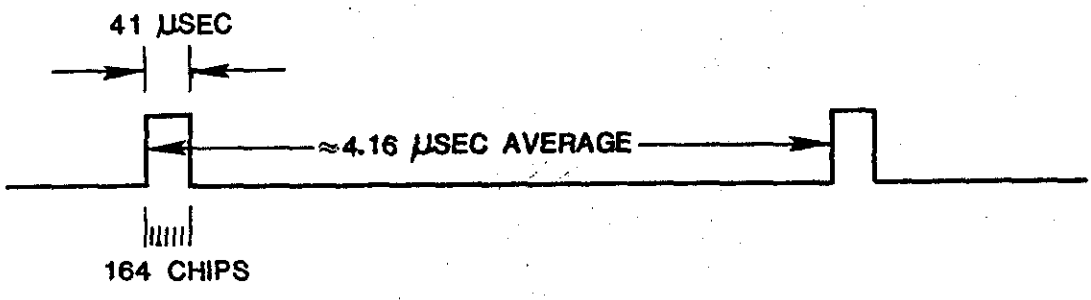
In the class of pulsed waveforms, where the short high energy pulse is adopted as the basic building block, there is a large variety of useful signal structures. One such was the TDMA Pulsed Design of Section IV. The different parts of the waveform were designed to correspond to specific functions: a pulse burst for ranging and synchronization followed by PPM for data transmission at a much lower duty cycle.

Although a short-term duty cycle is a good choice for a simple receiver, it puts demands on the transmitter. Another pulsed waveform, which relaxes the requirements on the transmitter, consists of a simple pulse train at a fixed PRF. Such a waveform will also reduce the demands on the user's clock in terms of stability. Each antipodal satellite pair uses a slightly different PRF to avoid repeated pulse overlap at the receiver, hence the name applied to this waveform.

A quick design study of one such PRFDMA waveform was made and is summarized below. The general conclusion was that, if computation costs keep decreasing at the rate experienced in the last decade, such a waveform may compete favorably with the TDMA pulsed system, because of the simplicity of the RF portions in transmitter and receiver. It was found to relax clock requirements but to be inferior in acquisition capability.

6.1 Ranging

The PRFDMA format is a signal format in which all satellites transmit a single pulse at a given PRF, rather than a burst of pulses over, say, 20 msec, repeated every .5 seconds as described in Section 4 . Figure 6-1 shows a PRF waveform choice. Assuming a 2 kW satellite transmitted peak power, the received power level, is 56.1 dBHz. For the chosen pulse length of 41 μsec, this corresponds to a SNR of 12.2 dB when no frequency offset is present. A 5 kHz frequency offset causes the SNR to drop to 11.6 dB and a 9.2 kHz offset results in a SNR of 10 dB. This is the minimum detection threshold for operation.



$PRF_i = 240 + i$ FOR 1.5 SEC

$= 240 - i$ FOR NEXT 1.5 SEC

Fig. 6-1. Sketch of the PRFDMA waveform.

The pulse length of 41 μsec was chosen to allow tolerance to a frequency offset of up to 9.2 kHz from all sources (relative user-satellite motion and clock frequency offset or drift), while still guaranteeing a SNR of 10 dB for pulse detection. This pulse is band-spread by means of a 4 MHz BPSK modulation.

The ranging accuracy is derived from the bandspreading, using PN encoding with a 164-chip code, identical for all satellites. Ranging based on a CCD matched filter implementation is expected to result in a standard deviation on a single pulse measurement of $\sigma_R = \frac{\Delta}{\sqrt{12}}$, where Δ is the chip length of 259 nsec or $\sigma_R \simeq 75$ ft. Pseudorange accuracy is not based on a single pulse, but on the output of a pseudorange tracker updated by range readings at the PRF rate.

A ranging pulse is received approximately every 4.2 msec. The time of transmission of these pulses (based on the satellite ID and the data) is also known. (This may involve counting the number of pulses with respect to a reference point in time such as the Z-count.) One can then track, for example with a simple fixed gain (α,β) tracker, the received pulse time t versus satellite transmit time s. The tracking allows very accurate prediction of the arrival time of the next pulse, thus helping in associating the correct pulse with the correct track, controlling the false alarm rate. But more importantly in the context of accuracy, tracking increases ranging accuracy and allows strobing of all pseudorange tracks at a common time t on the user's clock and determination of corresponding time s^i at which each satellite would have sent a pulse arriving at the user at time t. Correct user position determination can then occur, corresponding to time t even though the readings from the several satellites are asynchronously received.

The twelve PRFs needed could be defined by ($PRF_i = 240 \pm i$, for $i = 0,11$) where the minus value holds for 1.5 seconds and the plus value over the next 1.5 seconds for an overall average of 240 pulses per second over three seconds. Obviously with as many as eleven satellites visible, pulses will occasionally overlap (about one out of ten). However they must overlap to within two chips for them to overlap undistinguishably (probability of about one out of eight hundred). Such a condition can be detected and no decision made on data or ranging information derived from that single pulse.

It is easy to show that for the above PRF rates two indistinguishably overlapping pulses (aligned within two chips) from two satellites will, for the next pulse interval, be separated by about 68 chips. For a signal which had already been acquired, the pseudorange tracker would simply coast for one update. The data demodulation would be based on the remaining "clear" pulses.

The pseudorange tracker helps to sort the pulses into the right satellite file, the identity of which is re-affirmed once per half second by the data. If any contradiction in observed PRF_i and the identification exists, part of the data may have to be purged or re-acquisition procedures initiated.

6.2 Data Demodulation

The pulses are PN-encoded. One may choose a different code to represent a zero or a one or use PPM for data modulation. Pulse position data modulation would increase the computation load and make initial startup more complicated. Code division increases the hardware complexity. One bit would be transmitted per four pulses for a data rate of 60 bps which corresponds to the 50 bps of GPS, and another 10 bits of additional data, for example, for extra identification.

6.3 Satellite Identification

The PRF's of the several satellites are different; but in order to maintain synchronism with the P-signal at regular intervals (for handover and for simultaneous change-over to refreshed data once every hour) each satellite can be assigned two PRF's, averaging to the same value for all satellites, as was suggested above in Section 6.1.

6.4 Clock Stability and Computer Load Implications

In Section 9 an analysis is presented on the implications of the PRF update rate of 4 msec for this approach versus .5 second for the pulsed and tone-ranging TDMA candidates, on the clock stability requirements.

With the user-clock timing error contribution to ranging restricted to 13 nsec and for the assumption of white noise the clock stability requirement expressed in Allan variance can be relaxed roughly by a factor of five when using a 4 msec update interval instead of .5 seconds. In terms of computer

load, it was judged necessary to have dedicated computer chips, estimated to cost a few dollars each, to perform pseudorange tracking in order to free the microprocessor for other tasks.

In conclusion, a PRFDMA system, with relaxed requirements on the satellite transmitter and the user clock, can be designed with satisfactory performance. The receiver would be simple but the acquisition, pulse sorting and pseudorange tracking tasks put additional burdens on the microprocessor as compared with the TDMA pulsed design.

7. IMPLICATIONS OF THE TDMA FORMAT

In this section we discuss some advantages of a TDMA format for sharing the channel capacity amongst the 24 GPS satellites.

The low-cost GPS receiver will most likely be sequential, with the receiver cycling between a minimum of four signals needed for position and clock bias calculation, and possibly tracking a fifth or sixth signal as well, to cope with signal interruptions caused by maneuvers or to prepare for re-selection of satellites to maintain acceptable geometric dilution of precision.

In the case of four satellites, only during 25% of the time that signals are continuously transmitted, are they actually used by the sequential receiver. In the proposed TDMA format, a satellite transmits only during one-twelfth of the time, in a fixed timeslot. For the same average transmitted power the received signal level can be about 10 dB stronger. Also, it can be received during 100% of the time that it is transmitted and since satellites transmit only in assigned timeslots, separated by a time interval to avoid signal overlap, the satellite signals do not mutually interfere.

The advantages for data collection are obvious. In the case of GPS C/A and a sequential receiver, message portions are received piecemeal and out of sequence. In addition to this, it is hard to match the cycle time of the receiver against a particular subdivision of the messages. Due to the differential delay between satellites (worth up to 1 bit) and re-acquisition times (worth several bits), several data bits between pieces of the message will most probably be lost.

In the proposed TDMA format there is no difficulty in assembling the message portions in sequence and it is done without loss of bits.

The counter-argument could be that, although now transmission and reception are one hundred percent matched, the receiver is idle except during four out of twelve timeslots, or 75% of the time. And since data are received in a burst, the higher instantaneous data rates will imply a more powerful and therefore more costly processor. However, the latter argument does not hold, since the

circuitry for data demodulation for the pulsed system discussed here is simple, and unaffected by the high instantaneous data rate. The proposed instantaneous rate of about 4 kbits per second in the particular TDMA pulsed design of Section IV is within the capabilities of present-day processors. Let us point out that the 4 kbps rate is a consequence of our choice of a three times higher average data rate and should be correspondingly reduced when comparing the receivers for equal performance.

To understand better the discussion on GDOP which follows, we include some general conclusions concerning our short study on satellite selection based on optimal or sub-optimal GDOP.

A study on ephemeris behavior of GPS satellites and, in particular, of a constellation of four satellites selected for their good GDOP properties revealed that:

obtaining the "optimal" selection is a laborious task

optimality is usually short-lived (on the order of ten minutes)

the change of GDOP as a function of time for a selected constellation can be quite fast, leading possibly to a 50% deterioratation in a time span of ten minutes.

Often the best constellation is one that is made up of three low elevation satellites, as equally distributed as possible, together with one high elevation satellite. Usually, at least one is setting. Good GODP and longevity, therefore, do not go hand-in-hand.

The problem is that when a single satellite becomes undesired (sets below 5^{o} EL for example), it is most often true that no other satellite can replace it to maintain a satisfactory GDOP. Usually a completely new constellation must be acquired and tracked. The short-livedness (order of 10 minutes) of a good constellation makes this a difficult problem for the receiver. The proposed TDMA system would not have this problem.

For the pulsed TDMA system (Section 4), acquisition and tracking of a new satellite is virtually instantaneous by design. The signals of all visible satellites are available and can be tracked since they arrive in a strict non-overlapping time-ordered basis at the receiver. All signals have virtually equal strength and are modulated by the same PN code and are on the same frequency. There is no adjusting or re-configuring necessary on the part of the receiver.

The 75% of the cycle time during which the receiver is idle, when compared to the C/A GPS receiver tracking only four satellites, can be exploited to great operational advantage.

The idea of using four range measurements to four selected satellites based on their associated "good" GDOP can be totally dispensed with. The algorithmic complexity of selecting and continuously reselecting an optimal or at least sub-optimal set as often as every ten minutes, and the associated interruptions for re-acquisition of the new signals, are then unnecessary. Also the problem of being forced into unscheduled re-selection when one satellite signal disappears (sudden turn) need never occur.

One may use ranging information from all available (visible) satellites. When more than four ranges are used, position can be determined, for example, in a least squares sense instead of a deterministic sense. The resulting position coordinates are then used to update the user tracker. In many designs it is natural that range measurements are obtained at a higher rate than is necessary for the user tracker. One may then smooth pseudorange measurements before combining the pseudoranges, and extrapolate or interpolate the measurements to a common time at which user position is determined.

It should be pointed out that while the GDOP resulting from the use of all visible satellites is always truly optimal (and better than the optimal GDOP resulting from a subset of only four visible satellites), no such claim can be made in general for the GDOP based on only four. Indeed, it is almost unthinkable, for example, that with eleven visible satellites all 330 combinations of four satellities would be selected and compared based on their GDOP. Instead some algorithm will first reduce the number of combinations to a tractable number (maybe ten) and only then compare the remaining combinations based on

explicitly calculated GDOP values. Such algorithms are therefore at best sub-optimal. Also, the selected combination out of the reduced set, based on the best GDOP, is only best at a given instant of time.

It is quite complicated to predict how long the selected combination will remain best. In general, one could compare the remaining combinations based on the expected lifespan of the combination or on the average of its GDOP values now and at some future time. This makes the procedure of selection of a combination even more sub-optimal.

And finally, it should be pointed out that GDOP is a scalar measure expressing linear amplification of the error on a single range measurement, as expressed by the User Equivalent Ranging Error (UERE). In reality the geometric dilution is more pronounced in some directions than in others, and these directions and the amplifications suddenly change at the moment of re-selection of one or more satellites, even if the scalar value for GDOP remains the same.

If all visible satellites are used, the GDOP transitions, in scalar value or in the directionality of the main component of error, will be less severe since satellites are replaced automatically one at a time. Although the tracking algorithms are exercised more often, a sizable portion of software associated with satellite (re-)selection and the time to exercise it, can be saved.

8. HANDOVER FROM THE ALTERNATE WAVEFORM MODE TO THE P MODE

It is possible that a user of the pulsed C/A waveform may wish to acquire the normal GPS P-signal for increased accuracy of ranging. For this reason the pulsed signal should contain handover information equivalent to that provided in the normal GPS C/A signal. The basic requirement is the inclusion of a data word containing the so-called Z-count and a means of recognizing the instant of time at which this Z-count applies. The latter part of this requirement implies a basic synchronism between the P-waveform and the pulsed C/A signal, as discussed in more detail below.

The P-signal is formed from two shift-register sequences, each of which is short-cycled to 15,345,000 chips. Since the P-signal chipping rate is 10.23 MHz, the cycle corresponds to precisely 1.5 seconds, called a super-epoch. At GPS zero time (midnight Saturday night), the two codes, which are added (modulo 2) to form the P-code, are started together in the all-ones state. In satellite Number 1 they are clocked together, while in satellite Number 2 the second code is delayed one chip, and so on up to a delay of 23 chips for the last satellite. At the end of the first super-epoch, the second code is held fixed (clock pulses are inhibited) for 32 chips and then again clocked in synchronism with the first code. The second code marks time in this way for 32 chips at the start of each super-epoch, thus precessing with the first code to form a very long sequence for the resultant P-code.

One week contains 403,200 super-epochs, hence the second code has slipped by 12,902,400 chips by that time, after which the second code is reset to all-ones and the entire week-long code (a different one for each satellite) is repeated.

The Z-count is the number of super-epochs which have elapsed since zero-time, and it requires a 19-bit word for full representation. The normal C/A signal contains data (synchronous with the C/A PN code) organized into 6-second subframes. Each subframe begins exactly at the start of a super-epoch and lasts for four super-epochs. Each subframe also contains a Z-count word of 17 bits (the two lowest order bits are truncated) which identifies the super-epoch coincident with the start of the next forthcoming subframe. The user who has

acquired this information can therefore set the two code generators which will form his P-signal to the all-ones state and know exactly how many chips to inhibit the second after starting the first at the time of the next subframe start. He will be in step with the super-epochs of the arriving P-code to an accuracy of a fraction of a C/A code chip, i.e., to an accuracy of fewer than ten P-code chips. Each of the codes which form the P-signal can be slewed ahead or behind by a method analogous to the slewing of these codes relative to one another to complete the P-code acquisition.

Any C/A waveform can provide the same handover, provided it is synchronous with the P-code super-epochs in some way and also provides Z-count information. For example, in the pulsed design described above it is only necessary to make the 0.5 second frames synchronous with the P-code super-epochs so that every third frame begins exactly at the start of a super-epoch. A 19-bit full Z-count word can be transmitted during the triplet of frames corresponding to one super-epoch. Since each satellite begins transmitting at the start of a precisely-timed time slot, the user can preset his P-code generators for the start of the next forthcoming super-epoch and also allow for the delay expected between this super-epoch start and the arrival of the sync burst from the satellite in question in the first frame of this super-epoch. There is no requirement for a precise user clock to span a long interval of time in this procedure, since all the burden for precision of timing (e.g., time-slot start) is placed upon the satellite.

9. CLOCK STABILITY REQUIREMENTS IN TDMA SYSTEMS

In a conventional multichannel GPS receiver, pseudorange measurements from all tracked satellites are continuously available. By simultaneous sampling of these outputs, the range differences necessary for user position determination can be obtained without reference to a clock. However, in any system with intermittent transmissions from the satellites, such as the TDMA systems described above, pseudorange for each satellite is obtained only at intervals, and must be timed and tracked in some way so that simultaneous measurements can be synthesized from the track for positioning. This process of tracking and estimation brings with it a dependence on clock behavior and a requirement for short-term clock stability.

In the PRFDMA system, pseudorange tracking is necessary to sort the pulses from different satellites and also to perform smoothing, or integration, of the pulses from each satellite in order to build up the accuracy of the individual pseudorange measurements. A simple "alpha-beta" tracker has been suggested for this purpose, one for each satellite in track, and these tracker outputs can be used to predict each pseudorange ahead to a common measurement time. The predicted values are then used in a conventional user position calculation, just as though they had been obtained by "strobing" (simultaneously sampling) a set of continuous pseudorange outputs.

In the TDMA schemes, the situation is a bit different, since pseudorange is obtained once per frame (one-half second interval) from each satellite, and every measurement, being made from a pulse burst or tone-ranging signal, already has the accuracy required for positioning. The measurements are made at staggered times, by the basic nature of the TDMA format, and require interpolation (instead of smoothing and prediction) to a common measurement time for input to the position calculation algorithm.

In all cases, the pseudorange measurements actually used for location are estimates, based on timed, intermittent data, and hence carry errors due to clock imperfections in addition to the usual errors of TOA (time-of-arrival) measurement. A general analysis has been carried out of the effect

of clock instability on user position accuracy for several tracking and interpolation schemes, and for two standard models for clock behavior. The general conclusion is that clock timing errors are less serious in their effect on positions than TOA errors of equal magnitude, due to the correlated nature of clock errors and the tolerance of the tracker-interpolation schemes to low-frequency clock noise.

In this section, only the TDMA case, with simple two-point interpolation, will be discussed in detail. This is done because the TDMA systems, with their relatively long intervals between measurements, as compared to the PRFDMA system place the greatest demands on the clock. It also turns out that the analysis is considerably simpler than that of the alpha-beta tracker model, for which results will be stated without proof.

Taking the pulsed TDMA system first, consider the ranging bursts from a single satellite, and let the measured arrival times of these bursts be R_1, R_2,..., according to the user's clock. From the data sent by this satellite, the user knows the corresponding transmission times, S_1, S_2, ..., of the bursts, according to the satellite's clock. Both times are adjusted to refer to a common fiducial mark within the pulse burst, and no satellite clock error is assumed. In effect, the user can plot transmission time, S_n, versus reception time, R_n, and join the points with straight lines for each of the satellites being used. When position computation is desired, a single time, R, is chosen and the corresponding transmission times are read off of each plot, providing simultaneous pseudorange for location. Position computation is asynchronous with the TDMA frame, being exercised on demand from the user position tracker.

Since the data points were joined by straight lines, the process represents simple linear interpolation, for each satellite, between the two measurements which straddle the measurement time, R. Linear interpolation is appropriate, since smoothing by curve-fitting is not required to improve accuracy, and the effect of acceleration in user-to-satellite range (which is dominated by the user's own acceleration) is negligible over the half-

second intervals involved here, amounting to no more than one foot per g of acceleration.

In practice, it is only necessary to save the two most recent measurements (S and R for two successive frames) for each satellite in ping-pong fashion until a position update is required. If this request comes at time R, the data in store is retrieved, and the measurement time $R_o = R - \Delta$ is established, where Δ is the frame duration for the system. A moment's thought will show that the stored data will contain one pair with reception time earlier than R_o and one pair later than R_o, for each satellite. If these pairs for one satellite are called R_1, S_1 and R_2, S_2, then the transmission time estimate corresponding to the chosen reception time R_o is simply

$$\hat{S}_o = S_1 + \frac{S_2 - S_1}{R_2 - R_1} (R_o - R_1).$$

It might seem simpler to use the same data to predict ahead to time R, but the general analysis shows that clock error effects are minimized by using R_o in the middle of the data points, and that prediction outside the measurement interval causes significant degradation. The fact that the position fix is older, by Δ (.5 sec.), than otherwise would be the case, is of no importance since the user position tracker will be predicting position in any case in order to drive the user's display.

Suppose that true time is denoted by t, and that the corresponding time by the user's clock is

$$R(t) = t + E(t).$$

In other words, the clock error at true time t is E(t), and all other sources of error in TOA measurement, such as receiver noise, are ignored here. If the true times corresponding to R_o, R_1 and R_2 are t_o, t_1 and t_2,

then the actual transmission time that would correspond to reception at true time t_o will be given by

$$S_o = S_1 + \frac{S_2 - S_1}{t_2 - t_1} (t_o - t_1) ,$$

if range acceleration is neglected. It is easy to show that the resulting error $e = \hat{S}_o - S_o$ is given by

$$e = E_o - \frac{t_2 - t_o}{t_2 - t_1} E_1 - \frac{t_o - t_1}{t_2 - t_1} E_2 ,$$

where $E_o = E(t_o)$, etc, and the factor

$$\frac{S_2 - S_1}{t_2 - t_1} = \frac{ds}{dt} ,$$

which multiplies the entire expression for e, has been approximated by unity. This is permissible, since ds/dt is the user-satellite range rate, divided by the velocity of light, and is hence only a few parts in 10^6. By the same argument, $t_2 - t_1$ is essentially equal to Δ, and we write e in the final form

$$e = E_o - (1-X)E_1 - X E_2$$

where

$$X \equiv \frac{t_o - t_1}{t_2 - t_1} = \frac{t_o - t_1}{\Delta} .$$

The true receive times are then $t_1 = t_o - X\Delta$ and $t_2 = t_o + (1-X)\Delta$, so that the error is completely parametrized by the quantity X, as shown in Fig. 9-1. where the ranging bursts are shown as simple pulses. This error in transmission time will be used instead of pseudorange error, since the reception time will be common to all satellites, cancelling out of the position calculation, and the sign is of no importance.

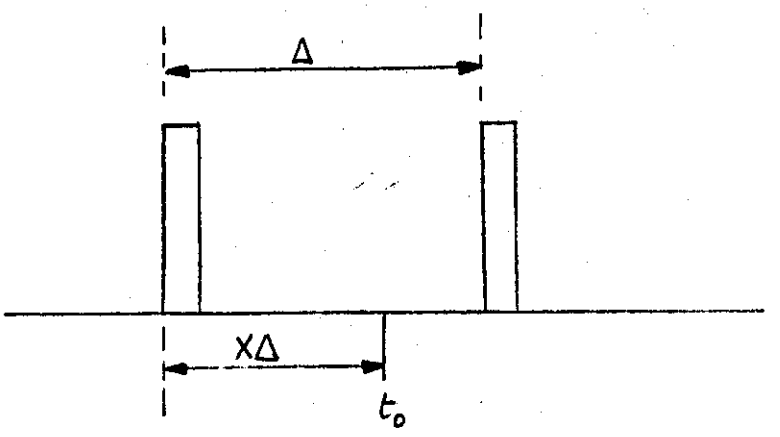


Fig. 9-1. Definition of X.

Each pseudorange error is of the form given above, each with its own value of X and corresponding receive times, while all share the error term E_o. If the clock had only a deterministic error, consisting of a bias and a rate error, then E(t) would have the form $a + bt$, and it is easily seen that each individual pseudorange error is zero in this case. Physically, this is so because linear, proportional interpolation is independent of the origin and scale of the axis against which the data is plotted. In our problem this has the deeper consequence that pseudorange errors are insensitive to the low-frequency components of random clock noise. In fact, we shall find that they are more tolerant of flicker noise than of white noise in the clock output, as defined below.

This brings up the question of clock noise models, and we follow the usual practice in the frequency standards field by representing the clock as an oscillator whose instantaneous phase at time t is

$$\phi(t) = 2\pi f_o \int_o^t [1 + y(s)]\, ds \quad ,$$

where f_o is the nominal frequency of the oscillator. The quantity y is the normalized instantaneous frequency error:

$$y(t) = \frac{(\dot{\phi}(t)/2\pi) - f_o}{f_o} \quad .$$

A clock driven by this oscillator will be used to time an interval by measuring the accumulated phase difference over the interval, and dividing by $2\pi f_o$. In terms of our previous notation, the error in timing the true interval (t_1, t_2) will be

$$E(t_2) - E(t_1) = \int_{t_1}^{t_2} y(s)\, ds \; .$$

The "frequency noise", y(t), is modelled as a stationary random process with zero mean (a constant mean corresponds to a constant frequency offset, which has no effect in our problem).

Two models are used here for the y(t) process. In the first, y(t) is taken to be "white noise", with a constant power spectral density, while in the second, the spectral density is assumed to vary inversely with frequency, corresponding to "flicker noise". The effects of white noise are easily analyzed in the time domain, as will be done here, but flicker noise requires a tedious analysis in the frequency domain, and only the results will be quoted.

Clocks are usually characterized by the behavior of their Allan variance*, $\sigma_a^2(\tau)$, a dimensionless quantity which measures normalized frequency stability over an interval of length τ, as a function of τ. If the frequency noise is white, with single-sided power spectral density G_o, then it is known that the Allan variance is

$$\sigma_a^2(\tau) = \frac{G_o}{2\tau}$$

varying inversely with τ. Since the clock frequency noise is white, the errors in timing non-overlapping times are uncorrelated, and the error $E(t') - E(t)$ has the mean squared value

$$[E(t') - E(t)]^2 = \frac{G_o}{2} \; |t' - t|.$$

(The error function $E(t)$ is therefore a Wiener process). The expression allows us to evaluate the mean squared values of pseudorange errors, and also their correlations, in terms of G_o and hence, ultimately, in terms of Allan variance.

Flicker noise has spectral density $G_{-1}f^{-1}$, and the corresponding Allan variance is independent of τ

$$\sigma_a^2(\tau) = 2 \ln 2 \; G_{-1} .$$

---

*D. W. Allan, "Statistics of Atomic Frequency Standards," Proc. IEEE, Vol. 54, No. 2, February 1977, pp. 221, 230.

This connection again allows us to express our results directly in terms of Allan variance, after an analysis in the frequency domain. It should be noted that white and flicker noise models are in general agreement with the Allan variance behavior of quartz crystal clocks in the range of time intervals between tens of milliseconds and seconds which are relevant to our problem.

Returning to our expression for pseudorange error, we note that it can be written in the form

$$e = (1-X)\ (E_o - E_1) - X\ (E_2 - E_o).$$

Since the intervals (t_1, t_o) and (t_o, t_2) do not overlap, we easily find the mean squared value, for white noise:

$$\overline{e^2} = \frac{G_o}{2}\ [(1-X)^2\ X\ \Delta + X^2\ (1-X)\Delta] = \frac{G_o \Delta}{2}\ X\ (1-X).$$

In terms of the Allan variance at $\tau=\Delta$, we can put

$$\overline{e^2} = \sigma_a^2\ (\Delta) \cdot \Delta^2 \cdot X\ (1-X)\ .$$

Notice that the error vanishes when X equals zero or unity, which makes sense, because interpolation is unnecessary when a pulse arrives exactly at the measurement time. A similar calculation yields the mean of the product of two pseudorange errors e and e':

$$\overline{ee'} = \sigma_a^2(\Delta) \cdot \Delta^2\{XX'(1-X') + (1-X)(1-X')X\}$$

where X is the smaller of the two parameters, X and X'.

With these formulas we can evaluate the moment matrix for the clock-related errors for any number of pseudoranges, corresponding to the satellites currently being used for location. This moment matrix, in turn, fixes the contribution of clock errors to find position uncertainty. Unfortunately, such a moment matrix depends in a complicated way on all the X-parameters which describe the detailed positions of the received pulses relative to the chosen measurement time. This measurement time, however, bears no special relation to the pulse arrival times and, averaged over many position updates, using different sets of satellites, the X-values will appear randomly distributed over their range, from zero to unity. Since we only require the effect of clock errors on mean squared position error, it is reasonable to average out these X-parameters, yielding a much simpler, but representative, pseudorange error moment matrix. By integrating over X and X', from zero to unity, we easily find the mean values

$$\overline{e^2} = \frac{1}{6}\sigma_a^2(\Delta)\Delta^2$$

and

$$\overline{ee'} = \frac{1}{10}\sigma_a^2(\Delta)\Delta^2 .$$

The averaging over the X-parameters is now included in the significance of the overbar.

According to our previous discussion, $\sigma_a(\Delta) \cdot \Delta$ is the rms clock error in timing an interval of length Δ, and hence might be compared with the rms error, σ_R, in measuring time of arrival due to receiver noise and quantization effects. We would like to be able to say that a clock error, $\sigma_a(\Delta) \cdot \Delta$, of one nanosecond is equivalent to a TOA error of some other number of nanoseconds in its effect on user position accuracy. However, TOA errors are independent among pseudorange measurements, corresponding to the values

$$\left.\begin{aligned} \overline{e^2} &= \sigma_R^2 \\ \overline{ee'} &= 0 \end{aligned}\right\} \quad \text{TOA errors.}$$

The desired comparison is made possible by comparing the moment matrices in the two cases for pseudorange differences, rather than pseudoranges directly. User position can always be computed, in the case of N satellites, from N-1 pseudorange differences, one satellite being chosen as reference for all the others. Let e_o be the error in the reference pseudorange, and let e and e' be two other pseudorange errors. The (N-1)-dimensional moment matrix of differences will have diagonal elements of the form

$$\overline{(e - e_o)^2} \equiv A,$$

and off-diagonal elements of the form

$$\overline{(e - e_o)\ (e' - e_o)} \equiv B.$$

We can work these out in terms of the quantities we already know:

$$A = \overline{e^2} + \overline{e_o^2} - \overline{2ee_o}$$

$$B = \overline{ee'} - \overline{ee_o} - \overline{e'e_o} + \overline{e_o^2} \quad .$$

For TOA errors, the cross-product terms are zero, and all pseudorange errors have variance σ_R^2, hence

$$A = 2\ \sigma_R^2$$

$$B = \sigma_R^2 = A/2$$

(the diagonal elements are all equal to A; the off-diagonal elements equal to B). For clock errors, using our averaged values, $\overline{e^2}$ is the same for all satellites, and $\overline{ee'}$ is the same for all pairs, which means that

$$A = 2\ (\overline{e^2} - \overline{ee'})$$

$$B = -\overline{ee'} + \overline{e^2} = A/2 \ .$$

This general result holds for all noise models; for white noise we have

$$A = \frac{2}{15} \sigma_a^2(\Delta) \Delta^2$$

$$B = \frac{1}{15} \sigma_a^2(\Delta) \Delta^2 .$$

Since these moment matrices (for pseudorange differences) have the same form, except for a proportionality constant, clock errors will behave exactly like TOA errors in their effect on position accuracy, regardless of the number of satellites used for location on the geometry of the constellation in use.

The proportionality between these two error sources is expressed by the relation

$$(\sigma_R)_e = \frac{\sigma_a(\Delta) \cdot \Delta}{\sqrt{15}} = 0.258 \sigma_a(\Delta) \cdot \Delta \text{ (white noise)}$$

which gives the rms TOA error equivalent to the effects of a clock Allan variance $\sigma_a^2(\Delta)$.

The entire analysis can be repeated for flicker noise, which results in more complicated formulas for $\overline{e^2}$ and $\overline{ee'}$ in terms of the X-parameters. When these expressions are averaged over the parameters, we obtain the proportionality relation

$$(\sigma_R)_e = 0.159 \sigma_a \cdot \Delta \quad \text{(flicker noise)}$$

where σ_a, in this case, is independent of Δ.

Taking white noise, the worst of the two models, and the frame length of .5 seconds for Δ, we see that a clock whose short-term stability (apart from bias and frequency offset) is characterized by an Allan variance of one part in 10^7 (this is actually the square root of the Allan variance), for time intervals of the order of one second, contributes position errors equivalent to TOA errors of rms value 13 nanoseconds. Since these errors are combined on a "root-sum-squared" basis with other errors of the order 30 nsec or more, it may be concluded that the clock stability requirements of the TDMA system is a few parts in 10^7, for one-second time intervals. Although the discussion was carried out in the context of a pulsed TDMA system, the results also apply to a tone ranging system, if we consider only errors in timing due to the clock.

The PRFDMA system presents a similar, but more difficult, analysis problem. In steady state, each tracker is smoothing over the whole past of pulse arrivals, spaced Δ seconds apart (Δ being a representative pulse repetition period). The chosen measurement time will be $X\Delta$ seconds after the latest pulse arrival time, where X varies for each satellite, being in all cases between zero and unity. The transmission time error, e, for one satellite is then a weighted sum of the clock errors at all previous arrival times. As before, it can be shown that e vanishes for a deterministic clock error consisting of bias and rate offset.

For white noise it is possible to evaluate $\overline{e^2}$ and $\overline{ee'}$ in terms of Δ and the tracker parameters, and the time-phase parameters X and X'. Again, X and X' are averaged out to yield an effective TOA ranging error of the form

$$(\sigma_R)_e = K\,\sigma_a(\Delta)\cdot\Delta \quad ,$$

where K is given in terms of the tracker parameters by the expression

$$K^2 = \frac{2\alpha - (\alpha-3)\ (\beta/4) + (\beta^2/5)}{3\alpha\ (4 - 2\alpha - \beta)} \quad ,$$

values of α and β corresponding to a tracker "integration time" of about 10 pulses are

$$\alpha = 0.34545 \ ; \ \beta = 0.05455,$$

which results in $K = 0.438$. For smoothing over 5 pulses, the corresponding K is 0.363. These K-values are larger than these of the two point interpolation, mainly because transmission time is being extrapolated outside the measurement interval, but the difference is small. In our application, the pulse interval, Δ, is only a few milliseconds, and this results in much less demanding clock requirements than in the TDMA case, far overcoming any difference in K-factors (flicker noise effects have not been evaluated for the alpha-beta tracker, but should result in smaller K-factors than white noise).

As an example, taking $\Delta = 4$ msec and $K \simeq 0.5$, we have

$$\sigma_a(\Delta) = \frac{1}{K\Delta}\ (\sigma_R)_e$$

or

$$\sigma_a(.004 \text{ sec}) = 6.5 \times 10^{-6} \ .$$

To be fair, this same clock will have a smaller Allan variance at $\tau = .5$ sec

(the specification point for the TDMA case discussed above), with our white noise model, namely

$$\sigma_a(.5 \text{ sec}) = \left(\frac{4}{500}\right)^{\frac{1}{2}} \times 6.5 \times 10^{-6} = 5.8 \times 10^{-7};$$

hence, the clock requirements are relaxed here only by a factor of about five.

APPENDIX A

ANALYSIS OF PERFORMANCE OF GPS C/A WAVEFORM WITH MORE POWER

This appendix provides some analytical details to back up the summary statements of Section III. The outline of this appendix is the same; first we discuss code synchronization issues, then carrier phase lock issues and finally data demodulation and issues related to data gathering.

A.1 Code Synchronization

A.1.1 Sequential Code Synchronization

The average time required for code synchronization is a function of the assumed signal-to-noise power density. This study assumes an RF transmitted power by the GPS satellites of from 14 dBW (present level) to 29 dBW. Table A-1 summarizes our estimates of link losses and the ratio of signal power to noise power density at the input of the receiver.

The two issues to be considered are: the time required for sync detection, and the sensitivity to doppler. If the system is sensitive to doppler then one must not only perform a code search, but a search in the frequency dimension as well. In a serial code search, the replica code is tested against the incoming code with steps of one half chip length. Detection is based on the power level of the despread carrier falling in the narrow bandpass filter of the sync detector. With T_s denoting the signal integration time, it is assumed that the incoming signal is not doppler shifted in frequency as compared to the local oscillator by more than $\frac{1}{2T_s}$ Hz. This frequency offset causes a loss of no more than 3.8 dB in output signal to noise ratio at the correlation-peak. Another 1.25 dB is lost in case the two codes are one quarter chip misaligned and the testing is done with one half chip increments.

The signal-to-noise ratio must be 15 dB, i.e.,

$$\frac{C}{N_o} T_s > 15 \text{ dB} ,$$

to guarantee 10 dB under the worst conditions of maximum frequency shift and code misalignment. The 10 dB corresponds to a probability of false alarm of less than one percent combined with a probability of correct detection of 90%.

TABLE A-1. HIGH POWER GPS BUDGET

RF Sat. Transmitted Power	14 to 29 dBW
Sat. Antenna Gain	11.5 dB
Pathloss	-185.5 dB
Atmospheric Loss	-1 dB
At Receiver Antenna	-161 to -146 dBW
Receiver Antenna Gain	0 dB
Receiver Loss	-2.5 dB
Margin for Other Loss	-2 dB
	-165.5 to -150.5 dBW
Receiver Noise for 5 dB Noise Figure	-199 dBW/Hz
C/N_o; Ratio of Carrier Power to Noise Power Density	33.5 to 48.5 dBHz

Since $\frac{C}{N_o}$ is variable from 33.5 to 48.5 dBHz, the integration time ranges from 14.3 to .45 msec. Since the code epoch of the presently proposed GPS C/A signal is 1 msec long, one will prefer coherent integration times which are multiples of 1 msec. If a different GPS signal format with shorter epochs were used in combination with more power, coherent integration times less than 1 msec could be exploited. We expand on this in a later paragraph.

The above integration times obviously depend on the threshold set for reliable sync detection. If a higher degree of false alarm and missed detections is acceptable, shorter times will result.

It would generally be unwise to have coherent integration times of over 19 mseconds to reach the required detection energy, as would be necessary at the lower scale of the power level since this would reduce the filter bandwidth to less than 75 Hz. This makes the code synchronization function much more sensitive to doppler and frequency drift from the oscillator.

It would not be useful to have integration times less than 1 msec for the present C/A code since appropriate threshold levels for partial code auto correlations are difficult if not impossible to set reliably.

A coherent integration time of 1 msec allows roughly a frequency band of ± 500 Hz. For a sequential receiver cycling through four satellite signals with cycle times on the order of 1 to 10 seconds, the frequency drift accumulated from relative user and satellite motion and clock drift will generally be well within the ± 500 Hz band. This avoids a sequential code sync search over several frequency bands.

The situation is entirely different when the receiver is trying to acquire the signal from a new satellite. The number of frequency bands to be searched depends on the a priori frequency uncertainty. Even with a perfect oscillator, the carrier frequency can be off as much as ± 4342 Hz (worst case at the earth poles) from satellite motion with respect to a stationary user and ± 1000 Hz from user motion. If sufficiently accurate satellite ephemeris data are available, for example, in the form of pre-stored almanac data, only the latter

frequency error is important, implying a search in at most 2 or 3 frequency steps. If no almanac data are available, as would be the case in a "cold start-up" mode, as many as 10 frequency steps may be necessary.

Since we have assumed that the code synchronization search is done sequentially in one-half chip steps, it could require up to 2046 steps. The integration time discussed above must then be multiplied by this number and by the number of required frequency steps which can be from 2 to 10, depending on power level.

In summary, the effect of more power on sequential code synchronization is that integration time (inversely proportional to power) can be significantly reduced, but has a practical lower limit of one msec (code epoch). The number of frequency steps can also be reduced, since that number equals the assumed frequency uncertainty present, divided by the bandwidth (inverse of integration time). While the reduction can be dramatic for start-up conditions when the code is acquired the first time, the number of frequency steps is more likely to be two in normal operations with no reductions due to increased power possible when cycling times between satellite signals are from 1 to 10 seconds. The largest factor in synchronization is the number of code phases to be tested. That number is not directly dependent on power level, certainly not in the acquisition phase. However, in the re-acquisition mode (when cycling between satellite signals) the number of code phases to be searched is a function of the range-uncertainty accumulated during the time other satellites were visited. More power can provide better ranging (see Section 3.1.4) and more accurate range prediction and therefore will reduce the number of code phases to be searched in the re-acquisition mode.

The above analysis points to shorter synchronization times with increased power. This is an important operational benefit. Hidden benefits are in the potential for use of a cheaper IF oscillator, elimination of software to perform incoherent integration and eventually elimination of software associated with frequency searching in normal operations. Frequency searching would be done only when a new satellite is being acquired.

The possibility of elimination of code phase searching is suggested in Section A.1.3 on matched filtering.

A.1.2 Possible Use of a Shorter Code

The previous section suggests that if available signal-to-noise power density ratio exceeds 45 dBHz, then with the same detection thresholds as above the integration time for the test on one code phase in one frequency band could be less than 1 msec. The objection against lesser time was that the code epoch was 1 msec and constituted a lower bound. That objection is removed if one is willing to redesign the GPS waveform with a shorter code of, let's say, 511 or even 127 chips.

The potential of the short code of 127 chips at the present 1.023 Mbps clocking rate is that one can coherently integrate over one epoch and add the results incoherently to reach a sufficient energy level for detection and be totally independent of the maximum expected doppler. Tolerable frequency offsets are ± 7900 Hz for a 127 chip code. Circuitry and algorithms corresponding to frequency stepping can be dropped and a low quality oscillator is permissible.

But all this has a price. Shorter codes result in a lower margin between the auto correlation peak and the cross correlation peaks. That margin is difficult to calculate except for Gold codes. The margin can also be affected by the carrier frequency offset existing between the cross correlated signals if that offset is larger than several inverse integration times. For example, for a 1023 bit long Gold code with 1 msec integration time, doppler offsets of several kHz among the signals of the (at most) eleven satellites do exist. The margin discussed above is in the worst case about 3 dB lower than the margin given by the Gold bound expressing maximum cross correlation power P at zero frequency offset:

$$P = \left(\frac{2^{\left(\frac{10 + 2}{2} \right)} + 1}{2^{10}} \right)^2 \qquad \text{or } -23.9 \text{ dB.}$$

Obviously, given a set of 24 Gold codes, the exact margins for the applicable frequency offset range can be obtained.

For a code as short as 127 chips and corresponding short integration times, no applicable doppler offset can exceed several inverse times and the above effect does not exist. The margin expressed by the Gold bound then holds for all pairs of satellite signals; for a seven-register code, P equals:

$$P = \left(\frac{\left(\frac{7+1}{2}\right)}{\frac{2}{2^7} + 1} \right)^2 \quad \text{or } -17.5 \text{ dB.}$$

On the other hand, insensitivity to doppler means that now more satellite signals are capable of interfering. Since as many as eleven can be simultaneously visible more signal interference exists. Several dB in margin can be lost due to multiple signal interference.

In summary, the negative aspects of using a shorter code are that the margin between auto correlation peak and cross correlation peak is lowered. But, even for a short 127 chip code, a healthy 17.5 dB margin continues to exist. Even with more signals being capable of interfering, multiple code accessing capability is sufficient.

We conclude this section on shorter codes with some comparison curves for code acquisition in the worst case, i.e. all code phases must be searched and all frequency bins must be searched for the maximum doppler that possibly can occur and the threshold for detection of synchronization is a guaranteed 10 dB (even with the code time misalignment of one quarter chip and the frequency offset between incoming signal and locally generated signal equal to one-half the frequency band or equal to the inverse of the coherent integration time).

Figure A-1 presents the time required for a detection decision on a single code phase for:

a) Curve A: A short 127 chip code and incoherent sequential decision. The coherent integration time is taken to be about 125 μsec (corresponding to a 127 chip code length at 1.023 MHz clock rate). The result of each integration is the sum of in-phase and quadrature component energy and is added N times. The number N is chosen to represent a final probability of

detection of 0.9 and a probability of false alarm of 5×10^{-3}. The average value for N, denoted $\bar{N}$, can be derived from radar considerations and is given roughly by

$$-8 \log_{10} \bar{N} = \frac{E}{N_o} - 11 \quad ,$$

where $\frac{E}{N_o}$ is the ratio of energy per pulse from coherent integration to the one-sided noise spectral density expressed in dB. For example, for $C/N_o = 43.8$ dB Hz (which is 10 dB above the present GPS signal level) and $T_I = 0.125$ msec, $E/N_o = 4.8$ dB and $\bar{N} = 6$. Effectively one needs at least one more step to exceed the threshold. Plotted is therefore $T_I (\bar{N} + 1) = 0.9$ msec.

b) Curve B corresponds to the coherent integration time needed to obtain the required E/N_o when a filter bank is used with as many frequency bins as needed in parallel to cover the total doppler D assumed to be ± 4 kHz (user motion induced doppler and maximum doppler at mid latitude). The required E/N_o was held at 13 dB to obtain similar detection and false alarm capabilities.

c) Curve C is the total test time if only one filter exists which sequentially tests over all frequency bins. If T_I' is the coherent integration time the total time T_s equals

$$T_s = 2 T_I'^2 \, D.$$

d) Curve D simply represents the number of frequency steps corresponding to curve B and is inversely proportional to assumed C/N_o.
It is clear that short incoherent sequential detection is comparable to detection with coherent multiple velocity filter detection as far as detection time on a single code phase. The sequential nature of the detection is traded against multiplicity of filters albeit with a slight loss in efficiency.

Figure A-2 presents again curve A of Fig. A-1 compared to curve E representing the more realistic GPS case of coherent integration of over 1 msec and incoherent sequential integration to satisfy the E/N_o requirement of 13 dB combined with the use of a single velocity filter.

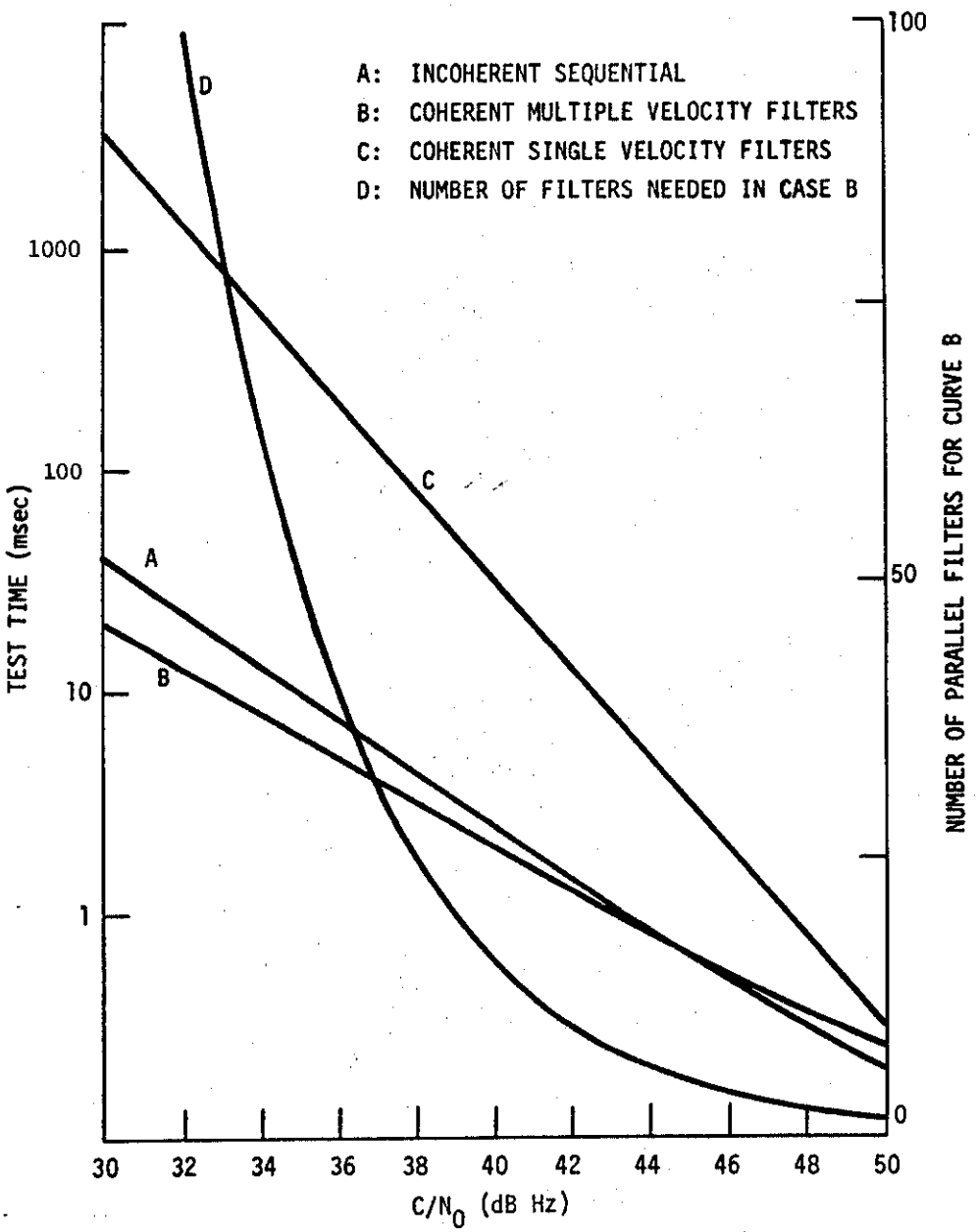


Fig. A-1. Test time on one code phase for initial code acquisition for maximum frequency uncertainty.

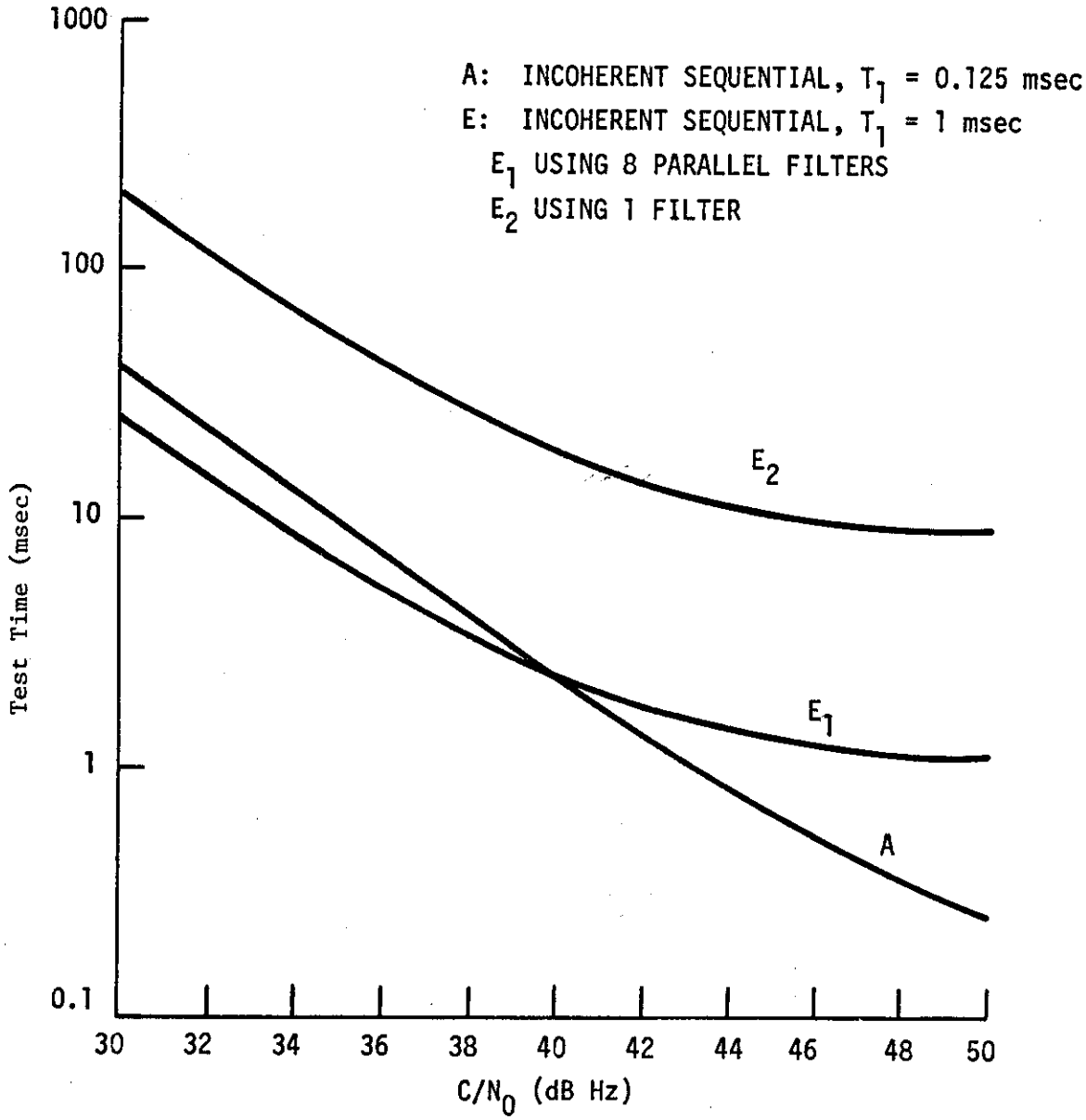


Fig. A-2. Test time for one code phase for initial code-acquisition for maximum frequency uncertainty.

It is clear that short codes are advantageous for code acquisition and for simplicity of acquisition circuitry due to their independence of doppler. The cost is a possible false lock due to the reduced cross correlation margin and higher sidelobes of the auto correlation function. Also, the range ambiguity interval is smaller for shorter codes and may require more software manipulations.

A.1.3 Use of Matched Filter

If instead of the sequential digital filter one uses a matched filter such as a CCD correlator, one does not have to test every code phase in one half chip increments and there is a corresponding saving in acquisition time. Programmable CCD devices are available for code lengths of 127. Programmable CCD devices for code lengths as long as 1023 will not be available in the near future. The short code and sequential incoherent detection method therefore can benefit from this technology to great advantage.

The whole operation not only of code synchronization but also of code tracking the PLL and demodulation could be done at baseband.

A.1.4 Improved Code Tracking

Once synchronization is established, despreading of the signal can be done for the purposes of demodulation of the data and obtaining a doppler reading. A feedback from the PLL "tunes" the locally generated signal to maximize the error readings for any choice of discrimination curve. These error readings are also a function of signal-to-noise. The chip duration Δ can be "split" more accurately with increased C/N_o or in other words more accurate range measurements can be expected.

Ranging error expressed as the variance of chiplength Δ in terms of C/N_o is (Ref. 9).

$$\sigma_\Delta = \sqrt{\frac{B_n}{C/N_o}\left(0.5 + k\ \frac{B_w}{C/N_o}\right)}^{1/2} ,$$

where $k = 0$ for coherent tracking; $= 1$ for a non-coherent DLL

B_n is one-sided bandwidth of the code-tracking loop

B_w is the inverse of the integration time of the predetection bandpass filter in the non-coherent DLL.

This equation shows for example, that a 10 dB increase in C/N_o gives a chip splitting capability at least three times better.

A.2 Performance of the Carrier Tracking Loop

The PLL is used principally to demodulate the data coherently. A higher C/N_o level provides better performance and allows larger bandwidth. The latter makes the system less sensitive to relative user/satellite dynamics. For a 2nd order loop the acceleration lag θ_{Lag} in response to a constant acceleration step $\Delta\dot{w}$ (rad/sec^2) is

$$\theta_{Lag} = \frac{\Delta\dot{w}}{w_m^2} \quad .$$

For a critically damped loop the loop bandwidth relates to the eigenfrequency w_m by $B_L = 0.53\ w_m$. The noise term contribution

$$\sigma_\theta = \sqrt{\frac{B_L}{C/N_o}} \quad .$$

It can be seen that the value E of total phase error

$$E = k\theta_{Lag} + \sigma_\theta$$

will be optimized for different B_L based on the weighting k and the assumed dynamics, but is strongly affected by the assumed C/N_o. Increased C/N_o reduces the phase error E and allows the loop bandwidth, which is optimized for the noise contribution and the lag component, to be larger. This is shown more clearly in Fig. 3 1 of Section 3 for the indicated weighting. The choice of weighting is obviously up to the designer. An acceptable choice for k is 2, allowing the PLL to cope well with dynamics as high as several g.

A.3 Data Demodulation

High C/N_o will allow better data demodulation. First, the larger bandwidth for the Costas loop will mean fewer phase slips of that loop when subjected to a dynamic situation while maintaining a small phase tracking error. The cosine of

that error affects the detection energy per bit collected over 20 msec. The energy is obviously proportional to C/N_o which can then be directly translated, for a DPSK data modulation, to bit error rates.

A.4 Satellite Scheduling

In the previous sections the sensitivity of acquisition and re-acquisition time to SNR was demonstrated. In a sequential receiver the latter depends also on cycle time. Let C denote the cycle time during which the four satellites selected for user position determination have been "visited".

For shorter C, less range uncertainty to a particular satellite will have accumulated while the signals from the three others are being tracked. In the software algorithms range and range rates are assumed to be tracked. Obviously shorter C means higher track update rate and correspondingly more accurate estimates of ranges, range rates, clock bias, user position and velocities due to the smoothing process applied to the data. The uncertainties on predicted range (predicted code delay) and range rate (doppler offset) are then mostly determined by unexpected accelerations.

Accelerations from satellite motion are less than one wavelength per second squared. The biggest accelerations derive from user motion and are limited to a few tens of meters per second squared. Range uncertainty accumulated over the three-quarter cycle time C therefore will seldom exceed one chip length (300m) for C up to 5 seconds.

Therefore re-acquisition may require only the checking of one to three half chip code phases. Fig. A-1 indicated the test time for high SNR still only to be of the order of tens of msec. Accelerations also cause frequency offset to accumulate linearly with cycle time during the three-quarter cycle time. To this one must add the oscillator error during the three-quarter cycle time. For a master oscillator of 10^{-7} Hz/Hz the frequency error accumulates at the rate of 160 Hz/sec. The DLL for code synchronization, for high SNR, can be designed with a wide enough bandwidth (e.g., 1 kHz) so that no frequency steps must be considered. The PLL will then require AFC aiding to cope with such large initial frequency offsets.

What is saved on oscillator cost therefore may well have to be paid in a more complicated PLL unless one reduces the cycle time drastically to, for example, two seconds. The disadvantage here is that the effective time for acquiring data becomes very small and message assembly of out-of-sequence portions, with no guarantee of message completeness due to missing data bits, becomes difficult.

Since tracking and data collection must occur simultaneously, the better approach is for longer cycle times, and a better oscillator (e.g., 10^{-8} Hz/Hz).

Another point to be emphasized is, that if ranging accuracy can be relaxed the DLL which provides the ranging potential can be simplified. Instead of using parallel channels to provide early/late/on time outputs of the local code generator, a well known timesharing method called tau dither can be used with only slight degradation in performance.

APPENDIX B

DETAILS OF TDMA PULSED GPS RECEIVER

In Section IV a design was described which represents the class of TDMA Pulsed waveforms. Several small studies were performed to determine the feasibility and simplicity of the approach.

In this Appendix we include some details on the sync processing and data processing logic and how the logic translates into hardware. The flow charts and diagrams are otherwise self explanatory.

Figure B-1 presents the processing waveform. Fig. B-2 shows the sync processing flow chart translating into logic cards D_1 to D_3 in Figs. B-3 to B-5. Fig. B-6 shows the Data Demodulator counter states and Fig. B-7 summarizes the layout of the GPS receiver.

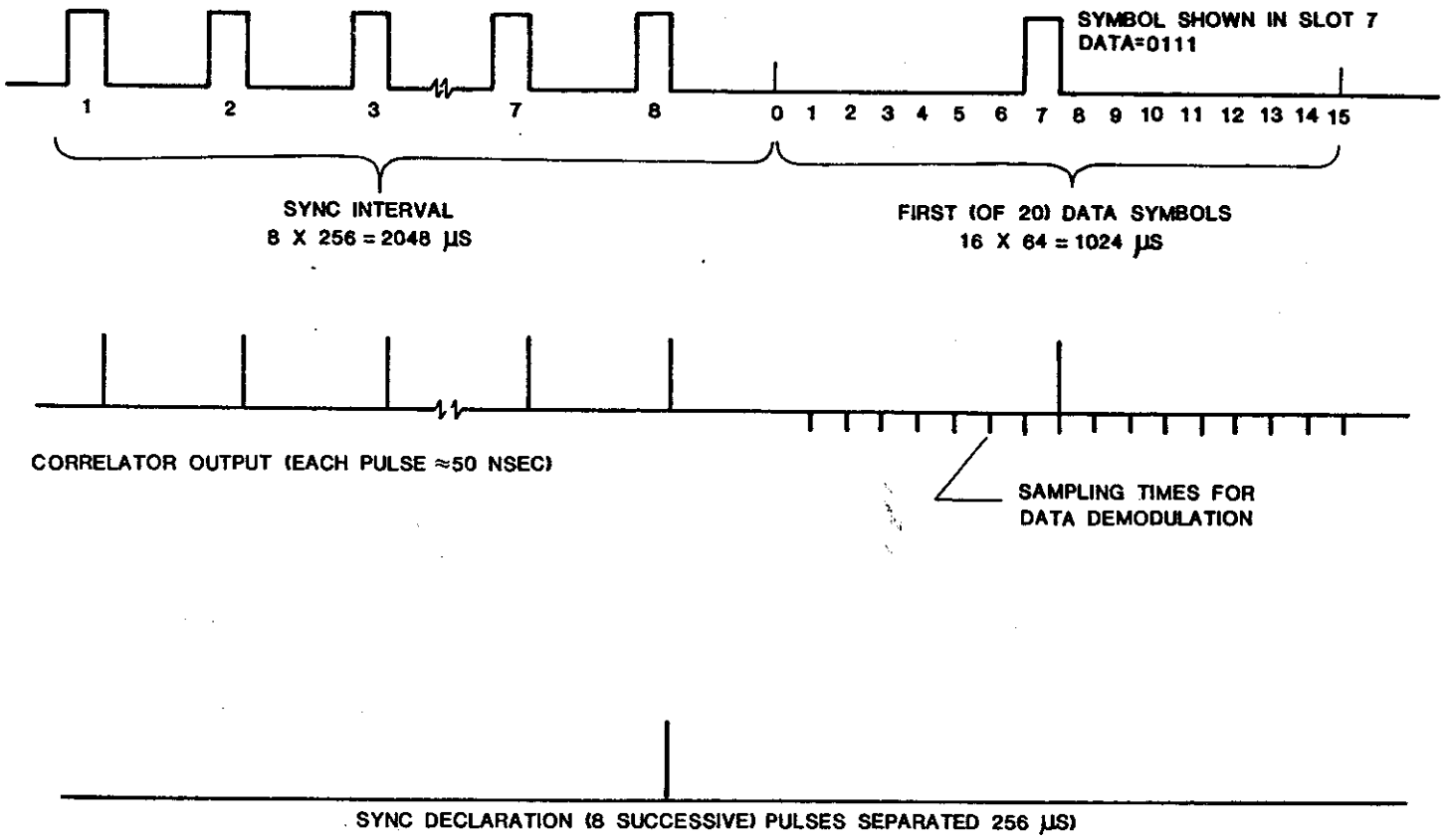


Fig. B-1. Processing waveforms.

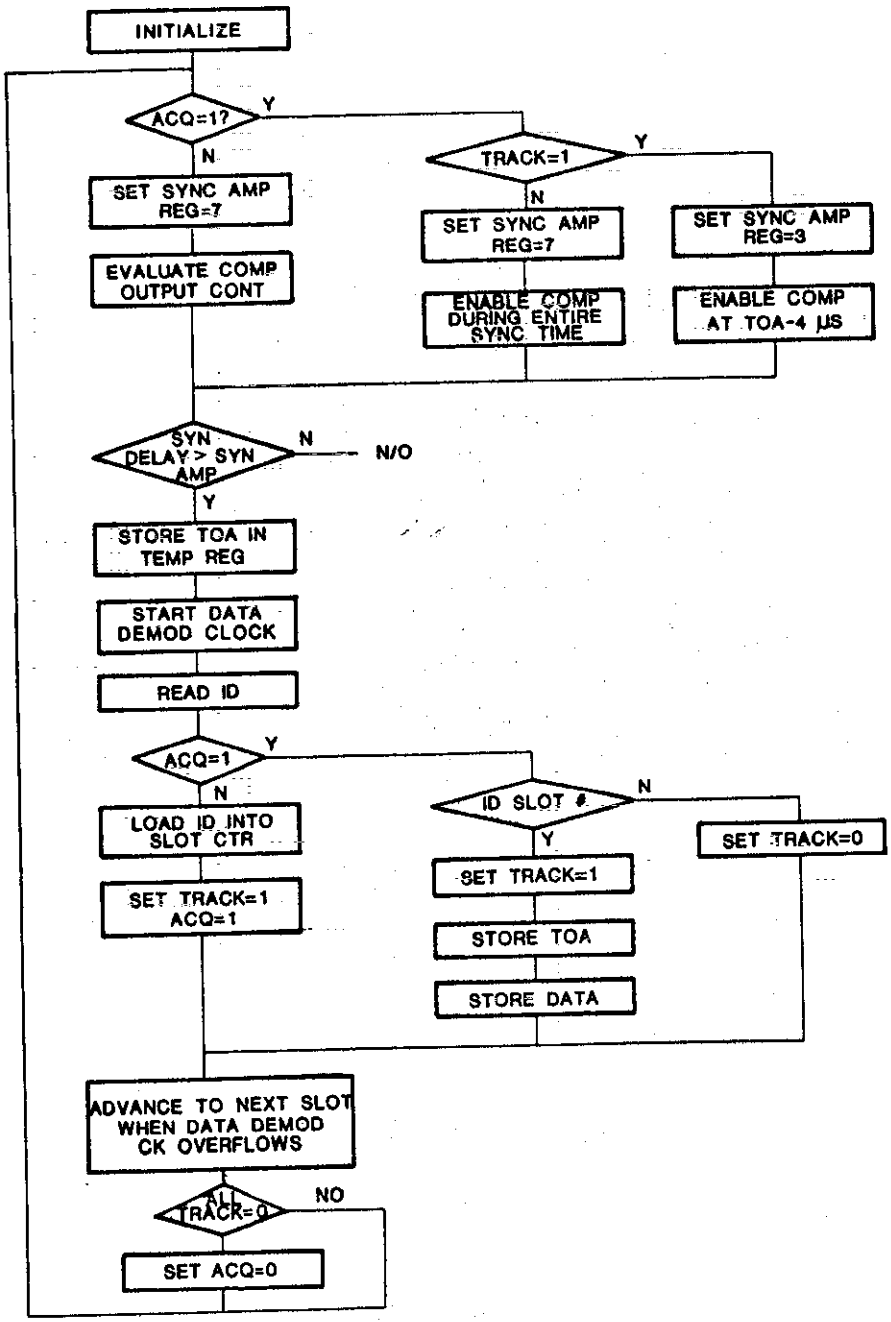


Fig. B-2. Sync processing flow chart.

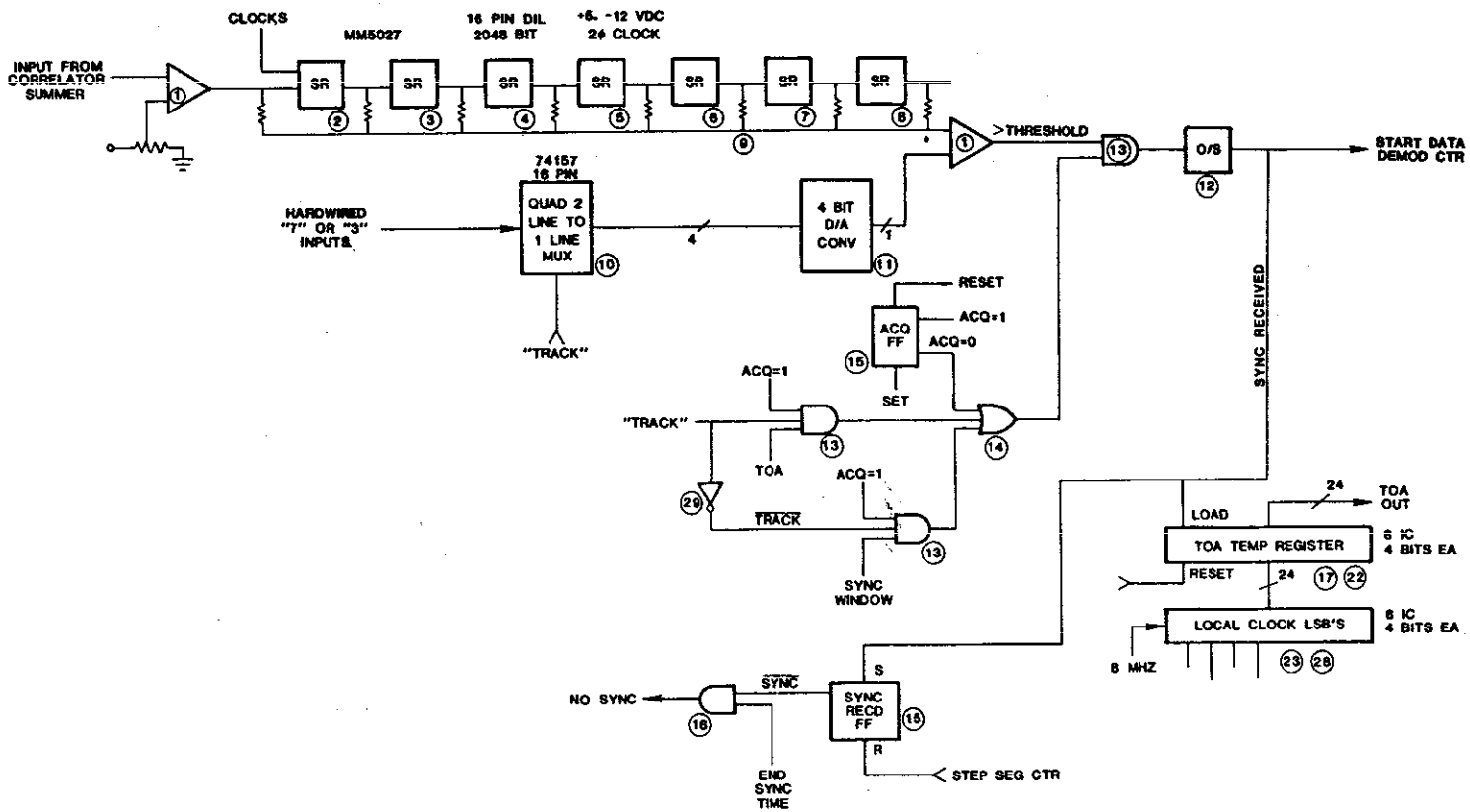


Fig. B-3. Special purpose logic card D-1.

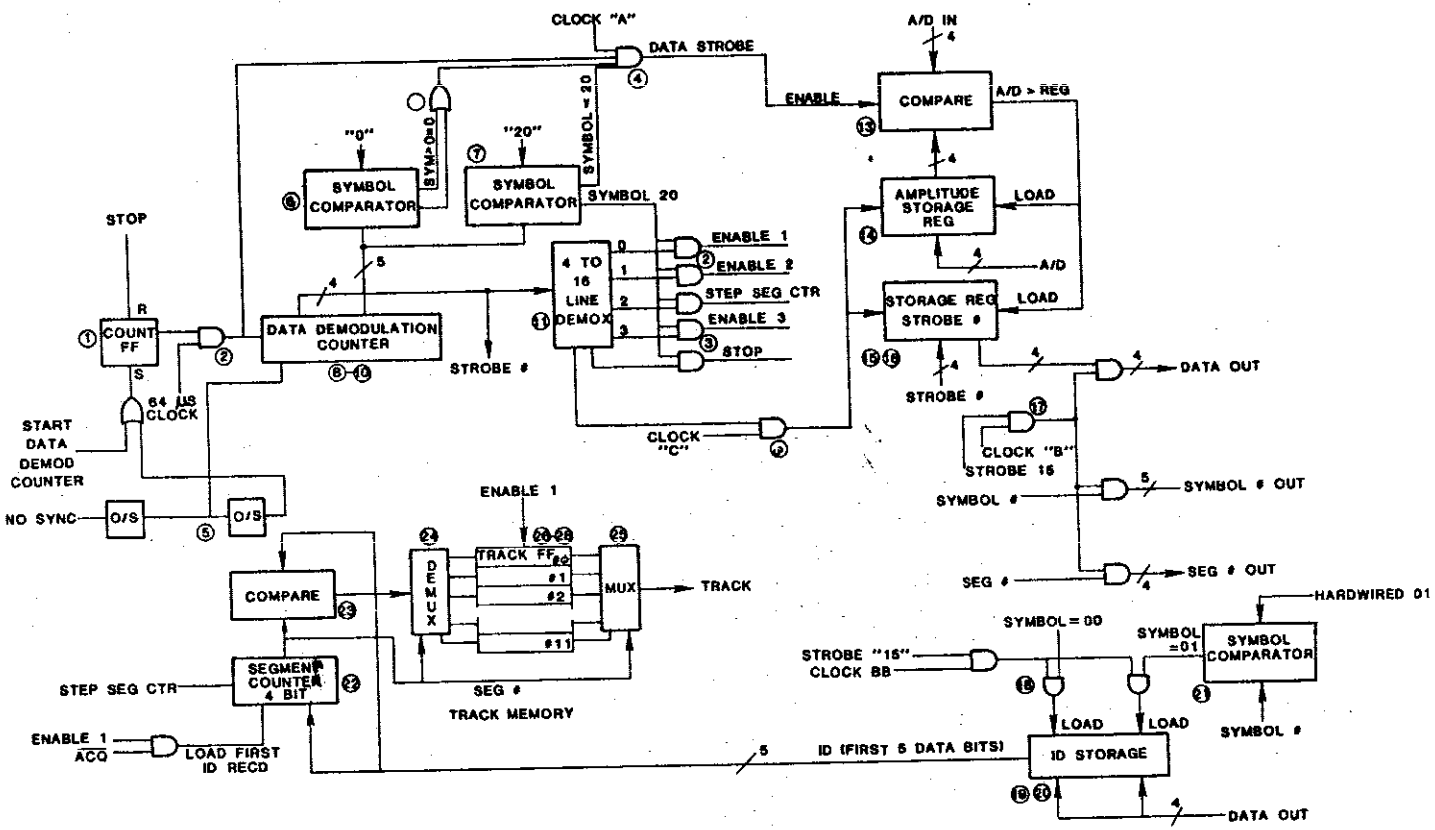


Fig. B-4. Special purpose logic card D-2.

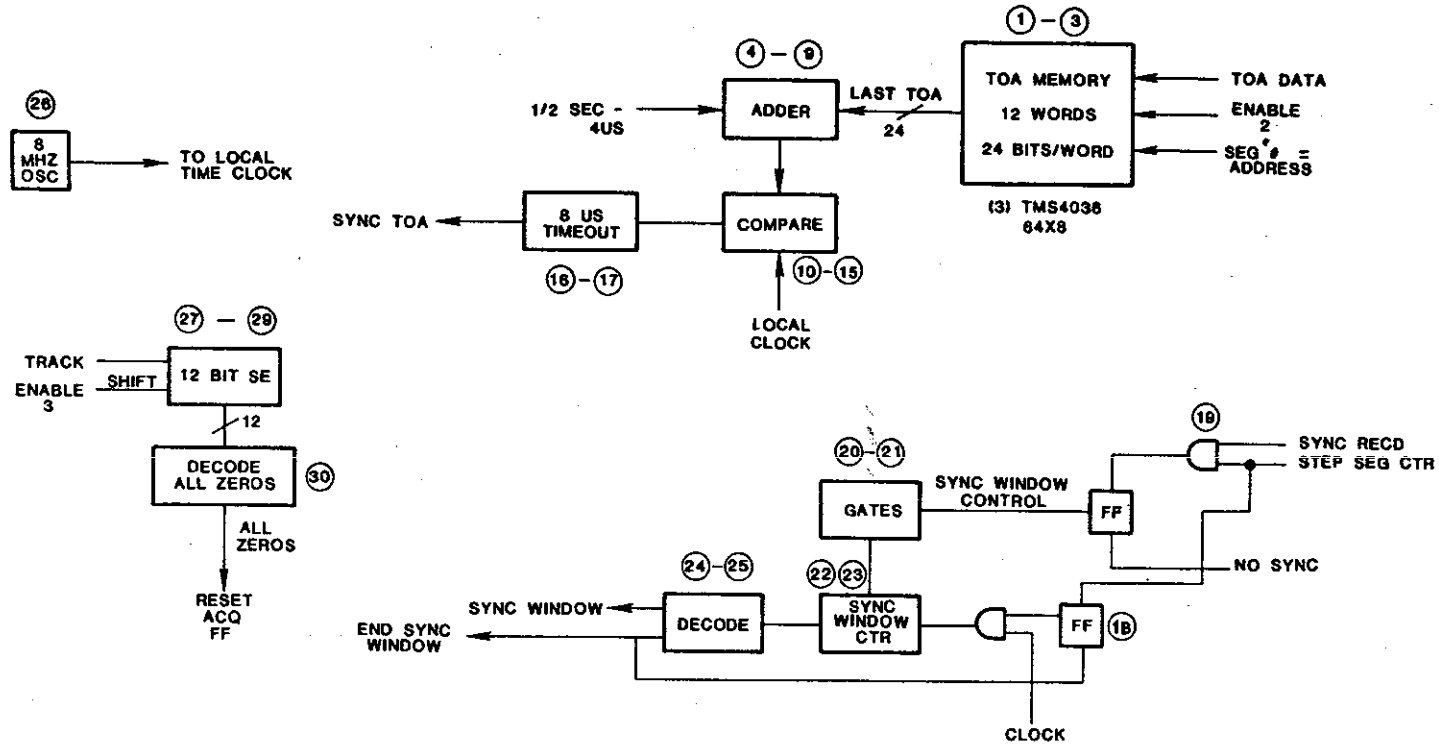


Fig. B-5. Special purpose logic card D-3.

COUNT BASE 10	COUNT	BINARY	ACTION
	SYMBOL	STROBE #	
-4	11111	1100	START
0	00000	0000	DATA STROBE ZERO, SYMBOL 1
1	00000	0001	ONE
2	00000	0010	TWO
⋮			
15	00000	1111	FIFTEEN
16	00001	0000	DATA STROBE ZERO, SYMBOL 1
17			ONE
⋮			
32	00010	0000	DATA STROBE ZERO, SYMBOL 2
⋮			
319	10011	1111	DATA STROBE FIFTEEN, SYMBOL 19
320	10100	0000	ENABLE 1
321			2
322			STEP SEGMENT COUNTER
323			3
324			
⋮			STOP

Fig. B-6. Data demodulator counter states.

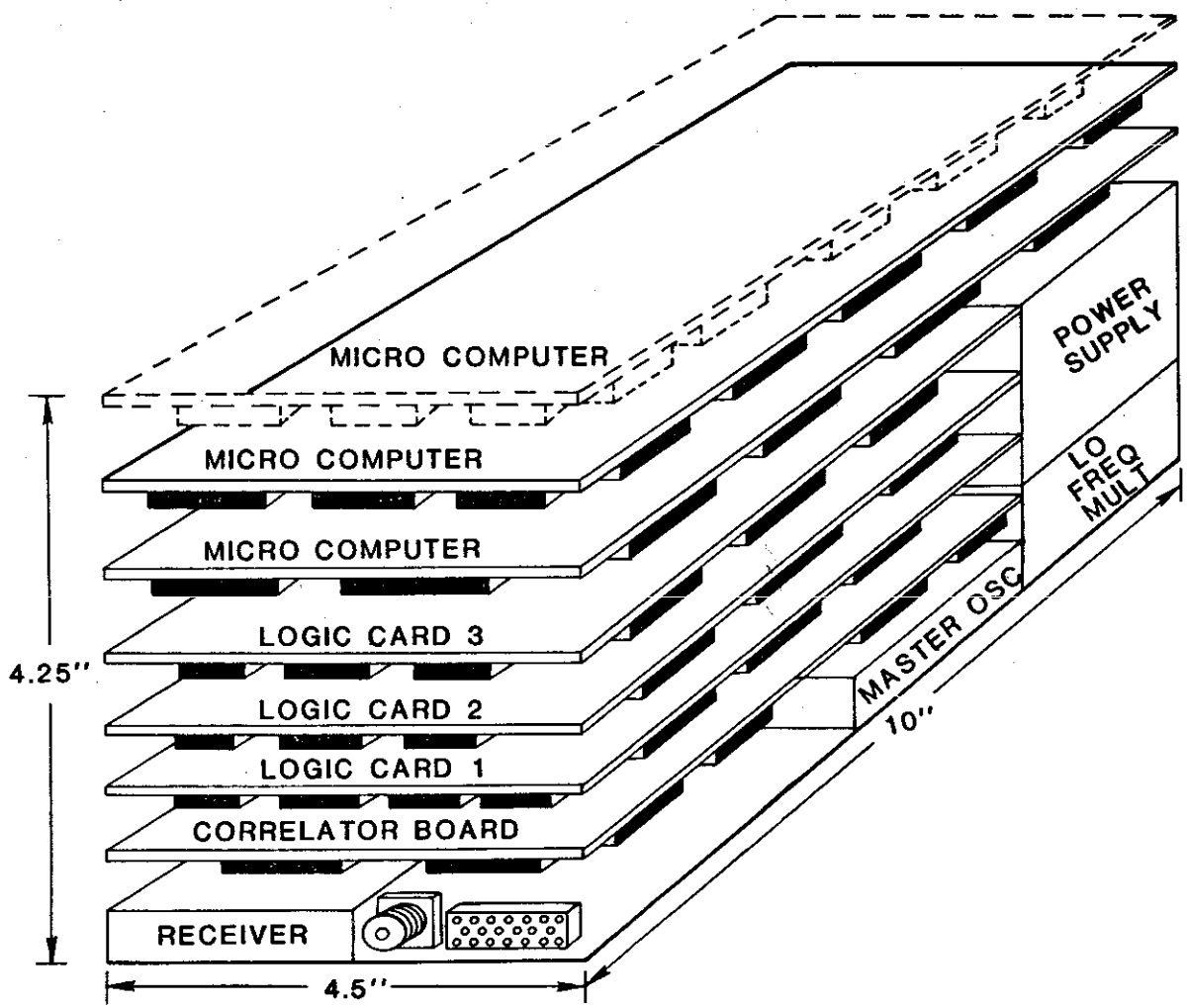


Fig. B-7. Pulsed GPS receiver: case removed.

APPENDIX C

POWER DISTRIBUTION IN TONES FOR GPS TONE RANGING SYSTEM

In the tone ranging system described in Section 5 the FM or PM modulation index is chosen so that the power in the fundamental spectral lines is maximized. In the tone ranging method, no conventional FM or PM demodulation is used. Instead, a PLL is applied to recover the specific tone, after heterodyning the modulated carrier to baseband. That allows us to realize an increase of 3 dB in loop SNR over the SNR at IF. It is important to maintain sufficient power in the carrier frequency spectral line during tone modulation in order to allow the carrier PLL to track with sufficiently small phase-error. The signal amplitude of the tone is proportional to the cosine of that phase error.

In this appendix we derive an expression for the SNR in the tone PLL as a function of modulation index. The index is then chosen to maximize the ratio $\frac{T}{N_o}$, signal-to-noise power density ratio at the input of the tone PLL to the ratio $\frac{C}{N_o}$, signal-to-noise power density at the input of the receiver.

In the following derivation C(t) will represent the received RF waveform, the set f_i' are the tones, f_c is the carrier frequency, n(t) is the noise component of the RF signal.

Let the carrier, frequency f_c, be phase modulated by the sum of N tones f_i'. The received RF waveform is then

$$C(t) = A \cos\left[2\pi f_c t + \phi + \sum_{i=1}^{N} \Delta \cdot \sin(2\pi f_i' \cdot t)\right] + n(t)$$

where $$n(t) = n_c(t)\cos\left(2\pi f_c t\right) - n_s(t)\sin\left(2\pi f_c t\right)$$

The noise components n_c and n_s are low pass, gaussian with zero mean, and standard deviation equal to standard deviation of n. If n has single-sided spectral density N_o in the RF band B, then n_c and n_s have density $2N_o$ in lowpass band $\frac{B}{2}$. ($N_o = kTN_F$, N_F is noisefigure). In terms of the received waveform the signal-to-noise power density ratio $\frac{C}{N_o}$ equals:

$$\frac{C}{N_o} = \frac{A^2}{2N_o} \quad .$$

Let $Q \triangleq A \prod_{i=1}^{N} J_o(\Delta_i)$, where J_i are Bessel functions of i^{th} order,

then $$C(t)-n(t) = Q\cos(2\pi f_c t + \phi) + Q \sum_{i=1}^{N} \frac{J_1(\Delta_i)}{J_o(\Delta_i)} \cos\,[2\pi(f_c + f_i')t + \phi]$$

$$- Q \sum_{i=1}^{4} \frac{J_1(\Delta_i)}{J_o(\Delta_i)} \cos\,[2\pi f_c - f_i')t + \phi] + \ldots.$$

higher order terms.

The higher order terms are harmonics and intermodulation products at spectral frequencies

$$f_c + \sum_{i=1}^{N} k_i f_i'$$

for all possible combinations $[k_1, k_2, k_3, \ldots k_N]$ for positive and negative integers k_i .

Assuming that an accurate carrier reference is obtained during acquisition, a quadrature reference

$$b(t) = a \sin\,(2\pi \hat{f}_c + \hat{\phi})$$

can be formed. For a PLL $f_c = \hat{f}_c$, $\hat{\phi} = \phi + \delta\phi$, $\delta\phi$ being the carrier phase tracking error.

The received signal C(t) is heterodyned to baseband by this quadrature reference b(t). Let $C'(t) \triangleq C(t)-n(t)$; then

$$< C'(t)\, b(t) >_{\text{Low Pass}} = - \frac{aQ}{2} \sum_{i=1}^{N} \frac{J_1(\Delta_i)}{J_o(\Delta_i)} \sin\,(2\pi f_i' t - \delta\phi)$$

$$+ \frac{aQ}{2} \sum_{i=1}^{N} \frac{J_1(\Delta_i)}{J_o(\Delta_i)} \sin(-2\pi f_i' t - \delta\phi)$$

$$= - aQ \sum_{i=1}^{N} \frac{J_1(\Delta_i)}{J_o(\Delta_i)} \cdot \cos\delta\phi \sin(2\pi f_i' t).$$

Observe that the amplitude of the tone fundamental at baseband is, for tone f_i' ,

$$\cos\delta\phi \cdot aQ \frac{J_1(\Delta_i)}{J_o(\Delta_i)} = aJ(\Delta_1).J_o.(\Delta_2)..J_1(\Delta_i)..J_o(\Delta_N) \cdot \cos\delta\phi$$

We now discuss the term n(t) heterodyned to baseband :

$$< n(t) \quad b(t)>_{\text{Low Pass}} = \frac{a}{2} (n_c \sin\phi - n_s \cos\phi)$$

This is low pass noise over bandwidth $\frac{B}{2}$ and spectral density $\frac{a^2}{4} 2 N_o = \frac{a^2 N_o}{2}$.

The tone f_n' is now isolated in a PLL with tone-to-spectral noise density $(\frac{T}{N_o})_n$; of

$$\left(\frac{T}{N_o}\right)_n \triangleq \frac{\frac{1}{2}\left(aQ \frac{J_1(\Delta_n)}{J_o(\Delta_n)}\right)^2 \cos^2\delta\phi}{\frac{a^2 No}{2}} = \frac{A^2}{N_o} J_o^2(\Delta_1)..J_1^2(\Delta_n)..J_o^2(\Delta_N) \cos^2\delta\phi ,$$

or, since $\frac{C}{N_o} = \frac{A^2}{2N_o}$,

$$\left(\frac{T}{N_o}\right)_n = 2 \frac{C}{N_o} J_o^2(\Delta_1)..J_1^2(\Delta_n)...J_o^2(\Delta_N)\cos^2\delta\phi .$$

A similar analysis shows that the $\left(\frac{C}{N_o}\right)_{\text{carrier}}$ at the RF frequency f_c is

$$\frac{C}{N_o} \prod_{i=1}^{N} J_o^2(\Delta_i) \cdot \cos^2\delta\phi .$$

This is the power ratio available to the carrier PLL for carrier tracking and must be sufficient to maintain accurate phase track since the cosine of the phase error scales the power available to the tone PLL.

The method of clustered tones allows the range accuracy to be determined from the average of all tone phases while ambiguity resolution is based on the sum or difference of two tones implying a range error which is a factor $\sqrt{2}$ worse than the error associated with a single tone. This effectively favors the finest of the fundamental set of tones compared with all other tones and is a desirable feature. For simplicity one may allow the clustered tones to have equal modulation indices Δ.

For any tone PLL one may then write

$$\left(\frac{T}{N_o}\right)_n = k_n \frac{C}{N_o} = k \frac{C}{N_o}$$

where

$$k = 2 \cos^2\delta\phi. \quad J_o^{2N-2}(\Delta)\, J_1^2(\Delta)$$

For $\delta\phi$ = o and N = 4, Fig. 5-3 in Section 5 shows k as a function of modulation index Δ. Factor k is sometimes called the suppression factor. The optimum Δ is .7 for minimum suppression. However, as can be seen, the minimum is fairly flat for Δ between 0.5 and 0.9. One may favor smaller Δ to limit the maximum phase deviation (2.8 radians for 4 tones and Δ = 0.7) or limit the effective bandwidth which is occupied. The approximate bandwidth of the RF signal is given by the Carson rule ,

$$2(\Delta_1 + 1)f_1 \ ,$$

since all tones are clustered around and below f_1. The SNR in the tone PLL, for $\frac{C}{N_o}$ equal to the projected 45 dBHz, and Δ = .7 would equal 38.1 dBHz. The design assumed 35 dBHz to allow about 3 dB of margin.

REFERENCES

1. H. B. and B. Goode, "Technical Assessment of Satellites for CONUS Air Traffic Control," Project Report ATC-26, Vol. III, Lincoln Laboratory, M.I.T. (19 February 1974), DOT-TSC-RA-3-8-3.

2. "Global Positioning System Spartan Receiver/Processor," Rockwell Space Division No. SD75-GP-006, April 1975.

3. "Design Development Study for a Global Positioning System Spartan Set," NRL 85001042 4 Sept. 1975.

4. A. Rihaczek Principles of High-Resolution Radar (McGaw-Hill, New York, 1969).

5. A. Rihaczek "Radar Waveform Selection A Simplified Approach," IEEE Trans. Aerospace Electron. Systems Vol. AES-7, 6 (1971).

6. G. C. Kronmiller and E. J. Baghday, "The Goddard Range and Range Rate Tracking System: Concept, Design and Performance," Intern. Space Electron. Symposium, Miami, Fla. 1965.

7. H. Salwen, "Data Analysis and Design Studies in Support of the Aerosat Program" TSC File Material, 1976.

8. H. Van Trees, Detection, Estimation and Modulation Theory, Part I (Wiley, New York, 1968).

9. C. R. Cahn and E. H. Martin "Design Considerations for a Spread Spectrum Navigation Receiver," Intern Conf on Electron. System and Navigation Aids, Paris, 14-18 November 1977.