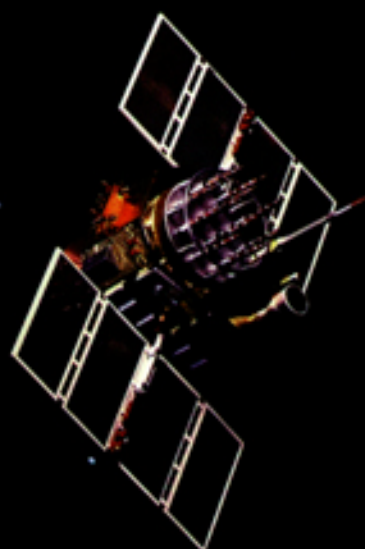


UNDERSTANDING

THE GPS

An Introduction to the
Global Positioning System

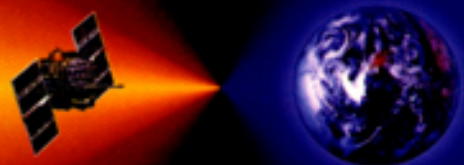
What It Is and How It Works



G109.5
F73
1996

GREGORY T. FRENCH





UNDERSTANDING THE GPS is a comprehensive, easy-to-understand, fully illustrated presentation of the Global Positioning System (GPS). Emphasizes practical, need-to-know information for real-world applications.

Whether a professional or student, technician or manager, this is your best introduction to the subject, written by a widely published Master GPS Instructor and Senior Project Manager with GeoResearch Inc.



Published by
GeoResearch, Inc.

*The leader in integrated GPS/GIS applications
and field data collection systems.*

ISBN 0-9655723-0-7



UNDERSTANDING

THE
GPS

An Introduction to the
Global Positioning System

What It Is and How It Works

By
GREGORY T. FRENCH
GeoResearch, Inc.

FIRST EDITION

Copyright ©1996, Gregory T. French. All rights reserved.

No part of this work may be reproduced in any form, or by any means, without the permission of the publisher. Exceptions are made for brief excerpts to be used in published reviews.

Published by

GeoResearch, Inc.
8120 Woodmont Avenue, Suite 300
Bethesda, MD 20814

Library of Congress Catalog Card Number: 9680018

ISBN: 0-9655723-0-7

Printed in the United States of America

This book is available at quantity discounts for bulk purchases.

Supplemental materials for instructors and trainers are available in various media.

For information, call 800-GEOLINK or 301-664-8000.

Preface

There is an ever-growing supply of information about the Global Positioning System. Unfortunately, these new (and now, some not so new) documents seem to be located at each end of the comprehension scale: either at the “gee-whiz” level which basically describes how interesting and useful this new utility is, or at the engineer’s level which starts out with Keplerian orbits and Hopfield Modeling. What seems to be missing is a comprehensive, yet easy to understand, presentation of the Global Positioning System (GPS) for people who may have a very real need to apply this new technology but lack the basic understanding necessary to make important, and often expensive, decisions about it. Thus this book.

This book is designed to support an introductory course on the fundamentals of the Global Positioning System based on a series of graphic representations and distilled concept-bullets. Math is scrupulously avoided—that level of information is readily available through numerous highly technical publications and is no more necessary for most users than is a textbook on electronics necessary for the purchaser of a television set.

Each concept is presented in one to four graphics found in this book on the left page of each page-pair. The opposing right page presents a brief discussion of the concept. While much more could be said on each of the topics presented, only those highlights considered by the author to be of most immediate value to the geographer, project manager, field technician, or others needing to learn the fundamentals of the GPS are included here. At the end of the book, there is a list of suggested readings for those who are interested in gathering more in-depth and detailed information on most of the topics covered.

Errata

Page 12. Graphic shows VOR, Transit, ILS, and GPS incorrectly located along the electromagnetic spectrum. This has been corrected in the presentation packages (overheads and 35mm slides).

Page 83. Paragraph three should read:

Although that is the *theoretical* maximum resolution possible in carrier-phase positioning, modern geodetic surveying receivers are regularly achieving testable and repeatable accuracy in the area of one to two centimeters, or 10 to 20 millimeters, at a 95% probability level. Some claim even higher accuracy.

Page 103. Paragraph two, first sentence should read:

PDOP, or *Position Dilution of Precision*, probably the most commonly used, is the dilution of precision in three dimensions.

Page 144, 145. NOAAJCORS has recently changed the web pages to make navigation easier. Therefore, the graphic and navigation instructions no longer accurately represent the current pages. The address remains the same. This has been updated in the presentation packages (overheads and 35mm slides).

Page 168. Graphic should read:

THE LATEST AND GREATEST BEST FIT ELLIPSOID IS
The World Geodetic System of 1984

This has been corrected in the presentation packages (overheads and 35mm slides).

Page 169. First sentence, first paragraph should read:

The latest and greatest best-fit ellipsoid is the World Geodetic System of 1984, or WGS84.

Contents

Part I Introduction and Background

Introduction	3
Topics	7
What is GPS?	9
Radio-Navigation Systems	11
Evolution of the GPS	15
GPS Civil Applications	19
GPS Segments	21
Control Segment	23
Control Segment Locations	25
Space Segment	27
Orbits	29
Launch History	31
How Does GPS Work?	33
Two-Way vs. One-Way Ranging	35
Single Range to Single Satellite	37
Two Ranges to Two Satellites	39
Three Ranges to Three Satellites	41
Why Four Satellites?	43
Clock Timing Error	45

Part II Basic Signal Structure and Error

Levels of GPS Service	53
Basic Signal Structure	55
Pseudo-Random Codes	59
Where Are the Satellites?	65
GPS Signal Structure Map	67
Signal Strength	69
GPS Resolution C/A-Code	71
GPS Resolution P-Code	73
Anti-Spoofing (A/S)	75

Carrier-Phase Positioning	79
GPS Resolution - Carrier-Phase	83
GPS Velocity	85
GPS Error Budget	87
Ionospheric / Tropospheric Refraction	89
Satellite Mask Angle	91
Multi-Path Errors	93
Selective Availability	95
Dilution of Precision (DOP)	97
Project Planning	105
Position Offsets	109
Almanacs	117
Absolute Accuracy	119

Part III Data Correction Techniques and High-Resolution Accuracy

Differential Correction	123
Post-Processed Corrections	127
Real-Time Corrections	131
Post-Processing vs. Real-Time	135
Differential Data Sources	137
C.O.R.S. Network	139
U.S. Coast Guard Beacons	145
U.S.C.G and A.C.O.E Radio-Beacon Coverage	147
W.A.A.S	149
Commercial Geostationary Satellites	151
Real-Time FM Sub-Carriers	153
Predicted Coverage for FM DGPS	155
Other Improvement Techniques	157
Accuracy	159
Error Terms	161

Part IV Basic Geodesy, Data Collection Techniques and GPS Applications	
Geodetic Coordinate Systems	165
Ellipsoid vs. Geoid	167
WGS84	169
What's So Special About GPS Heights?	171
Geodetic Heights	173
Data Collection Techniques	175
Points vs. Positions	179
Lines From Points	181
Areas From Points	183
Differential Applications	185
Geographic Information Systems	187
GPS/GIS Applications	189
Aerial Photo Control	191
Satellite Imagery, GPS and GIS	193
Geographic Features	195
GPS GIS Point Data Capture	197
GPS GIS Line Data Capture	199
Areas From Points	201
External Data Source	203
GPS Surveying	205
GPS Navigation	207
IVHS	209
Receiver Types	219
The Future of GPS	221
Appendix I Glossary	227
Appendix II Suggested Readings	247
About the Author	255

PART I

INTRODUCTION

AND

BACKGROUND

Since Earliest Times, We've Been Trying To
Figure Out Where We Are And Where We're Going



Introduction

Since earliest time, humankind has concerned itself with where it's at and where it's going. Some of the earliest techniques that travelers used were simple rock cairns marking the trail, either for finding their way back, repeating their path, or for others to follow. This technique is still used today. The problems with it, however, are obvious. What do you do if snow covers them? How do you identify one path vs. another? In any event, the vagaries of nature insure that the markers are not likely to last very long unless they are indeed substantial (as many were).

A better method was to record this spatial information on a clay tablet or piece of parchment which could be copied and handed from one person to another. We call these maps. The first recorded maps date back to the Mesopotamians some 5,000 years ago, constituting a revolution in geographic positioning that has enjoyed widespread use ever since. While the technology behind cartographic techniques has improved many orders of magnitude over the centuries, conceptually they remain fundamentally the same even today.

Today, we live in a world of precision. We expend great amounts of intellectual and monetary currency on ever-smaller units of measurement. Knowledge of where we are and where we are going has, for the past several thousand years, relied on highly trained and skilled surveyors. The science of surveying has achieved phenomenal levels of precision. But, unfortunately, only for those very few whose needs have outweighed the ever-increasing cost necessary to achieve that precision.

The Ultimate Achievement of That Desire
To Define Location Is The

Global Positioning System

Or Simply

GPS

The Ultimate Achievement

The ultimate achievement of humankind's urge to know where he or she is at, at extraordinarily high levels of precision, is manifested in today's Global Positioning System. Those of you who have grown up with *Star Trek* may find the idea of simply flipping open a small device to locate where you are on the planet something of a yawner. You're already used to the idea. The fact is this technology represents a true revolution, comparable in scope to the invention of the accurate ship-board clock that heralded the age of global circumnavigation of the 1700's.

Today, GPS is causing a renaissance of the navigation, surveying and mapping professions and may, within only a few years, completely replace conventional methods of transportation navigation and land surveying. The uses and implications of the GPS system are yet to be fully realized, and new applications are being found at an ever-increasing rate. Such diverse areas as natural resource management, mineral exploration, transportation, fleet management, agriculture, shipping, utilities, disaster mitigation, and public safety are all areas where GPS is rapidly becoming critically important. GPS is even being used to test Einstein's theory of relativity, as well as a tool to measure gravity to previously unheard of levels of precision and accuracy. Clearly, there is a geographic revolution underway, and the instrument of that revolution is what this book is about.

Topics

Part 1

**Introduction to the GPS
Background and History**

Part 2

**Basic Signal Structure
Basic Error and Accuracy**

Part 3

**Data Correction Techniques
High Resolution Accuracy**

Part 4

**Basic Geodesy
Data Collection Techniques
GPS/GIS Applications**

Topics

This book is broken into four broad sections with each topic building on the one before.

Part I Introduction and Background

This first section will introduce you to the basic concepts of what the GPS is, what it's meant to do, and the fundamentals of how it works. We will also take a brief look at the events that have led to the development of the Global Positioning System as it exists today.

Part II Basic Signal Structure and Basic Error and Accuracy

In this section, we will examine the actual signal structure that the GPS satellites (frequently referred to as SVs, or Space Vehicles) use to determine a user's position. In addition, basic sources of error and consequent real-world accuracies will be examined.

Part III Data Correction Techniques and High-Resolution Accuracy

This section will explore some of the more sophisticated methods by which GPS errors can be corrected and what levels of high-resolution accuracy can be expected as a result.

Part IV Basic Geodesy, Data Collection Techniques, and GPS Applications

In this final section, you will be introduced to the basics of Geodesy, or the study of the shape of the Earth, necessary to understanding what the GPS measurements are referenced to. We will also look at some of the techniques of GPS data collection used in the "real" world, as well as some of the ways GPS is being used today and what we might expect of it in the near future.

What Is GPS?

**GPS Is A Space-Based Navigation System
Designed By The U.S. Military To Provide:**

- **Autonomous Geo-Positioning**
- **10 - 20 Meter Accuracy**
- **Worldwide Coverage**
- **Availability 24 hours per day**
- **Military Security**
- **Low End-User Cost**
- **Receivers To Each Soldier**
- **Installation On Every Vehicle**

What Is GPS?

We begin with the most basic question: What is the Global Positioning System? The Global Positioning System is a space-based navigation and positioning system that was designed by the U.S. Military to allow a single soldier or group of soldiers to autonomously determine their position to within 10 to 20 meters of truth. The concept of autonomy was important in that it was necessary to design a system that allowed the soldier to be able to determine where they were without any other radio (or otherwise) communications. In other words, with a single, one-way receiver whose use could not be detected by potential hostiles.

Since the U.S. Military is truly a global force, it was further necessary that the system provide worldwide coverage, and that the coverage be available 24 hours a day. At the same time, it had to be militarily safe in that the U.S. Military had to have the ability to deny any hostiles' use of the system without degrading their own use.

Ultimately, it is planned that each soldier and each military vehicle will be equipped with a GPS receiver. Therefore, it was necessary that the receivers be sufficiently low in cost to meet this end. Once all soldiers are so equipped, dependence on all other systems could eventually be phased out.

Radio-Navigation Systems

Ground-Based

System	Freq.	W/L	Error 2D	Coverage
● <i>Omega</i>	10-13 KHz	26 Km	3-6 Km	Global
● <i>Loran</i>	100 KHz	2.5 Km	460 M	~10%
● <i>VOR/DME</i>	108-118 MHz	~2.5 M	60-180 M	N.A. Tot
● <i>ILS</i>	330 MHz	~1 M	5-10 M	Limited

Space-Based

System	Freq.	W/L	Error 2D	Coverage
● <i>Transit</i>	150-400 MHz	2M-73Cm	460 M	Global
● <i>GPS</i>	1575.42 MHz	19 Cm	<100 M	Global

There Are Others But Generally Not For Public Use

Radio-Navigation Systems

GPS is far from being the only radio-navigation system that exists. Even before the Second World War, various schemes were attempted to provide crude positioning for ships and airplanes. Each new system built on the previous system, with each increasing the accuracy, and/or range of usability. Several systems developed during World War II are still in use today, albeit much more refined than in their earlier incarnations.

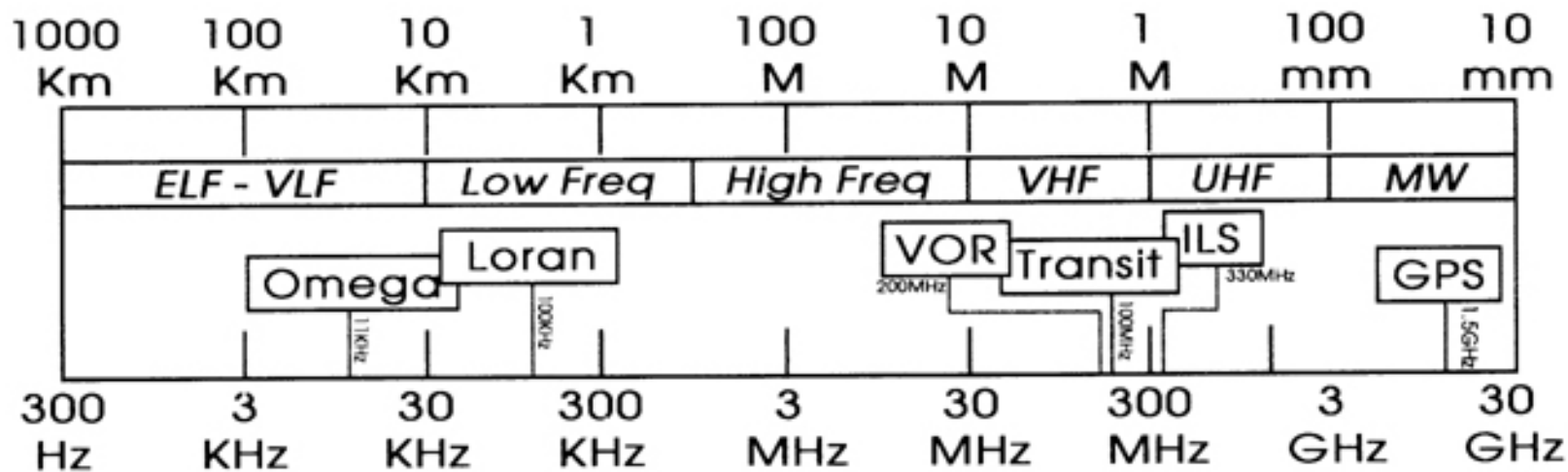
Today, there are at least a half-dozen different radio-navigation systems including *Omega*, *Loran*, *VOR/DME*, *ILS*, *Transit*, and, of course, the GPS. The first four are ground-based systems; the Transit and GPS systems are both space-based. The Russians also operate a system called *GLONASS* that is similar to GPS but has so far been far less reliable. Though slowly gaining in importance, it will not be covered in this book.

The ground-based Omega and Loran systems are very similar in that they both employ difference-of-arrival techniques, with Omega measuring the phase difference and Loran measuring the time difference of the signals from two or more transmitters. These transmitters send out very low frequency carrier waves that are very long—26 kilometers for Omega; 2.5 kilometers for Loran. The advantage is that the long wavelength is able to “tunnel” through the atmosphere by “bouncing” off of the bottom of the ionosphere (a layer of electrically charged particles in the upper atmosphere) for great distances. This phenomenon is known as “*Wave-Form Ducting*.” In fact, this phenomenon is so effective, full global coverage is achieved by Omega with only eight transmitters. The disadvantage is low precision due to the long wavelength: six kilometers of potential error for Omega. While Loran’s precision is as high as 450 meters, only some 10% of the globe is covered by Loran “Chains.”

Aviation systems such as the *VOR/DME* (Very High Frequency, Omnidirectional Ranging/Distance Measuring Equipment) and *ILS* (Instrument Landing System) systems operate at much higher frequencies and consequently provide much higher precision; on the order of 60-80 meters for *VOR/DME*, to less than 10 meters for *ILS*.

Radio-Navigation Systems

Wavelength



Frequency

Frequency and Precision

Higher frequency produces higher precision. However, it also requires line-of sight since the higher frequency wavelengths “punch” right through the ionosphere rather than bounce off of it as do the longer wavelengths. The VOR/DME system covers essentially the entire United States, but this line-of-sight requirement makes it only useful in the air because the transmitters are all ground-based. The ILS is much more precise, but also suffers from the line-of-sight requirement and, in addition, provides only very limited coverage. Since it’s designed for landing aircraft, *and* is very expensive, it’s only located at the higher traffic airports.

Ever since the first Soviet Sputnik satellite in 1957, there have been attempts to use space-based platforms for radio-navigation to eliminate the line-of-sight requirement of high frequency, high accuracy systems. The U.S. Transit system, first launched in 1959, was the first successful such system and is still in operation today. The system includes six satellites (frequently referred to as SVs or Space Vehicles) in polar orbits some 360 kilometers high, and provides precision on the order of $\frac{1}{2}$ kilometer or better, which is fine for coarse navigation and positioning, such as for ships at sea. The system relies on measuring the Doppler shift in the transmitted signal as the satellite passes from horizon to horizon. The drawback is that this occurs only about once an hour and requires some 15 minutes of reception to derive a fix. In addition, the system only provides two-dimensional fixes and gives no elevation information.

Enter the GPS, the highest frequency, shortest wavelength, and most precise system to date, with its full constellation of satellites providing total global coverage.

Evolution Of The GPS

1960's

USN: *Transit* and *Timation*; USAF: 621b

1973, Apr

Sec of Defense Orders *Timation* & 621b
Merged Into the **Navstar GPS**

(NAVigation System with Timing And Ranging)

1973-1979

Phase I: Concept Validation

1977, Jun

First Launch of *NTS-2* (Navigation Technology Satellite 2)

1978, Feb

First *Block I* Navstar GPS SV Launched

1979-1985

Phase II: Full Scale Development & Tests

1985-Present

Phase III: Production and Deployment

1989, Feb

First *Block II* Navstar GPS SV Launched

1993, Dec

DOD Declares IOC

1995, July

DOD Declares FOC

Evolution of the GPS

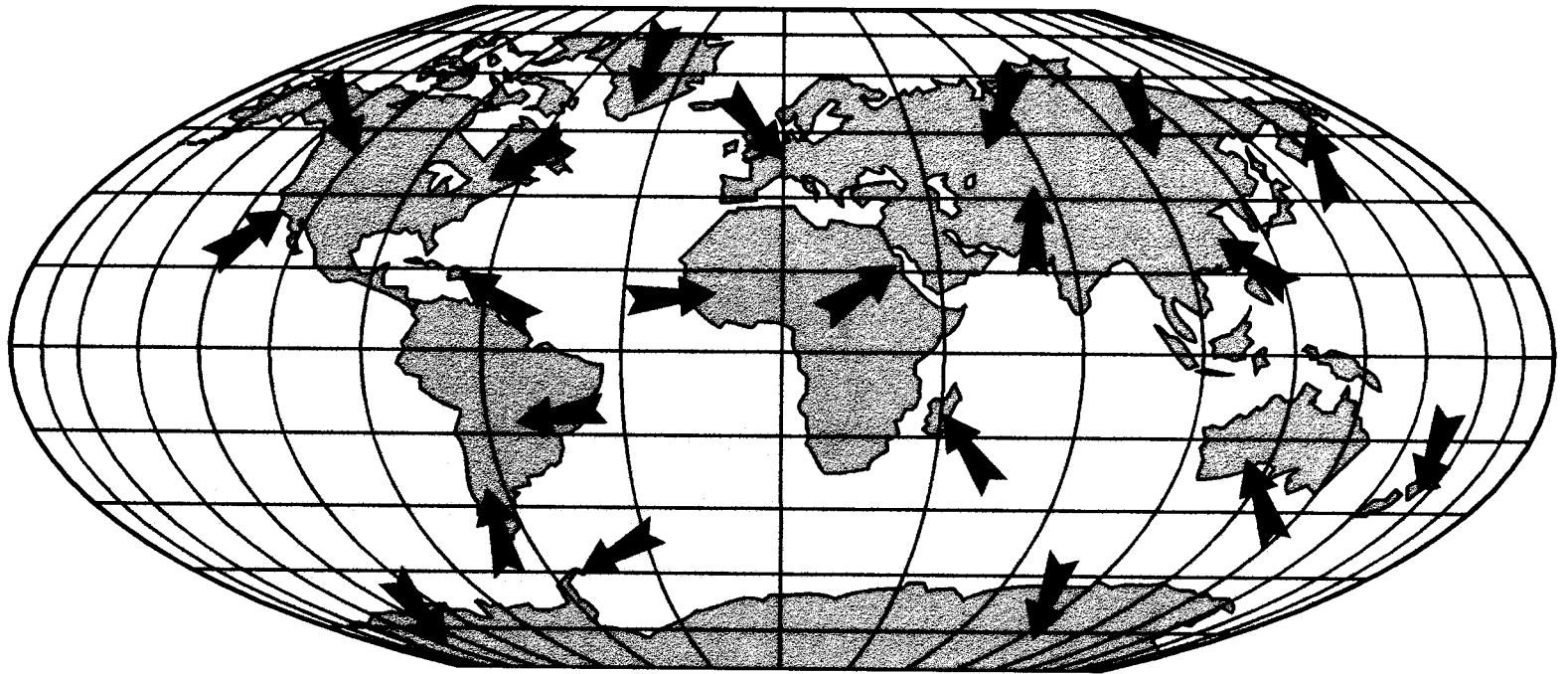
During the late 1950's and early 1960's, the U.S. Navy sponsored two satellite-based positioning and navigation systems: *Transit* and *Timation*. The Transit system became operational in 1964 and was made available to the public in 1969. Timation was a prototype system that never left the ground.

Simultaneously, the U.S. Air Force was conducting concept studies for a system called the *System 621B*. Ground tests were performed to validate the concept but before the system could be implemented, the U.S. Deputy Secretary of Defense, in April 1973, designated the Air Force as the executive service to coalesce the Timation and 62 1B systems into a single *Defense Navigation Satellite System (DNSS)*. From this emerged a combined system concept designated the **Navstar** (for Navigation System with Timing And Ranging) *Global Positioning System*, or simply **GPS**.

The 1970's saw the implementation of **Phase I**, the concept validation phase, during which the first prototype satellites were manufactured and tested. The first functional Navstar prototype satellite launch occurred in June 1977, and was called the NTS-2 (Navigation Technology Satellite 2, which was actually a modified Timation satellite).

While the NTS-2 only survived some 7 months, the concept was shown to be viable, and in February 1978 the first of the **Block I** Navstar satellites was launched. In 1979, **Phase II**, full-scale development and testing of the system, was implemented with nine more Block I satellites launched during the following six years. This was followed in late 1985 by **Phase III**, the full-scale production and deployment of the next generation of **Block II** satellites. Civilian access to the GPS signal, without charge to the user, was formally guaranteed by President Reagan in 1984 as a direct response to the shoot-down of the Korean Airline Flight KAL-007 in 1983, when it strayed over the Soviet Union. The launch of the first of the production Block II satellites occurred five years later, in February 1989.

**WITH GPS EVERY POINT ON THE EARTH
HAS A UNIFORM AND UNIQUE "ADDRESS"**



AND IT' S BASED ON THE WGS84 DATUM

GPS Addresses

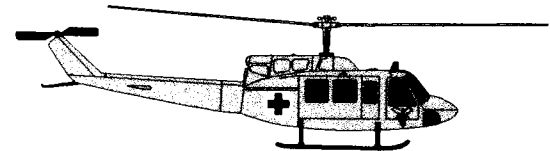
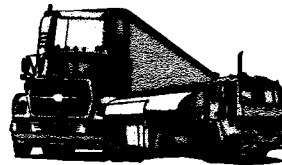
In December 1993, the Department of Defense declared Initial Operational Capability (IOC) for the system, with the minimum combined total of 24 Block I and Block II satellites in their proper design orbits and fully functional.

Finally, in July 1995, with a full constellation of 24 Block II satellites operating in orbit, the DoD declared Full Operational Capability (FOC) for the system.

Today, the system is fully operational, providing positioning and navigation service to virtually anyone anywhere on the globe. In a sense, it has allowed us to give every centimeter of the surface of the planet its own unique address that can be understood by anybody through the use of a universal geocoordinate system. It could be in the not too distant future that you'll find yourself inviting a friend to your home by saying something like ". . . sure, come on over. My address is 39°45' 16.174634"N by you can't miss it." And the fact is they couldn't, because *on the entire planet there is no other place that shares that same address*. It is yours, yours alone, and there's no mistaking it. Seem far-fetched? We'll see. It's hard to argue with the level of success that the global positioning system is currently enjoying. As we'll discuss later, the costs of receivers are plummeting. They have become consumer items that, at the low end, cost less than the typical low-priced VCR. So... why not?

GPS Civil Applications

- Land / Sea / Air /Space Navigation
- IVHS - Intelligent Vehicle Highway Systems
- Search and Rescue
- Mapping / GIS
- Surveying
- Recreation



GPS Civil Applications

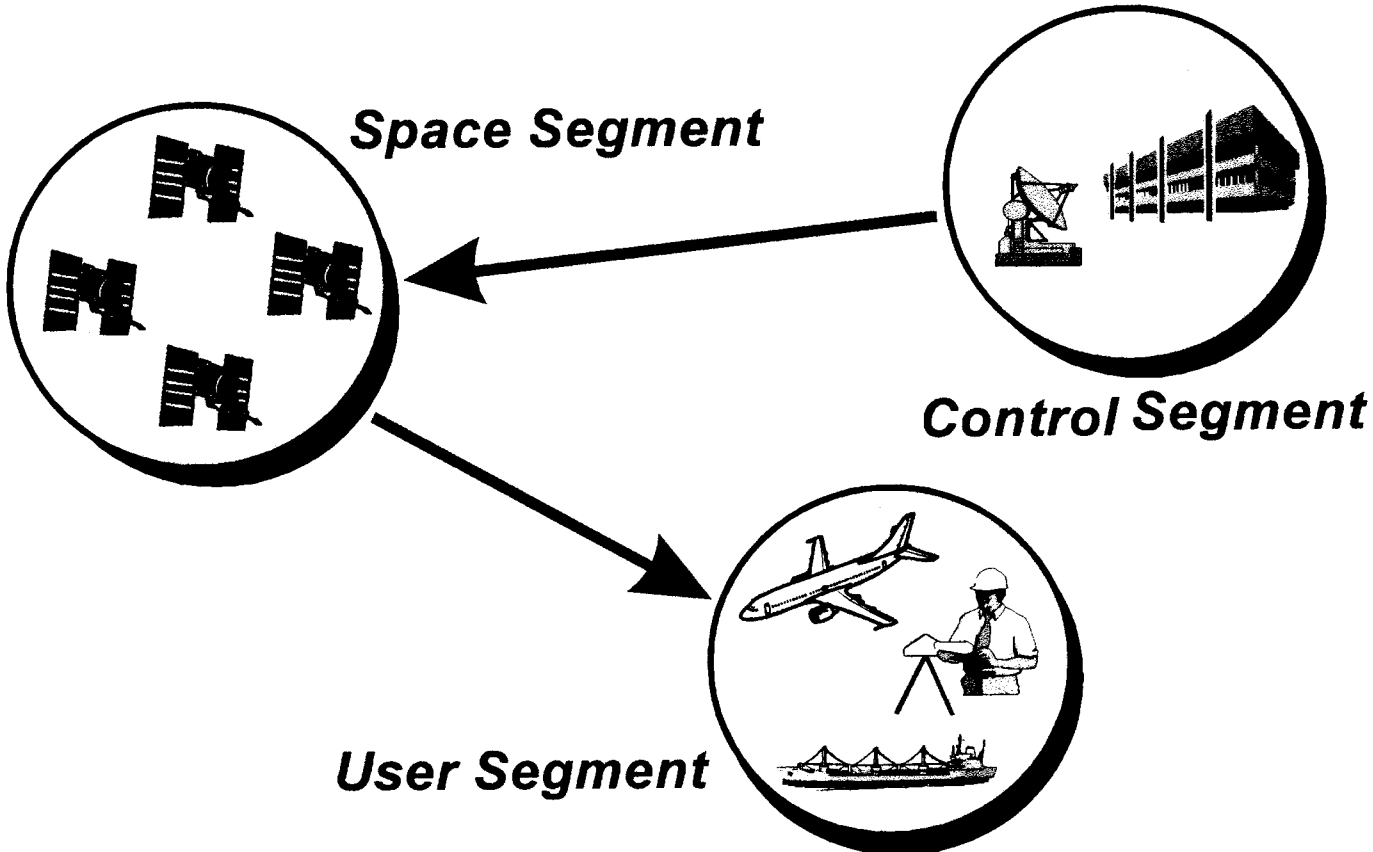
The global positioning system is one of the few big-budget government projects that has come in ahead of schedule, under cost, and works better than the designers had ever dreamed...which has been both a boon and a bane for the military. Clearly, any administrator would be delighted with this kind of outcome for one of their projects. However, the military has a very different agenda than that of the widely varied civilian users of the system. And that's the problem.

Civilian uses for GPS have far out paced the military's. The civilian applications have proven so useful that there has been a growing dependency on the system that is expected to quickly move into critical areas such as airline navigation. This creates a problem for the military. That is, how do they maintain military security over the system when civilian lives may now depend on their free and continuous access? For the moment, there **are** ways, as we will discuss later, but the problem still exists and will only get more complicated with time.

So who's using GPS? Almost anyone who needs to know where they are and where they're going-and, anymore, that includes almost everyone. With low-end receivers costing less than \$200 (and falling), virtually everyone **can** use the system. Receivers connected to map displays are already available to new car buyers, insuring they'll never get lost again. Delivery companies are optimizing their routes on a minute-by-minute basis. Mapping is largely becoming a matter of simply going someplace, automatically creating a map on the way. High-precision surveying can be done in minutes instead of days. Perhaps even more importantly, you'll never again forget where that great fishing hole was. The list is almost endless.

GPS Segments

The Global Positioning System Consists of Three Major Segments



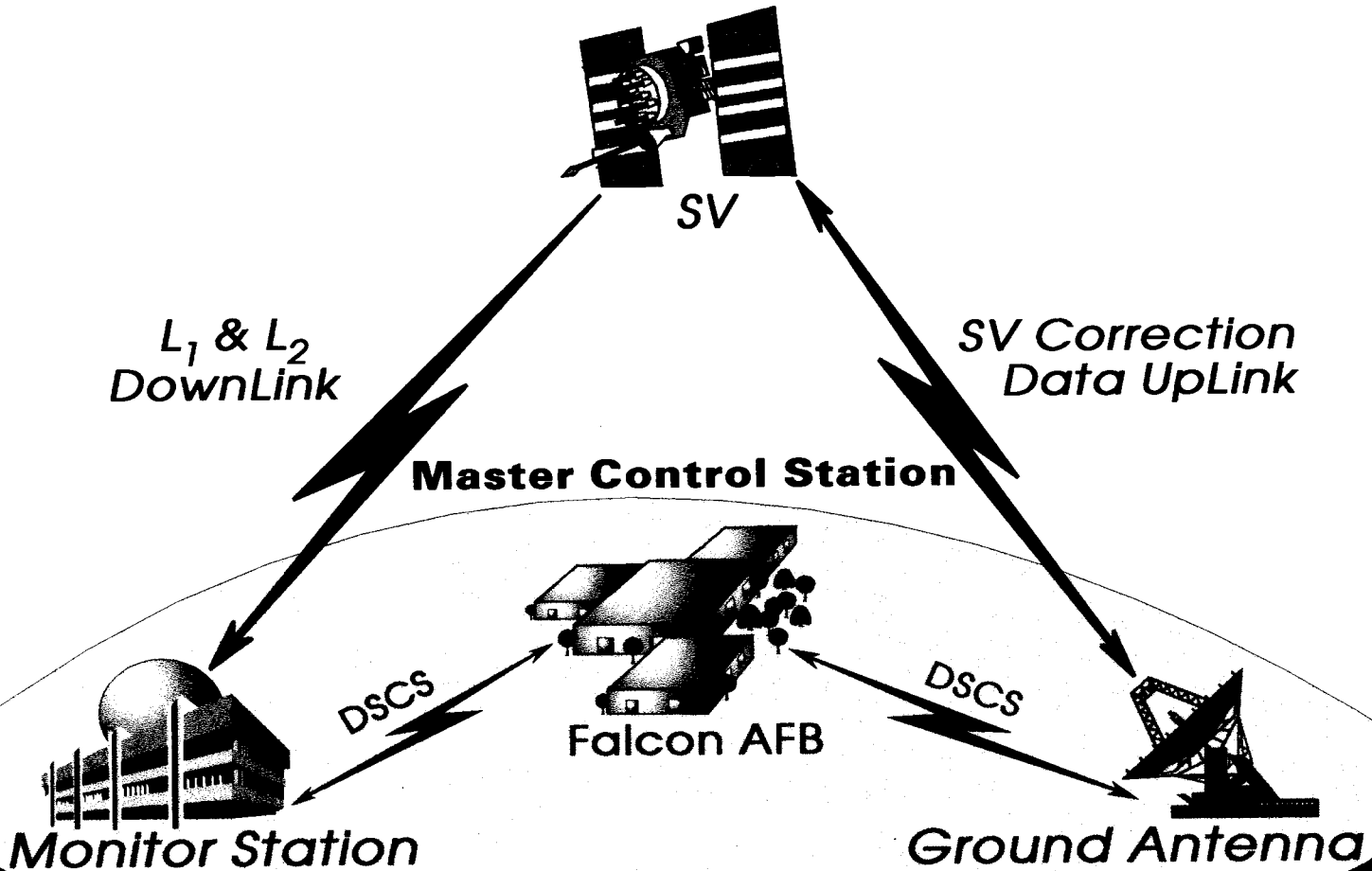
GPS Segments

The Global Positioning *System* consists of three major segments: the *Space Segment*, the *Control Segment*, and the *User Segment*. The space and control segments are operated by the United States Military and administered by the U.S. Space Command of the U.S. Air Force.

Basically, the control segment maintains the integrity of both the satellites and the data that they transmit. The space segment is composed of the constellation of satellites as a whole that are currently in orbit, including operational, backup and inoperable units.

The user segment is simply all of the end users who have purchased any one of a variety of commercially available receivers. While the user segment obviously includes military users, this book will concentrate on the civilian uses only. Each of the segments will be examined more closely in the following pages.

Control Segment



The Control Segment

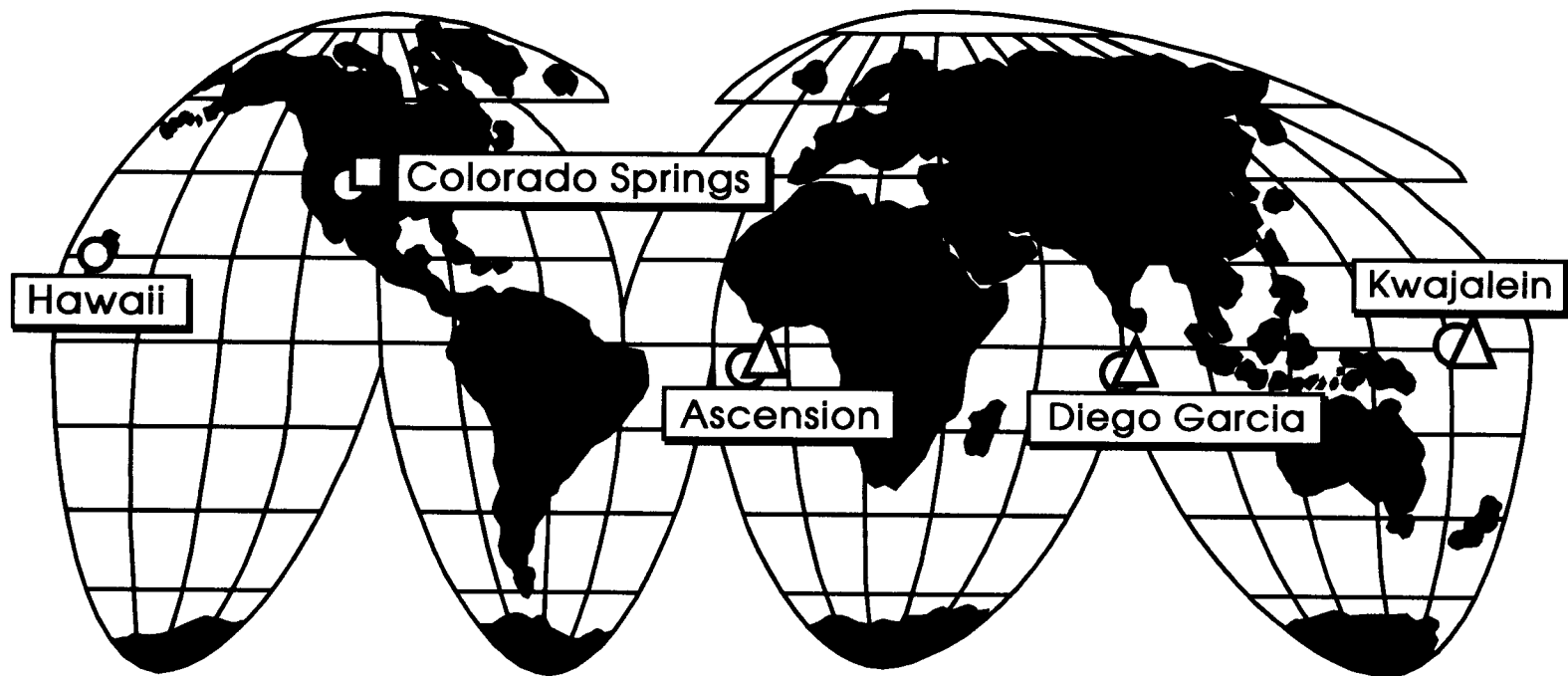
The control segment of the Global Positioning System consists of one Master Control Station (MCS) located at Falcon Air Force Base in Colorado Springs, Colorado, and five unmanned monitor stations located strategically around the world. In addition, the Air Force maintains three primary ground antennas, located more or less equidistant around the equator. In the event of some catastrophic failure, there are also two back-up Master Control Stations, one located in Sunnyvale, California, and the other in Rockville, Maryland.

The unmanned monitor stations passively track all GPS satellites visible to them at any given moment, collecting signal (ranging) data from each. This information is then passed on to the Master Control Station at Colorado Springs via the secure DSCS (Defense Satellite Communication System) where the satellite position (“ephemeris”) and clock-timing data (more about these later) are estimated and predicted.

The Master Control Station then periodically sends the corrected position and clock-timing data to the appropriate ground antennas which then upload those data to each of the satellites. Finally, the satellites use that corrected information in their data transmissions down to the end user.

This sequence of events occurs every few hours for each of the satellites to help insure that any possibility of error creeping into the satellite positions or their clocks is minimized.

Control Segment Locations



- *Master Control Station*
- △ *Ground Antenna*
- *Monitor Station*

Control Segment Locations

This map illustrates the locations of each of the control segment components. The single Master Control Station (MCS) is located at Colorado Springs, Colorado. That facility is co-located with a monitor station that continuously observes the positions and clock settings of all satellites that happen to be in view at any given time.

There are four other unmanned monitor stations located at strategic spots around the world. One is located at Hawaii, another at the tiny Ascension Island off the West Coast of Africa (population 7 19), another at Diego Garcia off of the southern tip of India, and the fourth at Kwajalein, part of the Marshall Islands group in the Western Pacific.

The three upload ground antennas are co-located with the monitor stations at Ascension Island, Diego Garcia, and Kwajalein.

The Global Positioning Satellite

Physical Characteristics

Name: NAVSTAR

Manufacturer: Rockwell International

Weight: ~900 kg

Size: ~5 meters wide (panels extended)

Lifespan: 7.5 years

Number Built:

11 Block I Prototype

28 Block II Production

Orbital Parameters

Altitude: ~ 20,200 km

Orbital Period: 12 hr

Orbital Plane: 55° to equatorial plane

Number Of Orbital Planes: 6

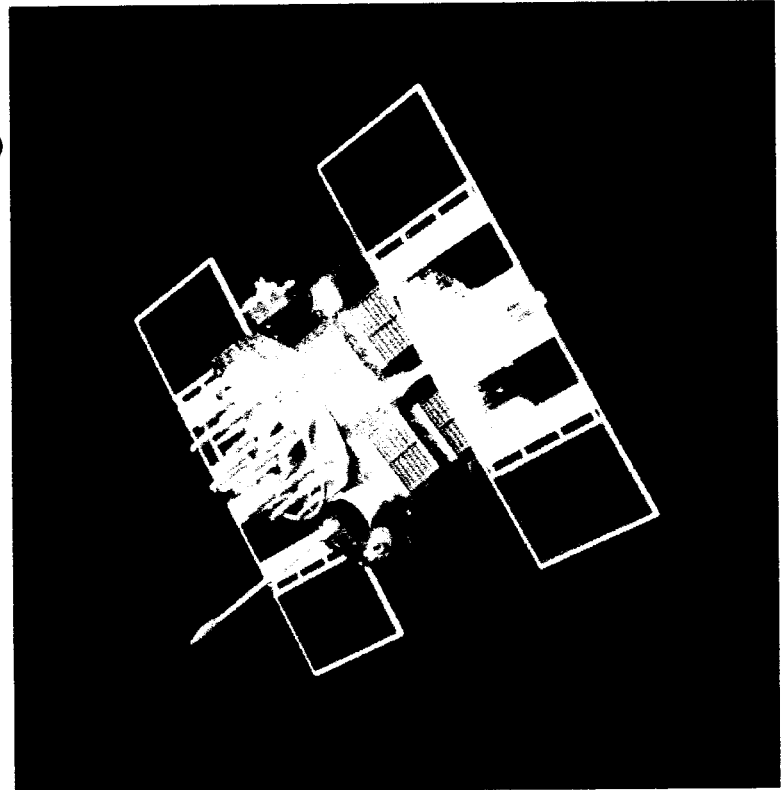
Orbital Plane Spacing: 60° apart

Number Satellites Per Orbital Plane: 4

Total Number In Constellation: 24

Number Operational: 21

Number of Backups: 3

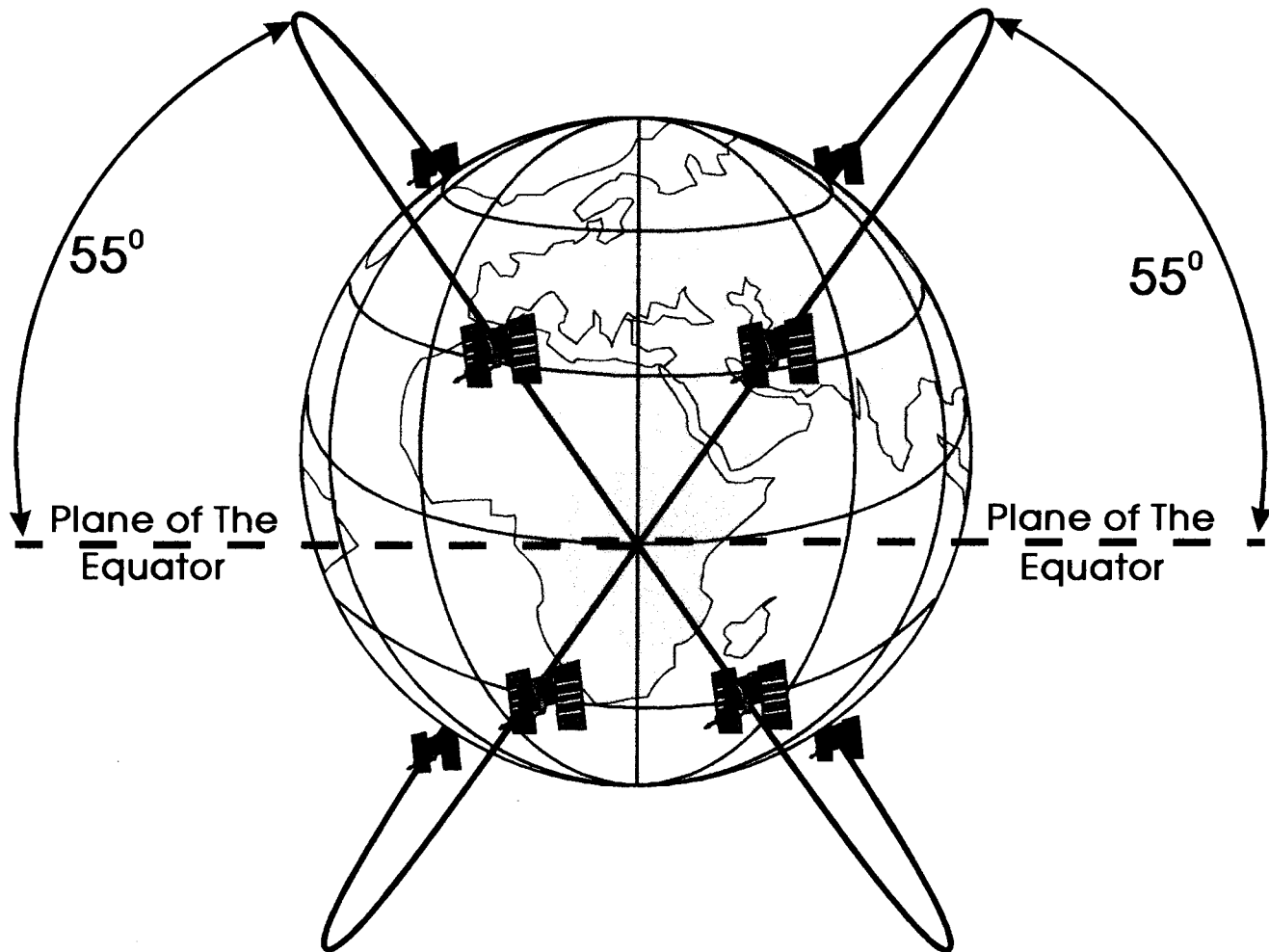


The Space Segment

The space segment consists of the complete constellation of orbiting Navstar GPS satellites. The current satellites are manufactured by Rockwell International and cost approximately \$40 million each. To each satellite must be added the cost of the launch vehicle itself which may be as much as \$100 million. To date, the complete system has cost approximately \$10 billion.

Each satellite weighs approximately 900 kilograms and is about five meters wide with the solar panels fully extended. There were 11 Block I prototype satellites launched (10 successfully), followed by 24 Block II production units. Currently, only one of the Block I satellites is still operational, while four Block II backups remain in ground storage.

The base size of the constellation includes 21 operational satellites with three orbiting backups, for a total of 24. They are located in six orbits at approximately 20,200 kilometers altitude. Each of the six orbits are inclined 55 degrees up from the equator, and are spaced 60 degrees apart, with four satellites located in each orbit (see diagram on next page). The orbital period is 12 hours, meaning that each satellite completes two full orbits each 24-hour day.



Orbits

This diagram illustrates two of the orbital planes of the space segment. For clarity, only two orbits are shown, spaced 180° apart, whereas in reality there are six planes, spaced 60° apart. Each of the orbits has three or four satellites more or less equally spaced, for a total of 24. The Master Control Station can move any of the satellites at any time within their own orbits. They cannot, however, move a satellite from one orbit to another.

The orbits are steeply inclined to the equator at 55° , being more than “halfway up.” This is opposed to the polar, or “straight up” (north to south) orbits of the much lower orbiting Transit satellites.

The satellites orbit at an altitude of approximately 20,200 kilometers, or about half the altitude of a geostationary satellite. A geostationary satellite, orbiting at about 40,000 kilometers altitude, circles the Earth every 24 hours, the same time period that the Earth takes to complete one full rotation (one day). Therefore, a geostationary satellite always remains over the same spot on the Earth (thus “*geostationary*”), essentially following that “spot” on the surface as the Earth rotates. The GPS satellites, at one-half that altitude, complete one orbit every 11 hours, 58 minutes (its “orbital period”). Since the Earth is rotating underneath the orbiting satellites, any given satellite’s orbit slowly “moves” slightly westward with each rotation.

Launch History

Seq	SVN	PRN Code	Internat. ID	NASA Catalog Number	Orbit Plane Pos'n	Launch Date (UT)	Available	Decomm
Block I								
	01	04	1978-020A	10684		78-02-22	78-03-29	85-07-17
	02	07	1978-047A	10893		78-05-13	78-07-14	81-07-16
	03	06	1978-093A	11054		78-10-06	78-11-13	92-05-18
	04	08	1978-112A	11141		78-12-10	79-01-08	89-10-14
	05	05	1980-011A	11690		80-02-09	80-02-27	83-11-28
	06	09	1980-032A	11783		80-04-26	80-05-16	91-03-06
	07					81-12-18	Launch failure	
	08	11	1983-072A	14189		83-07-14	83-08-10	93-05-04
	09	13	1984-059A	15039		84-06-13	84-07-19	94-06-20
	10	12	1984-097A	15271	A-1	84-09-08	Rb 84-10-03	
	11	03	1985-093A	16129		85-10-09	85-10-30	94-04-13
Block II								
II-1	14	14	1989-013A	19802	E-1	89-02-14	Cs 89-04-15	05:02 UT
II-2	13	02	1989-044A	20061	B-3	89-06-10	Cs 89-08-10	20:46 UT
II-3	16	16	1989-064A	20185	E-3	89-08-18	Cs 89-10-14	20:21 UT
II-4	19	19	1989-085A	20302	A-4	89-10-21	Rb 89-11-23	03:13 UT
II-5	17	17	1989-097A	20361	D-3	89-12-11	Cs 90-01-06	03:30 UT
II-6	18	18	1990-008A	20452	F-3	90-01-24	Cs 90-02-14	22:26 UT
II-7	20	20	1990-025A	20533	B-2	90-03-26	Rb 90-04-18	23:13 UT
II-8	21	21	1990-068A	20724	E-2	90-08-02	Cs 90-08-22	15:00 UT
II-9	15	15	1990-088A	20830	D-2	90-10-01	Cs 90-10-15	00:39 UT
Block IIA								
II-10	23	23	1990-103A	20959	E-4	90-11-26	Cs 90-12-10	23:45 UT
II-11	24	24	1991-047A	21552	D-1	91-07-04	Rb 91-08-30	04:44 UT
II-12	25	25	1992-009A	21890	A-2	92-02-23	Cs 92-03-24	11:00 UT
II-13	28	28	1992-019A	21930	C-2	92-04-10	Cs 92-04-25	20:32 UT
II-14	26	26	1992-039A	22014	F-2	92-07-07	Cs 92-07-23	19:43 UT
II-15	27	27	1992-058A	22108	A-3	92-09-09	Cs 92-09-30	20:08 UT
II-16	32	01	1992-079A	22231	F-1	92-11-22	Rb 92-12-11	14:49 UT
II-17	29	29	1992-089A	22275	F-4	92-12-18	Cs 93-01-05	16:39 UT
II-18	22	22	1993-007A	22446	B-1	93-02-03	Cs 93-04-04	05:20 UT
II-19	31	31	1993-017A	22581	C-3	93-03-30	Rb 93-04-13	20:53 UT
II-20	37	07	1993-032A	22657	C-4	93-05-13	Cs 93-06-12	16:15 UT
II-21	39	09	1993-042A	22700	A-1	93-06-26	Cs 93-07-20	12:54 UT
II-22	35	05	1993-054A	22779	B-4	93-08-30	Cs 93-09-28	19:29 UT
II-23	34	04	1993-068A	22877	D-4	93-10-26	Cs 93-11-22	18:20 UT
II-24	36	06	1994-016A	23027	C-1	94-03-10	Cs 94-03-28	14:20 UT
II-25	33	03			C-2	96-03-28	96-04-09	

Launch History

The launch history of the GPS program dates back to the first NTS-2 launch in June 1977, which was the first space-based platform to transmit a GPS signal to Earth. However, the program's roots actually dated almost two decades earlier.

Although the NTS-2 only survived some seven months, it proved that the system could do what it was intended to do. This fueled full program implementation and on February 22, 1978, the first Block I satellite was successfully launched.

In all, 11 Block I satellites were launched with 10 successes. Unfortunately, Block I SV number 7 was destroyed in a launch failure on December 18, 1981. The last of the Block I satellites was launched on October 9, 1985, marking the end of Phase I of the program. As of this writing, only one of the original Block I satellites is still functional. That single remaining Block I SV is the number 10 unit, launched on September 9, 1984. At approximately 12 years of age, it has survived almost three times its design specification of four and one-half years which speaks well of its design.

The first Block II satellite, SV number 14, was launched on February 14, 1989, and was followed by an unbroken series of successful launches culminating with SV number 33 which was launched on March 28, 1996. (Satellite 13 was designed as a ground-test unit and was never intended to fly.) Design life-span for the Block II satellites is seven and one-half years. As of this writing, every Block II satellite is still operational.

Eventually, as the Block II satellites begin to fail in the years to come, they will be replaced by the Block IIR, and Block IIF, or replacement and follow-on satellites, respectively, that will be even more robust and longer-lived than the first generation of Block IIs. The first of these is expected to launch in mid to late 1996.

How Does GPS Work?

- GPS Is A Distance / Ranging System
- Operates On The Principal Of Trilateration
- Satellites Transmit Unique Radio Waves
- Receivers Passively Receive SV Signal
- Receivers Meas. Time For Signal To Reach It
- Distance Computed by $D = V \times T_{\Delta}$
- $V = C = 300,000 \text{ Km/Sec (186,000 Mi/Sec)}$

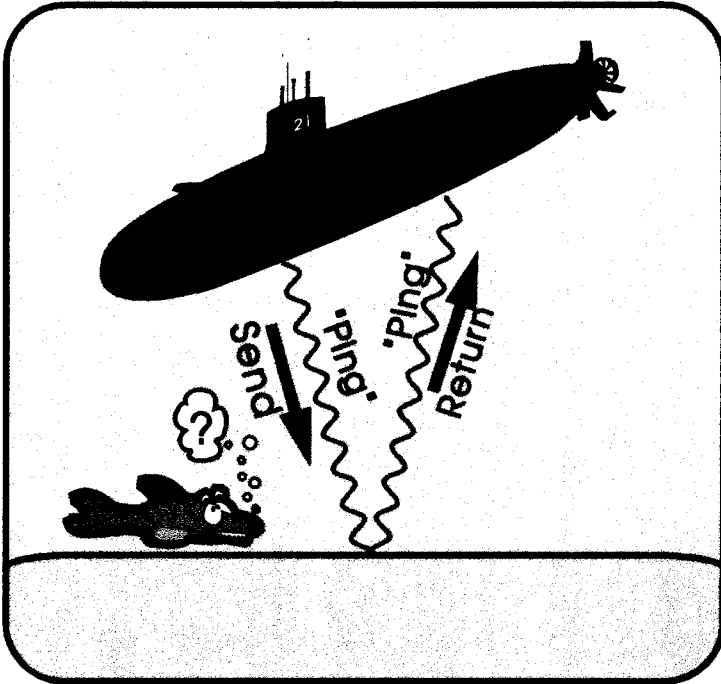
How Does GPS Work?

How the Global Positioning System works is, conceptually, really very simple. All GPS is, is a distance (ranging) system. This means that the only thing that the user is trying to do is determine how far they are from any given satellite. There is no inherent vector information, which implies azimuth (compass direction) and elevation, in the GPS signal. All that the GPS satellite does is shoot out a signal in all directions, although there is a preferential orientation toward the Earth.

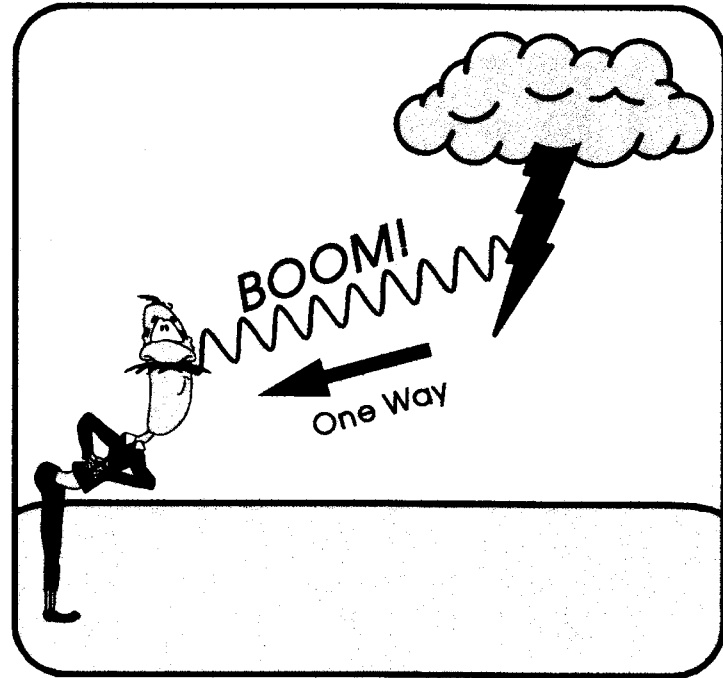
In essence, the GPS operates on the principle of trilateration. In trilateration, the position of an unknown point is determined by measuring the lengths of the sides of a triangle between the unknown point and two or more known points (i.e., the satellites). This is opposed to the more commonly understood triangulation, where a position is determined by taking angular bearings from two points a known distance apart and computing the unknown point's position from the resultant triangle.

The satellites do this by transmitting a radio signal code that is unique to each satellite. Receivers on the ground passively receive each visible satellite's radio signal and measures the time that it takes for the signal to travel to the receiver. Distance is then a simple matter of computing $D = V \times T$, or deriving distance (D) by multiplying the time in transit (T) of the signal by the velocity of transit (V). This is the old "if a car travels a 60 mph, how far will it travel in two hours?" Since radio waves travel at the speed of light, which is essentially fixed at 300,000 kilometers per second, the velocity is a given. Therefore, the only thing needed by the user to calculate distance from any given satellite is a measurement of the time it took for a radio signal to travel from the satellite to the receiver.

Two-Way Ranging



One-Way Ranging



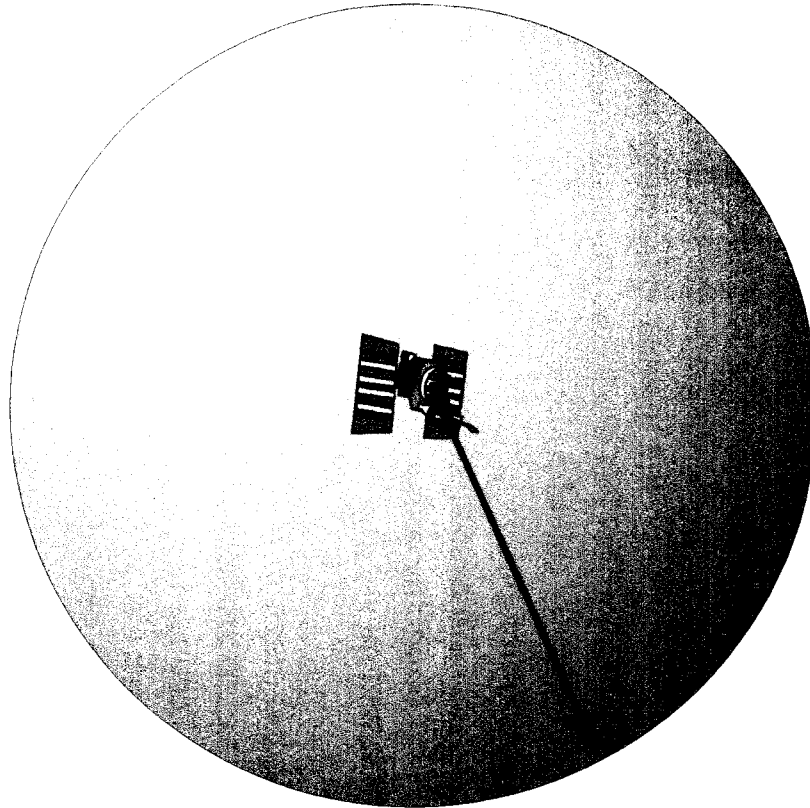
Two- Way vs. One- Way Ranging

The two diagrams to the left illustrate common examples of the two principal types of ranging, *One-Way Ranging* and *Two-way Ranging*, that most of us are familiar with.

We've all seen those WWII submarine movies where the SONAR (SOund NAvigation and Ranging) man intently listens to the "Ping, Ping, Ping" of the destroyer above that is trying to locate and sink the sub. While this is seldom done anymore, it serves well to illustrate the concept of two-way ranging. In the case of the diagram at left, the submarine sends out a unique and recognizable sound (the "ping") and measures the time it takes to reach something (in the diagram, the sea floor) and bounce back up to the listener. Essentially, the listener is listening for and timing the echo. The listener knows how fast the sound travels through the water and so can quickly and easily calculate how far away that something (the sea floor) is. More contemporary examples can be seen in modern EDM's (Electronic Distance Measuring equipment) which measure how far away something is by bouncing either a laser beam or, in some cases, sound waves, off of it and measuring the time it takes to return.

The second diagram illustrates the concept of one-way ranging in a way that most of us are familiar with-the thunderstorm. We know that by counting the seconds that it takes for the thunder to reach us after the flash of lightning, we can determine how far away the storm is. We know that it takes about five seconds for sound to travel one mile and we know precisely when the lightning occurred. Even though the light from the lightning does take a finite span of time to reach us, considering how (relatively) close the storm is and how fast light travels, for all intents and purposes, we see the flash the instant it occurs. This is, conceptually, how GPS works. The difference is that GPS measures radio-wave transit time rather than sound.

Single Range Known



20,000 Km Radius

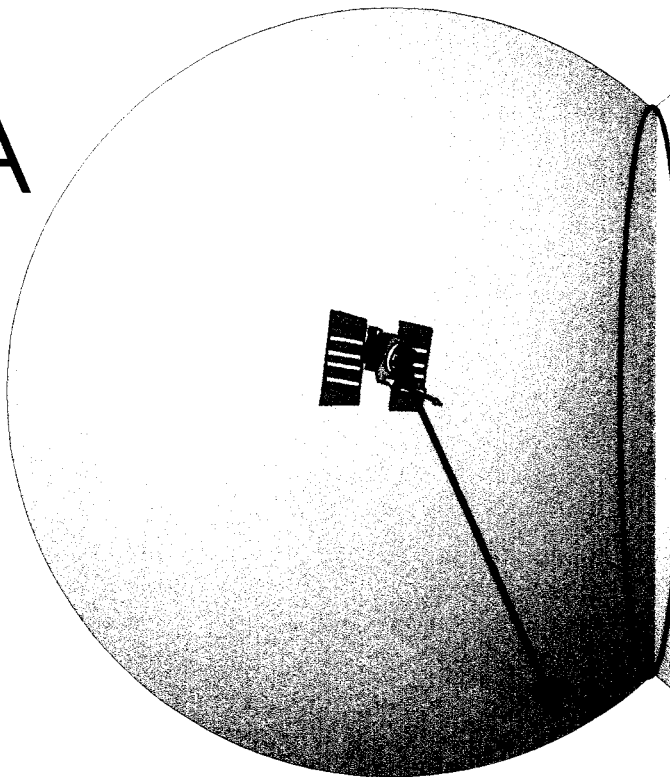
Single Range To A Single SVKnown

The GPS Navstar satellite transmits a radio signal unique to each individual satellite. The signal is essentially omnidirectional, although there is a preferential orientation toward the Earth since the satellite's antennas are located on one side of the vehicle which is, of course, aimed at the Earth. For simplicity's sake, let's assume that the signal is truly omnidirectional and that the satellite broadcasts its signal uniformly outward in all directions.

If we happen to know that the range (distance) to a particular satellite is precisely 20,000 kilometers (for example), then the only place in the universe which is that precise distance from the satellite is somewhere on the surface of an imaginary sphere that has a radius of 20,000 kilometers. With only this amount of information there is no way to know *where* on the sphere we might be located, only that we're no closer than 20,000 kilometers and no farther than 20,000 kilometers. It could be in *any* direction. Remember, there is no direction information given in the satellite's signal.

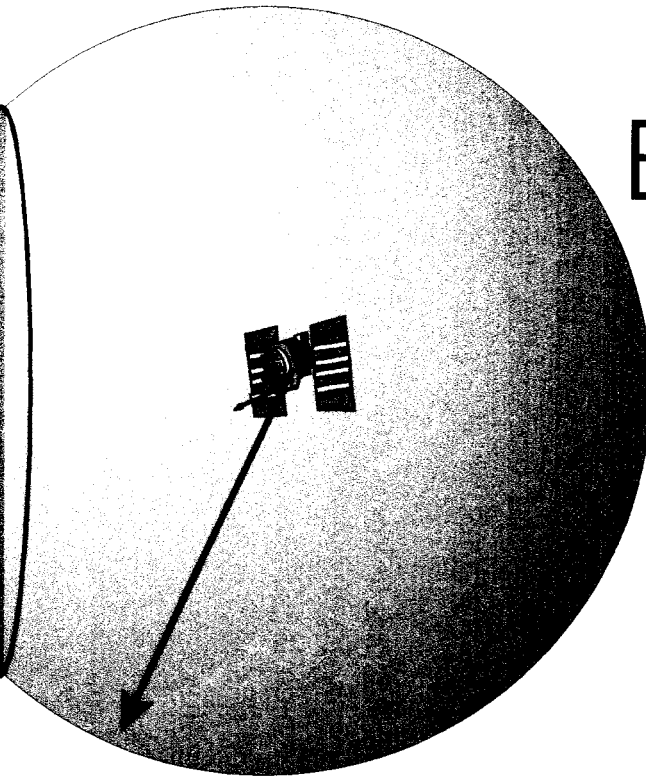
Two Ranges Known

A



20,000 Km Radius

B



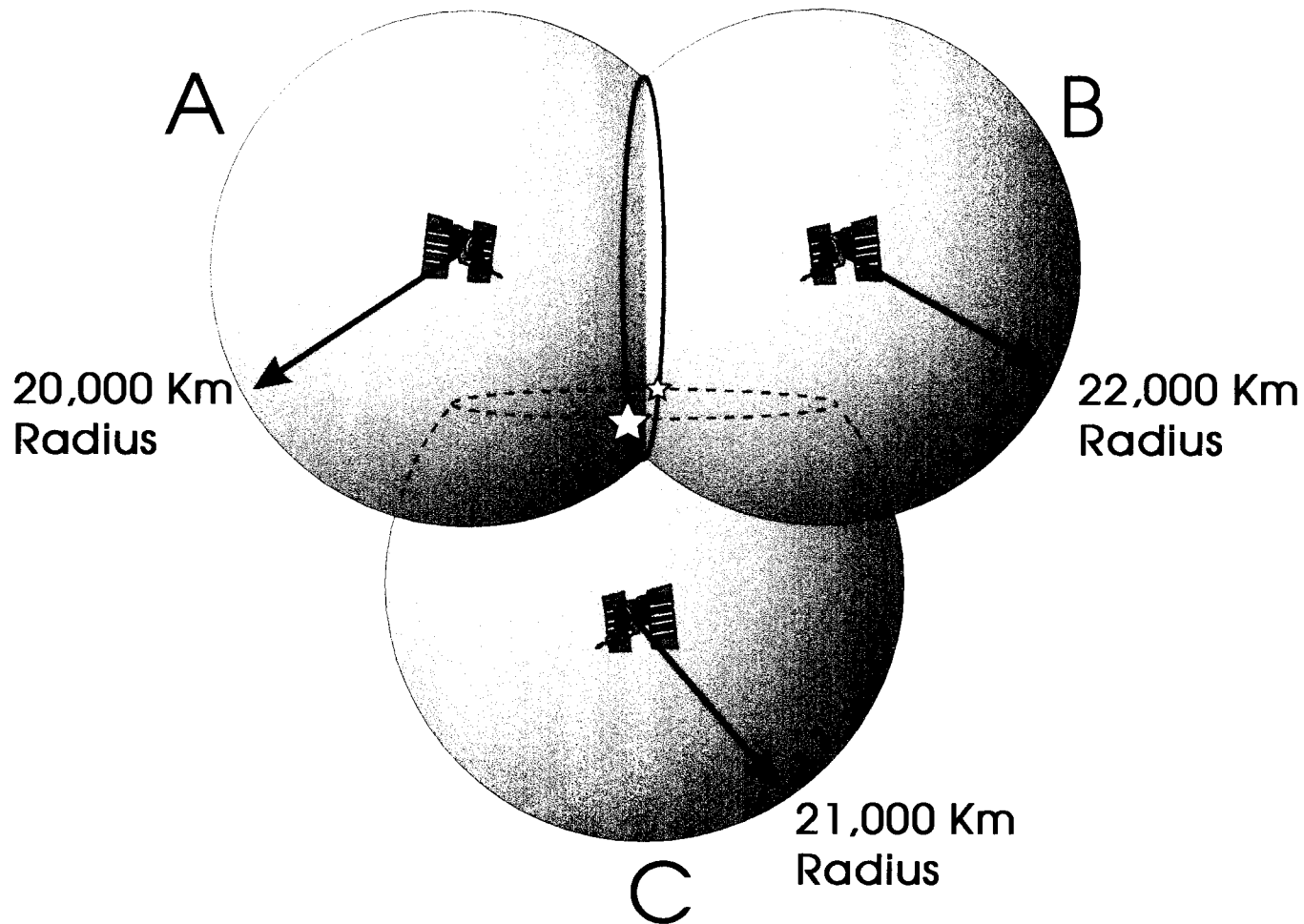
22,000 Km Radius

Two Ranges To Two SVs Known

We can narrow down this positional ambiguity considerably by adding a range to a second satellite. In this example, we already know that we're 20,000 kilometers away from the first satellite (satellite "A"). We just don't know in what direction. If we determine that we're *also* precisely 22,000 kilometers from another, second satellite (satellite "B"), we find that the only place in the universe which is that distance away from satellite "B," *and* is still 20,000 kilometers away from satellite "A," is located somewhere on a circle where the two respective spheres intersect (shown as the black ellipse in the diagram).

While this has narrowed down our position considerably, we still don't know *where* on the sphere-intersection-circle we are. And *that* positional ambiguity is still really big. What we need is a range to yet another satellite.

Three Ranges Known



Three Ranges To Three SVs Known

If we add a third satellite with a known range of (for example) 21,000 kilometers, we'll almost be there. Now, the only place in the universe which is, *at the same time*, 20,000 kilometers from satellite "A," 22,000 kilometers from satellite "B," *and* 21,000 kilometers from satellite "C," is at the *only* two points where all three of the spheres happen to intersect.

We've now narrowed down our position in the universe considerably. We now know where we are precisely—that is, at either one of two possible points. We don't know which one is the right one, but from here it's fairly easy to figure out. In fact, one of the two points is almost always out somewhere where it makes no sense, like thousands of kilometers out in space. The receivers are smart enough to know that one of the two positions will be wrong and to reject the one that makes "no sense." To further insure a reliable choice, most receivers require that, upon initialization, the user input their approximate location, usually to within 500 kilometers or so, which can be gotten from virtually any ordinary map.

So, there it is. Three satellite ranges have given us our precise location in the universe. Well, not exactly. Actually, it turns out that four satellites are really needed to insure an accurate position.

Why 4 Satellites?

- Accurate Positioning Requires Very Precise Meas.
- Only Takes $6/100$ Sec. For SV Signal To Reach Ground
- At 300,000 Km/S $1/1,000,000$ Sec Error \Rightarrow 300 M Pos. Error
- Satellites Have Very Precise Atomic Clocks
- Receivers Have Only "Inexpensive" Quartz Clocks

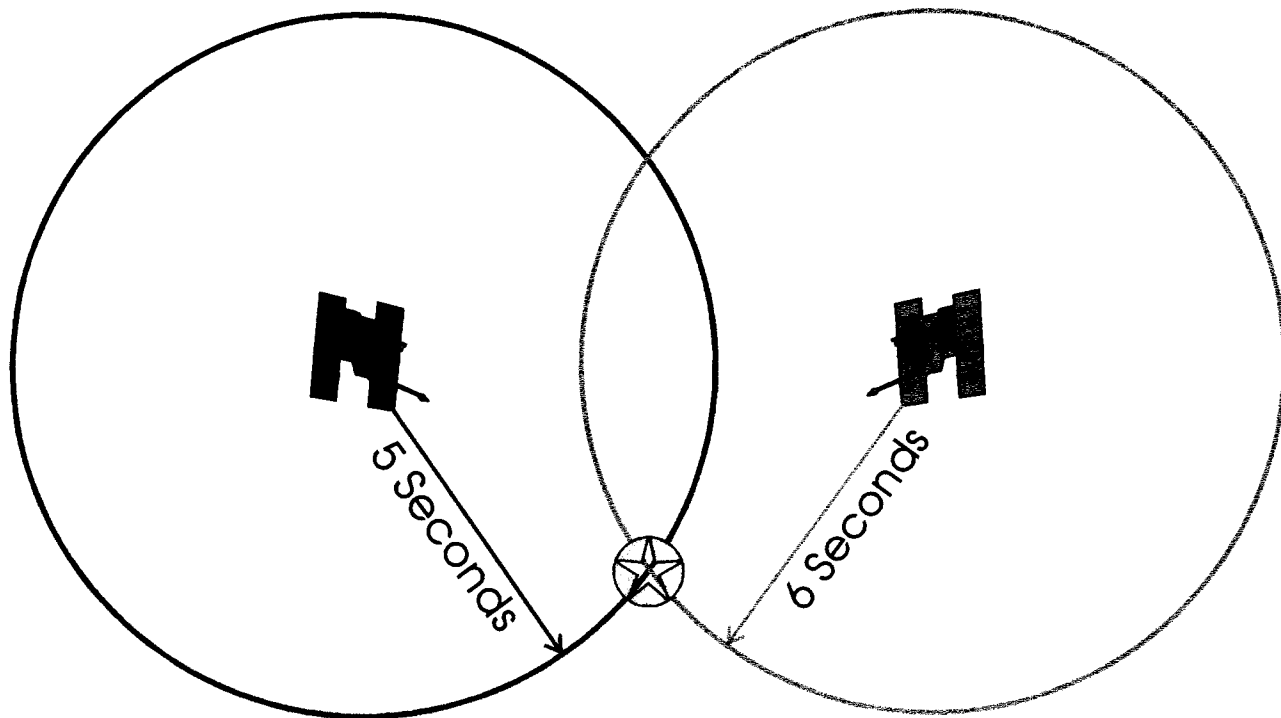
Why Four Satellites?

Why, when three satellites can determine our three-dimensional position so precisely, do we need four satellites? Remember that what we're measuring is the time it takes a radio signal to travel from a satellite transmitter down to our receiver. To acquire an accurate position, we have to make very, very precise time measurements. It turns out that it only takes something like $1/15$ of a second for a satellite signal from orbit to reach our receiver on the ground. With radio waves traveling at some 300,000 kilometers per second, only $1/1,000,000$ (one one-millionth) of a second of error in measuring the travel time translates into approximately 300 meters of error in our position. There is, however, a way to largely eliminate this problem.

It starts at the satellites themselves. To keep very accurate time, each satellite carries four atomic clocks on board, two rubidium and two cesium. These clocks are accurate to within billionths of a second per month. This is certainly accurate enough for our needs, but not really practical for our ground-based receivers. Besides weighing hundreds of kilograms each, each clock costs something like \$200,000.

Each receiver, on the other hand, only carries "inexpensive" quartz clocks with much lower accuracy. Nevertheless, it is critical that the satellite and receiver both start "counting time" at exactly the same moment and continue to count time at the same rate since it's the time it takes for a signal to reach us that we're trying to measure. It turns out that we can insure this by adding a fourth satellite that acts as a time "referee."

True Time / Range



Assumption: Receiver Clock & SV Clock In Perfect Sync

No Clock Timing Error

For ease of illustration, we can look at the problem of clock timing error (called “*clock bias error*”) as a two-dimensional problem. We *could* illustrate the concepts as three dimensional (as is the case in reality) but it would make the diagrams unwieldy and more confusing than they need to be.

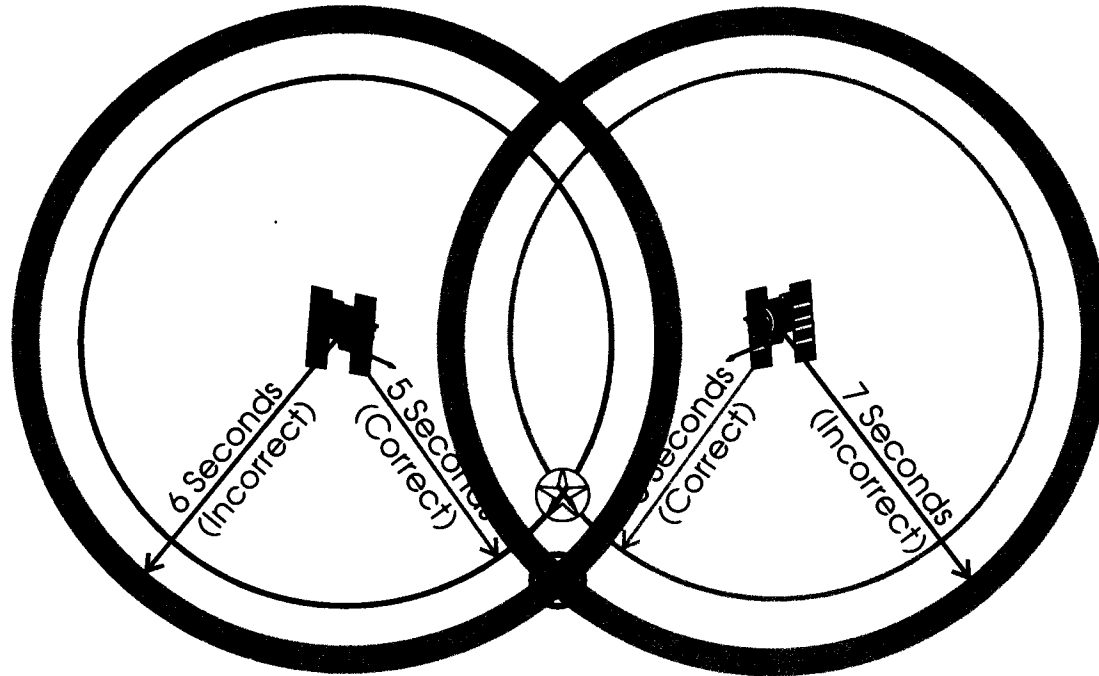
We’ll start by making several assumptions: First, that the clocks on board the satellites are absolutely, exactly right on. This is not too unreasonable an assumption since so much time and money went into them and the fact that they are constantly monitored and corrected by the Control Segment.

Another assumption for this diagram is that the receiver clock and the satellite clocks are in perfect synchronization. This is *not* a reasonable assumption, as we’ve already seen, but for the sake of this illustration, let’s just say that it’s so.

Also, for the ease of illustration, let’s say that the travel time is measured in whole seconds rather than in the millionths of seconds that are measured in reality.

In our two-dimensional diagram we know that, being five seconds from the left satellite and six seconds from the right satellite, we can only be at the two possible points shown in the illustration where the two circles intersect. We also know that the receiver is smart enough to know that one of those two points is not reasonable and rejects it. That, then, leaves only one possible point where we could be located, marked on the diagram with a star.

Receiver Time 1 Second Fast



⊛ = True Position

⊗ = Incorrect "Perceived" Position

Receiver Clock "Fast" By 1 Second - "Perceives" 2:59:59 PM As 3:00 PM. Time Is Distance, So Distance Off By One Second - Time Offset Same For Both SVs, so Receiver Doesn't "See" A Problem.

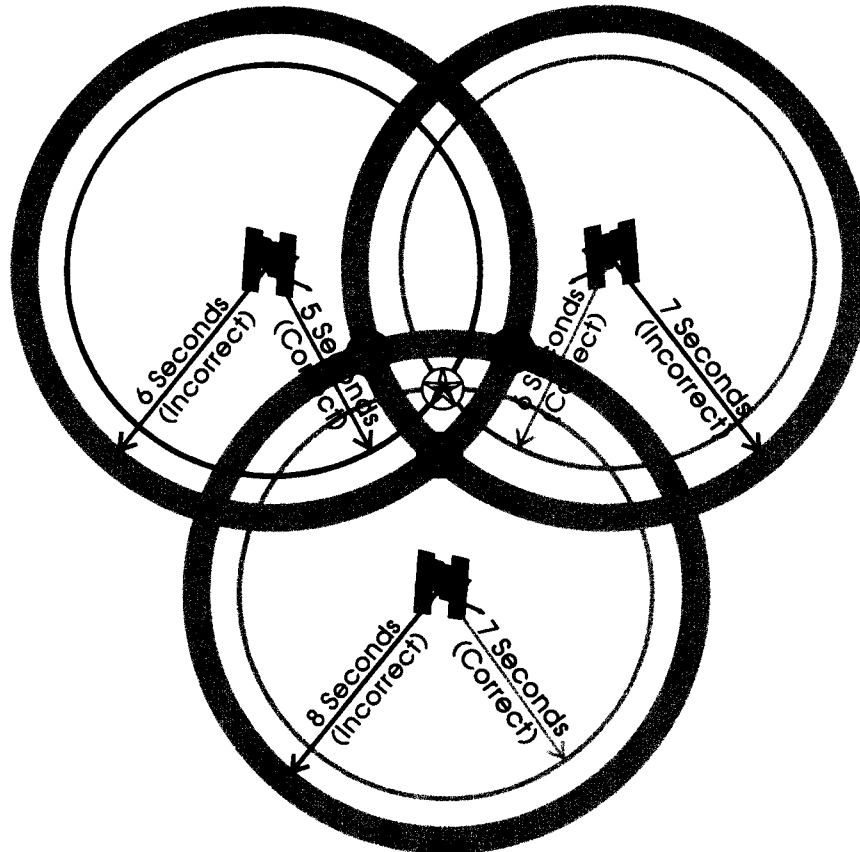
Receiver Time One Second Fast

The fact of the matter is that the satellite and receiver clocks are never perfectly synchronized. We also know that any error in synchronization between the clocks *must* be because of our receiver clock since we've paid so much to insure that the satellite clocks are as absolutely accurate as humanly possible. Since in this application distance is measured by time, we can further simplify things by just treating time as if it *were* distance.

For this illustration, we'll assume that the receiver clock is fast by one second. In other words, the receiver clock "perceives" the actual time of 2:59:59 PM as 3:00 PM. This means that, when measuring how long it takes for the signal to reach the receiver from the (accurately timed) satellite, it appears that the signal took one second longer than it really did (and so, therefore, "seems" that much farther away than it really is). Because the problem is with the receiver and not the satellites, this error will be identical for any satellite from which the receiver happens to collect a signal. Those *incorrect* "spheres" of distance around each satellite are shown on the diagram as grey bands, outside of the correct range "spheres."

With only two satellites in our illustration, the receiver doesn't "see" a problem. Instead it calculates what it believes to be an accurate position based on the incorrectly measured time/distance signals. That point is marked on the diagram by an "X."

Addition Of A SV Time / Range



With Additional SV, There Is No Place Where All Three Radii Intersect
Receiver "Sees" Problem And Slews Time Until They Do

Addition Of Another SV Time/Range

The problem becomes apparent to the receiver when an additional satellite is included in the calculations. Because the problem is in the receiver clock and not the satellite clock, the additional satellite time measurement will also be off by one second. In this case, the correct seven-second travel time to the third satellite is perceived as eight seconds.

It turns out that with three satellite ranges, there is no place in the universe that is six seconds from the first satellite, seven seconds from the second satellite and eight seconds from the third, as illustrated by the grey bands in the diagram.

As soon as the receiver recognizes this, it knows that the problem is with its own internal clock and so it “skews” its clock setting slightly forward and backward until all three ranges intersect. Actually, it just does this mathematically using the “four equations for four unknowns” algebraic technique.

This illustration is shown in only two dimensions but the concept remains the same in three dimensions. It is only necessary to add one more satellite, making four satellites necessary to determine a three-dimensional position.

(Actually, you could determine your three-dimensional position from only three satellites if you already happen to know one of your ranges. You could replace one of the four satellites with the Earth itself, with the center of the Earth kind of acting like the fourth satellite, and sea level as the surface of its “range sphere.” But this requires accurate knowledge of your elevation and is useful mostly only at sea level. Even so, accuracy will still be questionable because sea level isn’t really what we think it is, as we’ll see later on.)

Two GPS Services

SPS *Standard Positioning Service*

PPS *Precise Positioning Service*

Levels Of GPS Service

Two levels of navigation and positioning are offered by the Global Positioning System: The *Standard Positioning Service* (SPS), and the *Precise Positioning Service* (PPS). The Precise Positioning Service is a highly accurate positioning, velocity and timing service that is designed primarily for the military and other authorized users, although under certain conditions can be used by civilians who have specialized equipment.

The Standard Positioning Service offers a base-line accuracy that is much lower than the PPS, but is available to all users with even the most inexpensive receivers. As we will see, there are various techniques available that substantially increase the SPS accuracy, even well beyond that which is offered by the PPS.

Published specifications for the Precise Positioning Service are:

- 17.8 meter horizontal accuracy
- 27.7 meter vertical accuracy
- 100 nanosecond time accuracy

Published specifications for the Standard Positioning Service are:

- 100 meter horizontal accuracy
- 156 meter vertical accuracy
- 167 nanoseconds time accuracy

PART II

BASIC SIGNAL STRUCTURE AND ERROR

Basic GPS Signal Structure

Each Satellite Transmits On Two Frequencies

L1 Carrier 1575.42 MHz and **L2 Carrier 1227.60 MHz**

Superimposed On These Carriers Are Pseudo-Random, Binary, Bi-Phase Modulation Codes Called PRN (Pseudo-Random Noise) Codes Unique To Each Satellite

Coarse Acquisition Code (C/A-Code)

Precise, Or Protected Code (P-Code)

Basic GPS Signal Structure

Each satellite transmits its ranging signal on two different radio frequencies: 1575.42 Megahertz (or 1.57542 Gigahertz, part of the so-called “L-Band”) which is referred to as the *L1 Carrier*, and 1227.60 Megahertz (or 1.2276 Gigahertz, also of the L-Band) designated as the *L2 Carrier*.

Superimposed on these radio carrier wave signals are pseudo-random, binary, bi-phase modulation codes called PRN (Pseudo-Random Noise) codes that are unique to each individual satellite. This simply means that the carrier signal is modulated (varied) by changing its phase (up-down position of the waves) back and forth (bi-phase) at a regular and programmed rate and interval. This regular programmed variation in the signal carries important information, sort of like a Morse-code, “dash dot dot dash...” (binary) signal.

This modulation of the signal, which is really just a series of “dots and dashes,” is very long and complicated. So complicated, in fact, that if you were just to look at it without knowing what it was, it would simply look like a bunch of random noise that made no sense at all. But it really does make sense to those in the know. Thus the term pseudo-random noise.

There are two different pseudo-random code strings used by the GPS. They are the *Coarse Acquisition Code (C/A-code)*, sometimes called the “Civilian Code,” and the *Precise, or Protected Code (P-Code)*.

Basic Signal Structure

Basic Carrier Wave



String Of Ones and Zeros To Be Transmitted



Amplitude Modulation (AM)



Frequency Modulation (FM)



Phase Modulation (PM)



Basic GPS Signal Structure

When a radio transmits a signal, it is in the form of a simple sine wave that has a particular frequency (the number of “humps” on the sine wave that pass a fixed point per unit of time-usually given as Hertz, or times per second), *wavelength* (the distance between “humps” or any matching successive point on the sine wave), and amplitude (the “height” of the “humps”). A basic carrier sine wave is illustrated at the top of the diagram. Radio wavelengths can range from tens of kilometers down to tiny fractions of micrometer. Frequencies, intrinsically linked to wavelengths, also have wide ranges, from only a few per hour (low frequency) to billions per second (high frequency).

By itself, the carrier wave carries no information other than its frequency, wavelength, and amplitude. If we want to transmit any useful information on that carrier wave, we have to modulate or vary it at a regular rate. The second line in the diagram represents a string of zeros (offs) and ones (on’s) that we want to send on the carrier wave, much like Morse-code. There are several methods of transmitting that information on a carrier wave. The first is by varying (modulating) the amplitude, or how “high” and “low” the sine “humps” go. If you’ve ever listened to AM radio, you’ve heard Amplitude Modulation.

You could also vary, just slightly, the frequency of the carrier wave around a central “flat” frequency. That concept is illustrated by the line second from the bottom in the diagram. This is how FM, or Frequency Modulation, radio works.

Finally, you could modulate the *phase* of the carrier. The phase is the relative up/down position of the sine “humps.” By regularly reversing the ups and downs you can transmit your “Morse-code” information. This is how GPS transmits data on its two carriers. This is illustrated in the bottom line of the diagram.

Pseudo-Random Codes

Coarse Acquisition (C/A) Code: A Sequence of 1023 Bi-Phase Modulations of The Carrier, Repeated 1000 x sec. With A "Chip Rate" of 1.023 MHz (SPS)



Precise (or Protected) (P) Code: A Very Long Sequence of Bi-Phase Modulations of The Carrier, Repeated every 267 Days With A "Chip Rate" Of 10.23 MHz (PPS)



**Ranges From These Codes Are Called
"Pseudo-Random Ranges"**

Pseudo-Random Codes

Two Morse-code-like signal strings are transmitted by each satellite. They are the *Coarse Acquisition*, or *C/A-code*, and the *Precise*, or *Protected Code* - more commonly referred to as simply the *P-Code*.

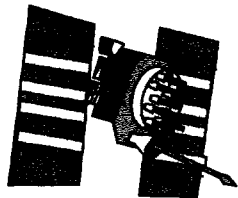
The C/A-code is a sequence of 1,023 bi-phase modulations of the carrier wave. Each opportunity for a phase-reversal modulation, or switch from a zero to a one, is called a “*Chip*” (whether or not the phase is actually reversed at that moment). This entire sequence of 1,023 chips is repeated 1,000 times each second, resulting in a “*Chip-Rate*” of 1.023 MHz or one (opportunity for a) phase switch (chip) every one-millionth of a second. Each satellite carries its own unique code string. The C/A-code is the code used for the Standard Positioning Service.

The Precise (P) code is similar to the C/A-code, but instead of a sequence of 1023 chips, the chip-count runs to the millions. As a result, the complete sequence for the P-code takes 267 days to complete, rather than the one one-thousandth of a second for the C/A-code. One-week segments of the 267-day string are assigned to each satellite and are changed weekly. The P-code is the code used for the Precise Positioning Service.

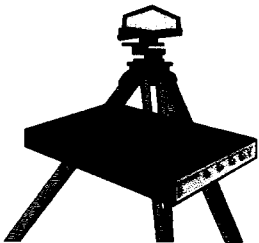
The chip rate of the P-code is an order of magnitude higher than for the C/A-code, running at phase-reversal chip rate of 10.23 MHz, or one phase switch (chip) opportunity every one ten-millionth of a second. This means that there are ten million individual opportunities for a phase reversal each and every second. Since distance is a direct function of time, the radio wave clearly can't travel very far in only one ten-millionth of a second. Consequently, the P-code is considerably more precise than C/A-code. As we'll see, this fact is critical in understanding how GPS determines distance and why one service is so much more accurate than the other.

The Code Is The Key

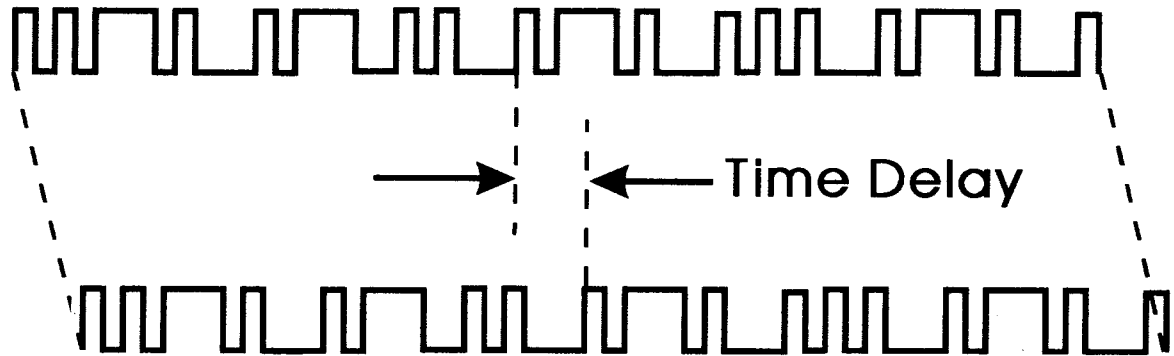
- Receiver "Knows" Each SV' s Unique Code
- Receiver Generates Replica Of Each Code Internally
- Receiver Then Measures The "Lag-Time" Of SV Code



Satellite



Receiver



The Code Is The Key

The code is the key to understanding how GPS determines distance between the satellite and receiver, both for the Standard Positioning Service as well as the Precise Positioning Service. Both use their respective codes essentially the same way: they simply derive different levels of precision by using different chip strings. Conceptually, both work identically.

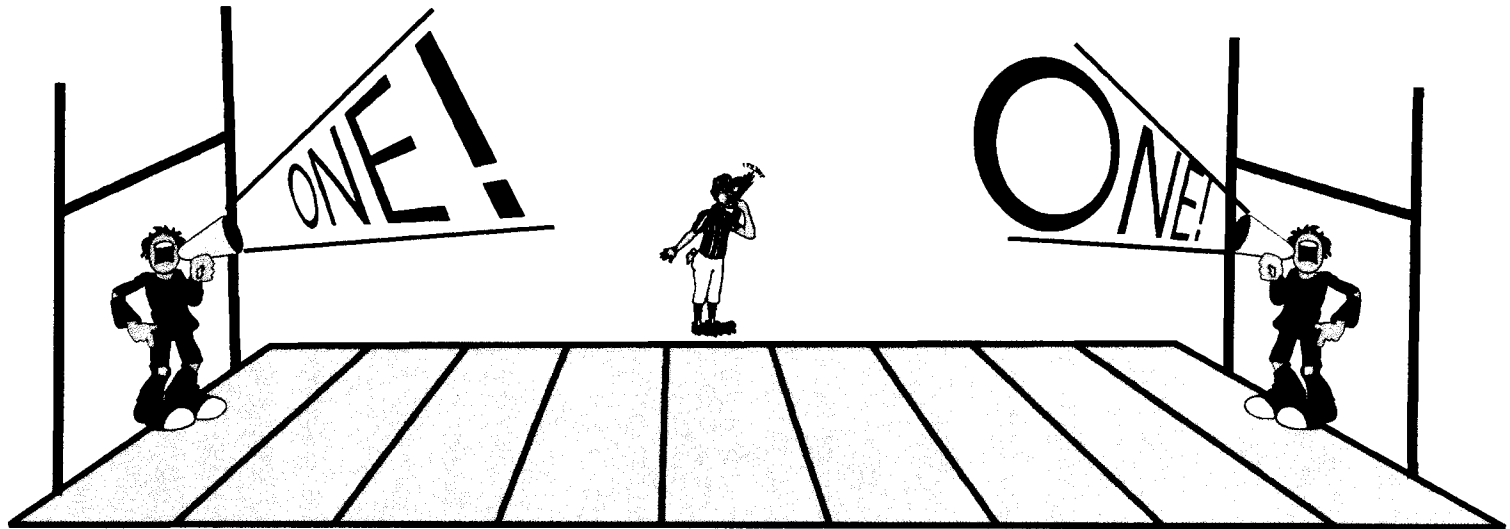
The basic concept is illustrated in the diagram. Each receiver has in its own memory each of the satellite's unique codes. The receiver uses this information to internally generate an exact replica of the satellite's code at the same instant that the satellite generates its "real" code.

Because it took some finite amount of time for the signal from the satellite to reach the receiver, the two signals don't quite match up — there's a tiny delay, or lag time. It's that time delay that is used to determine the distance between the satellite and receiver. This method of range measurement by comparing the delay between two copies of the code is called "*Code Correlation*." Distances derived in this manner, before any kind of error correction is applied to the signal (which we'll talk about shortly) are called "*Pseudo-Ranges*."

You might ask "Why the ultra-complex chip string? Why not a simple, regular 'beep' for example? Wouldn't that do the same thing?" Well, conceptually, yes, but it's really not that simple.

Imagine for a moment that you're standing on the goal line of a football field and a colleague of yours is standing 100 yards away at the other goal line. At the 50-yard line, there's a referee. It's agreed upon that at the exact moment the referee drops a flag, you and your colleague will begin yelling "HEY!" to each other at a pace of once per second. What would you hear at your end?

$$D = V \times T \blacktriangle$$



Hey, Which “HEY?”

Obviously, you would hear yourself yell “HEY!” A moment later you would hear your colleague’s “HEY!” You could then measure how long after you yelled your “HEY!” that your colleague’s “HEY!” got to you. Assuming that you knew the speed of sound under your current conditions, calculation of your distance from your colleague would be straightforward.

But there could be a problem here. How do you know that the “HEY!” you hear from your colleague is the right one to match your “HEY!”? In other words, what if, for example, it took $2\frac{1}{2}$ seconds for his “HEY!” to reach you? You wouldn’t hear his “HEY!” until between your second and third one. You could quickly lose track and might even think that he was only $\frac{1}{2}$ second away because, after all, *one* of his “HEY!”s did come only $\frac{1}{2}$ second after *one* of your “HEY!”s—just the wrong one!

Now imagine instead that at the same moment you both started yelling a count: “ONE!, TWO!, THREE!...” and so on. Now when you heard *any* “number” that he yelled, you’d instantly know which equivalent number of your own you would need to measure the time delay against. This would allow you to jump in anywhere and know right away where you were in the count-string. Conceptually, that’s how GPS measures distance with the C/A- and P-codes. Of course, GPS doesn’t use numbers; instead, it uses those unique strings of on’s and off s-zero’s and one’s.

In the “real” world of GPS, it’s easy to find out where you are in the C/A-code string since the whole string “passes by” in only $1/1,000$ of a second. There’s a problem, however, when trying to figure out where you’re at in the 267-day long P-code string. Thus the term: *Coarse Acquisition*—because P-code receivers use the C/A-code to “get close” to where they need to look in the P-code string, or to “ramp up” to P-code “lock-on.” If the C/A-code can tell the receiver where it’s at within a few hundred meters, then it only has to look at a very small part of the P-code.

Also Need To Know Where SV Is

- SV' s Are Above Earth' s Atmosphere
- Orbits Very Predictable
- SV' s Constantly Monitored 2X Day
- Orbital Variations Are Measured
- Corrections Uploaded To SV' s From Ground Antennas
- SV' s Send SV Position and Correction Data To Receivers
- Called The "NAV/SYSTEM" Data Code Or The "NAV-msg"
- SV Position Data Called "Ephemeris" (Ephemerides)

Where Are The Satellites?

It turns out that just knowing how far away you are from the requisite four satellites isn't enough. The ranges to the satellites only tell you where *you are relative to the satellites*. But where are the satellites? It is also necessary to know *where* each satellite is in space.

Fortunately, that's not too tough. In the first place, the military is very careful about where it sticks its very expensive space hardware. Once in place in space, the satellites' orbits tend to be very stable through time because they are far above virtually all of the atmosphere and the drag that it can induce. Variations in orbits that are due to gravitational forces are fairly easy to predict and compensate for.

To compensate for the inevitable unpredictable perturbations in the satellites' orbits, they are constantly monitored from the ground. Corrections for any orbital variations that are identified are quickly uploaded from ground antennas to the satellites which then send the information back down to each receiver that's tuned in to them. This satellite position and orbital information is called the "*Ephemeris*," or, as plural, "*Ephemerides*." (Orbital position is constantly changing, thus the term, based on the word "ephemeral," meaning lasting only a short time.)

The ephemeris is part of the Navigation/System data message (the "*NAV-msg*") that is also superimposed on the L1 and L2 carriers, in a sense acting as a modulation of the modulation that we've already talked about.

Finally, in addition to the corrected satellite orbital and position data (the ephemeris data), the NAV-msg also carries a correction for any clock bias, or error in the atomic clocks, on board the satellites so that the receivers on the ground can compensate for these errors.

GPS Satellite Signals

L1 Carrier 1575.42 MHz



C/A Code 1.023 MHz



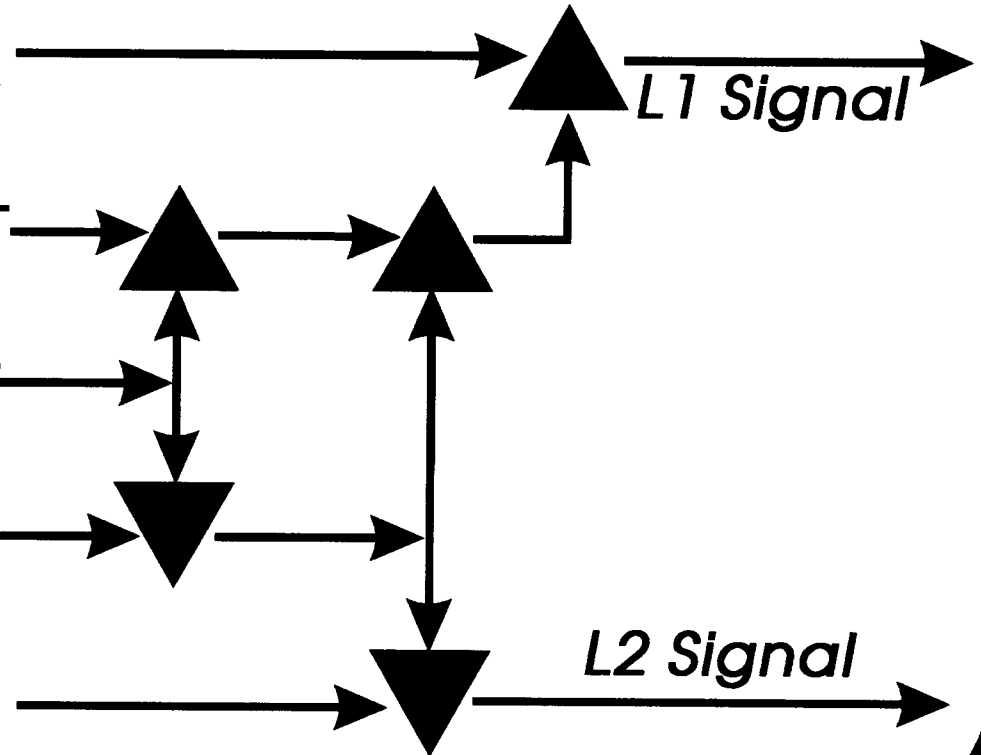
Nav/Msg Data 50 Hz



P-Code 10.23 MHz



L2 Carrier 1227.6 MHz



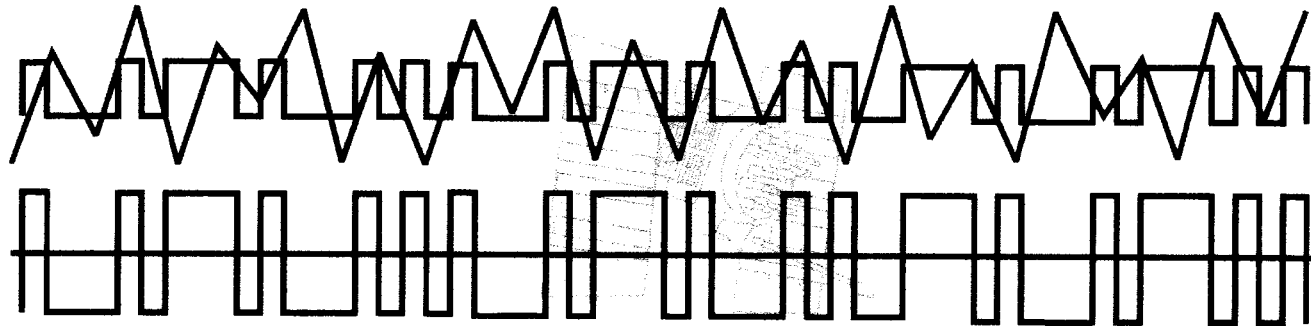
GPS Signal Structure Map

The diagram at left graphically illustrates the various codes that are transmitted on the two carrier frequencies. The 1575.42 MHz L1 carrier wave (top of the diagram) carries the C/A-code, the P-code, and the NAV-msg.

The 1227.6 MHz L2 carrier wave (bottom of the diagram) only carries the P-code and the NAV-msg. Therefore, while the P-code is available on both L1 and L2 frequencies, the C/A-code is only available on the L1. The NAV-msg is transmitted on both carriers.

Signal Strength

- Satellite Signal Strength Very Low
- Less Than The Background Noise
- **But** -
- Background Noise Is Random
- The PRN Code Is *Not*



Signal Strength

You'd think that with all of these radio waves raining down on us from dozens of satellites in space we'd all glow in the dark. Actually, the strength of the GPS signal is very small, equivalent to the tail light of a car seen from 2,500 kilometers away-halfway across the U.S.! Weaker, in fact, than the ordinary background radio noise that's all around us all of the time.

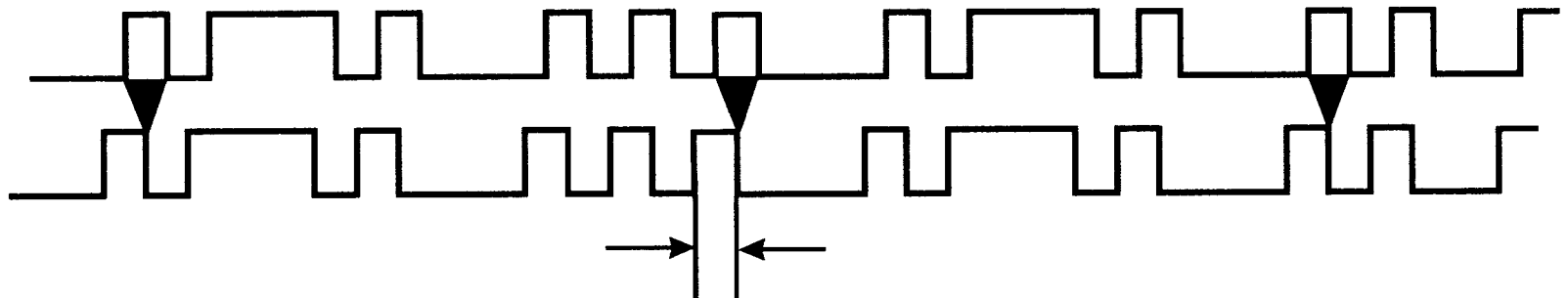
How to isolate a coherent signal from a louder background noise can be solved by an interesting little concept discovered in information theory. Because the background noise is truly random, you can take random segments of that noise and repeatedly "lay" them on top of each other. Because they are random, they would eventually cancel, or zero themselves out.

The pseudo-random code, while *seemingly* random, is not. So if you do the same thing with the code as you did with the random noise, you'll get a very different result. Remember, the receiver has an internal copy of the satellite's PRN (pseudo-random noise) code. The receiver can take its copy of that code and "lay it down" over the incoming noise (which contains the satellite code signal), and then "slew" its replica slightly back and forth. When the replica code and "hidden" satellite code align, they will reinforce each other resulting in a slightly stronger code signal. The receiver can then lay another copy of the code string and again slew it slightly back and forth until it lines up with the now slightly stronger satellite signal, and so on.

Because the electronics are operating essentially at the speed of light, a lot of the "overlays" can be done in a very short time, quickly canceling out the noise (or most of it, anyway) and at the same time magnifying by many times the strength of the desired code.

GPS Resolution - Code Phase

C/A Code 1.023MHz



1 "Chip" ~1 Microsecond
= 300 Meters

Signal Processing Allows Resolution To ~ 1 % Of Chiplength

= 3 Meters

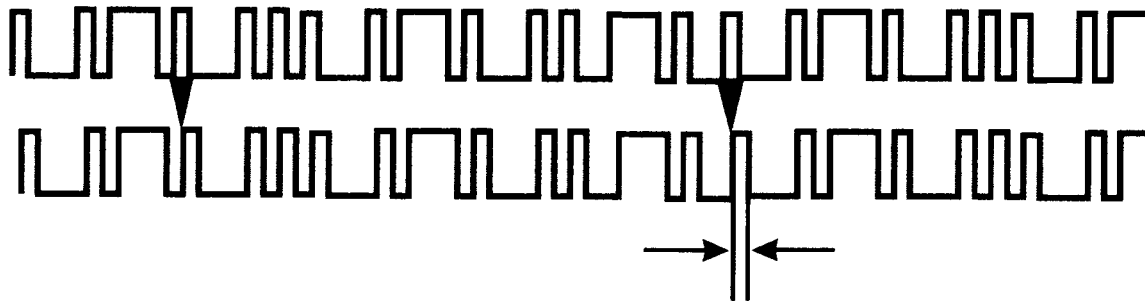
GPS Resolution — C/A-Code

The C/A-code, sometimes referred to as Code-Phase, is used to calculate distance by measuring the time delay between equivalent chips on the satellite's code string and the receiver's replica. There is a bit of a problem here, though, in that the comparison can only be made to within a single chip which lasts only about one microsecond, or one one-millionth of a second. Any one chip could line up anywhere within the matching chip, but there's no way to know *where* within the chip length. At the speed of the radio waves, one one-millionth of a second translates into a distance of some 300 meters.

Fortunately, as a general rule of thumb, signal processing techniques are able to refine the observation resolution to approximately one percent of a signal's wavelength (in the case of the C/A-code, the chip length) which translates to about three meters. That's the *theoretical* maximum resolution, or error range, that is possible, by design, of the C/A-code. As we'll see, actual resolution *can be* considerably higher.

GPS Resolution - Code

P- Code 10.23MHz



1 "Chip" $\sim 1/10$ Microsecond
= 3 Meters

Signal Processing Allows Resolution To $\sim 1\%$ C

= 0.3 Meters (30 Cm)

GPS Resolution — P-Code

The P-code is used to calculate distance by the same method as for the C/A-code. That is, by measuring the time delay between equivalent chips on the satellite's code string and the receiver's replica.

While the same problem of chip matching ambiguity still exists, the chip length of the P-code is only 1/10 that of the C/A's code, resulting in a chip length of only about 30 meters.

Again, as for the C/A-code, the general rule of thumb that observations can be resolved to approximately one percent of the signal's wavelength (or chip-rate length) also applies. In the case of the P-code, that translates to about 0.3 meters, or about 30 centimeters, which is a full order of magnitude of increased precision over that possible with the C/A-code. This represents the theoretical maximum resolution, or error range, that is possible, by design, of the P-code.

Anti - Spoofing (A/S)

No Question About It Then!

Use P-Code Over C/A Code, Right?

Right?

You Can Use P-Code And Get Higher Precision Than C/A Code

But...

Anti-Spoofing (A/S)

Well then, should you use C/A or P-code? At first thought, there's no question, right? After all, if both C/A-code and P-code operate the same way, then there should be no fundamental reason why one receiver would be any different than the other and so why not go for the higher accuracy P-code? Well, even though many texts on the subject state that the P-code is for military and other authorized users only, you *can*, indeed, use the P-code and get the expected higher resolution. However, there *is* a problem...

As we saw earlier, one of the military's principal criterion for GPS was that it be militarily safe. That is, that it couldn't be used against them by potential hostiles. They insure this security by two principal means: *Selective Availability* (SA) and *Anti-Spoofing* (AS). We'll talk a little later about Selective Availability, but for now we'll only look at the Anti-Spoofing technique.

One way that a potential hostile might interfere with U.S. military operations is by transmitting a "false" GPS signal that overrides the real GPS signal. By doing this, they would, in effect, be sending the U.S. military "off in the wrong direction." This is called signal "*spoofing*." The military anticipated this by designing in an Anti-Spoofing technique. With this technique, they can encrypt the P-code making it useless to anyone who doesn't have the proper decryption "keys." This means that anyone using the P-code for positioning or navigation without the proper decryption keys could suddenly find themselves "out in the cold" any time the military decided to turn on the encryption. When implemented, the P-code becomes designated as the Y code. The encryption would apply to the P-codes transmitted on both the L1 and L2 carriers. The C/A-code, however, would not be affected. The down-side is that if a hostile did implement some form of signal spoofing, users of the C/A-code would not be able to compensate for it.

Anti - Spoofing (A/S)

- P-Code Is The "Military Code"
- Designed To Be Secure Against Spoofing
- Military Can (And Does) Encrypt With No Warning
- P-Code Then Becomes Encrypted Y-Code
- Decryption Only To Authorized Users
- Allows Military To Severely Degrade Or Turn Off Completely The C/A Code But Retain Coded Positional Accuracy

Anti-Spoofing (A/S)

So, while the P-code does indeed provide much higher accuracy than the C/A-code, you can't trust it to be there when you need it. The fact is that it is much more likely that the military will arbitrarily and without notice encrypt the P-code than it is that some hostile will start sending out a spoofing signal. The likelihood of hostile spoofing is considerably greater overseas than it is within the United States. However, when the military initiates its Anti-Spoofing encryption to protect its signal, say in Saudi Arabia, the encryption affects the P-code worldwide. They can't (yet) geographically selectively apply the Anti-Spoofing technique. Currently, the P-code is now almost always encrypted to Y-code, with only infrequent periods when it's turned off. This may change, but for the foreseeable future, you can probably count on its being there.

There **are** civilian P-code receivers available. Sort of. Several manufacturers have developed proprietary receiver-software combinations that can, in fact, "see through" the encryption by "re-constructing" the code using various techniques that are product-specific. These receivers tend to be very expensive compared to the more ordinary C/A-code receivers. Their advantage is that they retain high accuracy over very long base lines (an important consideration which we'll talk about later). The military, of course, buys P (Y) code receivers with included direct decryption capability almost exclusively. (An exception was during the Gulf War, when there simply weren't enough P (Y) code receivers available. The military ended up buying every available commercial civilian hand-held C/A-code receiver they could find-creating a tremendous shortage of civilian receivers here in the U.S. Wives of servicemen were seen purchasing units off of the shelves of boating and marine stores to send to their husbands in the Gulf. Obviously, the military couldn't turn on the encryption during that period of time because too many of their own people in the field would be adversely affected. Kind of a backward logic, but it worked. Fortunately, the Iraqis were so far behind the technological curve that hostile use of the system was not a problem.)

Carrier Phase Positioning

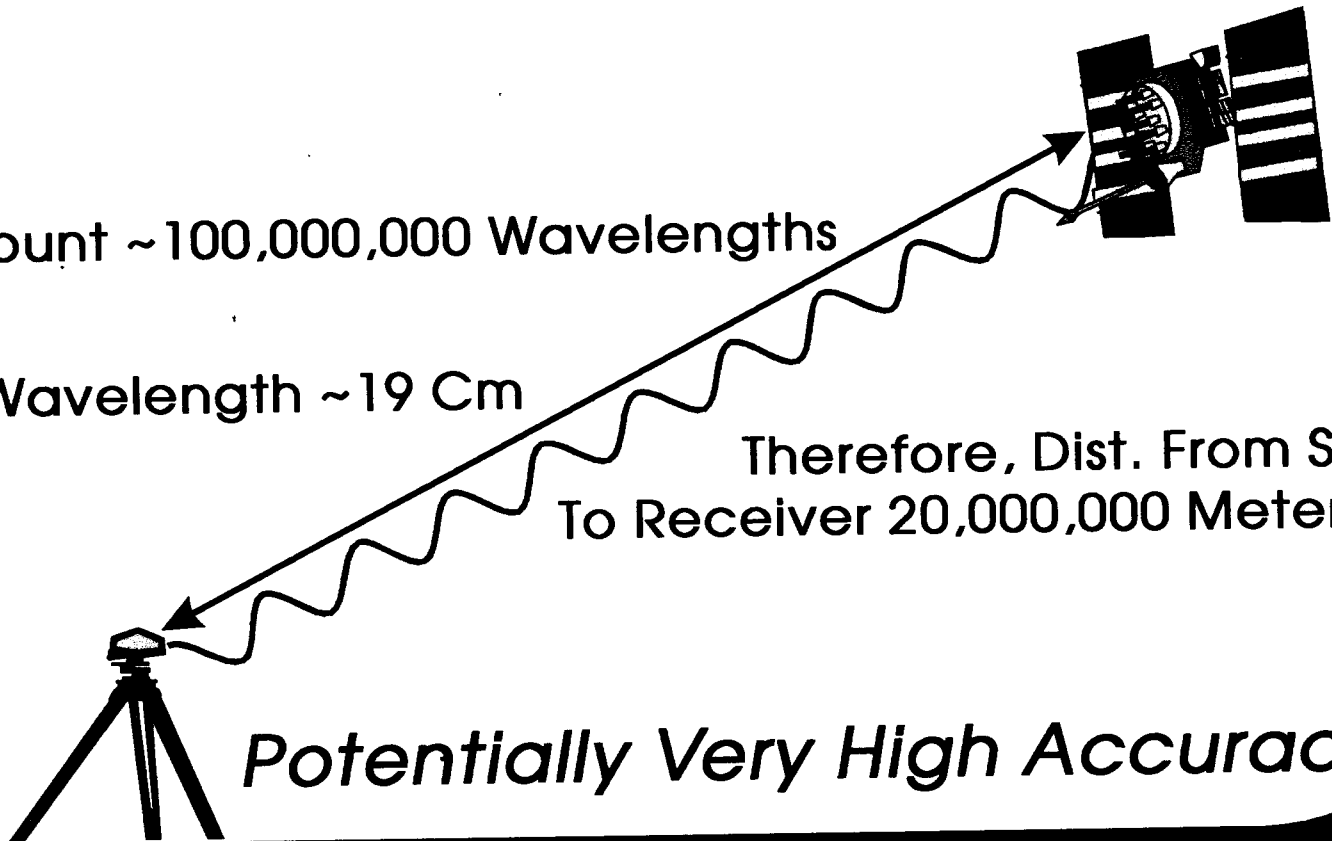
Higher Frequency Gives Higher Accuracy
Count Wavelengths Instead Of Chiplengths

Count ~100,000,000 Wavelengths

1 Wavelength ~19 Cm

Therefore, Dist. From SV
To Receiver 20,000,000 Meters

Potentially Very High Accuracy



Carrier-Phase Positioning

We've seen that the accuracy of a GPS-derived position is directly related to chip length. It didn't take too long for someone to figure out that, instead of measuring the code strings, much higher accuracy could be gained by measuring the wavelength of the carrier wave itself. The wavelength of the higher frequency L1 1575.42 MHz carrier is about 19 centimeters, which is much shorter than even the P-code chip length, and consequently is much more accurate.

Conceptually, the basic idea for carrier-phase measurements is similar to code measurements. Simply count wavelengths of the carrier wave instead of chip lengths of the code. For example, if you were to count 100,000,000 wavelengths between a given satellite and a receiver, and each wavelength is 19 centimeters, you could calculate that the distance was around 20,000 kilometers. Sounds simple, right? Well...

Carrier Phase Positioning

Conceptually Easy - Practically Very Difficult

- With PRN Signal, Receiver Immediately Knows Where It's At Along The Code String
- With Carrier Wave, No Way Of Knowing Where It's At Since All The Cycles Are Essentially The Same
- Called "Carrier Phase Ambiguity"
- Sophisticated Receivers Use C/A Code To "Ramp Up" To The Carrier Phase, To Within About 100 Cycles
- To Resolve Final ~100 Cycles (Ambiguity Resolution) Need To Continuously Observe Changes In Constellation
- Typically Takes 30-90 Min To Lock In The Last Few Cycles

Requires At Least Two Receivers

Carrier-Phase Positioning

Although the concept is straightforward, there are some very real difficulties that must be overcome. With a PRN (pseudo-random noise) code signal, the receiver immediately knows where it's at along the code string because the string's code sequence is unique and known by the receiver.

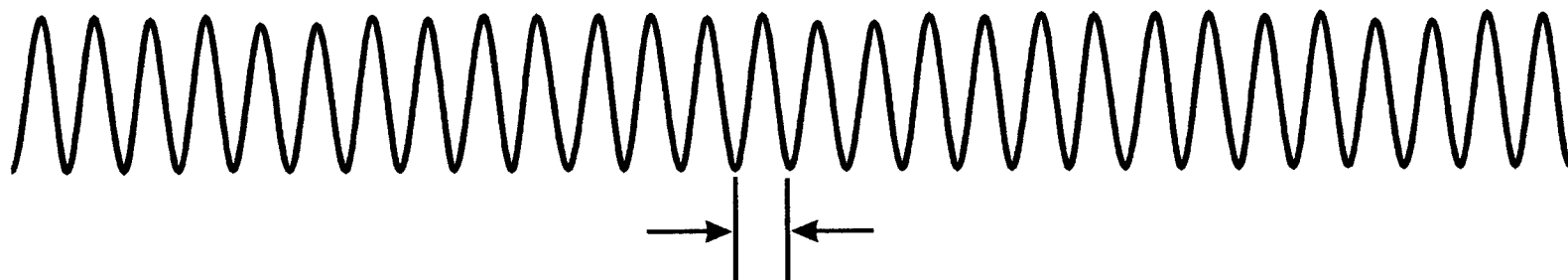
With the carrier wave, there is no way of knowing where you're at along the length of the signal since every wave (or cycle) is essentially identical to the next. This unknown is called the "*Carrier-Phase Ambiguity*." Sophisticated carrier-phase receivers can use the C/A-code to "ramp up" to the carrier-phase, or get to within about 100 or so cycles of the actual count of waves.

To resolve that final -100 cycles (called *Ambiguity Resolution*), it is necessary for the receiver to continuously observe the change in position of each of the observed satellites through time without any interruptions in the reception of the wave train (called "cycle *slips*," or, if the entire satellite radio connection is lost, even momentarily, it's called "Loss of lock," which nullifies the position resolution data set entirely). What the receiver is actually measuring is the continuous change in range (the "*delta-range*") through time of the carrier as the satellite moves through space. Once a series of delta-ranges for each satellite have been accurately measured over a span of time, any one of several different techniques can then be used to actually calculate the final solution. Different manufacturers have different mathematical models for their own systems.

Just a few years ago it took 60 to 90 minutes to sufficiently resolve the cycle ambiguity. Today, with what's called "*On The Fly Ambiguity Resolution*," or **OTF**, high-end receivers can get very good results with only a few seconds to a few minutes of data collection using only minimal satellite position shifts.

GPS Resolution - Carrier Phase

L1 Carrier 1575.42 MHz



1 Wavelength = 19 Cm

Signal Processing Allows Resolution To ~ 1 % Of Wavelength

= 0.0019 Meters (1.9 mm)

GPS Resolution — Carrier-Phase

Just like code phase measurements, accuracy using the carrier-phase measurements is directly related to wavelength. The wavelength of the L 1 carrier at 1575.42 MHz is only about 19 centimeters, or about 1/1 58th the 30-meter length of the P-code chip, or 1/1,579th the 300 meter length of the C/A-code chip. This is significantly higher accuracy than is available with either of the codes under the best of conditions.

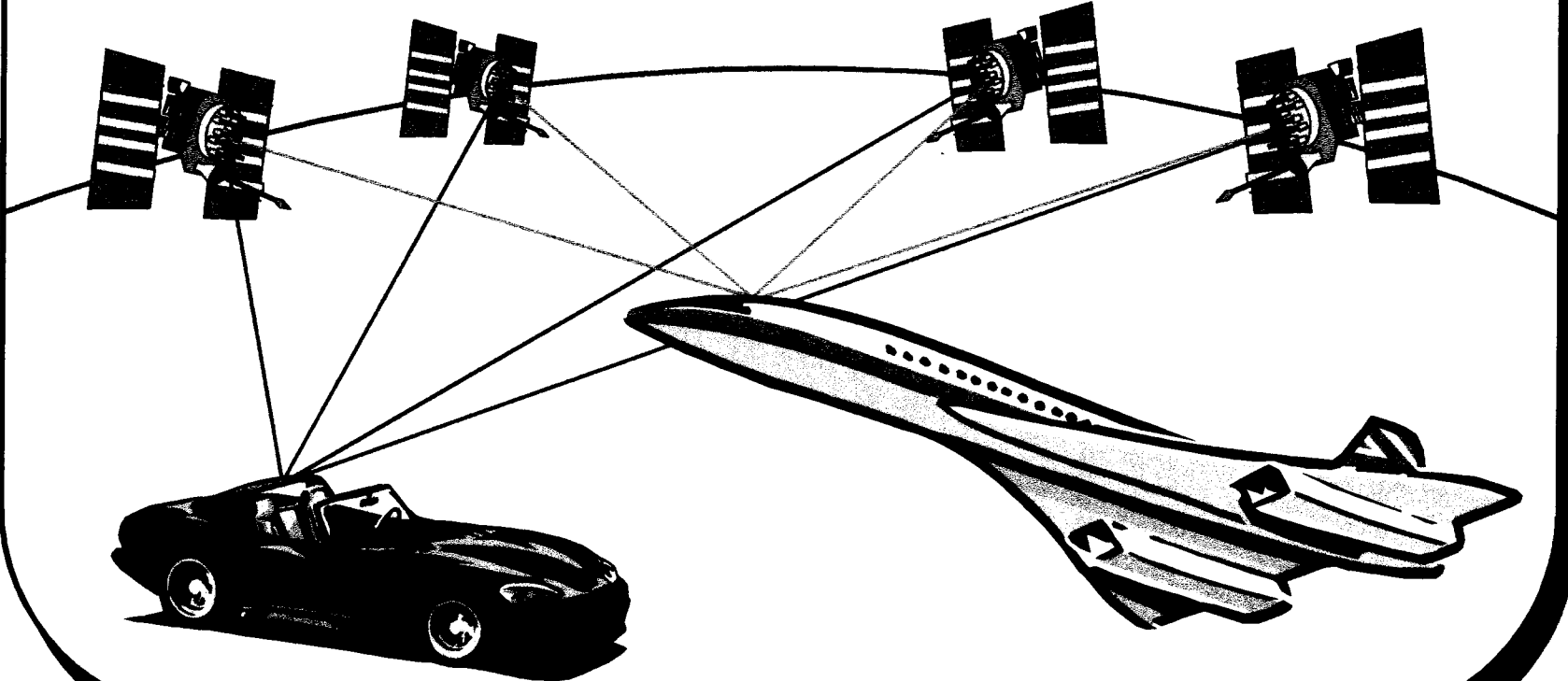
Since, as a general rule of thumb, signal processing techniques are able to refine the observation resolution to approximately one percent of the signal's wavelength, the resulting potential precision for a carrier-resolved resolution is down to 1.9 millimeters!

Although that is the *theoretical* maximum resolution possible in carrier-phase positioning, modern geodetic surveying receivers are regularly achieving testable and repeatable accuracy in the area of one-half centimeter, or around 30-50 millimeters. Some claim even higher accuracy.

As we will see next, this kind of accuracy does not come easily. In the first place, such precision cannot be achieved autonomously. Instead, at least two receivers must be used simultaneously. Why? To correct for a whole host of other sources of serious error.

Velocity

***GPS Gives More Than Just Positions
It Can Tell Us How Fast We're Moving, Too***



GPS Velocity

Positioning isn't the only thing that can be accomplished with the Global Positioning System. Another important function is for Navigation, or the measurement of instantaneous position, *velocity*, and heading. Instantaneous position is measured just as we've described. However, while velocity could, by extrapolation, be calculated by simply differencing the positions between time 1 and time 2, it is more frequently accomplished in a slightly different manner.

Because of the relative motion of the GPS satellites with respect to a receiver, the frequency of a signal broadcast by the satellites is always going to be "shifted," or slightly compressed or expanded, when received. This *Doppler shift* is proportional to the relative velocity between the satellite and receiver. The velocity of the satellites themselves as they move across the sky is known and is transmitted as part of the NAV-msg signal. Any *additional* Doppler shift that exists in the signal must, therefore, be due to motion of the receiver itself. From this, the receiver can deduce its own velocity from the measurement of any Doppler shift that is above and beyond that which is occurring as a result of the satellite's motions. This method of velocity calculation is virtually instantaneous and is extremely accurate. Typical velocity accuracy for a receiver with Selective Availability turned off (more about this later) is on the order of 0.5 kilometer per hour.

Heading, or direction of travel, is calculated in a more straightforward manner. Simply projecting a line from one position to another results in a direction of travel. By looking at a sequence of positions, the receiver can average out any individual position variation and produce a direction of travel, or heading, that is accurate to within one or two seconds of arc. In this manner, any moving GPS receiver can also be used as an accurate compass.

GPS Error Budget

● Satellite Clock Error	0-1.5 M
● SV Ephemeris Error	1-5 M
● Ionospheric Refraction	0-30 M
● Tropospheric Refraction	0-30 M
● Receiver Noise	0-10 M
● Multipath	0-1 M
● Selective Availability (30 Avg)	0-70 M

Dilution Of Precision: Above x 1 - 6

Total Error W/SA On: ~100 M

Total Error W/SA Off: ~28 M

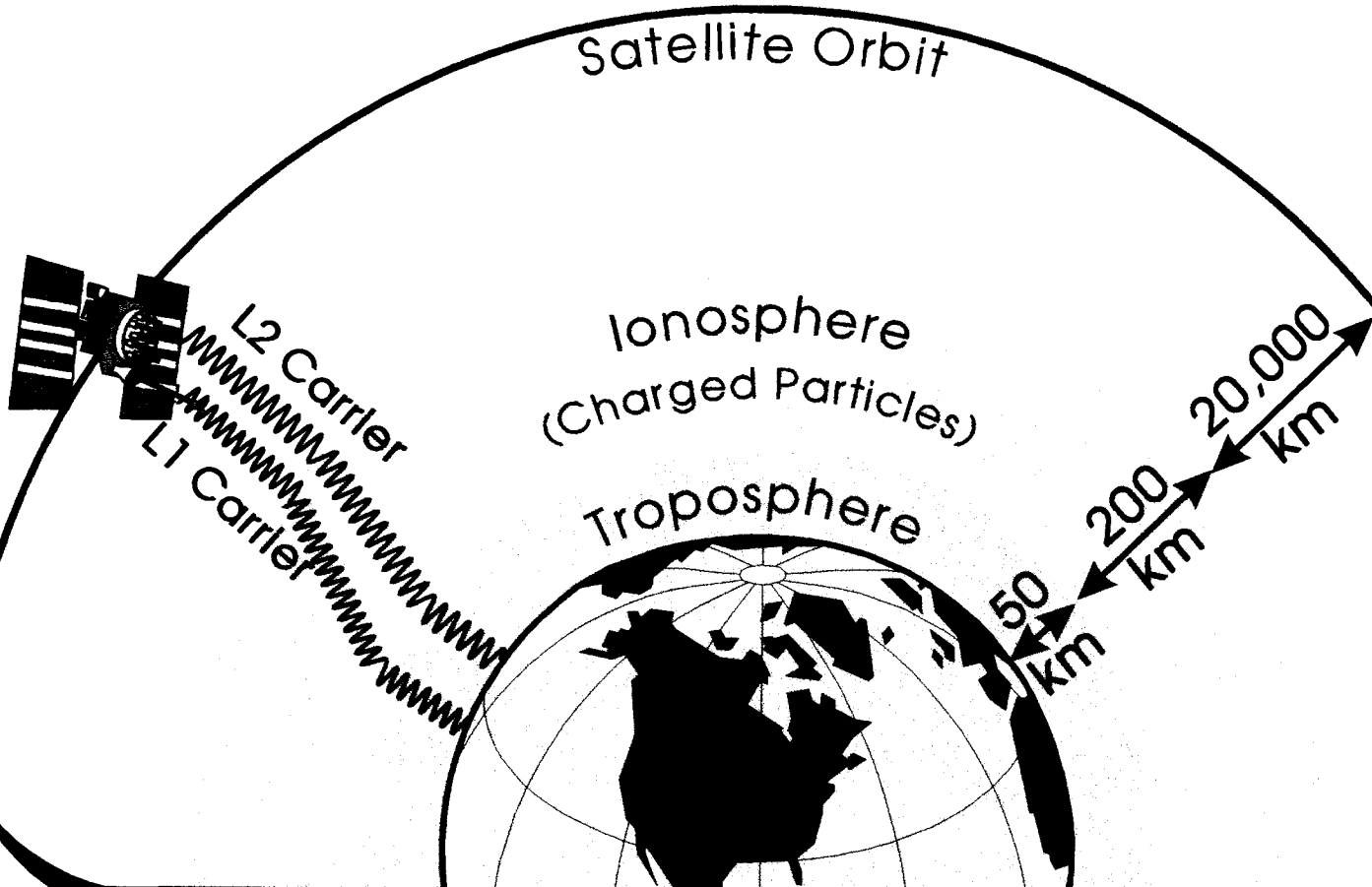
GPS Error Budget

Most of the error figures given thus far represent only best-case scenarios-which, unfortunately, never exists in reality. There are, in fact, several sources of error that severely degrade the accuracy of all forms of GPS positioning. They include satellite clock timing error, satellite position error (ephemeris error), ionospheric and tropospheric refraction, receiver noise, multipath, and, worst of all, something called *Selective Availability*, or SA which we'll get to shortly. Finally, the sum of the errors is multiplied by a factor of 1 to 6, a figure that represents the *Dilution of Precision*, or *DOP* (also to be discussed shortly).

Together they result in potentially very high error values, collectively referred to as the **UERE**, or *User Equivalent Range Error*. Fortunately, never are all of the factors operating at their worst at any given time and, in fact, vary widely. Without Selective Availability, these errors result in autonomous C/A-code positions that can be expected to be within around 28 meters or so (horizontal). With SA turned on, that figure is somewhere under 100 meters. We'll discuss each of these sources of error then follow with some techniques that minimize or effectively eliminate them altogether.

The first of these are the satellite position and clock timing errors. These errors are typically low since the Air Force constantly monitors each satellite and sends up correction data every few hours. It can happen, however, that something might have caused, for example, a satellite to shift slightly in its orbit. If you happened to take a series of positions using that satellite before the Control Segment could get a correction up to it, you could get a position error and not know it. You can, however, acquire "*precise ephemeride*" data after the fact from several private and public sources, including the government, which most higher-end GPS software packages can use to essentially eliminate this error. We'll look more closely at these and other data sources later on.

Ionospheric / Tropospheric Refraction

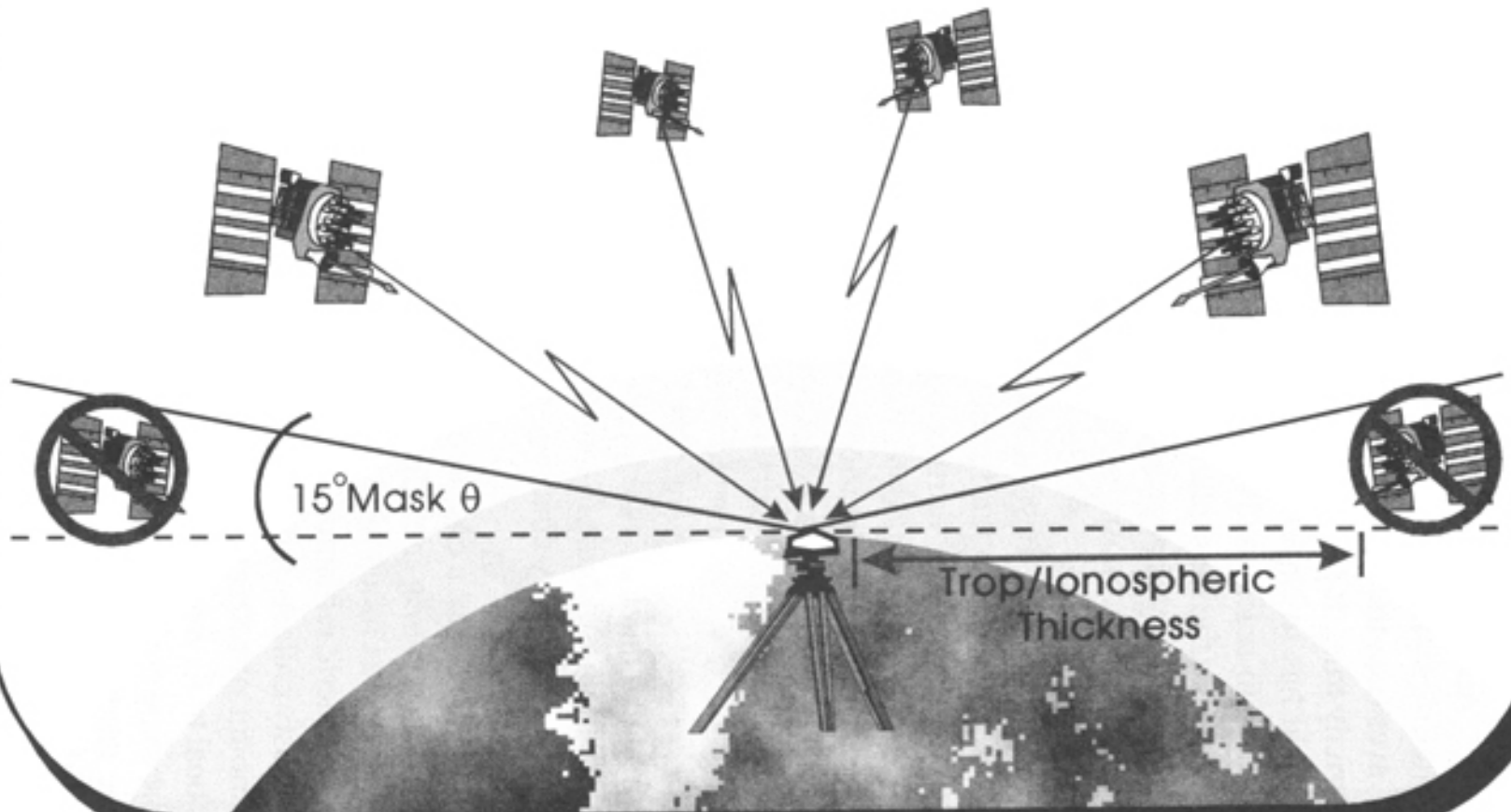


Ionospheric/Tropospheric Refraction

Another problem area is the atmosphere itself, through which the satellite signals must pass. If you recall, we said that the speed of radio waves (i.e., the speed of light) is a constant at 300,000 km/sec. That's not strictly true. It **is** true in the perfect vacuum of space. Unfortunately, the signal has to travel through some 300 kilometers of the Earth's atmosphere to reach us. The two most troublesome components of the atmosphere **are** the **ionosphere** and the **troposphere**. The ionosphere is a layer of electrically charged particles between around 50 and 200 kilometers altitude. The troposphere is simply what we **usually** think of as the atmosphere, extending from the surface up to between eight and 16 kilometers altitude. Each of these literally "drag" radio waves down, causing them to bend a tiny, but significant, amount. This "bending" of radio waves is called refraction. Further complicating the problem is the fact that the ionosphere and troposphere each refract differently. The problem with the ionosphere is the electrically charged particles that "drag" on the incoming signal. In the troposphere, the problem is with the water vapor content which does the same thing, just at a different rate. These problems are even further exacerbated when a satellite is low on the horizon. This is because a line tangent to the surface of the Earth (or nearly so) passes through a much thicker layer of atmosphere than if that line were pointing straight up. And just to muddy the waters a bit more, the amount of refraction is constantly changing with changing atmospheric conditions.

There are a couple of ways to deal with refraction. First, the satellite's NAV-msg includes an atmospheric refraction model that compensates for as much as 50-70% of the error. A more effective method for ionospheric refraction is to use a **dual-frequency** receiver which simultaneously collects the signals on both the L1 and L2 carriers. Because the amount of refraction that a radio wave experiences is inversely proportional to its frequency, using two different frequencies transmitted through the same atmosphere at the same time makes it relatively easy to compute the amount of refraction taking place and compensate for it. Unfortunately, tropospheric refraction is not frequency-dependent and so cannot be corrected by this method.

SV Mask Angle



SV Mask Angle

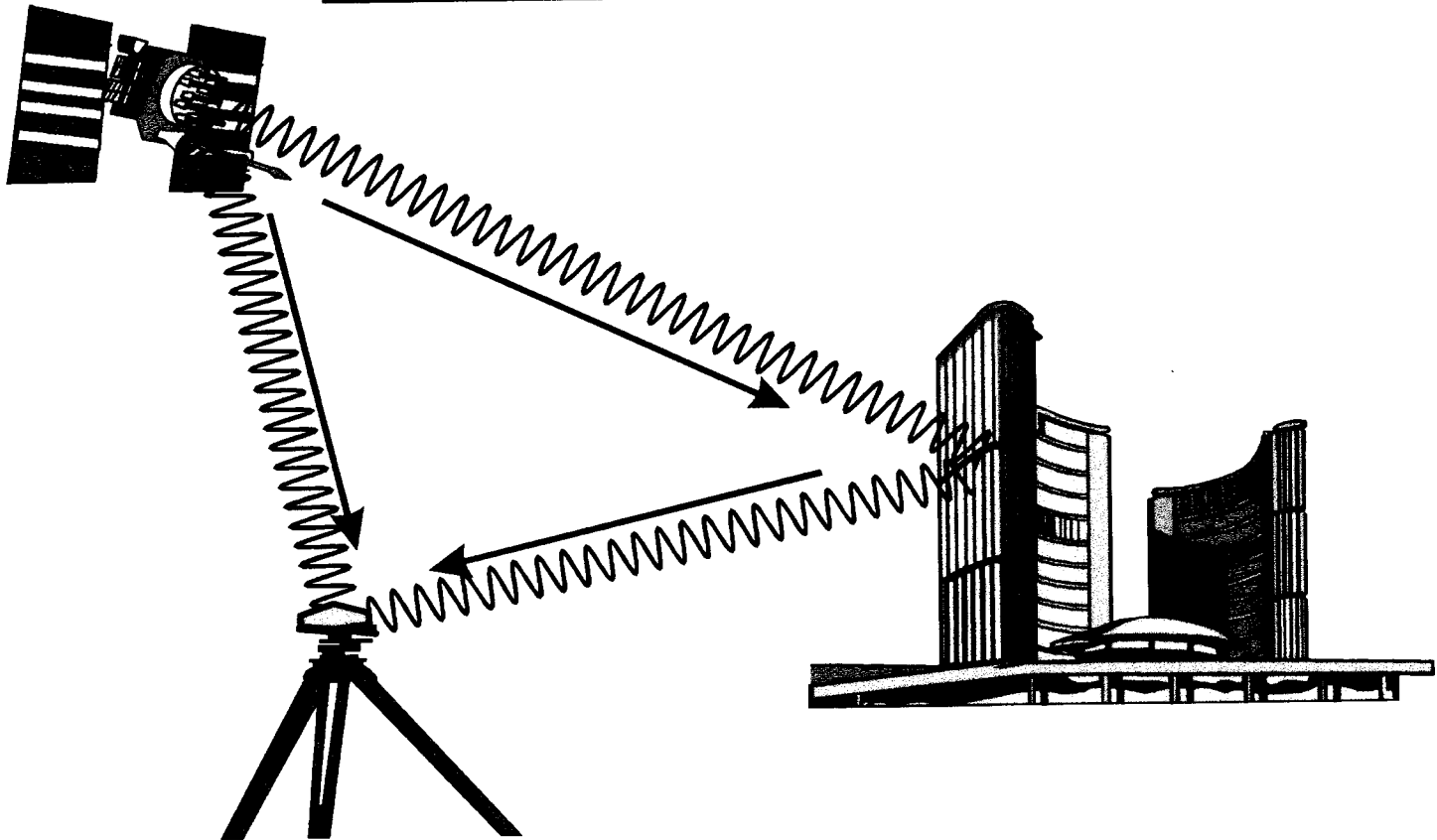
While dual-frequency receivers can virtually eliminate the ionospheric refraction problem, they're very expensive. However, the problem can be minimized with even the more commonly used single frequency receiver (likely receiving the L1 band alone). Nearly all GPS receivers, inexpensive or expensive, have a "*Mask Angle*" setting. This means that the receiver can be set to ignore any satellite signals that come from below a user-definable angle above the horizon, or "mask" them out. The most typical mask angle is usually somewhere between 10 and 15 degrees.

The drawback here is that setting the mask angle too high might exclude satellites needed to acquire the necessary minimum of four. It's a trade-off. Are you so desperate for a position at that exact time that you're willing to accept a degraded signal? It does happen. In that case, the mask angle could be set to maybe 5 degrees, or even to zero if there's a clear view of the horizon, such as at sea, and simply accept a degraded signal and possibly (probably) a poorer accuracy as a result.

In most cases it's better to keep the mask angle at that upper end of around 15 to (at most) 20 degrees and just wait for a sufficient number of satellites to become available above the mask. Now that the full GPS constellation is complete, there will rarely be times with too few satellites sufficiently high in the sky to get a good position.

Another potential source of error is receiver noise, or electronic noise produced by the receiver itself that interferes with the very weak incoming signal. While this error is highly variable among receiver brands, most have some kind of internal filtering designed to minimize the problem some better than others. Manufacturers of higher-end receivers have gone to great lengths to lower this source of error to where it is virtually insignificant.

Multi-Path Errors



Multi-Path Errors

Another potential, though relatively minor, source of signal error is *Multi-Path*. Multi-Path is simply the reception of a reflected satellite signal. With multi-path reception, the receiver collects both the direct signal from the satellite and a fractionally delayed signal that has bounced off of some nearby reflective surface then reached the receiver. This is the same kind of thing seen in television “ghosts.”

The problem is that the path of the signal that has reflected off some surface is longer than the direct line to the satellite. This can “confuse” some lower-end receivers resulting in an incorrect range measurement and, consequently, an incorrect position.

There are several ways to deal with this problem. Most receivers have some way of “seeing” and comparing the correct and incorrect incoming signal. Since the reflected multi-path signal has traveled a longer path, it will arrive a fraction of a second later, *and* a fraction weaker than the direct signal. By recognizing that there are two signals, one right after another, and that one is slightly weaker than the other, the receiver can reject the later, weaker signal, minimizing the problem. This ability is referred to as the receiver’s multi-path rejection capability.

Mapping and survey quality receivers also use semi-directional, ground-plane antennas to reduce the amount of multi-path that the receiver will have to deal with. Semi-directional antennas are designed to reject any signal below a tangent to the surface of the Earth, meaning that they are preferentially directional upward. This is usually seen as a large (up to 20 to 30 centimeters across) flat metal plate (usually aluminum) with the actual, much smaller, receiver antenna attached on top. The metal plate interferes with any signals that may be reflected off of low reflective surfaces below them, such as bodies of water.

Selective Availability (S/A)

- The Degradation of Position Accuracy By The DOD
- Done To Deny Hostiles High Accuracy Positioning

Done By:

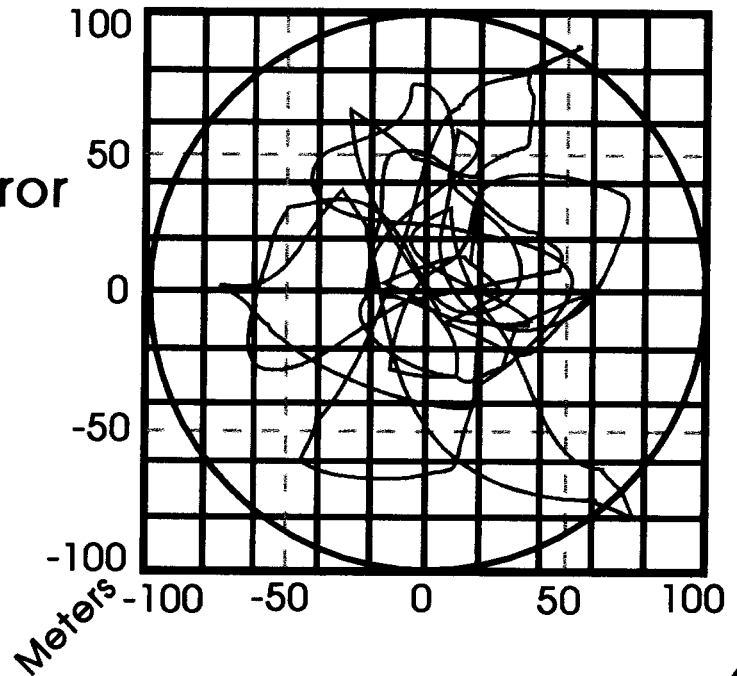
- Off-Setting SV Clocks
- Injection of Ephemeris Error

With S/A On:

~100 Meters (2drms)

With S/A OFF:

~20-40 Meters (2drms)



Selective Availability

Selective Availability is far and away the worst source of error in GPS positioning, producing up to 70 meters of positional displacement alone. And it's deliberate! Selective Availability is the intentional degradation of the GPS signal by either dithering the clocks or the orbital information to produce incorrect satellite positions and, thereby, provide incorrect receiver positions. The purpose is to limit accuracy for non-PPS-authorized users to a 95% probability of 100 meters or less. The amount of error induced into each satellite's clock and ephemeris data varies from satellite to satellite and is continuously varied in degree over periods of hours. This means that the error can't be averaged out with data collection periods of less than several hours, effectively eliminating any possibility of acquiring the higher potential accuracy of GPS in real-time.

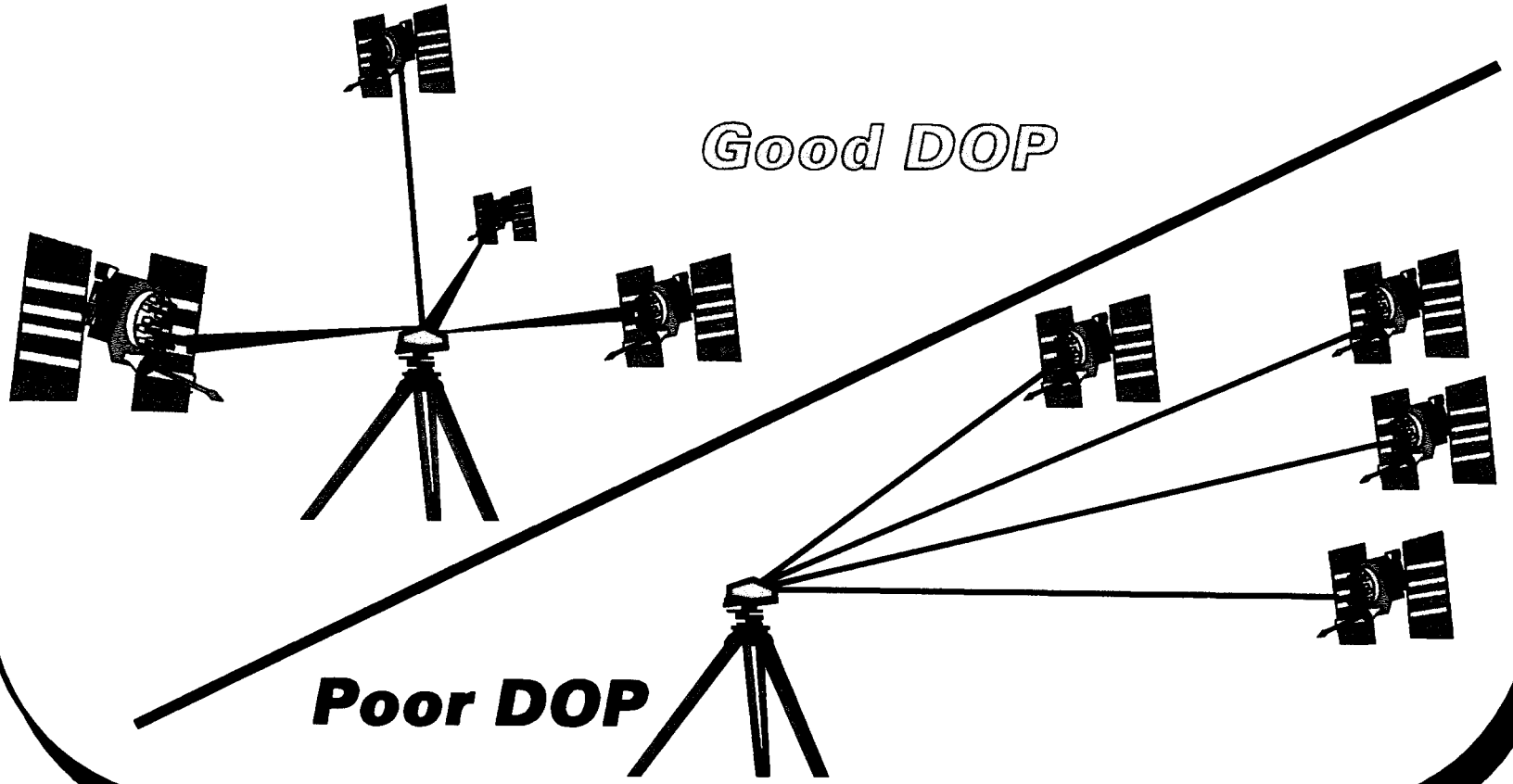
The diagram at left illustrates what a long-term (several hour) position-plot would look like with Selective Availability turned on. While the actual receiver position is located at zero-zero, the receiver-perceived position drifts over an area as wide as 100 meters (or possibly even more for as much as 5% of the time) from that true point.

It's interesting to note that during the Gulf War a large number of civilian receivers were fielded by the military which were unable "see through" the Selective Availability the way that military-designed receivers can. Consequently, the Department of Defense had to turn SA off for the duration of the war, which is directly *opposite* the intended purpose of SA! As soon as the war was over, they promptly turned it back on.

With the declaration of Full Operational Capability (FOC) in July, 1995, with a full complement of Block II satellites in orbit and operating, Selective Availability has been turned on continuously. However, on March 29, 1996, President Clinton released a Policy Fact Sheet that declared that the S/A would be turned off permanently within the next four years. When exactly can't be said although most industry analysts believe that it will be considerably sooner rather than later.

Dilution Of Precision (DOP)

A Measure Of The Geometry Of The Visible GPS Constellation



Dilution of Precision (DOP)

The cumulative UERE (User Equivalent Range Error) totals are multiplied by a factor of usually 1 to 6, which represents a value of the *Dilution of Precision*, or DOP. The DOP is, in turn, a measure of the geometry of the visible satellite constellation.

The ideal orientation of four or more satellites would be to have them equally spaced all around the receiver, including one above and one below. Because we're taking our position from only one side of the Earth, that's really not possible since that part of space is blocked by the planet itself.

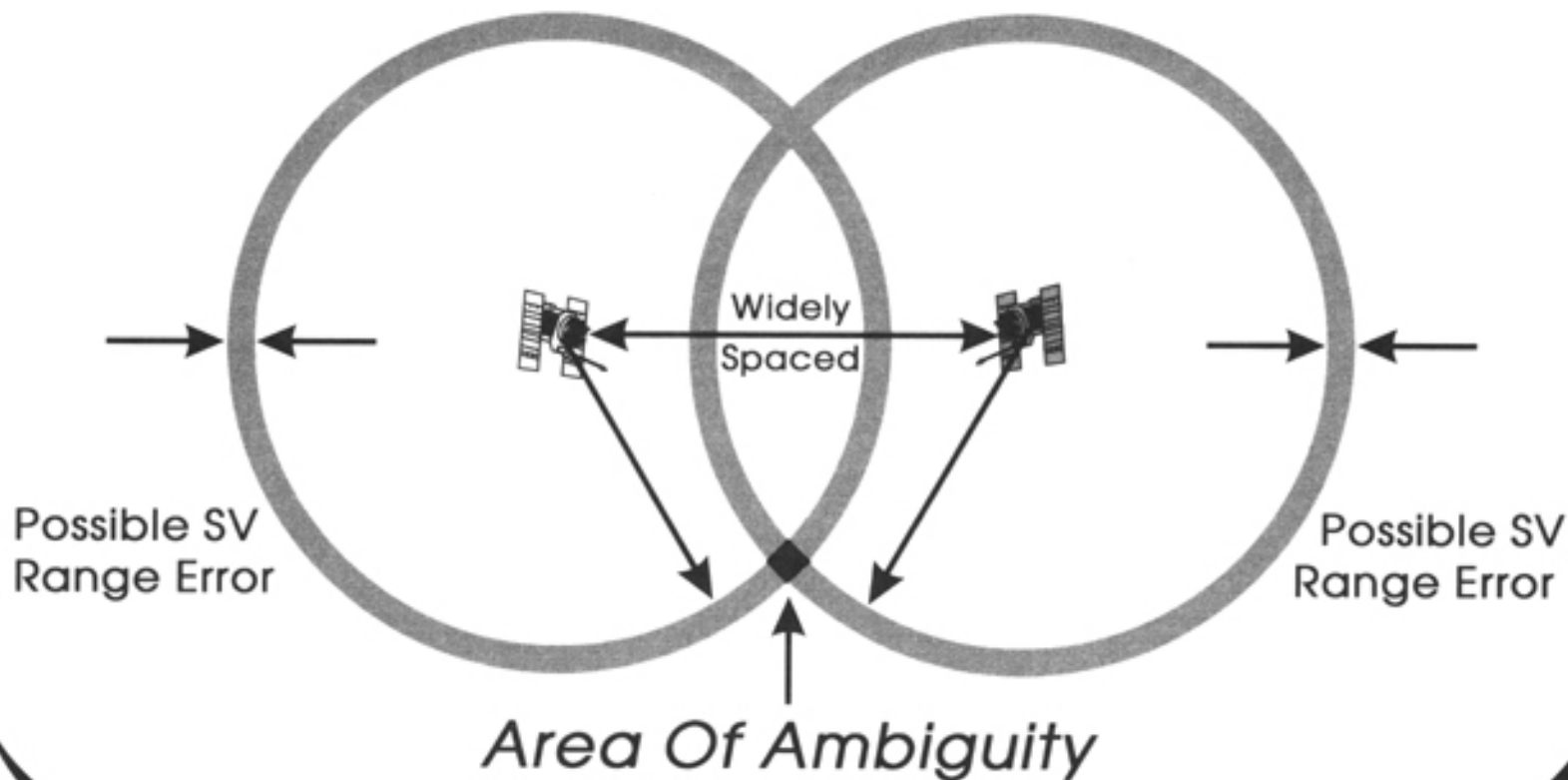
The upper diagram at left illustrates the next best orientation. That is, to have one satellite directly above and the other three evenly spaced around the receiver and elevated to about 25 to 30 degrees (to help minimize atmospheric refraction). This would result in a very good DOP value.

The lower diagram illustrates poor satellite geometry. In this case, all of the satellites are clustered together. This would result in a poor DOP value.

A low numeric Dilution of Precision value represents a good satellite configuration, whereas a higher value represents a poor satellite configuration. The DOP at any given moment will change with time as the satellites move along their orbits.

Dilution Of Precision

Good DOP



Dilution of Precision (DOP)

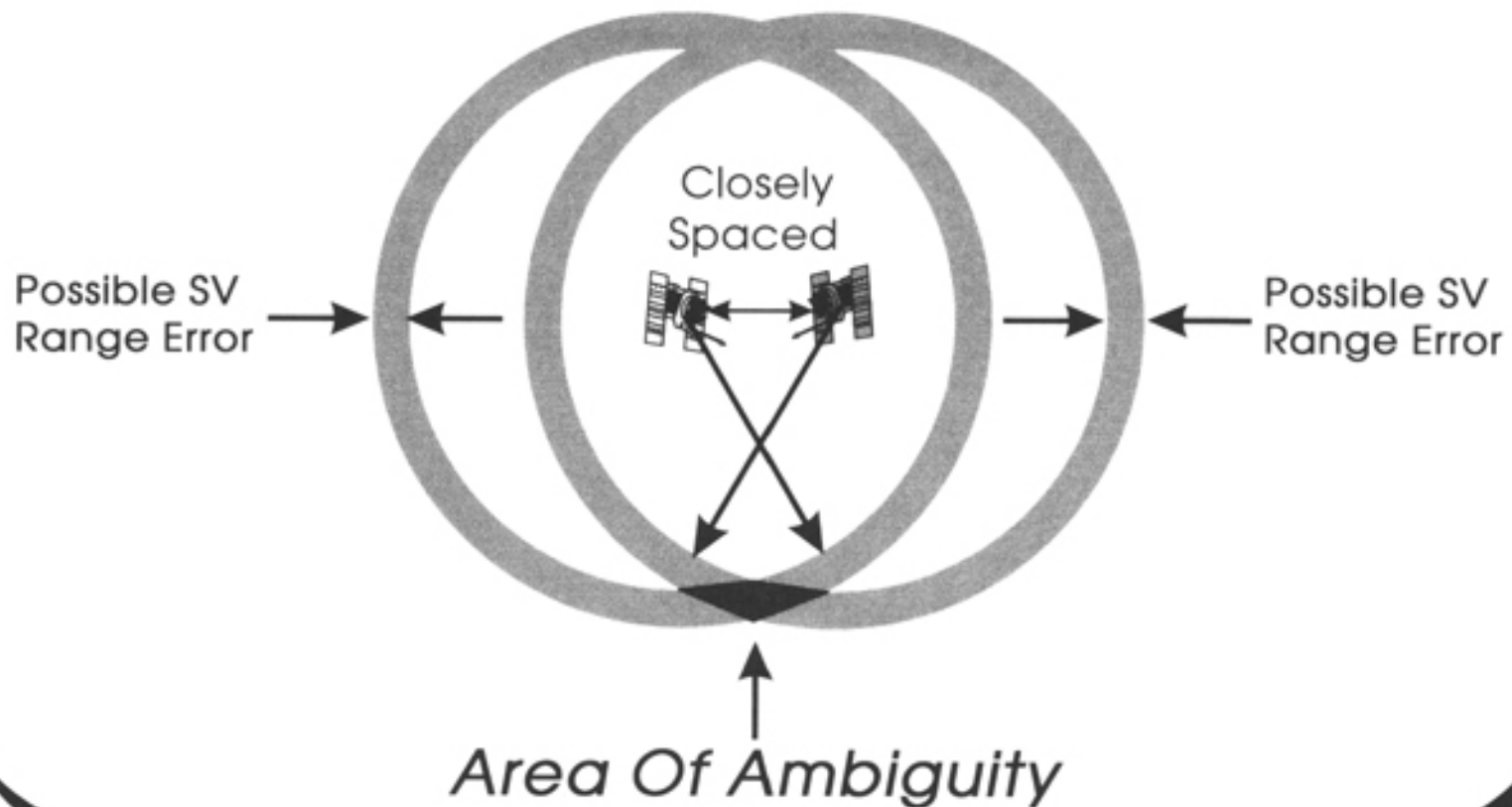
Why can satellite geometry so adversely affect accuracy? Because of the sources of error already discussed, there is inherently a certain range of possible error in the distance calculation from any given satellite. That “range error” is variable but applies to all ranges derived from all satellites.

When the satellites are widely spaced, the overlap area of the two zones of possible satellite range error is relatively small, called the “*area of positional ambiguity*.”

The diagram at left illustrates a pair of widely spaced satellites which would result in a good, or low Dilution of Precision value. The true position is somewhere in the area where the two “fuzzy position” zones overlap (indicated in the diagram as a small square). In this case, the area of positional ambiguity is relatively small.

Dilution Of Precision

Poor DOP



Dilution of Precision (DOP)

When the satellites are closely spaced, the overlap area of the two zones of possible satellite range error is considerably larger than when the satellites are spaced farther apart.

The diagram at left illustrates a pair of closely spaced satellites which would result in a poor, or high Dilution of Precision value. As in the case of the widely spaced satellite configuration, the true position is somewhere in the area where the two “fuzzy position” range zones overlap (indicated in the diagram as a small diamond). However, in this case, the area of ambiguity is large.

Dilution Of Precision (DOP)

- *PDOP* = Position Dilution of Precision
(Most Commonly Used)
- *VDOP* = Vertical Dilution of Precision
- *GDOP* = Geometric Dilution of Precision
- *HDOP* = Horizontal Dilution of Precision
- *TDOP* = Time Dilution of Precision

QUALITY	DOP
<i>Very Good</i>	1 - 3
<i>Good</i>	4 - 5
<i>Fair</i>	6
<i>Suspect</i>	> 6

Dilution of Precision (DOP)

There are a number of Dilution of Precision components. The overall **GDOP**, or *Geometric Dilution of Precision* includes:

PDOP, or *Precision Dilution of Precision*, probably the most commonly used, which is the dilution of precision in three dimensions. Sometimes called the *Spherical DOP*.

HDOP, or *Horizontal Dilution of Precision*, is the dilution of precision in two dimensions horizontally. This value is often lower (meaning “better”) than the PDOP because it ignores the vertical dimension.

VDOP, or *Vertical Dilution of Precision*, is the dilution of precision in one dimension, the vertical.

TDOP, or *Time Dilution of Precision*, is the dilution of precision with respect to time.

A DOP value of less than 2 is considered excellent-about as good as it gets, but it doesn't happen often, usually requiring a clear view of the sky all the way to the horizon. DOP values of 2 to 3 are considered very good. DOP values of 4 or below are frequently specified when equipment accuracy capabilities are given.

DOP values of 4 to 5 are considered fairly good and would normally be acceptable for all but the highest levels of survey precision requirements. A DOP value of 6 would be acceptable only in low precision conditions, such as in coarse positioning and navigation. Position data generally should not be recorded when the DOP value exceeds 6.

PROJECT PLANNING

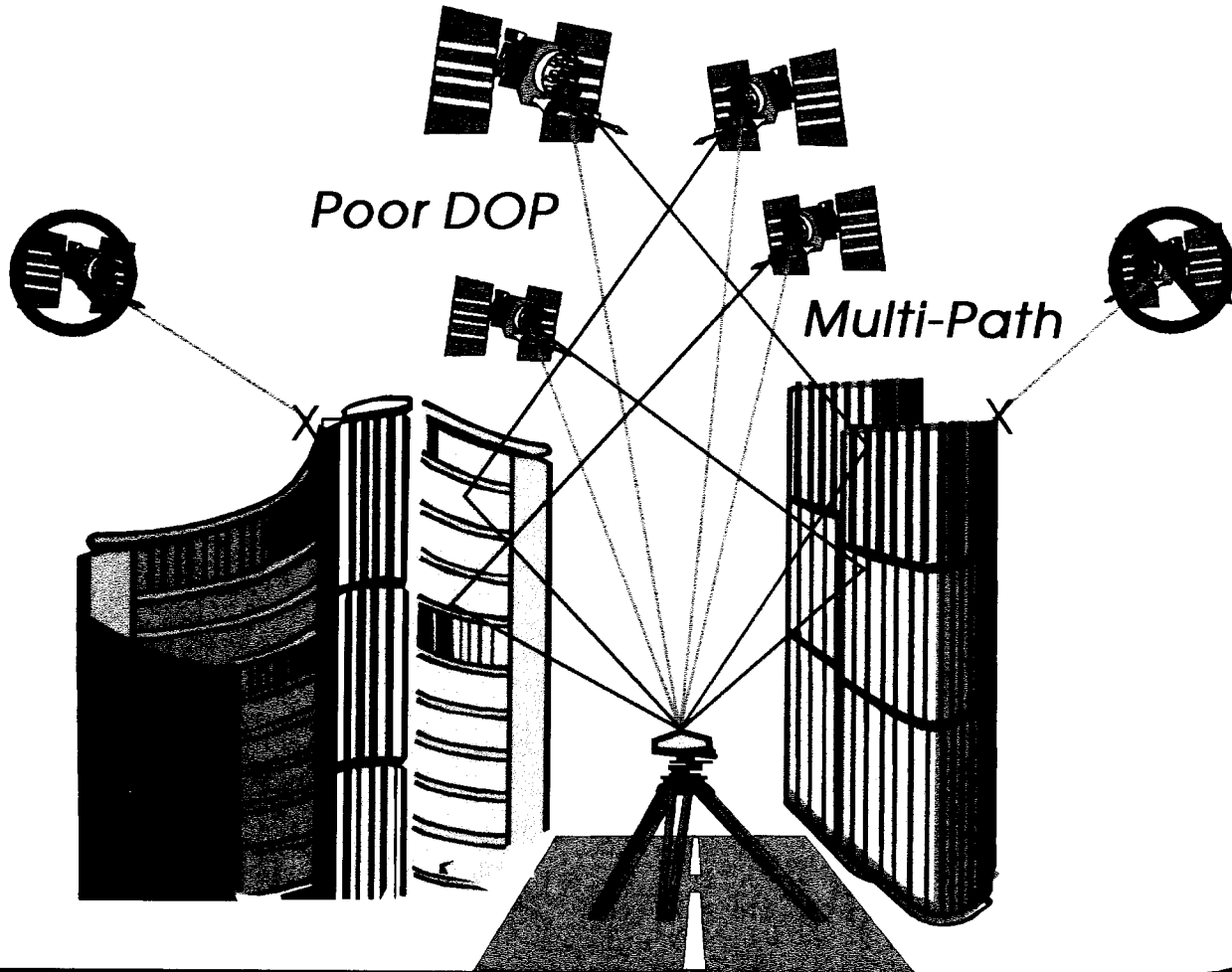


Project Planning

With the ever-growing ease-of-use of modern GPS receivers, it's all too easy to just throw everything into the trunk of a car and head out into the field whenever the "urge" to collect some positional data hits. However, as for any important project, pre-planning should always be mandatory. There are a number of things that can be done back at the office before ever heading out to the field, and unless the only interest is in very coarse positioning or navigation, the conditions under which the data are to be collected should always be carefully considered beforehand.

We've already seen that pre-setting the equipment to reject satellite signals that are too close to the horizon can help reduce atmospheric interference. If position data are to be collected in locations where there are numerous obstructions, setting the receiver's mask angle even higher than normal, say 20 or even 30 degrees, might be prudent. There are, however, also other potential problems to be considered before setting out for the field.

"Condo Canyon"



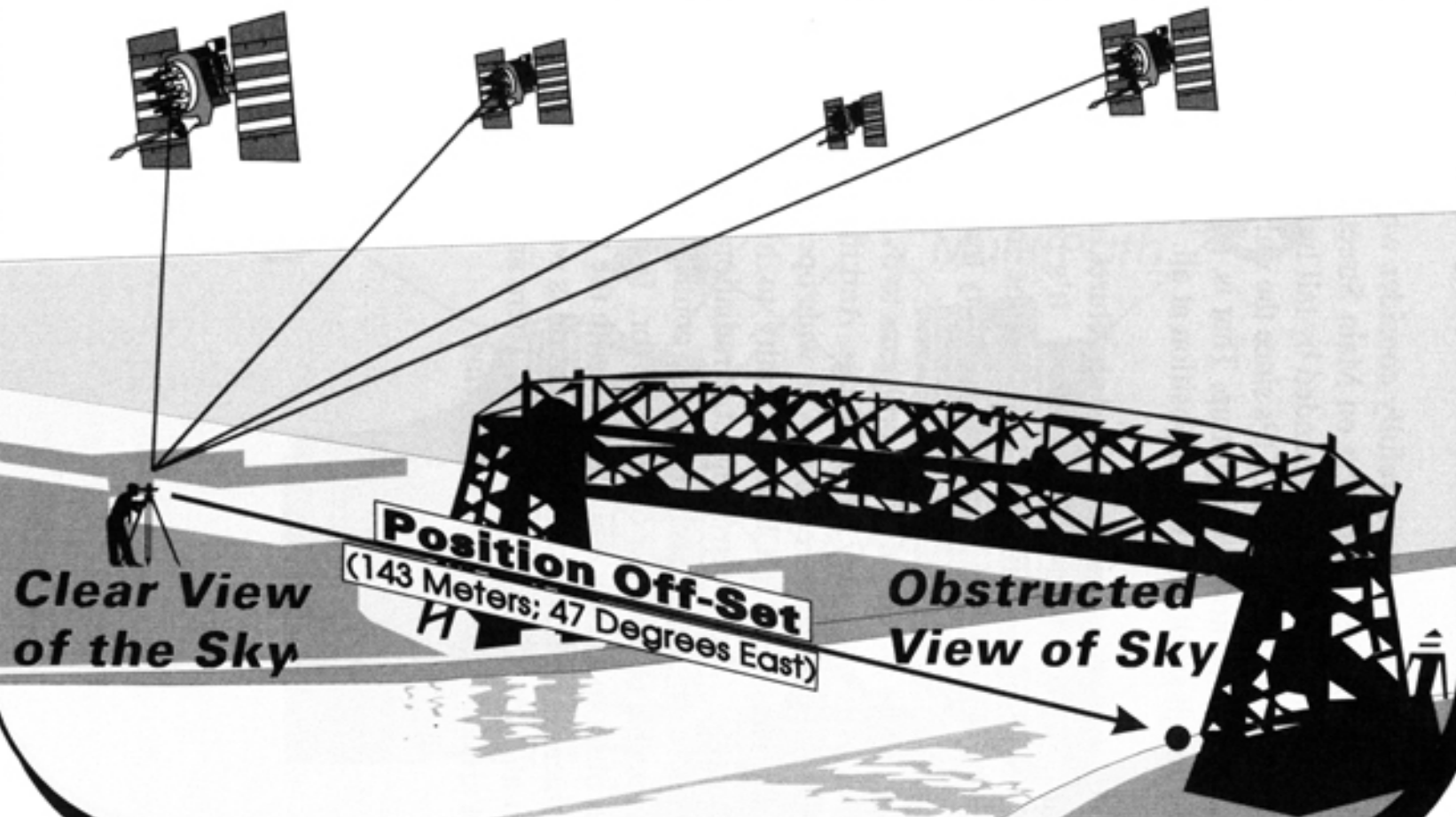
“Condo Canyons”

It's important to carefully consider where the data are to be collected. Is the area of interest on Main Street of a large city? If so, the receiver is likely to be surrounded by tall buildings that restrict satellite visibility resulting in poor DOPs since the only satellites that the receiver can see will be nearly straight up. That is, provided it's even possible to see enough satellites to get a position at all.

In addition, the glass-sided structures all around the receiver act as nearly perfect multi-path reflectors. It's possible that, because of the efficiency of the buildings to reflect the incoming satellite signal, the receiver's multi-path rejection capability may actually be overloaded.

These are very difficult problems to overcome, particularly in dense urban areas with many tall buildings. And the problems aren't just in the cities. Even out in the country with wide open spaces there are conditions to be considered. Close proximity to high-power lines could be a problem. The electromagnetic radiation surrounding the lines can interfere with the satellite signal, contributing an error that is nearly impossible to model or compensate for. Forests with dense canopy cover can obscure the sky and interfere with the incoming satellite signal. The problem is even worse if the vegetation is wet since the liquid water itself can also interfere with the signal. There are, however, some methods to get around these potential problems.

Off-Set Positions



Off-Set Positions

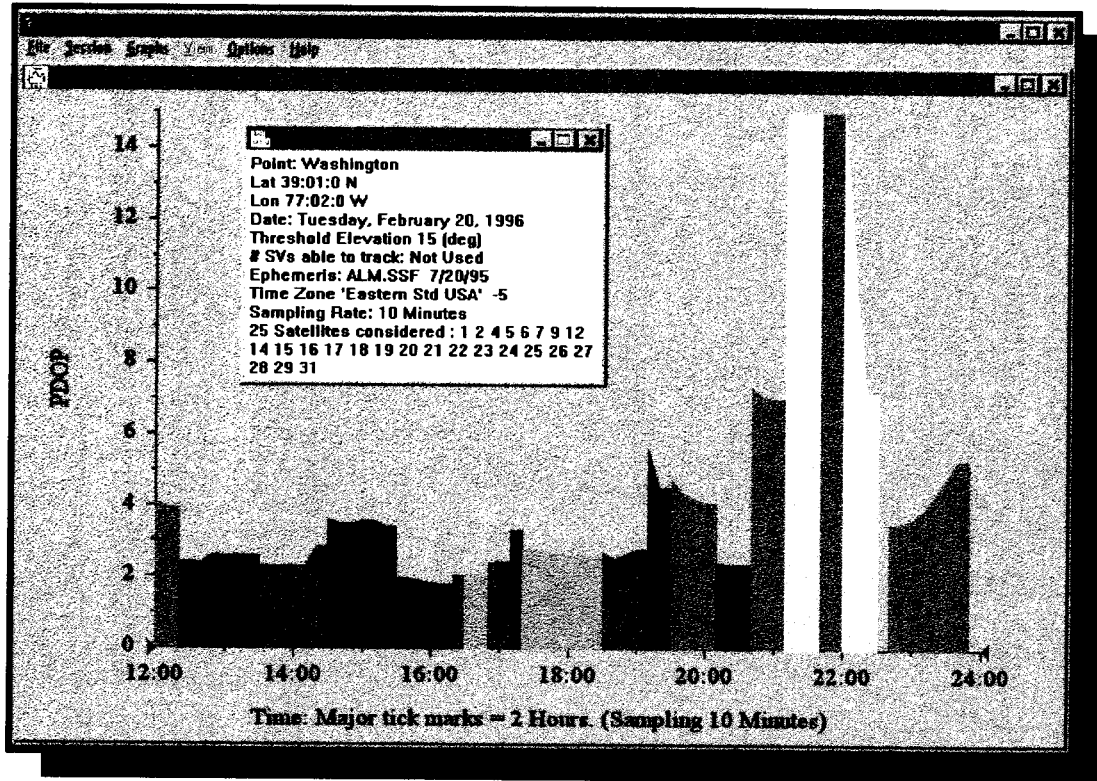
One method of getting around this problem is to “off-set” the position desired. What this means is to acquire the position for some place or object of interest without ever needing to occupy that location directly. This can be done by several methods.

One way is to tell the GPS receiver that the position you really want is “over there” somewhere, rather than where you actually happen to be standing with your receiver. Most receivers have software in them that allows the user to input such information directly. For example, if you happen to need a position of a fire hydrant that’s located under an overpass, you could set up your receiver some distance away where it has an unobstructed view of the sky. Then you could physically measure the ground distance and compass direction from the receiver to the hydrant and input that information directly into the receiver, usually with an attached keypad. The receiver then will automatically adjust for that offset and give the desired, actual position of the hydrant.

Another method is to use **an EDM**, or *Electronic Distance Measuring* device such as a laser range finder. This is a two-way ranging device that sends out a laser beam that bounces off of the target of interest (e.g., the hydrant) and returns to the device. Since the speed that the beam travels is known (the speed of light), it’s a relatively simple matter for the range finder to quickly calculate the distance based on the time in transit. Many mapping grade receivers have input ports that allow the user to directly connect such a device. In this case, the user has only to aim the range finder at the hydrant and the distance is instantly calculated and fed to the receiver, which then calculates the position of the hydrant by applying the range finder’s offset measurement to its current position.

The distances over which these units can operate can be considerable, some achieving high accuracy across many kilometers. Range and accuracy are, of course, a function of cost. In most cases, distances of only 100 to 200 meters are adequate and so EDM's that operate within that range tend to be quite reasonably priced.

Mission Planning



Mission Planning

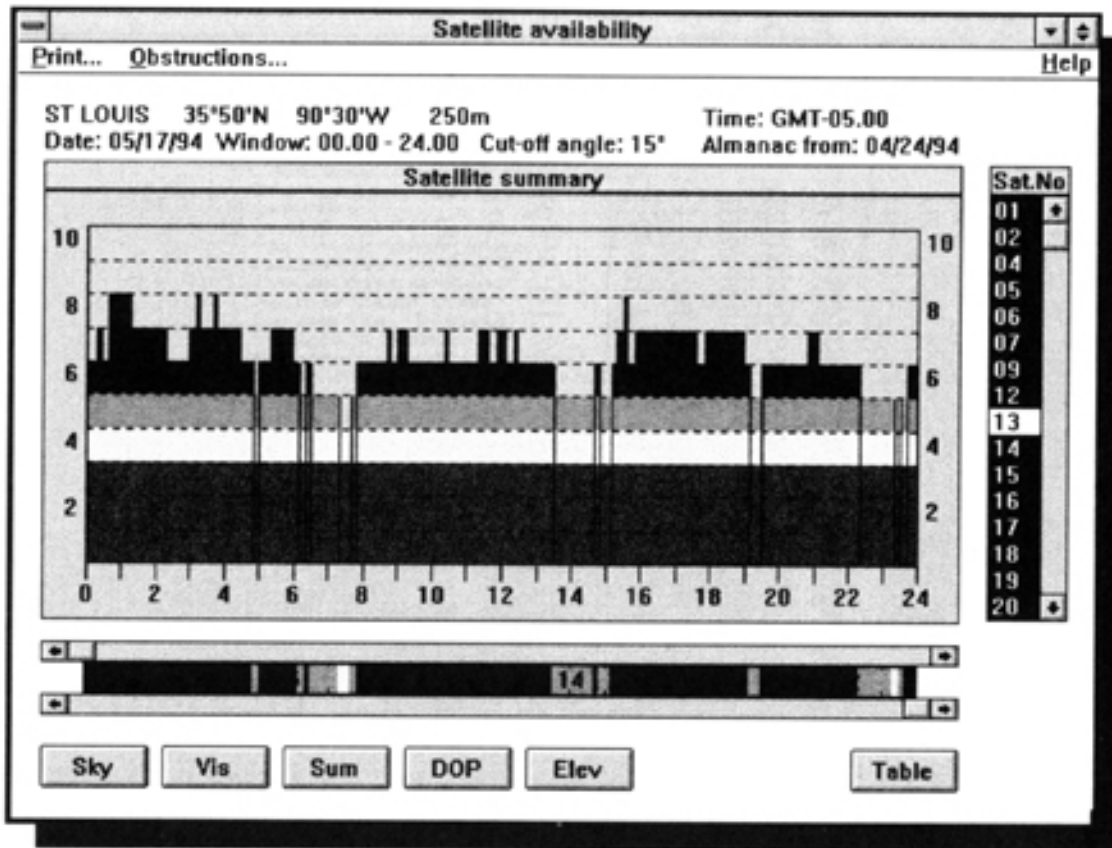
Mission planning should include a determination of what the satellite viewing conditions are going to be when the receiver is actually turned on in the field. Nearly all manufacturers of mapping and survey grade* receivers provide software that allows the user to predict what satellites will be available and where they'll be in the sky for the time period when the user will be in the field. By simply setting the approximate time and location of when and where the data are to be collected in the software, users can graphically see what kind of DOPs to expect (among other things).

The diagram at left is a program screen-dump showing the PDOPs for Washington, D.C., on Tuesday, February 20, 1996, between 12:00 noon and 12:00 midnight. DOP values remain below 4 until around 7:00 pm, when they jump up to almost 6. The PDOP actually drops down to 2 from 3:30 p.m. to about 4:30 pm, ideal for the most accurate positioning.

Between 9:00 p.m. and about 10:30 pm, the DOPs become unacceptably poor, actually going off-scale. Clearly, a user would not attempt to collect valid data during that time, even if trying to operate at night were not a problem in itself.

*In general, mapping grade receivers are those that are capable of accuracy under one meter; whereas, survey or geodetic receivers are those capable of centimeter accuracy.

Mission Planning



Mission Planning

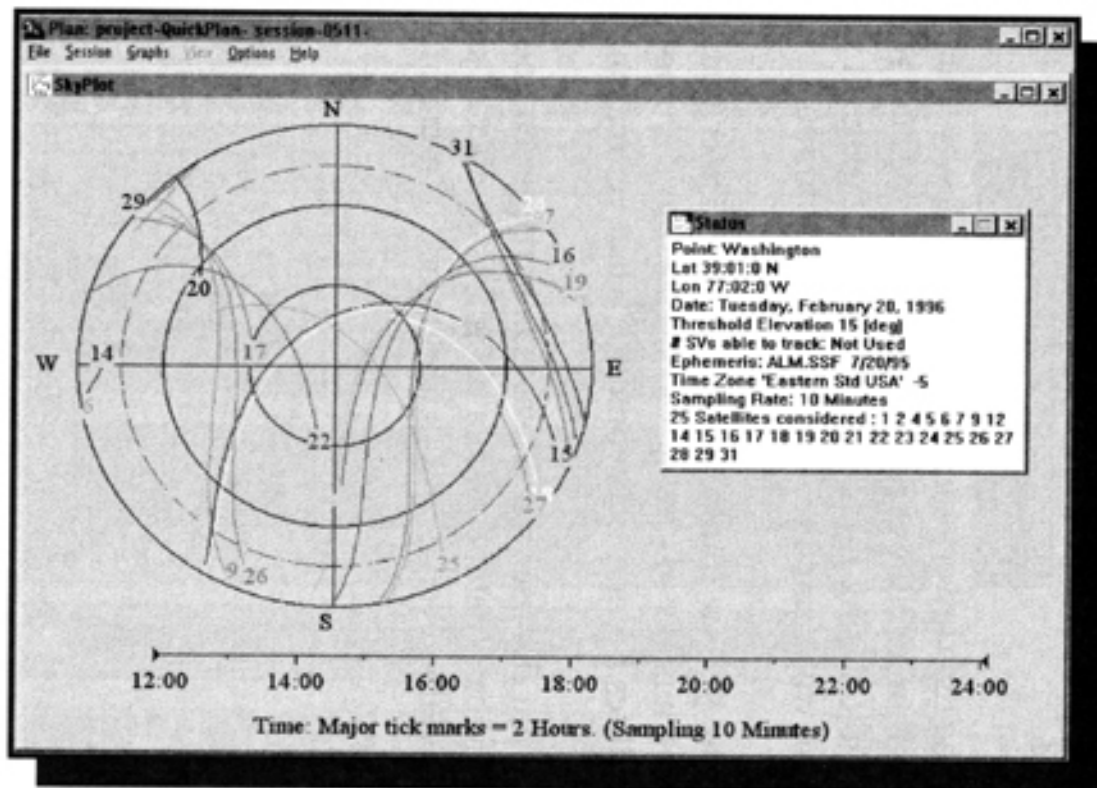
This diagram illustrates a screen-dump from another software program. Although this is a different program, all programs show essentially the same information, with variations only in how that information is presented. In this case, the screen shows how many and which actual satellites were available in St. Louis during the entire 24-hours of April 24, 1994.

The red (lower) zone indicates that fewer than the necessary minimum of four satellites are available. This does not occur in this diagram at any time during the 24-hour period, although it gets close at around 7:45 a.m. when satellite availability drops to four.

The yellow (middle) zone alerts the user that, while there are the minimum number of satellites available, there are no “backups” if one of the satellites becomes obscured for whatever reason.

What the user really wants is in the green and blue (upper) zones which indicate that there are a sufficient number of satellites with “back-ups” available even if lock is lost on one or more of them.

Mission Planning



Mission Planning

Another piece of information that mission planning software can provide are the actual plots of the satellites as they track across the sky during the selected time period.

The diagram at left, again for the Washington, D.C. area, illustrates the track of each of the visible satellites as they would move overhead. This makes it considerably easier for the field person to know exactly when the optimum times will occur during which enough satellites will be available and exactly how the satellite orientations will change through time.

Almanacs

***How Do You Know What's Where When?
And***

How Do The Receivers Know What SV's To Look For?

- Each SV Carries An Almanac Of Predicted Ephemerides
- Periodically Updated By The Control Segment
- Can Be Downloaded Direct From SV (~12.5 minutes)
- Can Be Down Loaded From Several Different BBS' s
- Usually Good For 2 - Months

Almanacs

All of this discussion about planning begs two questions: “How does the planning software *know* what satellites will be where and when?” and “How do the receivers know *what* satellites to look for when they’re turned on in the field?” Both of these questions are answered through the use of “*Almanacs*” which are libraries of satellite orbit data such as rise and set times, angles of elevation, positions in space, etc. Almanac data are periodically sent up to each of the satellites on an as-needed basis by the Control Segment and are good for 60 days.

Almanacs can be downloaded directly from the satellites by the receiver. This is done automatically whenever the unit collects data for extended periods of time. It takes approximately 12.5 minutes of continuous and unbroken lock on a satellite to complete the transmission, although in actuality, it often takes somewhat more than that because it never seems to fail that a critical satellite moves out of view during the download process, necessitating a restart. These operations are all transparent to the user. Most receivers automatically update the internal almanac whenever they collect a new one, although a few require some form of initialization to let them know that a new almanac is available.

Almanacs can also be downloaded to a PC from any number of sources such as the U.S. Coast Guard, the National Geodetic Survey, and numerous other private and public bulletin board servers. The PC project planning software will then use that information to make its predictions. The updated almanac can also be downloaded from the PC to the receiver so that, upon going into the field, the receiver can immediately know what satellites to begin looking for when it’s turned on. The opposite is also true. An almanac acquired by a receiver in the field can be downloaded to a PC for use in the mission planning software. Most programs will alert the user when a new almanac is required.

Autonomous Accuracy

	SA On	SA Off	Potential
C/A Code	<100 M	<30 M	3 M
P Code	<100 M	8 - 12 M	30 Cm
Y Code	8 - 12 M	8 - 12 M	30 Cm

Single Receiver - 95% Probable

Autonomous Accuracy

So, where does all this get us? For the most part, up until now, we've only been dealing with using a single receiver alone, or autonomously. This is frequently referred to as *Absolute Positioning*.

Using a single receiver with Selective Availability on, civilian users can only hope for positional accuracy *somewhere* under 100 meters regardless of what type of receiver is used. This is, however, widely variable depending upon the amount of dithering that is applied to the satellite signal at any given moment, with the degree of dithering constantly being varied. With Selective Availability turned off, that position figure can improve to somewhere between 10 and 30 meters (depending upon ionospheric conditions) with those receivers that have good noise filtering.

Civilian P-code receivers can achieve eight to twelve meter autonomous accuracy with Selective Availability off, although such receivers are rarely used autonomously because they tend to be very expensive. With SA on, the P-code signal is degraded to the same -100 meter level as C/A-code. Military P-code receivers, on the other hand, can "see through" the Selective Availability and still get to that published eight to twelve meter level. The term "published" is used since those are the figures commonly given in government documents. In fact, military P/Y-code receivers regularly achieve three to five meter accuracy under world conditions.

While the real-world accuracy of civilian C/A-code receivers is in the area of 10 to 30 meters (S/A off) up to as much as 100 meters (S/A on), the *potential* can be as low as 3 meters (remember, we said earlier that the phenomenal sub-centimeter carrier potential requires two receivers). The question is, how do we get to that potential?

PART III

DATA CORRECTION TECHNIQUES

AND

HIGH-RESOLUTION ACCURACY

Differential Correction

- Also Called "*Relative Positioning*"
- Requires 2 Receivers
- One Over A Known Position (i.e., NGS Monument), "*Base*"
- Other Receiver Over Unknown Point: "*Rover*"
- Base and Rover Receive Same SV Range Data
- Base Receiver Measures Vector *Difference* Between Received Position Data And Known Position
- Base Correction Value Applied To Rover Position Data

Differential Correction

Until now we've only discussed single-receiver or autonomous (absolute) operations. Using only a single C/A-code receiver we can never realistically hope to see accuracy better than around 10 to 30 meters or so. That's because of the combination of all of those User Equivalent Range Error factors. The problem is considerably worse when Selective Availability is factored in. Without any kind of reference, the receiver can't know that it is giving a positional error-it can only report what it sees, and if what it sees is in error, then its consequent position is also going to be in error.

The key here is the reference. If there were some way to tell the receiver that there was X amount of error in the signal that it was receiving, it could compensate that same amount. Well, there is a way-it's called relative or *differential* GPS positioning, or DGPS.

In differential positioning, two receivers are used. One of them is placed over a known point such as a National Geodetic Survey (NGS) survey monument. This is usually referred to as the base receiver.

The second receiver, generally referred to as the "rover," collects data from the unknown points in the field, pretty much the same as for autonomous position data collection. It is extremely important that both the rover and base receivers collect the exact *same* data from the *same* satellites at the *same* time.

The base receiver "knows" its exact location. It also receives a position from the satellites (usually with some component of error). From the difference between the known position and the GPS-derived position, a vector displacement, or *differential*, can be calculated, which can then be applied to the rover's position data.

Differential Correction

Base Station

Difference Between True
And GPS-Derived Position

GPS Position

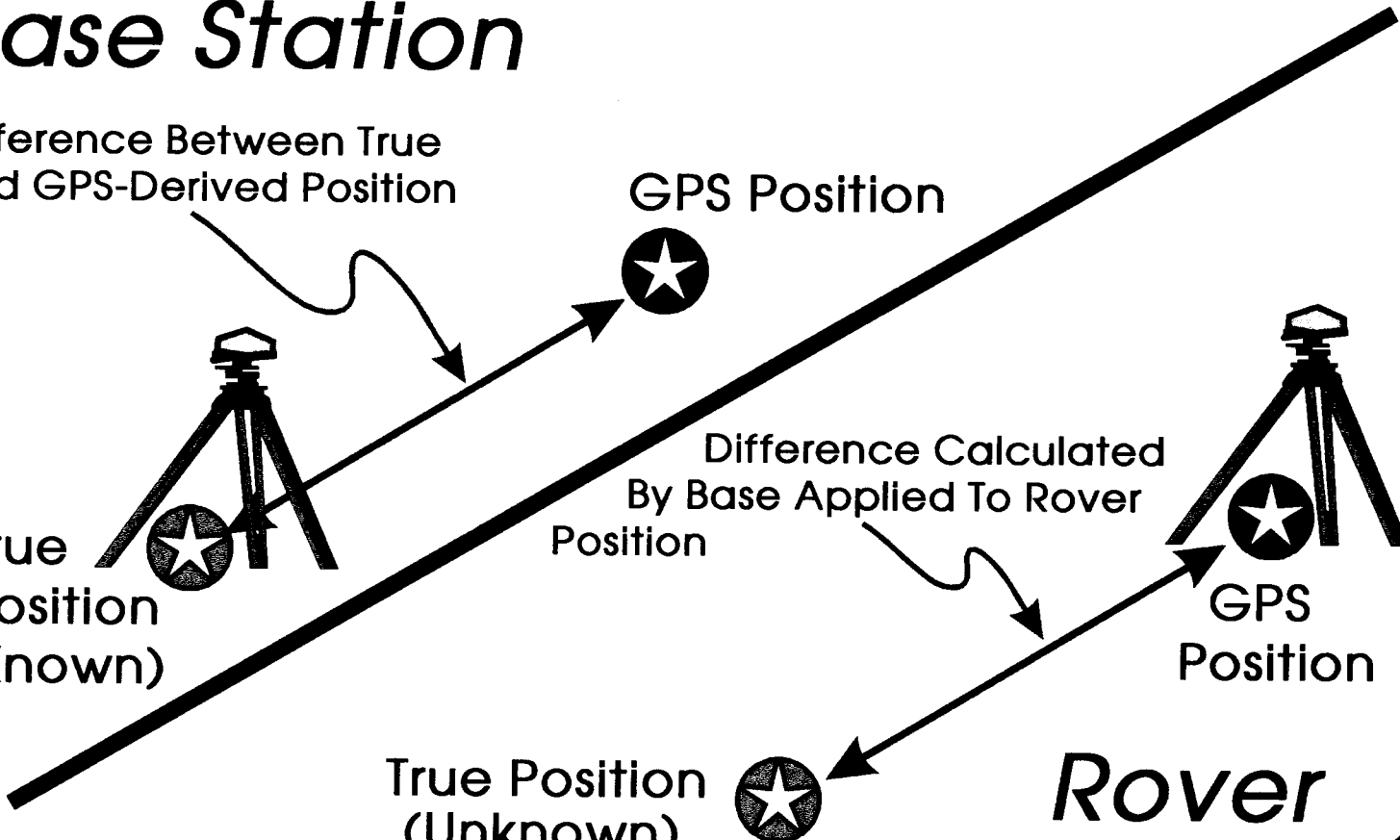
True
Position
(Known)

Difference Calculated
By Base Applied To Rover
Position

GPS
Position

True Position
(Unknown)

Rover



Differential Correction

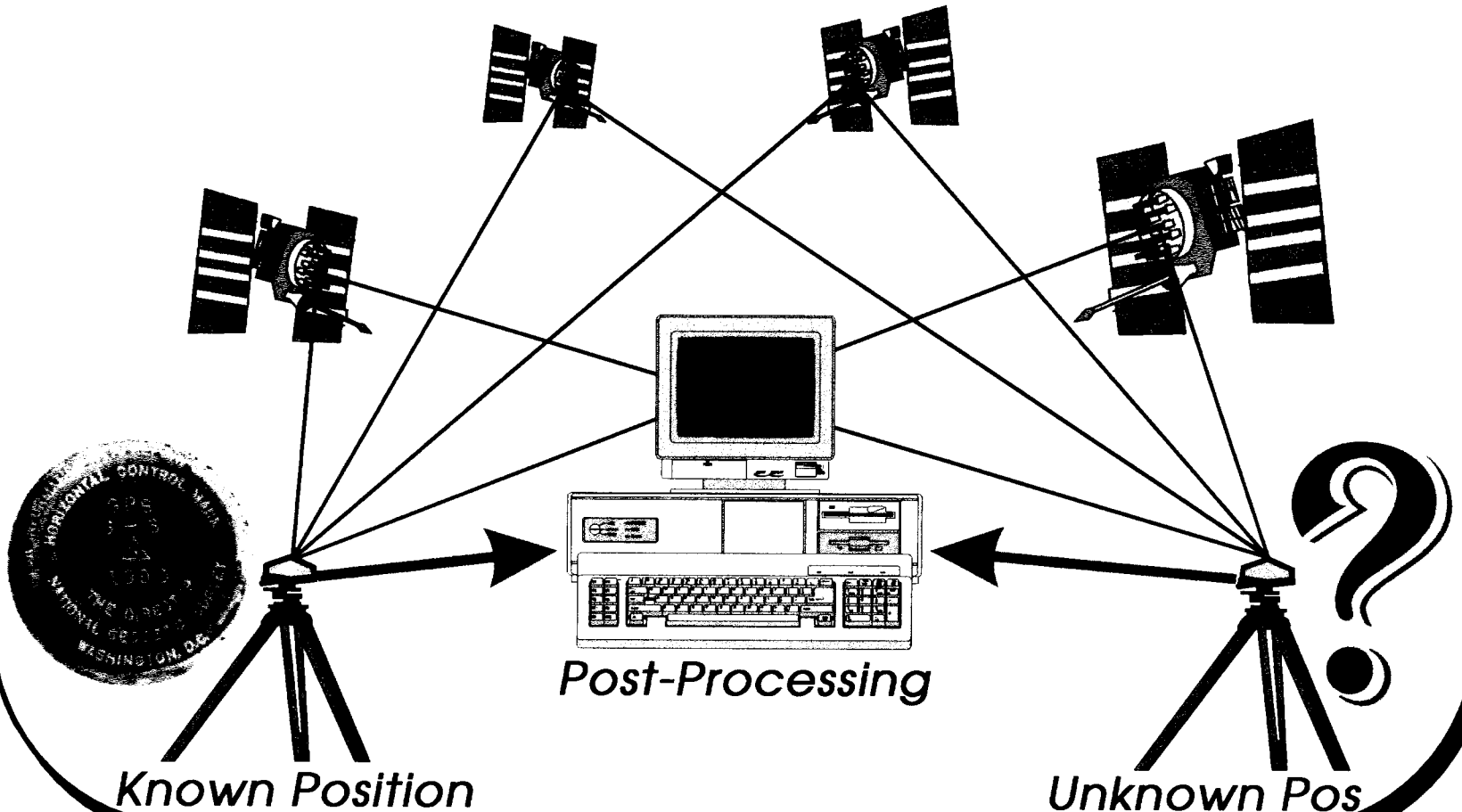
This diagram illustrates the differential concept. The upper part of the diagram shows the receiver antenna over its known point. The position that is recorded by that receiver from the satellite ranges is displaced some distance and direction away from the known point. From this, it's then a straightforward operation for software to calculate the displacement, or vector, represented in the diagram by the two-headed arrow.

At the same time, the rover in the lower part of the diagram has recorded a GPS range position of its own. Because most of the UERE components are essentially identical for both receivers, any correction differential calculated for the base unit could be applied to the rover's GPS range position, thereby canceling out those error factors. It's because of the fact that the error factors are continuously changing that it's so critical that the base and roving receivers record the same GPS data from the same satellites during the same periods of time.

Another very important consideration is base line length—that is, the distance between the base and rover. This is because if the rover is a substantial distance from the base, say 500 kilometers, it's possible that the two receivers might be observing one or more *different* satellites. It's also possible that ionospheric/tropospheric conditions might be different. At some distance apart, the magnitude of error will no longer be identical for both receivers and the accuracy of the differential correction will progressively decrease.

For code differential positioning, base line distances should be kept below 300 kilometers. For carrier-phase differential positioning, base line length is much more critical, adding between 1 and 10 parts per million (ppm) of error. That means that a 100 kilometer base line could add between ten centimeters and one meter of error. Therefore, carrier-phase base lines are best kept below 20 kilometers.

Post-Processing



Post-Processed Corrections

Post-processing is the most common method of differential correction, requiring no actual connection between the rover and base receivers during each of their respective periods of data collection. The correction data derived by the base is applied to the data collected by the rover at some later time. Again, it must be emphasized that both receivers must collect data from the same satellites at the same time. It's ok if the base unit begins collecting data before the rover does, and/or continues to collect data after the rover is finished. It is *not* ok for the reverse. In other words, the entire time-span of rover data collection must be fully matched by base data. If it does not, the rover data will not be fully correctable and must be retaken.

Typically, the base unit is turned on at the beginning of the day and continuously collects data until the rover returns. The rover is turned on only when position data are to be collected, turned off when done at that location, then turned on again to collect data at the next location, and so on throughout the day.

Once the rover is done collecting data, the base can also be stopped. The data from both receivers are downloaded into a computer where specifically designed software (usually product-specific) applies the determined correction vectors to the rover data sets. (Actually, only clock timing and ephemeris corrections are applied rather than actual position corrections. It's easier that way.)

Most programs have some form of rudimentary mapping function that allows the user to see a graphic representation of the corrected data. Some allow the user to simultaneously see both the uncorrected and corrected data.

PP Differential Data

*All Receivers Collect and Store Data
In Their Own Proprietary Format*

*Differential Post-Processing Base Data
Acquired From "Outside" Sources, Use The:*

RINEX

(Receiver Independent EXchange) Format

*GPS' s Equivalent To DOS' s ASCII
(Watch The Baseline Length)*

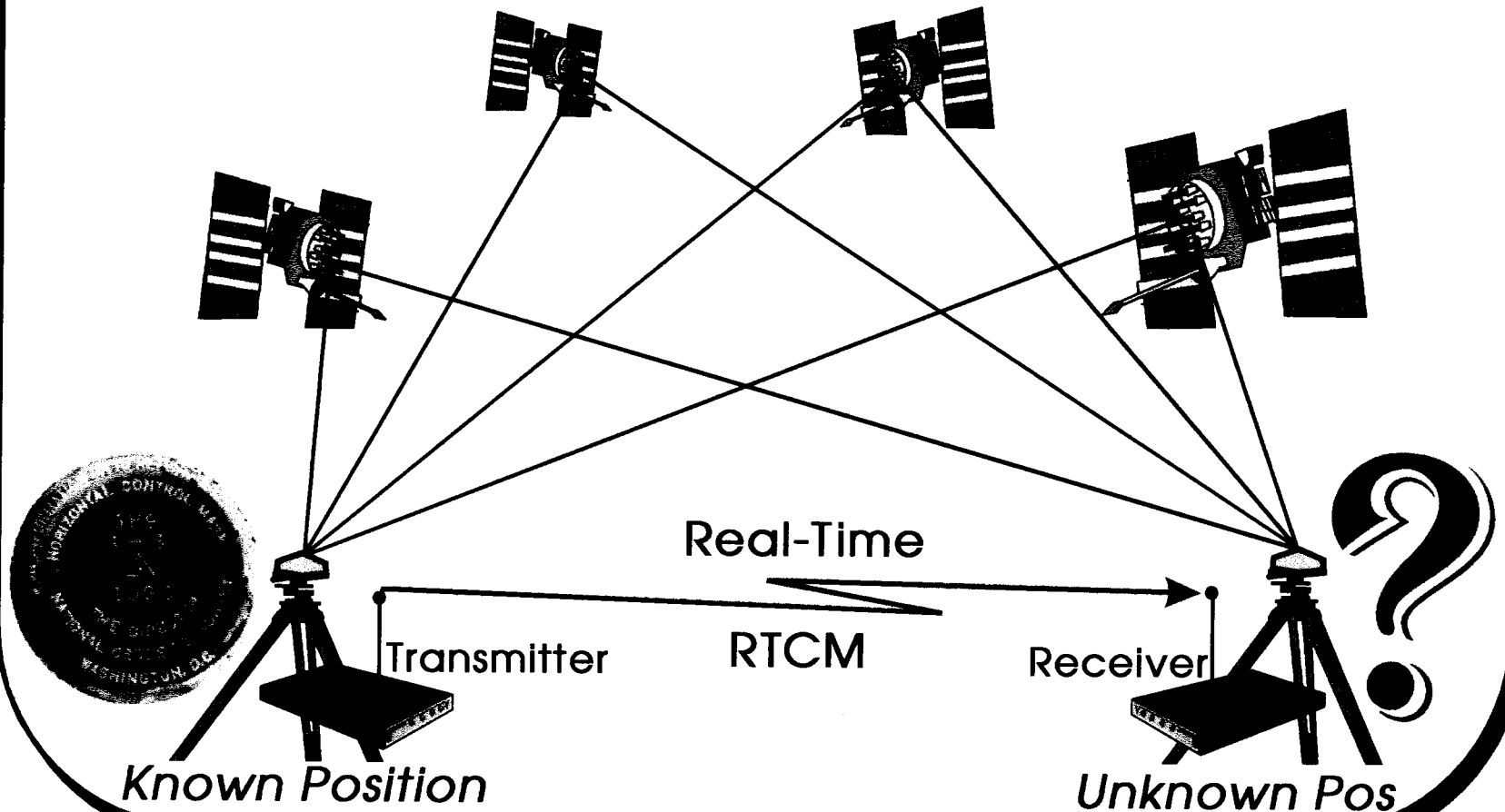
Post-Processing Differential Data

Raw pseudo-range data collected over many hours can quickly begin to use up enormous amounts of computer storage space. To get around this, the data are usually compressed. Unfortunately, all receiver companies collect and store their data in their own proprietary formats which results in data incompatibility between manufacturers.

To get around this problem, a set of standards was developed that would allow the transfer of data between different receivers. That format is designated **RINEX**, for Receiver INdependent EXchange. This format can be thought of as GPS's equivalent to DOS's ASCII. Virtually all manufacturers have modules in their processing software to allow the incorporation of RINEX data in their DGPS correction solutions, although some are easier to use than others. Ease-of-use is something one must be alert to when considering the purchase of a receiver that is expected to be used with outside DGPS correction sources.

The use of RINEX data allows a user to apply corrections to their data even when they might not have a base receiver of their own. A temptation is to acquire RINEX correction data from any reliable source (several of which we'll speak about at some length in a moment), wherever they might be. However, it's important to keep the base line length in mind since it can have a significant impact on the resultant accuracy.

Real-Time Correction



Real-Time Correction

At times it is necessary to have the position data corrected in real-time. For example, post-processed position data for an airliner coming in for a landing is of little use, regardless of the resultant accuracy. That airliner needs to know where it's at right at *that* moment.

The basic concept in real-time DGPS processing (often referred to as **RTDGPS**) is fundamentally the same as for post-processed corrections. What is different is that, rather than downloading the data from both receivers to a computer for later processing, the correction values are instantaneously calculated at the base and transmitted by some method directly to the rover. Typically, that method is a dedicated one-way radio transmission from the base position to the rover which is equipped with an appropriate dedicated radio receiver. It's also possible to carry the correction data over a cellular telephone line, although spending hours collecting data in this manner can quickly become more expensive than purchasing a dedicated radio in the first place.

Until very recently, this form of processing has had a price to pay—that was in accuracy. Typically, real-time data correction degraded the accuracy to two to three meters. Today, however, real-time C/A-code accuracy is now regularly achievable at sub-meter levels. Additionally, until recently, it was simply not possible to correct carrier-phase data in real-time at all since it required up to several hours of data collection to observe the satellite constellation shifts. That has changed dramatically. Modern geodetic receivers have now been designed to achieve real-time carrier-phase accuracy on the order of two to three centimeters with On-The-Fly Ambiguity Resolution or **OTF**. The down-side is that these receivers tend to be very expensive, typically exceeding \$50,000.00.

RT Differential Data

*Most Good Receivers Allow The Input Of
Real-Time Correction Data*

*RT Data Can Come From A Separate But
Connected Radio Receiver, Modem, Etc.
And Will Be Transmitted Using The:*

RTCM

(Radio Technical Commission, Marine) Format

Real-Time Differential Data

Real-time differential correction data can be transmitted by several means. However, it is necessary that the GPS roving receiver itself be specifically designed to accept and apply that correction signal.

Most good receivers are designed to allow the input of real-time correction data. But, as was the case for post-processing correction data, it was necessary early on to develop a common language for the real-time signal. That “universal” format was developed by the Radio Technical Commission for Maritime Services Special Committee 104 (currently Version 2.1), and is called the **RTCM SC-104** format, or simply RTCM. Although incorrect, it’s much easier to remember RTCM as simply the Real Time Correction Method.

The RTCM SC-104 is not the only correction transmission format that exists, although it is the most widespread and most universally applicable. Some others include the RTCA-WAAS (Radio Technical Commission for Aviation, Wide Area Augmentation System, which we’ll talk about in a moment), and the RTCA-LADGPS (RTCA, Local Area Differential GPS), along with a wide variety of proprietary, manufacturer-specific formats.

Post-Processing vs Real-Time

POST-PROCESSED

- + Highest Accuracy
- + Lower Cost W/Free Diff. Corr. Data
- Higher Cost W/2nd Rcvr
- 2 Persons w/2nd Rcvr
- + 1 Person w/Free Diff.
- + Less "Stuff" In Field W/Free Diff. Corr. Data
- More "Stuff" W/2nd Rcvr

REAL-TIME

- Not Carrier
- Lower Accuracy
- + Immediate Correction
- Higher Cost - 2nd Rcvr
- Extra "Stuff" In Field
- + Always One Person
- Must Maintain Lock On Two Systems

Post-Processing vs Real-Time

There are a number of advantages and disadvantages to the use of either of the two general DGPS methods and a number of factors must be considered. Post-processing gives the highest accuracy-as low as only a few millimeters. Real-time processing *generally* does not use carrier-phase measurements and therefore it only gives around one-half to one meter accuracy. Carrier-phase real-time *can* give two to three cm accuracy but at very high equipment costs. Post-processing can be lower in equipment cost when using a number of “free” DGPS correction sources (which we’ll talk about next) since only one receiver is required. On the other hand, it can be *higher* in equipment cost if using your own second receiver as a base.

When using one of the various sources of DGPS correction data, only one person is needed in the field and there is less “stuff” that has to be toted around. However, if using a second receiver as a base,. then that equipment must be moved to and from the base location and a second person must “stand watch” over it to prevent theft and/or vandalism.

RTDGPS has the obvious advantage of immediately providing a corrected position. But, with the exception of navigation applications, this is not often required. RTDGPS from commercial or government sources doesn’t require more than one person in the field. It does, however, entail the additional cost of having another radio receiver to collect the RTCM signal which also has to be toted around the field. The cost and logistics problems are compounded if using your own second receiver as a base station in that it will be necessary to field an RTCM radio *transmitter* as well. Even further complicating the problem is the necessity of acquiring an FCC license for your own base RTCM/RTDGPS transmitter.

Clearly, the decision to use one or the other method can be complicated and confusing, and requires much consideration. Which to use is totally dependent on the user’s positioning requirements and budget.

Differential Data Sources

- *Your Own Second Receiver (RT/PP)*
- *Universities (PP)*
- *Federal Government*
 - C.O.R.S. (PP)
 - Beacons (RT)
 - WAAS
- *Private Vendors*
 - GPS Manufacturers (RT/PP)
 - Satellite (RT)
 - FM Sub-Carrier (RT)

Differential Data Sources

There are a number of sources from which differential correction data can be acquired, both for post-processing and for real-time DGPS. The first is simply to use your own base receiver and locate it over a known surveyed point. This has the advantage of retaining all aspects of the positioning job to the user. The disadvantages are generally related to equipment and man-power costs. This method can be used for both post-processing and real-time (provided that real-time communication between the receiver and base is available).

Many engineering departments at universities, such as the University of Texas, collect and store their own DGPS data. In most cases, the data are downloadable from bulletin boards and are free, although finding such a source near enough to the desired location to minimize base line error can be logistically difficult.

There are a number of private vendors of differential GPS correction data who use a variety of methods to get the data to the user. Some of those methods include using their own satellites and receivers, GPS manufacturers themselves providing real-time and/or post-processing data, and FM radio sub-carriers. Each of these entails some cost, some more than others. We will be examining some of these options in a moment.

Finally, there are several Federal Government sources, all of which provide data without charge to the user. Two examples are the CORS network for post-processing, and U.S. Coast Guard DGPS Beacons for real-time correction. We will examine these next in some detail.

C.O.R.S. Network

*C*ontinuously *O*perating *R*eference *S*tation

- Consists Of A Group Of GPS Reference Stations
- Coordinated By NGS (National Geodetic Survey), NOAA (Nat. Oceanic And Atmospheric Admin.)
- Provides Survey-Quality Code Range And Carrier Phase Data For User Post-Processing
- Initially For Coastal Applications
- Ultimately, CORS Network Will Consist of 100-200 Stations Nation-Wide
- Currently 47 Stations In Operation (2/22/96)

CORS Network

The National Geodetic Survey (NGS) offers high precision, L1 and (usually) L2 code and carrier-phase differential correction data through a network of *Continuously Operating Reference Stations* located around the United States. The CORS system provides only post-processing capabilities.

Accuracy standards for the CORS data are specified at two centimeters horizontal using carrier-phase (when L1 and L2 frequencies are available and base lines are less than 20 kilometers), and one-meter horizontal with C/A-code on the L1 frequency. No encryption or user access limitations are built into the CORS specifications.

The majority of the stations are coordinated and operated by the National Oceanic and Atmospheric Administration (NOAA) and are co-located with U.S. Coast Guard beacon stations. Currently, there are 47 (as of 2/22/96) operating which includes the entire coastline of the Continental U.S., and most of Alaska. Originally, the U.S.C.G. set out to install only real-time DGPS beacons for shipping navigation. However, it was found out that for only approximately \$10,000 more per station (“peanuts” for the Federal Government), it was possible to provide post-processed data at geodetic levels of precision. Consequently, all U.S.C.G. beacon stations are to meet CORS standards.

Currently, the U.S. Army Corps of Engineers is installing CORS stations along all of the major U.S. inland waterways. Ultimately, the CORS network will consist of over 100 stations providing geodetic quality differential data for nearly the entire United States.

C.O.R.S. Network

*C*ontinuously *O*perating *R*eference *S*tation



Current Installations

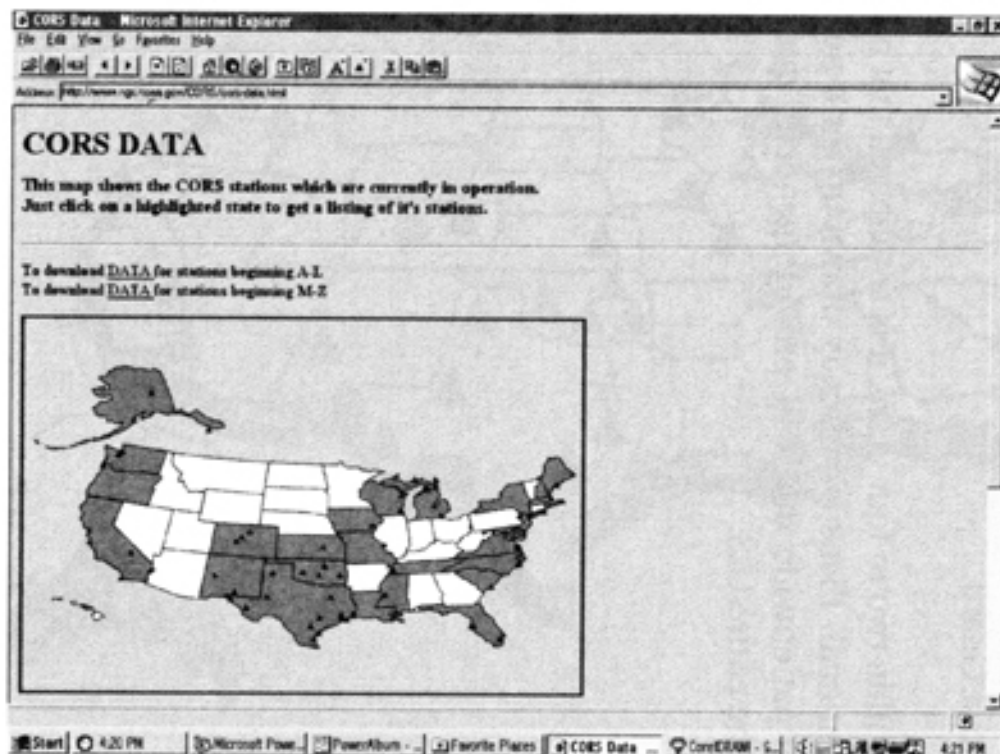
CORS Network

The map at left illustrates the locations of the 47 currently operational (as of 2/22/96) *Continuously Operational Reference Stations*. This figure is likely to have increased substantially by the time you read this due to the very rapid implementation of the network through the combined efforts of the National Geodetic Survey, the National Oceanic and Atmospheric Administration, the U.S. Army Corps of Engineers, and the U.S. Coast Guard.

Within two to four years, it is anticipated that over 100 stations will be operational. Plans are already drawn up for a network of nearly 200 stations nationwide that will provide geodetic quality DGPS data for nearly the entire U.S.

C.O.R.S. Network

Available On The The WWW @
<http://www.ngs.noaa.gov/CORS/cors-data.html>



Also via FTP: [Proton.NGS.NOAA.GOV](ftp://Proton.NGS.NOAA.GOV)

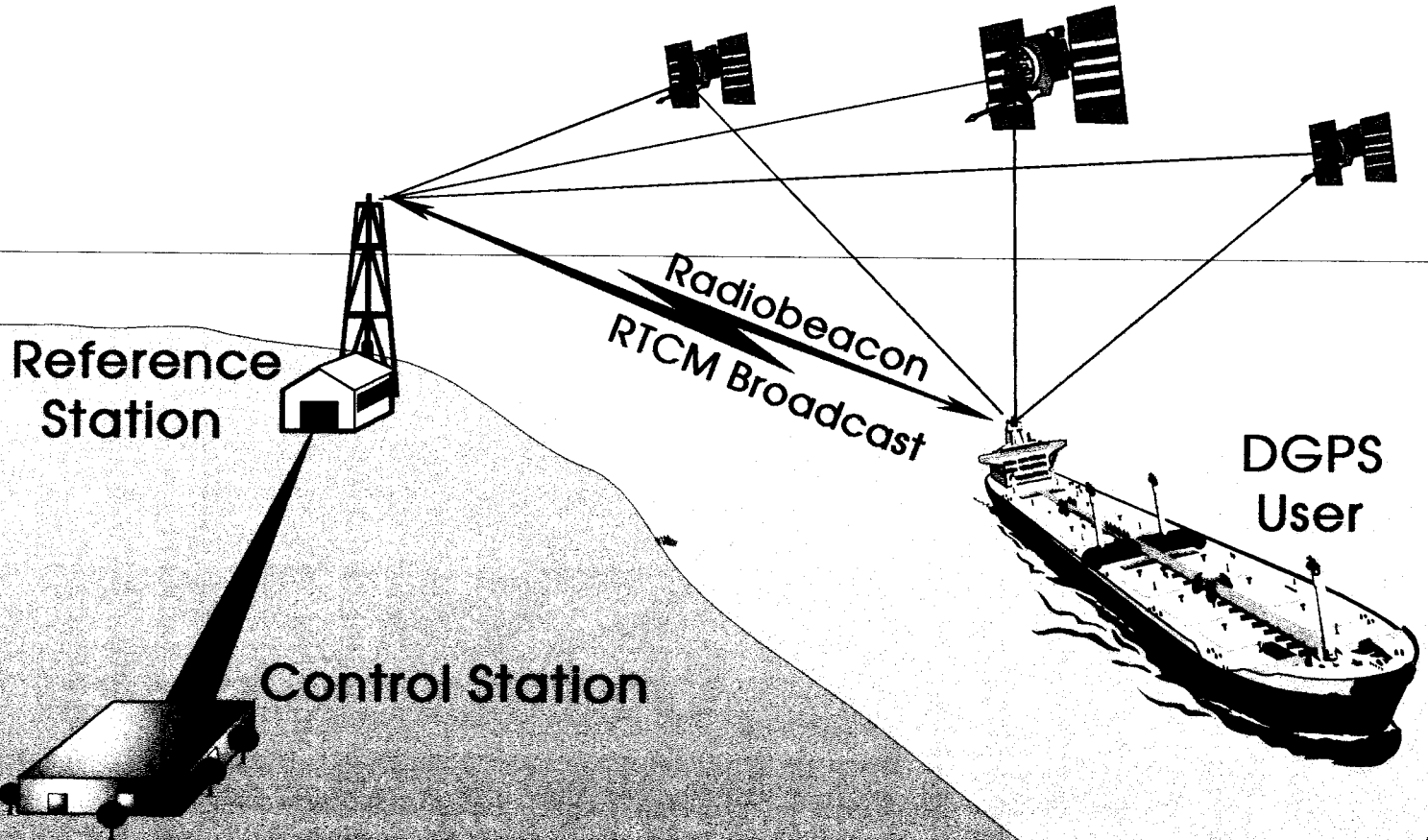
CORS Network

So where do we get all of this great free stuff? It turns out that there are exceptions to the rule “if it seems too good to be true, then it probably is.” In this case, it’s not only as good as it seems, it’s easy to get! All of the differential correction data from all over the country, are available on the Internet, free for the downloading. The illustration at left is a screen-dump of the National Geodetic Survey’s CORS Web page. The address is: “<http://www.ngs.noaa.gov/CORS/cors-data.html>”

There *are* a few steps involved. It is necessary to know the identifier of the desired CORS station. Simply clicking on the map will give that information. Clicking on the “**Information**” link below the map will take **you** to the “<ftp://proton.ngs.noaa.gov/cors/README.txt>” file which also includes the station location information along with general instructions on how to use the system. Date is in a year/day-of-year format, as in: 95272 for the 272nd day of 1995. Time of day is given in a simple code form, the “*UTC Time Block Identifier*” (important - Universal Time, *not* local time), found in the **RINEX.TXT** file. Finally, it will be necessary to acquire the fixed-base coordinates for the CORS station of interest which is found in yet another file using the “COORD” link, also under the CORS map. Once these are determined, the appropriate DGPS file is located through the “**DATA**” link and downloaded (at least two files will be needed for any one time period). The files will be compressed and will require decompression. The program for *that* can be downloaded from the “**PC-Utilities**” link. All of this is also available via “<ftp://Proton.NGS.NOAA.GOV/cors>” for non-Web Internet users.

There’s a bit more yet and so you’ll want to read all of the available files. Everything you’ll need will be there. In addition, your software may not be particularly “RINEX-friendly” and consequently require some extra “data-messaging.”

U.S.C.G. Beacons



U.S. Coast Guard Beacons

The U.S. Coast Guard transmits a real-time differential GPS correction beacon signal for free public use in all major harbor, harbor approach and critical waterways areas of the U.S. The Great Lakes, Puerto Rico, and most of Alaska and Hawaii are also covered. The purpose of the beacon system is to provide real-time navigation correction in important coastal and inland waters and is not used for post-processing. The co-located CORS stations provide that.

The RTDGPS corrections are broadcast as a modulation of standard marine radio-beacons operating in the 285-325 KHz frequency range using the RTCM-SC 104 format. The signal can be acquired with a modestly priced (< \$500) U.S.C.G. marine radio-beacon receiver. Provided that the GPS receiver is RTCM compatible, it's simply a matter of plugging the beacon receiver into the GPS's RTCM port and telling it that an RTCM signal is available (some receivers detect the RTCM signal automatically).

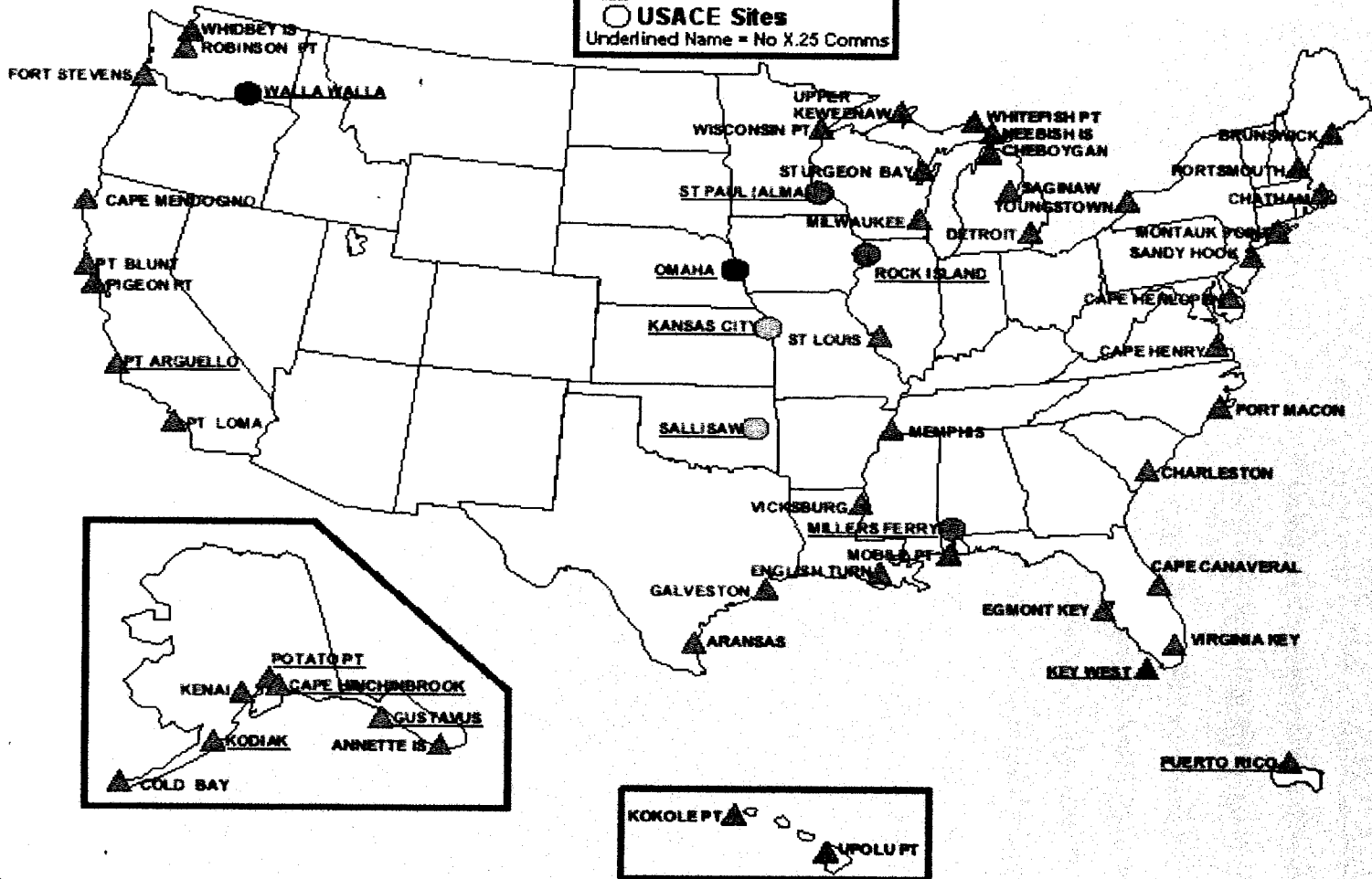
Complete coverage of all coastlines of the continental United States is currently provided to at least 92.6 kilometers (50 nautical miles) out-to-sea (the **CCZ**, or Coastal Confluence Zone) for better than 99.7% of the time. Harbor Approach Navigation requirements specify coverage out to 20 kilometers seaward of the beacon, although users frequently receive valid corrections out to 300 kilometers and more.

While harbor approach accuracy requirements specify 10 meters horizontal or less, the system is actually designed to provide a horizontal accuracy of five meters. In practice, beacons are consistently providing horizontal accuracy of three meters under real-world conditions. Vertical accuracy on the order of five meters can be expected.

55 DGPS SITES

03/08/96

- Transmitting DGPS
 - ◻ Under Construction
 - Future
 - ▲ USCG Sites
 - USACE Sites
- Underlined Name = No X.25 Comms



U.S.C.G./A.C.O.E.

Differential Radio-Beacon Coverage

The map at left illustrates the current locations of the combined U.S. Army Corps of Engineers and U.S. Coast Guard DGPS radio-beacons. As of March 8, 1996, 55 beacons are either in operation, are under construction (and will undoubtedly be completed by the time you read this), along with three planned sites for which funding has already been secured.

On January 30, 1996, the system was declared IOC, or "*Initial Operational Capability*." During the IOC period, enhancements to control station software and hardware will be accomplished and radio beacon antennas will be upgraded. Upon completion of IOC, the DGPS service will be declared FOC, or "*Full Operational Capability*." Completion of the system is expected sometime in 1996. The system is nearly fully operational as of this writing.

Discussions are currently ongoing with several Federal agencies for additional sites west of the Mississippi. If the full, planned 102-beacon system is implemented, a real-time, three-meter differential GPS correction (RTDGPS) broadcast will be available for nearly the entire United States and much of the southern tier of Canadian Provinces.

W.A.A.S.

What Is WAAS?

Wide **A**rea **A**ugmentation **S**ystem

- Planned By And For The FAA
- Uses Geo-Stat SV' s and Ground Stations
- Broadcasts Differential Corrections
- Accuracy Spec: 7.5 m (Cat III ILS: 4.1M H; 0.6M V)
- Availability Specified @ 99.999%
- Coverage: Entire U.S. To 30,500 m (100,000 ft.)
- No Encryption or Screening
- Expected Availability: End of 1997? (Phase I)

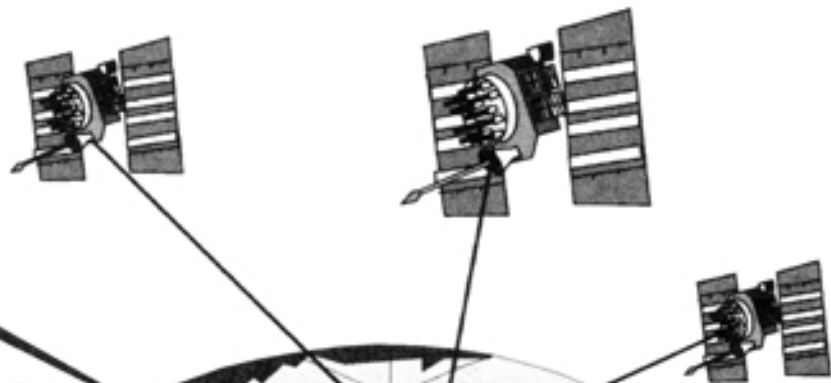
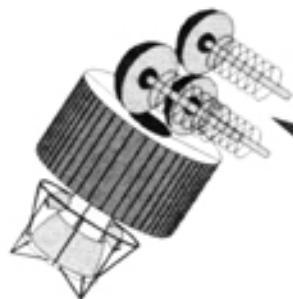
What Is W.A.A.S. ?

It was seen early on that one of the chief beneficiaries of the GPS system would be those using en-route navigation and high precision approach and landing systems such as airlines. Consequently, the Federal Aviation Administration moved quickly to begin implementation of a *Wide Area Augmentation System*, or WAAS. The WAAS is one part of a *Wide Area System*, or WAS. The other part is the LADGPS, or *Local Area Differential GPS* system. WAAS coverage is from the surface to 30,500 meters above sea level (ASL) over the U.S. out to and including, Alaska, Hawaii, Puerto Rico, and much of the Gulf of Mexico, and satisfies the accuracy requirements for all phases of flight up to Category I precision approach (which are: 17.1 meters horizontal; 4.1 meters vertical). Category II (5.2 meters horizontal; 1.7 meters vertical) and Category III (4.1 meter horizontal; 0.6 meter vertical) precision approach requirements are met by the LADGPS component of the system.

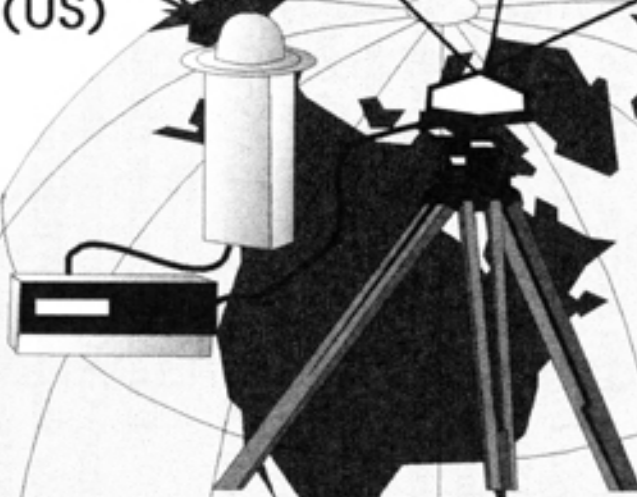
If implemented as planned, the FAA WAAS will consist of geostationary GEO communication satellites (space-based), *Wide Area Reference Stations* (WRSs; ground-based), and *Wide Area Master Stations* (WMSs). GPS data will be received at WRSs which will forward the data to WMSs. The WMS's will then determine the differential corrections and uplink them to the GEO satellites where the DGPS signals will be retransmitted back down to the user on the GPS L1 frequency. Implementation is expected sometime in 1998.

The LADGPS will serve airports using ground-based transmitters, called pseudo-satellites *or pseudolites*, that emulate GPS satellite signals. The system will transmit a precise local DGPS signal for use in high precision Category II and III landing systems. Because of the transmitter's low elevation, the line-of-sight DGPS signal will be relatively short range at ground level and so, unlike the space-based WAAS component, it will be of limited use for most ground-based applications. In addition, since they will not be RTCM-compliant, each of the components of the WAS will require their own receivers as they will transmit their signals on different frequencies than those used for the GPS itself.

Commercial Geostationary SV



- Coverage Everywhere (US)
- Only 1 Person Needed
- Reasonably Portable
- Expensive



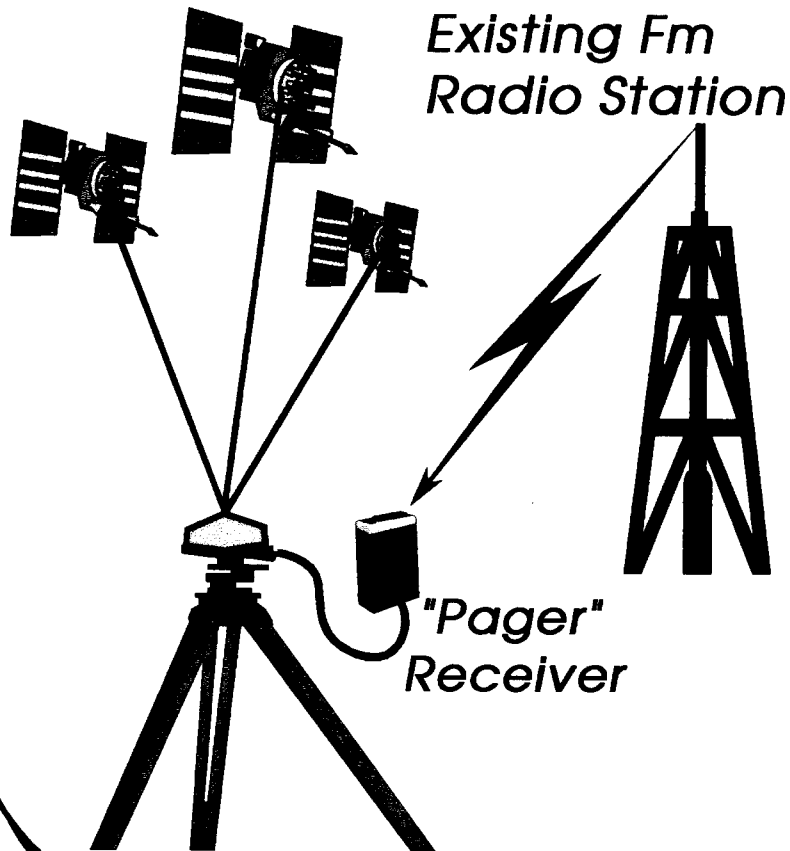
Commercial Geostationary SVs

Commercial companies currently offer several different means of real-time differential correction. The most sophisticated are the space-based systems where the private vendor has launched their own satellites which transmit correction signals back down to the subscribed user. Most use geostationary satellites orbiting at around 40,000 kilometers altitude. As a result, the same satellite is always available over large areas all the time. However, there is a disadvantage—that is if, for some reason, the signal from the differential satellite is obscured or interrupted (i.e., fails), then the user is stuck without a correction signal, unlike the GPS signal itself, which is supplied by as many as 12 satellites at any one moment, allowing the GPS receiver a “selection” to choose from.

The receivers and antennas are relatively portable and as such pose no problem when installed in vehicles such as airplanes or automobiles. They can be more difficult to handle when carried into the field for real-time surveying, etc.

The real drawback is that they are relatively expensive. A typical receiver can cost between \$4,000 and \$5,000 and annual subscriptions to the service can cost in excess of \$3,000. This is in addition to all of the other “normal” costs of GPS receivers and associated hardware and software. However, it *is* necessary at times. In very remote areas such as central Africa or South America, the nearest base station may be thousands of kilometers away, seriously degrading the correction. The differential satellite receiver acts as a “virtual base station” with a base line only the distance between the GPS receiver and the differential receiver, which is usually only a few meters at most. This can give to “autonomous” potential accuracy on the order of 30 to 50 centimeters, depending on the GPS receiver used.

Real-Time FM Sub-Carriers



- T/mits via Existing FM Sta.
- Inexpensive
- Pager-Size
- Sub-Meter Real-Time
- Limited Coverage
- Short Broadcast Range

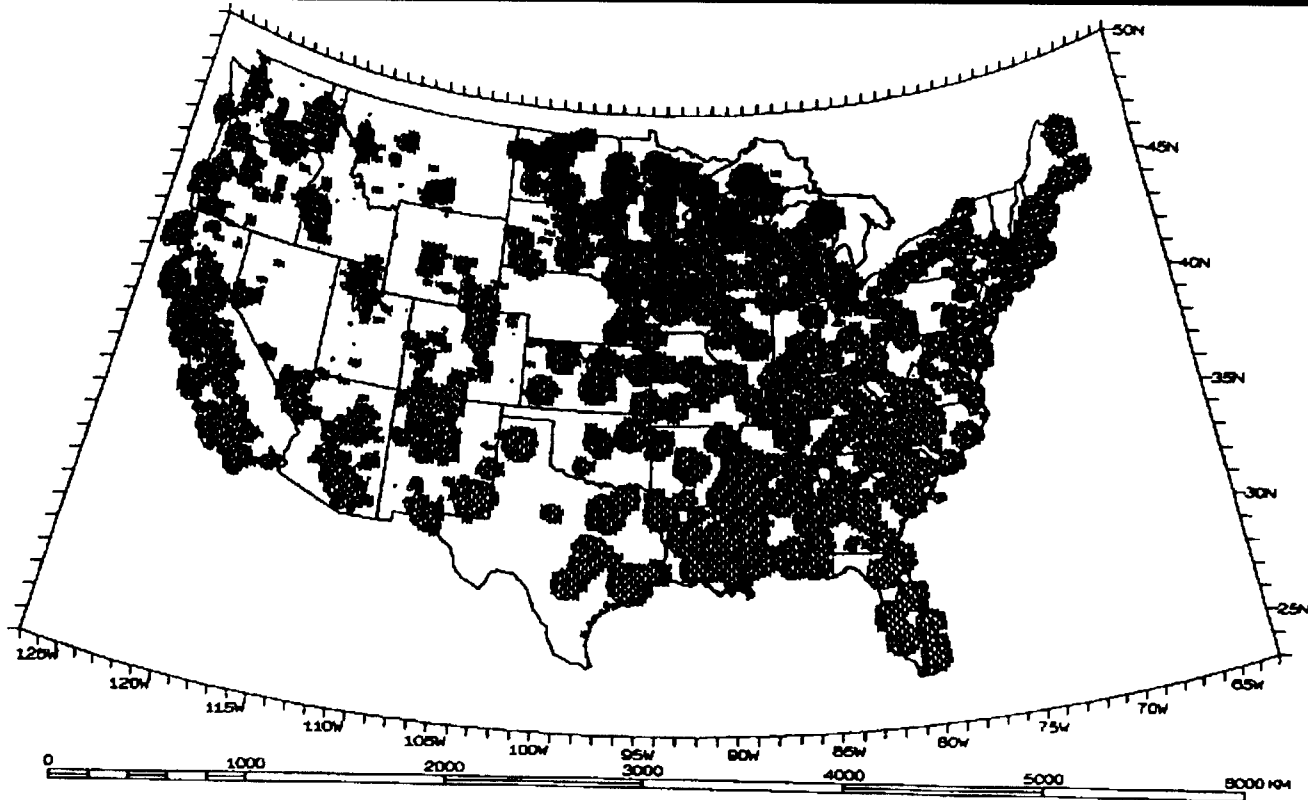
Real-Time FM Sub-Carriers

Real-time differential corrections can also be sent via FM radio signals. By using existing FM radio stations as a communications link, an RTCM differential correction signal can be provided to users within range of an RTCM-capable transmitter.

These systems use the Radio Data Systems (RDS), or Radio Broadcast Data Systems (RBDS-US) standard for the transmission of digital data using subcarrier modulation on existing broadcast FM radio stations. GPS correction data are multiplexed with other user services which are also placed on the FM subcarrier such as traffic alert and civil defense systems.

A drawback is that the range is often rather short, on the order of 50 to 75 kilometers, leaving many signal gaps in the system. The two major suppliers of this type of system, *Accpoint* and *Differential Corrections, Inc.(DCI)*, are rapidly expanding their coverage. Between the two, there are already over 1,000 FM radio stations that currently transmit a DGPS signal in the United States and Canada. The companies claim to presently cover over 90% of the population (*not* the surface area) north of Mexico. Both express confidence that they will achieve 80-90% surface coverage within the next two to three years. If you've ever tried to tune in an FM radio station out in the middle of the Sonora Desert, you may greet this optimism with a grain of salt. However, the 1994 U.S. Department of Commerce/Department of Transportation document "*A Technical Report to the Secretary of Transportation on a National Approach to Augmented GPS Services,*" predicts that full system implementation will, in fact, provide FM broadcast coverage to over 96% of the population and 80% of the land area of North America.

Predicted Coverage For FM Subcarrier Sys



*Per 1994 Technical Report to the Secretary Of Transportation
on a National Approach to Augmented GPS Services*

Predicted Coverage For FM DGPS

The map at left illustrates the report's predicted coverage of FM subcarrier systems when they're completed. Note the extensive areas that are not, nor are expected to be, covered, which means that there is a very real potential that there may not be coverage where you might need it. If you're not where "96% of the population" happens to be, you will likely be out of luck.

Another potential disadvantage to these systems is that, unlike the space-based DGPS correction services, base line errors can add an unknown but possibly significant amount of error. Base lines can include the distance from the receiver to the FM transmitter, *and* the distance between the FM transmitter and the actual DGPS source which is frequently many kilometers away from the transmitter.

There are, however, several significant advantages to these systems. The first is their low cost. At only around \$400 for the receiver, and subscription fees on the order of \$600 per year for sub-meter accuracy, the systems are quite affordable. In addition, receivers are pager-sized and weigh only a few ounces making fielding them a non-issue. Accuracy for these systems continues to improve, and as of early 1996, some companies have begun offering premium services at the decimeter level for selected higher-end GPS receivers.

Even with their limitations, for those real-time differential correction requirements that *do* fall within the geographic areas of coverage, this DGPS method is worth careful consideration. Cost and logistics are all low, and the companies *are* expanding coverage and increasing accuracy at a very rapid pace.

Other Improvement Techniques

C/A Code: Carrier Smoothing

- A Technique For Improving C/A Code Positioning
- Uses Carrier Phase To Smooth The C/A Position
- A Kind Of Backwards Running Mean

Carrier Phase: Dual Frequency Receivers

- Calculates Ionospheric/Tropospheric Refraction
- Receives Both L1 and L2 Frequencies
- Each Freq. Refracts Different Amounts

Pseudo-Satellites ("Pseudolites")

- Ground-Based GPS Transmitters
- Improves DOP With 3-D Geometry
- Mostly Used For Aviation

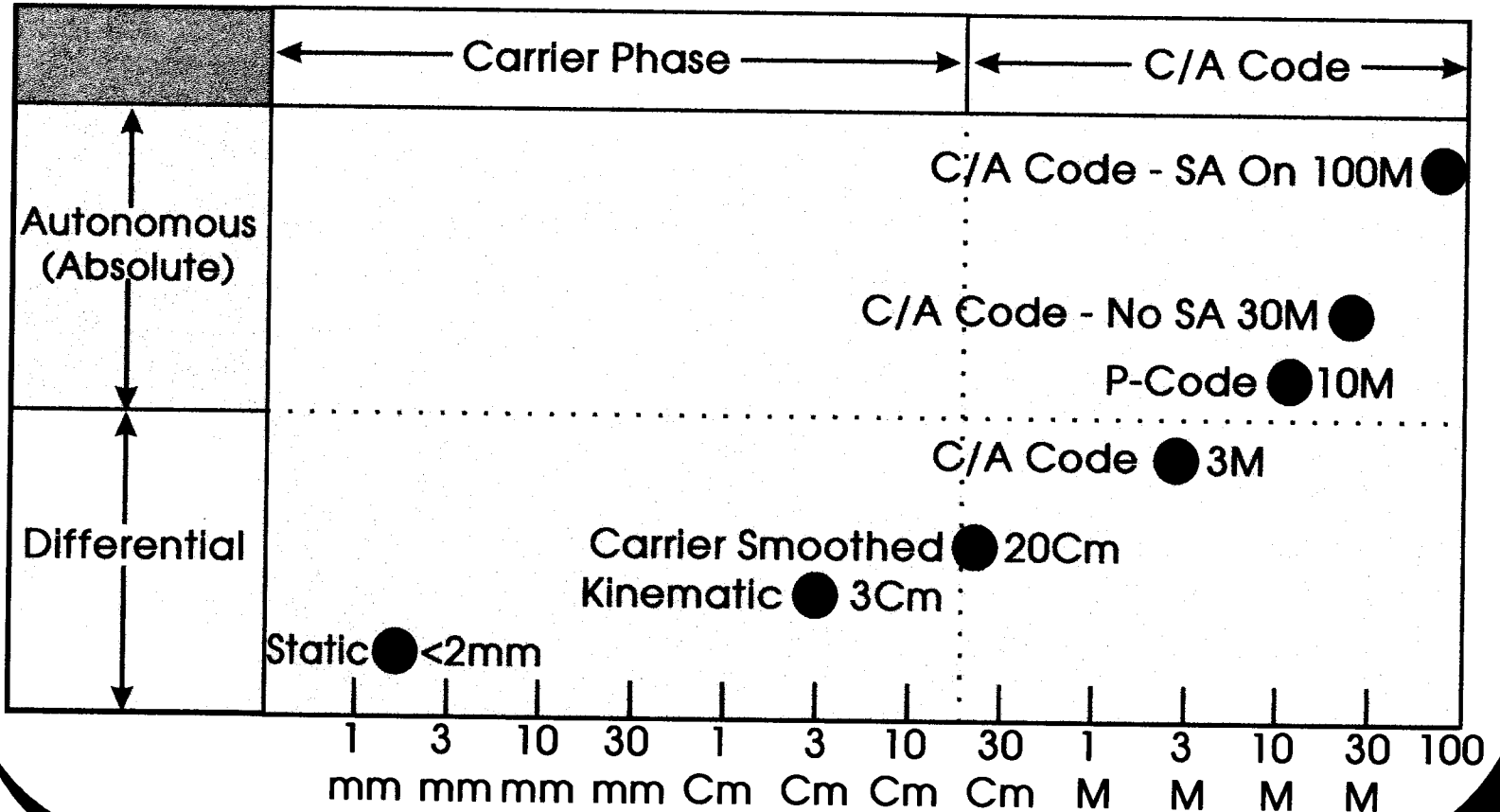
Other Improvement Techniques

There are an ever-growing number of methods to improve the accuracy of GPS for both C/A-code and carrier positioning. *Carrier Smoothing*, for example, is a technique that uses a limited measurement of the carrier wave to “smooth” out the amplitude of variation in C/A position ambiguity. As a satellite moves across the sky, the receiver-satellite range is continually changing. Because of this, the carrier wave is always Doppler-shifted to some degree. A code-smoothing receiver can’t measure the carrier pseudo-range, but it *can* measure the Doppler shift from one moment to the next. From this it can calculate the *difference* in range from time 1 to time 2 (or: $\Delta\Phi [T_1, T_2]$). By subtracting ($\Delta\Phi [T_1, T_2]$) from T_2 , the position of the first T_1 position can be calculated, in a sense mapping T_1 to T_2 . The process is then repeated for T_2 to and so on in a continuous “smoothing” action. The effect is cumulative, improving with the number of positions collected. Typically, this requires 30 to 90 seconds of collection to reach maximum smoothing, so a user should wait at least 30 seconds before recording a position after turning on a carrier-smoothing receiver. This technique is now used on all but the most inexpensive receivers.

Virtually all ionospheric refraction error can be eliminated by using *dual-frequency* carrier-phase receivers. Because the amount of ionospheric refraction is inversely proportional to frequency, measuring the different degrees of refraction simultaneously experienced by the L1 and L2 frequencies can be used to determine the actual amount of error induced. Such receivers are, however, quite expensive.

As discussed earlier, *Pseudo-Satellites*, or *Pseudolites*, are simply ground-based GPS transmitters. Such units can significantly improve DOPs by locating a transmitter below the receiver. Since the receiver must be above the transmitter, it’s really only applicable for aircraft.

Accuracy



Accuracy

So, where does all of this get us? There is a wide and growing range of equipment available to the end user, depending on the positioning requirements and how much the user is willing to pay as there is an almost direct relationship between accuracy and cost.

At the high end of differential carrier-phase positioning, users can achieve accuracy well under one-half centimeter in static, or stationary mode, and two to three centimeters in while in motion. But this comes at high cost, usually exceeding \$50,000 per receiver for the so-called *Real-Time Kinematic* (RTK) units.

Low-cost consumer receivers, often costing less than \$200, can't be expected to achieve accuracy much under 100 meters with Selective Availability turned on, or around 30 meters with SA turned off. However, for most recreational users such as boaters and hikers, this is more than adequate, and is orders of magnitude better than anything that was available before. With their small size (some as small as a cigarette pack), and easy affordability, sales of these units are exploding.

In between these two extremes are the carrier-smoothed C/A-code receivers that can regularly achieve sub-meter differentially corrected accuracy. Some of the more basic carrier-phase receivers, capable of 20 to 30 centimeter accuracy, can now be acquired for under \$8,000, complete with the required proprietary software.

These only represent the current state-of-technology and state-of-cost. Like all electronics, prices tend to fall rapidly when production goes up. As we'll see shortly, demand, and consequent production, is going up faster than anyone had imagined. We can fully expect prices and sizes to continue to drop, as accuracy and ease of use increases.

Error Terms

BE CAREFUL!

DRMS: Usually Given As 1DRMS, 2DRMS, Same As...

Sigma: Usually Given As 1 Sigma (1σ), 2 Sigma (2σ)

Both Are Simply The:

Standard Deviation: 1 SD = 68%, 2 SD = 95%

CEP: Circular Error Probable = Any Given Position Has 50% Prob. of Being Within The Distance Given Horizontally

SEP: Spherical Error Probable = Any Given Position Has 50% Prob. of Being Within The Distance Given Spherically

Error Terms

Up until now we've been discussing various levels of accuracy without thinking too much about what those terms actually mean. There are, in fact, a number of different ways that GPS accuracy can be represented and the user should be cautioned about what each means. Basically, when we speak of "error" or "accuracy," we're referring to how far from the true position we might be, rather than "precision," or the size of increment to which a measurement can be made. This is a statistical term and is subject to all of the varied "interpretations" that any statistic is ("Statistics don't lie, statisticians do...").

The basic error unit used in GPS is **DRMS**, or *Distance, Root Mean Square*. The RMS value represents the distance within which 68% of a group of positions will fall. It is, perhaps, better thought of as a 68% probability that any single position will fall within that RMS value. By inference then, there's a 32% chance that a position will **be farther** than the RMS value. Two RMSs, or **2DRMS**, means that there's a 95% probability of a single point falling within the RMS error distance. This is simply the standard deviation of a normal distribution. As such, it is sometimes given as **Sigma**, usually in the form of 1 **Sigma** (1 σ), or 2 **Sigma** (2 σ).

There are other ways of presenting error values to which the user should be alert. **Circular Error Probable**, or **CEP**, represents the 50% probability of a single point falling within that given figure of truth in the 2D horizontal plane. Again, by inference, there is a 50% chance that the point will be *farther* away than that. **SEP**, or **Spherical Error Probable**, is similar except that there is a 50% probability of being within the given distance spherically, or in 3D.

The most widely accepted (and honest) term is **2DRMS**, or the 95% probability. Sales brochures, though, frequently quote CEP or SEP values to present their products in the best light. While not dishonest, it puts the onus on the user to truly understand what the figures mean. *Caveat Emptor!*

PART IV

BASIC GEODESY

DATA COLLECTION

TECHNIQUES

AND

GPS APPLICATIONS

Geodetic Coordinate Systems

A Set Of Rules For Specifying How Coordinates Are To Be Assigned To Positions On The Surface Of The Earth.
Defined BY X,Y,Z On An Ellipsoid

What We
Like To Think
The Earth Is
Shaped Like

Sphere

Equipotential
Gravimetric
Surface
(Sea-Level)

Geoid

Best Fitting Math-
ematical Shape

Ellipsoid

Geodetic Coordinate Systems

We speak of accuracy and error as being relative to “truth,” which, in turn, begs the question “what is truth?” (certainly an eternal question). In other words, a position measured on the Earth is accurate relative to what? For GPS, that “what” is a *Geodetic Coordinate System*. A geodetic coordinate system can be defined as a set of rules for specifying how coordinate values are to be assigned to positions on the surface of the Earth, defined in the X, Y, and Z axes.

We may like to think of our Earth as a nice, round sphere, sort of like a billiard ball. The fact is, its general shape is more like a flattened egg, with a “bulge” toward the bottom. To make matters even worse, the actual surface is more like the surface of a potato, with a virtually infinite number of bumps and valleys. The varying make-up of the body of the Earth itself leads to a wide range of gravity potential across its surface, with areas of higher density materials having higher gravity, and areas of lower density materials having lower gravity.

We’re used to thinking in terms of height as so much distance above sea level. There are, however, problems with that reference line when trying to determine precise positions. If we were to imagine that the surface of the sea (sea level) were able to completely cover the Earth, extending right through mountains and filling valleys, it turns out that we would still end up with a very lumpy surface similar in concept to the middle diagram at left. This equipotential (equal-gravity), or Gravimetric surface is called a **Geoid**.

Because this geoidal surface is so infinitely complex, it can’t be directly modeled. Instead, a best-fit mathematical shape called an **Ellipsoid** is assigned to sort of “average” the surface out.

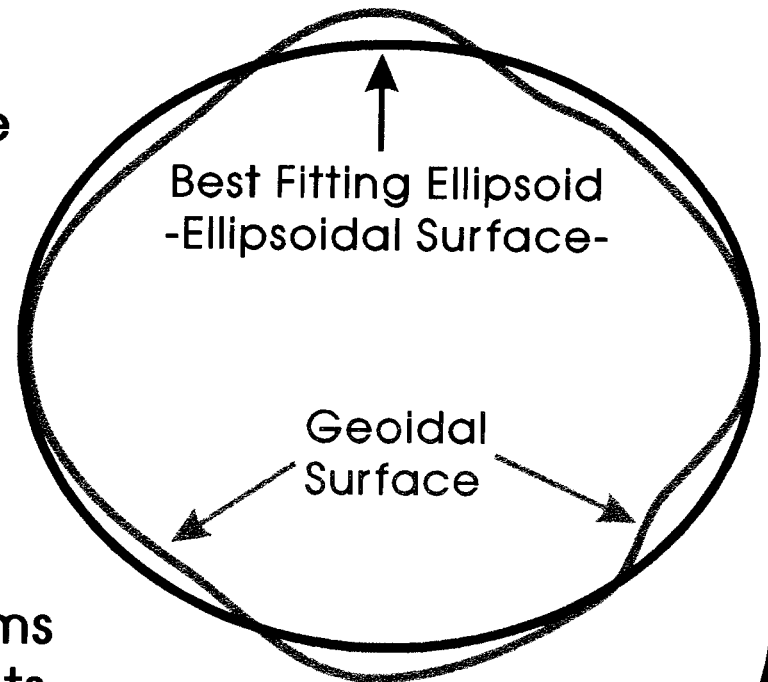
Ellipsoid vs Geoid

Ellipsoid

- Mathematical Definition
- Simple Geometrical Surface
- Described By 2 Parameters
- Cannot Be Sensed By Instru.

Geoid

- Physical Definition
- Complicated Surface
- Described By Infinite # Params
- Can Be Sensed By Instruments



Ellipsoid vs. Geoid

An **ellipse** is a two-dimensional curve where the sum of the distances from any point of the curve to two fixed points, called *foci*, is constant. Rotating an ellipse around its minor axis results in an ellipsoid of revolution, or, simply, **an ellipsoid**. An ellipsoid is a simple geometric surface that can be defined by a mathematical equation. Because of its simplicity, an ellipsoid can be defined by only two parameters, the semi-major axis (*a*) and the eccentricity (*e*). The eccentricity is, in turn, defined by:

$$e = \sqrt{\frac{a^2 - b^2}{a^2}} \quad \text{where (b) = the semi-minor axis.}$$

Ellipsoids are used as a reliable and repeatable base line from which position measurements can be made, but, since it is simply a mathematical construct and doesn't really exist, it can't be detected by instruments. The **Geoid**, on the other hand, is a physically measurable surface. And we measure it all the time, as when we speak of our height as being X number of meters above sea level, for example.

It turns out, though, that the geoidal surface undulates in a virtually infinite number of ways across the surface of the Earth. Because the geoid, or equipotential surface, is so complex, it is nearly impossible to globally model or to make reliable predictions for it. Instead, an ellipsoidal shape that comes closest to fitting the geoid is used. Heights are then measured as X distance above the ellipsoid, usually given as **HAE**, or Height Above the Ellipsoid. It happens here and there that the height is actually *below* the ellipsoidal surface. In those cases the HAE is simply given as a negative number.

Geodetic Coordinate Systems

THE LATEST AND GREATEST BEST FIT GEOID IS THE
The World Geodetic System of 1984



**WGS84 Ellipsoid Acts As An Earth-Centered 3D Datum
Essentially Identical To The North American Datum of
1983 (NAD83)**

WGS 84

The latest and greatest best-fit geoid is the World Geodetic System of 1984, or WGS 84. The origin of the WGS 84 reference frame is the Earth's center of mass. The WGS 84 ellipsoid serves as datum, or reference surface, that is essentially identical to the North American Datum of 1983 (NAD83).

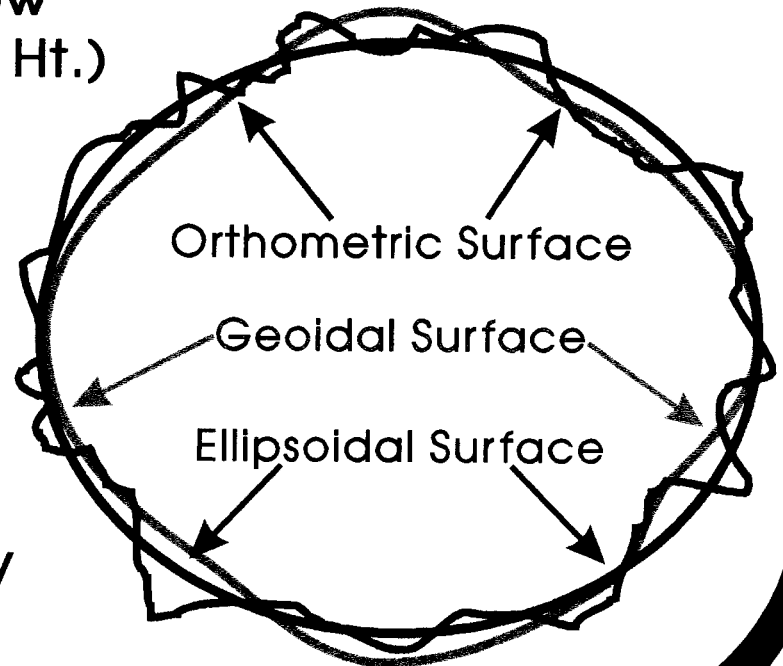
The Defense Mapping Agency (DMA) has been actively involved since 1960 in the development of World Geodetic Systems. To date, four such systems, WGS 60, WGS 66, WGS 72, and WGS 84, each successively more accurate, have been developed. The latest, WGS 84 represents DMA's state-of-the-art modeling of the Earth from a geometric, geodetic, and gravitational standpoint using the data, techniques, and technology available through early 1984. We can fully expect to see another WGS in the future as our understanding of the shape of the Earth continues to improve, no doubt in part to the amazing precision and accuracy of GPS.

What's So Special About GPS Heights?

- GPS-Heights Are Defined With Respect To The Ellipsoid
- NOT The Same As The Surface (Orthometric Height)
- We're Concerned With How High We Are ASL (Ortho. Ht.)

PROBLEM...

- No SV's Below
- One-Sided Geometry 1/2 Optimal Configuration
- Vertical Accuracy Typically 1/2 That Of Horizontal



What's So Special About GPS Heights?

Heights, as measured by GPS, are of particular concern as they are not what we ordinarily think of when we think of elevation. GPS heights are defined with respect to the ellipsoid and not to the geoidal surface, or MSL (Mean Sea Level), which is what we use in ordinary day-to-day life.

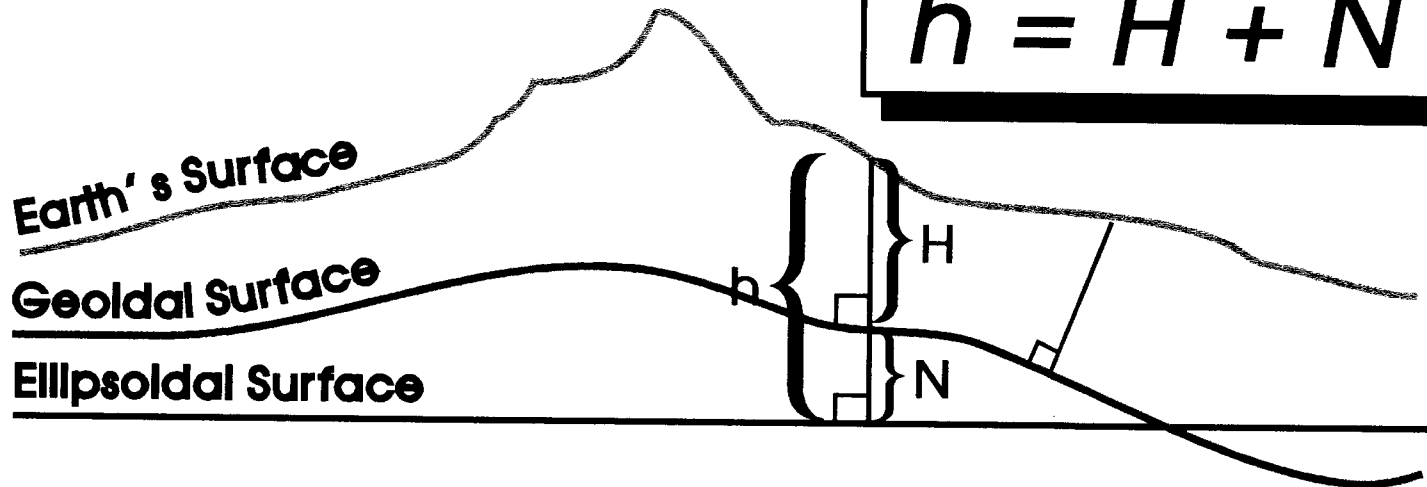
The height of the surface of the infinitely lumpy actual surface of the Earth above (or occasionally below) sea level is referred to as the **Orthometric Height**, or the height orthogonal (“square-to”) to the geoidal surface, regardless of which way that surface happens to be “tilting.” As we’ll see in a moment, the two ways of measuring height are not always easily converted from one to another.

It is important to note that GPS heights tend not to be as accurate as the GPS horizontal positions. This is principally because the satellite geometry is essentially one-half the optimal configuration since the Earth itself blocks visibility to any satellites below (with the exception of pseudolites for aviation). Consequently, as a rule of thumb, potential vertical error for any GPS position will be around two times greater than the horizontal error for that same position.

Geodetic Heights

- Orthometric Height (H) Referenced To The Geoid
- Ellipsoidal Height (h) Referenced To The Ellipsoid
- (h) Is What GPS Measures

$$h = H + N$$



The Height Difference Between The Geoid And The Ellipsoid (Geoid-Ellipsoid Separation) Called Geoidal Undulation (N)

Geodetic Heights

The diagram at left illustrates the three surfaces that have to be considered when determining heights. The geoidal surface is defined as X meters above or below the WGS 84 ellipsoid (+N or -N, respectively). That "N" value is known as the *geoidal height* or *geoidal undulation*.

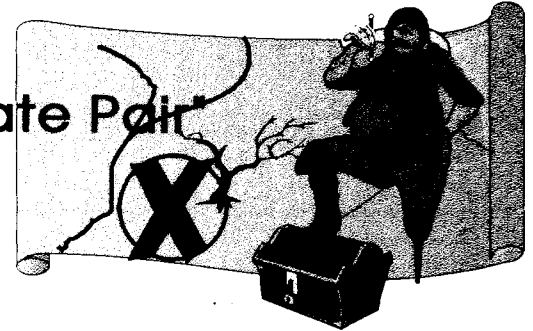
The Orthometric height (H) is the height of the surface of the Earth above the geoidal surface and is what we most commonly refer to as height above sea level. That height is always measured at a right angle to the geoidal surface (orthometrically). Therefore, if the undulating surface of the geoid happens to be "tilted" at the location where the height is to be measured, the height will still be measured perpendicular to the "tilted" surface, making the measurement also "tilted."

In general, the ellipsoidal height (h) is equal to the sum of the orthogonal height and the geoidal undulation or: $h = H + N$. If this were strictly true, there would be no problem in making exact direct conversions from one to the other (assuming that the necessary values were known). The difficulty is that because the orthogonal height (H) remains perpendicular to the geoid, it can happen that the orthogonal height is actually slightly greater than the sum of $H + N$. The variation is usually quite small and is rarely a problem. However, for high-precision geodetic leveling, the problem can influence the truth of the determined position by an amount that is, indeed, significant at that demanding level of work. In any event, accurate conversion of ellipsoidal height requires precise knowledge of the geoidal undulation.

Data Collection Techniques

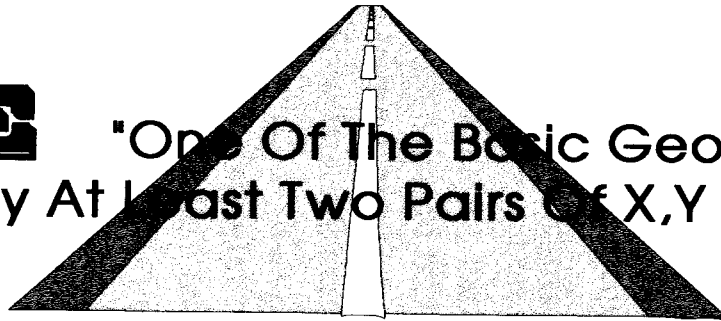
POINT

"A Single X,Y Coordinate Pair"



LINE

"One Of The Basic Geographical Elements, Defined By At Least Two Pairs Of X,Y Coordinates"



AREA

"A Series of Addressable Data Elements In The Form Of A Grid Or Matrix"



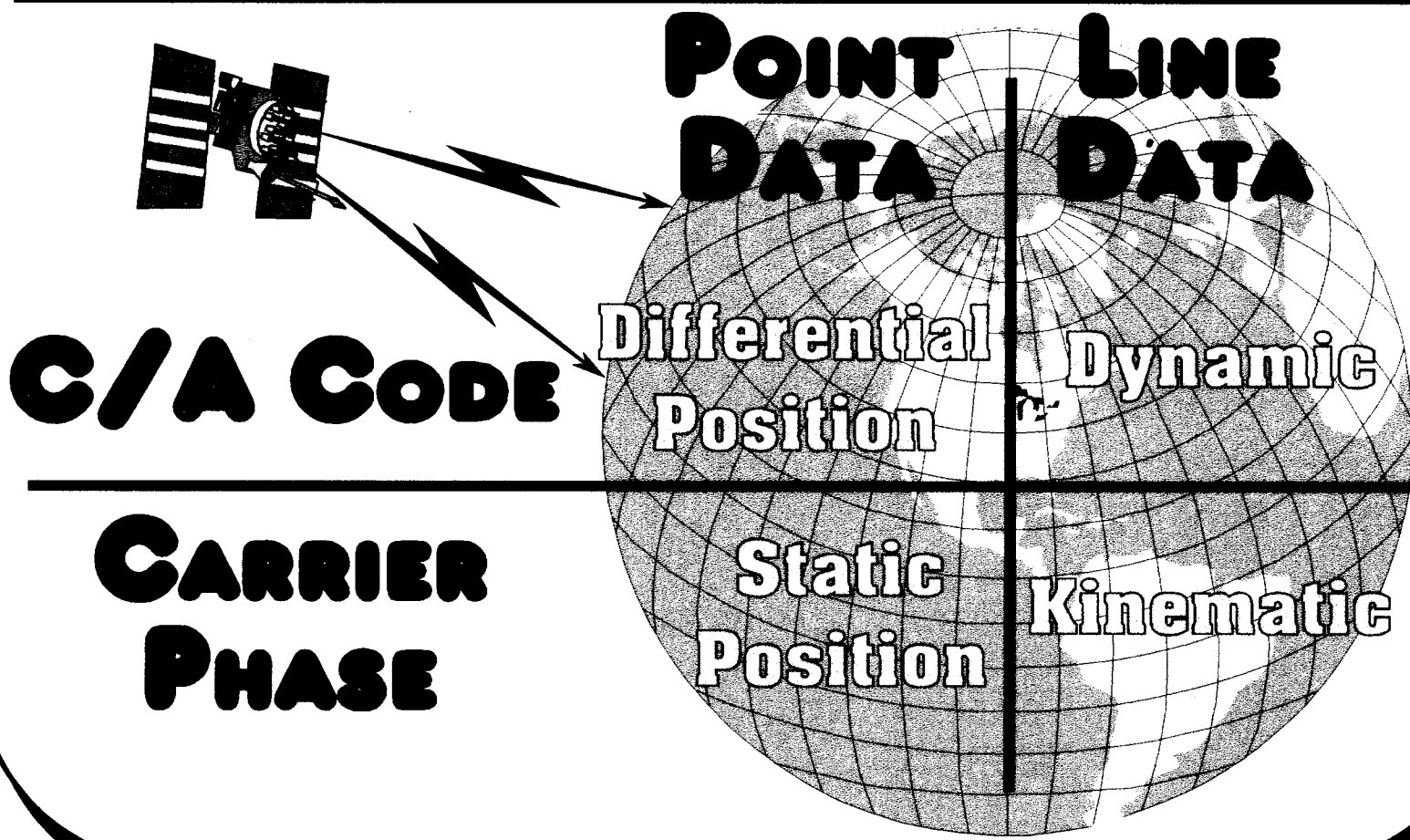
Data Collection Techniques

All spatial data can be represented by either a point, line, or area. (Actually, volume could be a fourth, but that's not necessary for this discussion.) Each builds on the one before. A point is defined by P.A. Burrough in his *Principles of Geographical Information Systems for Land Resources Assessment* (1988), as a single X, Y coordinate pair. This defines a point in two dimensions. Since GPS operates in three dimensions, a better definition would be a single X, Y, Z coordinate set. A line is simply a feature defined by at least two such points.

But-rough defines an area as a series of addressable data elements in the form of a grid or matrix. While his definition of an area is strictly correct, it is, at the least, cumbersome. Perhaps a more simple and direct definition would be a closed feature bounded by at least three points.

GPS is well suited to each of these three data elements. Points (or, as we'll see, *positions* is a more accurate term) are its forte'. Connect up a couple of GPS points and you've got a line. Add a third point (or more), sequentially connect and close them, and you've got an area. This is what GPS does. There are, however, a number of techniques that are employed to achieve these ends.

Data Collection Techniques



Data Collection Techniques

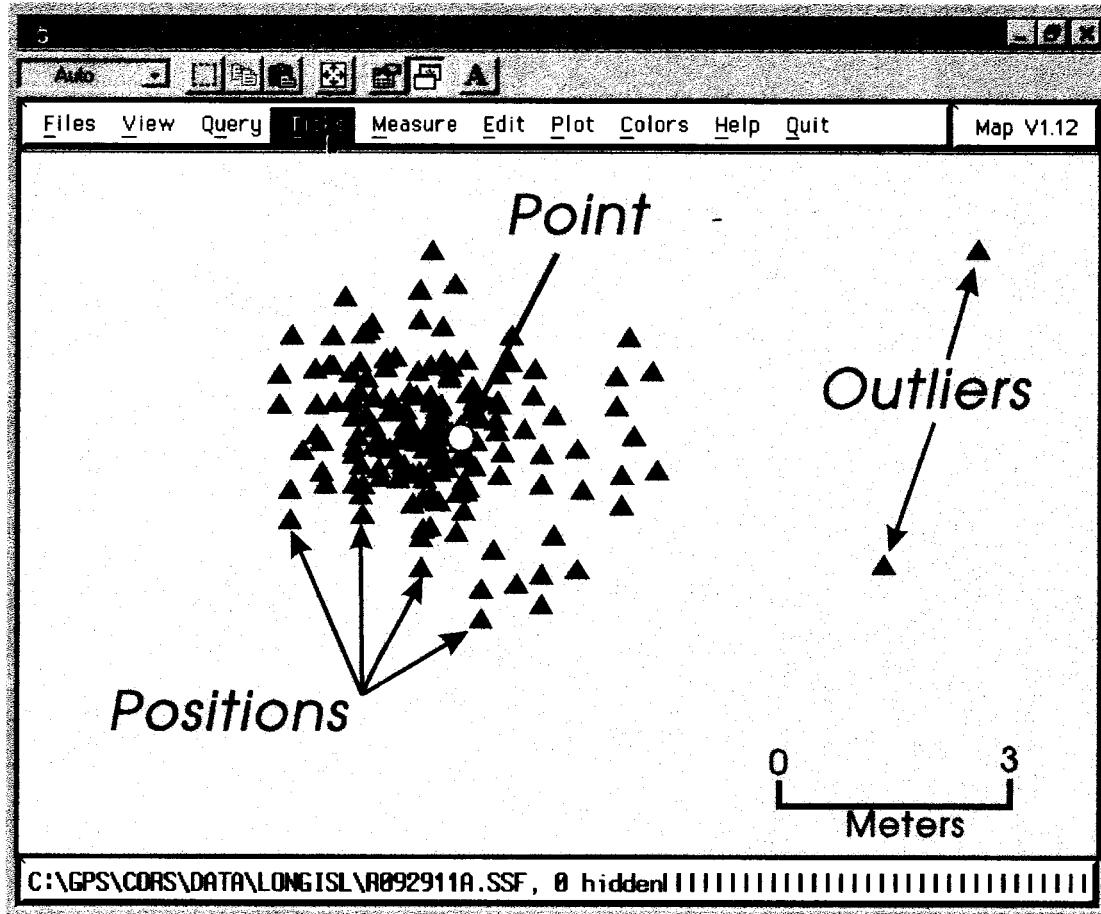
There are several techniques that can be employed in determining point and line features with GPS, and each has a specific term that implies the method used. Note that area is not included here. This is because area in the context of GPS positioning terminology is only an extension of point/line data acquisition and as such is without its own specific term.

Differentially-corrected (either real-time or post-processed) C/A-code positions are normally referred to as simply Differential Positions. If the position is not differentially corrected, the term most frequently used is *Code*, or *Code-Phase Position*. While it's possible (and sometimes desirable) to acquire line data by acquiring a series of individual code positions, line data are most commonly acquired while actually in motion (as we'll see, it's really not so different). In-motion line data collection using C/A-code is termed *Dynamic Data Collection*. This is the case whether or not the line data are post-processed or corrected in real-time.

When dealing with carrier-phase point data collection, the term used is *Static Position*. By definition, carrier-phase positions are differentially corrected, either through post-processing or real-time correction so the term is used for either. As for C/A-code, carrier-phase line data could be collected through the acquisition of individual, discrete positions. But, also as for C/A-code, linear data are most usually acquired while in motion. When in-motion carrier-phase data are collected, it's referred to as *Kinematic data collection*.

If you read widely through the growing wealth of documentation about GPS, you'll no doubt come across many misuses of these terms. In the "early days" of GPS (say, five to ten years ago), terms such as Kinematic and Dynamic were frequently used interchangeably. Today, however, they have come to have very specific meanings.

Points vs Positions



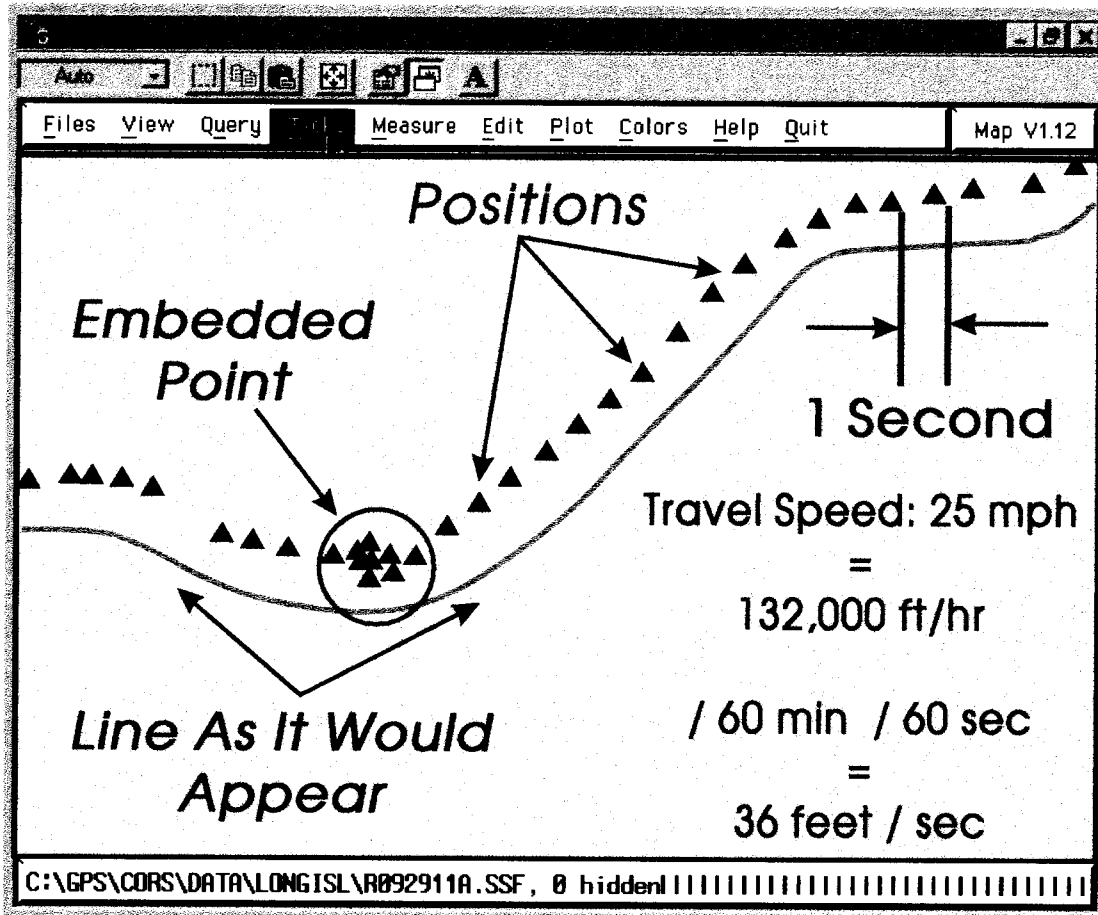
Points vs. Positions

Throughout these discussions, the term “*positions*” has carefully been used rather than the term “*points*.” That’s because in GPS the two have specific meanings that are different. GPS receivers don’t collect data in a smooth continuous fashion. Instead, the data are collected in discrete “packets” called *Epochs*. An epoch is an instantaneous event of measurement that is sometimes, incorrectly, referred to as the *interval* between the measurements. For example, “10-second epochs” (incorrect), rather than “10-second epoch intervals” (correct). The epoch interval in all but the lowest-end consumer receivers is user-definable. The most common interval is one second. Therefore, one minute of data collection would involve 60 epochs. Each epoch results in a single position.

When a location is occupied for any length of time, a number of positions are collected at one per epoch. The processing software averages the collection of positions and gives coordinates for that “point.” By this method the variability in a series of single positions is minimized resulting in a point coordinate that is typically much more accurate than that given for any single position. The technique is used for both code and carrier-phase positioning. The difference is in the amount of variability among the group of positions. For C/A-code that variability may be many meters; for carrier-phase, it may only be a few centimeters.

In some software this averaging of positions is transparent to the user. In others, it is possible to examine each of the point’s constituent positions. This has the advantage of allowing the removal of any obvious outliers. Normally, about five minutes (300 positions at one per second) is the maximum span for code collection. Any further collection is redundant. It’s less relevant for carrier positioning since it’s the change in constellation through time that is most important. Obviously, navigation applications do not have this advantage, thus the lower accuracy of dynamic or kinematic data over static or code-phase points. It is from this group of position data that an RMS error value can be calculated.

Lines From Points



Lines From Points

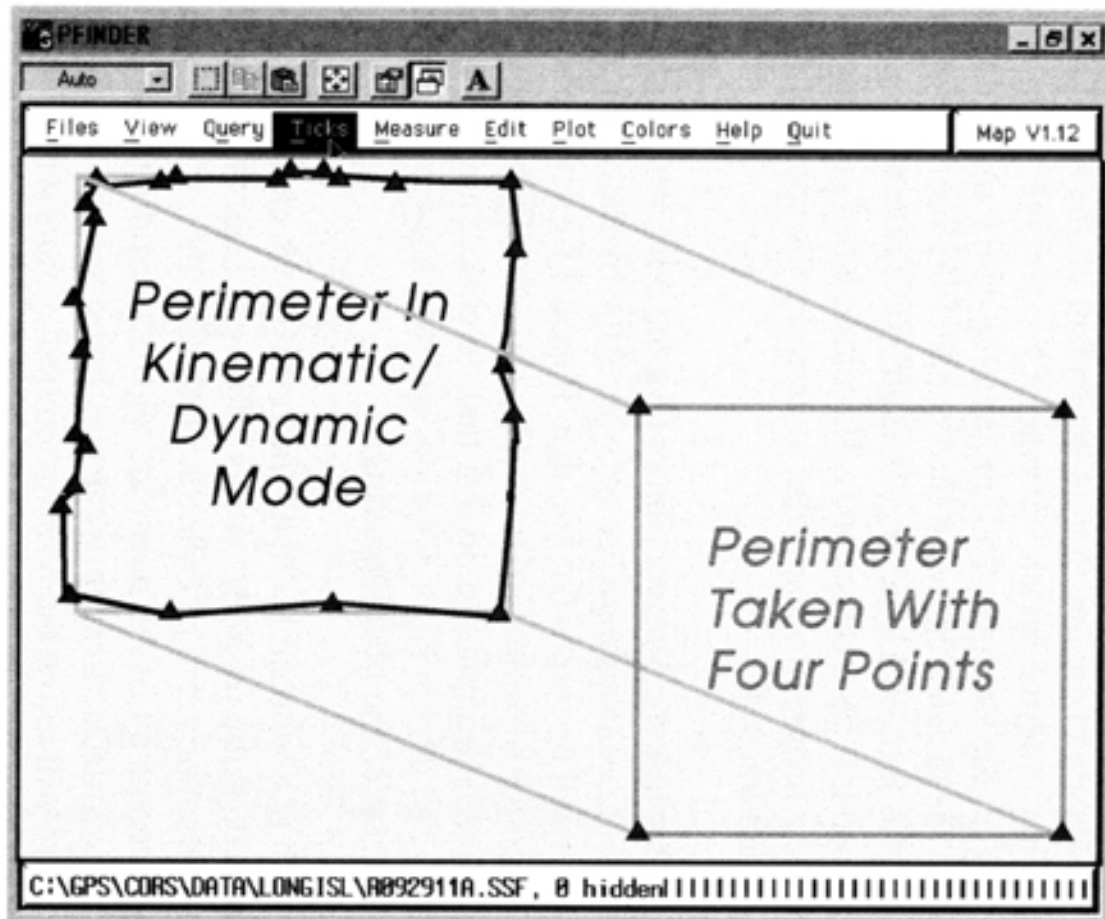
As stated earlier, lines are simply strings of points. This is also true for GPS line data collection. Basically, as a receiver is moved, it is collecting a series of individual positions that can be connected together by the software.

Density of the line then is dependent on the speed that the receiver is moving, and the epoch interval. If the receiver is on a vehicle that's moving at, for example, 25 mph, and the epoch interval is set at one per second, the distance between the "connect-the -dots" positions would be around 36 feet, or 11 meters (25 mph = 132,000 feet/hour, / 60 minutes / 60 seconds = 36 feet / second). Therefore, when planning data collection of a linear feature, it's important to consider the speed of the receiver and the epoch interval.

Because the receiver cannot take the time, while moving, to average a group of positions together anywhere, the overall accuracy of the resultant line will be somewhat less than if each 11 -meter example point were occupied for some span of time. This can partially be compensated for by periodically stopping to do just that; to collect a group of positions that will give a better position solution, and then use those points as "anchors" for the rest of the position string.

It is important to "tell" the receiver which form of data are to be collected. If the receiver "thinks" that it's collecting point data when, in fact, it's supposed to be collecting line data, then it'll automatically (usually) average all of the positions together resulting in a very weird point that has little to do with what was being sought. Most manufacturers' software will allow the user to "embed" points within linear features to give either the "anchors" just described, or to describe particular features located along the line. Again, it is important that the receiver be told what's going on or some very peculiar data can result.

Areas From Points



Areas From Points

Areas can be defined by a closed boundary of three or more connected points. Most GPS processing software has some mechanism by which an area can be derived from such a group of points.

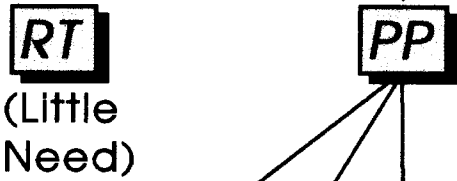
There are two basic methods by which an area can be taken with GPS. The first is simply to turn on the receiver, tell it that it's about to collect a linear feature (so that it doesn't try to average the points) and head out along the boundary of the area to be mapped. The software is later directed to calculate the internal area. That method is illustrated on the left side of the diagram. Because of the variability between successive positions, the line may well exhibit some random variability around the true line. When collecting the perimeter of an irregular area., such as a stand of a particular species of tree, this variability is probably acceptable.

When collecting the perimeter of an area with regular sides, such as a parking lot, it would probably make better sense to take a series of positions at the vertices, or corners of the area. The software can then be told to "connect the dots" and then calculate an area. This has the advantage of not requiring the user to follow precisely the perimeter line, as well as resulting in a "cleaner" line without unnecessary variation. This is illustrated in the lower right side of the diagram.

In addition, by collecting position data at each of the vertices, individual point data will be collected over a period of time (typically several minutes each). As a result, the accuracy is likely to be higher than would be the case for the dynamically/kinematically derived line positions. Static/C/A-code data collection is and remains more accurate than dynamic/kinematic data.

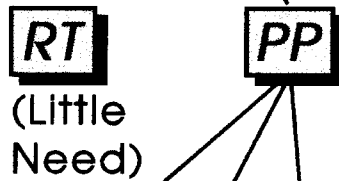
Differential Applications

GIS/Mapping



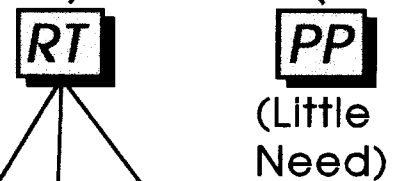
Point Line Area

Surveying



Point Line Area

Navigation



Pos. Vel. Heading

Differential Applications

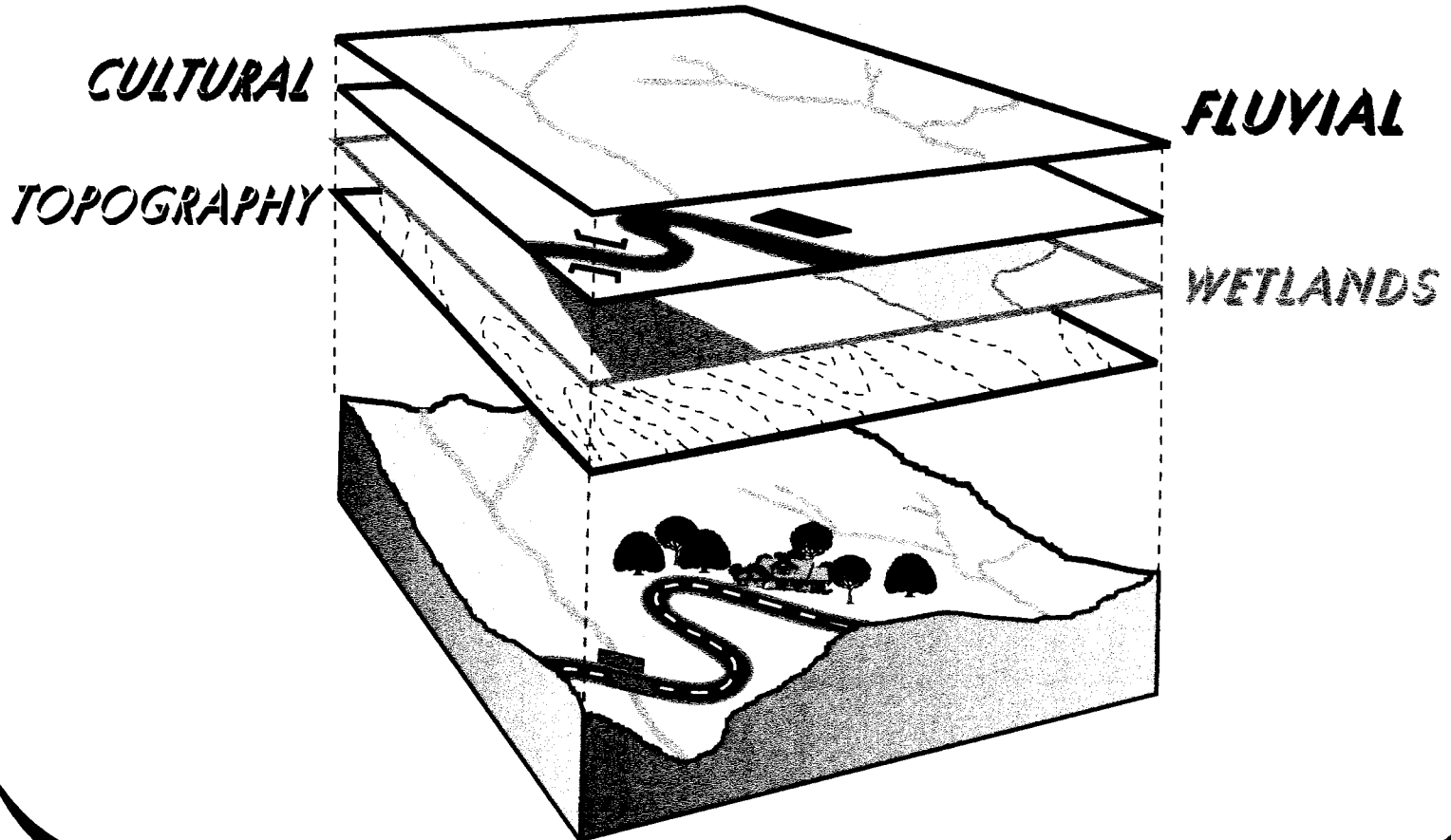
We pay particular attention to differential applications rather than autonomous applications because of their wider and generally more professional applications. Most of those applications can be broken down into three broad areas: **GIS/Mapping**; **Surveying**; and **Navigation**. These three broad classifications can further be broken down into real-time (RT) and post-processed (PP) applications depending upon the requirements of each area.

Geographic Information Systems (or GIS, which we'll speak about more in a moment) and Mapping applications generally deal with products that are not immediately time-dependent. Therefore, there is usually little need for real-time correction of GPS spatial data. The information requirements of GIS/Mapping applications can be met by post-processing and include the point, line, and area data discussed earlier.

Precision surveying applications also don't normally require real-time data correction. There are exceptions though. For example, the recovery of a survey monument that has been lost through time and the vagaries of nature is immensely helped when the surveyor, knowing the exact coordinates of the lost monument, can simply plug those values into a GPS receiver and then be guided to that exact spot. Like the GIS/Mapping applications to which it is intrinsically related, surveying deals with point, line, and area data.

Navigation is different in that it's usually necessary to know where the navigator is at that exact moment since the position is constantly changing. An airliner coming in for a landing, for example, has little use for knowing where it was yesterday. Navigation applications, therefore, are more interested in instantaneous position, velocity and heading-where it is *now*, where it's going, and how fast it's getting there.

Geographic Info Systems



Geographic Information Systems (GIS)

Geographic Information Systems, or simply GIS, can be defined as “A system of computer hardware, software, and procedures designed to support the capture, management, manipulation, analysis, and display of spatially referenced data for solving complex planning and management problems” (Logsdon, T., 1995. *Understanding the Navstar GPS, GIS, and IVHS*, Van Nostrand Reinhold Press, New York, P. 302). In a sense, it is simply a method of organizing and presenting virtually any kind of information about “stuff” in the world around us. What “stuff?” “Stuff” can be anything from fluvial systems (rivers), topography (shape of the surface), land use (cities or fields), and “ordinary” road maps, to city sewer lines, telephone poles, fire hydrants and more. Basically, if it’s out there, it occupies space. If it occupies space then it has spatial characteristics such as *where* it’s at, and individual characteristics like “big” or “small” or an individual serial number. This is all useful information to those who, in some form or another, plan, administer and manage such “stuff”.

GPS fits hand-in-glove with the requirements of GIS. All of the information about “stuff” that’s required for a computerized GIS can easily and accurately be acquired with GPS technology. The beauty of GPS for GIS applications is the simplicity and directness of data acquisition. For all intents and purposes, a GPS receiver can collect the necessary information and feed it directly into a computerized GIS with little effort expended in between.

The combination of GPS and GIS is truly a “match made in heaven” (or at least in high-Earth orbit). Where GIS provides spatial information about a wide variety of objects and features, GPS provides the real-world “connection” necessary for the GIS information to be meaningful.

GPS/GIS Applications

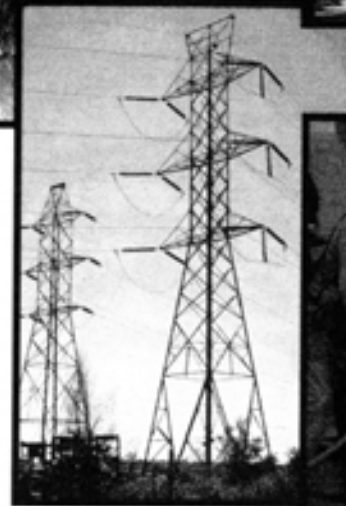
Resource Management



Postal Routing



Police Incident Reporting



Utilities Management

GPS/GIS Applications

Beyond just simple map-making, typical GPS/GIS applications include utilities management, police incident reporting, postal addressing, a wide range of demographics, resource management, and much more. For example, a county's natural gas company might want an accurate and detailed map of all of their pipelines so that anyone who wants to dig up the ground can quickly determine if it's safe to do so simply by entering the geocoordinates into the utility GIS.

By recording the GPS positions of any number of types of incidents in a GIS, police can instantly identify "hot spots" of criminal activity that might need extra patrols. Knowing the types of crimes that occur at a particular geographic location is invaluable and easily available through their GIS. Their GIS can tell them where a high number of car accidents occur, suggesting that there might be a problem with the road or with the lights or signs there, or perhaps the speed limit needs to be changed.

The U.S. Post Office needs to know the locations of all of the postal residents in a given area so that the most efficient carrier routes can be calculated by their GIS. Their system could even identify the geographic location of where every aggravating dog lives!

The field of resource management has been a major beneficiary of the mating of GPS and GIS. Resources such as wetlands and forest stands need to be mapped and monitored so that changes can quickly be identified. Since the GIS stores the data digitally, any part can be updated at any time with little effort. Update information gathered in the field with the GPS is quick and inexpensive which, in turn, allows frequent updating at reasonable cost. The list is almost endless.

The way GPS is used to provide GIS data can be categorized in two broad areas. The first is simply applying a controlling reference to some other source of data from which the GIS can collect the information that it needs. The other is to use GPS to acquire the GIS information directly.

Aerial Photo Point Control



Aerial Photo Point Control

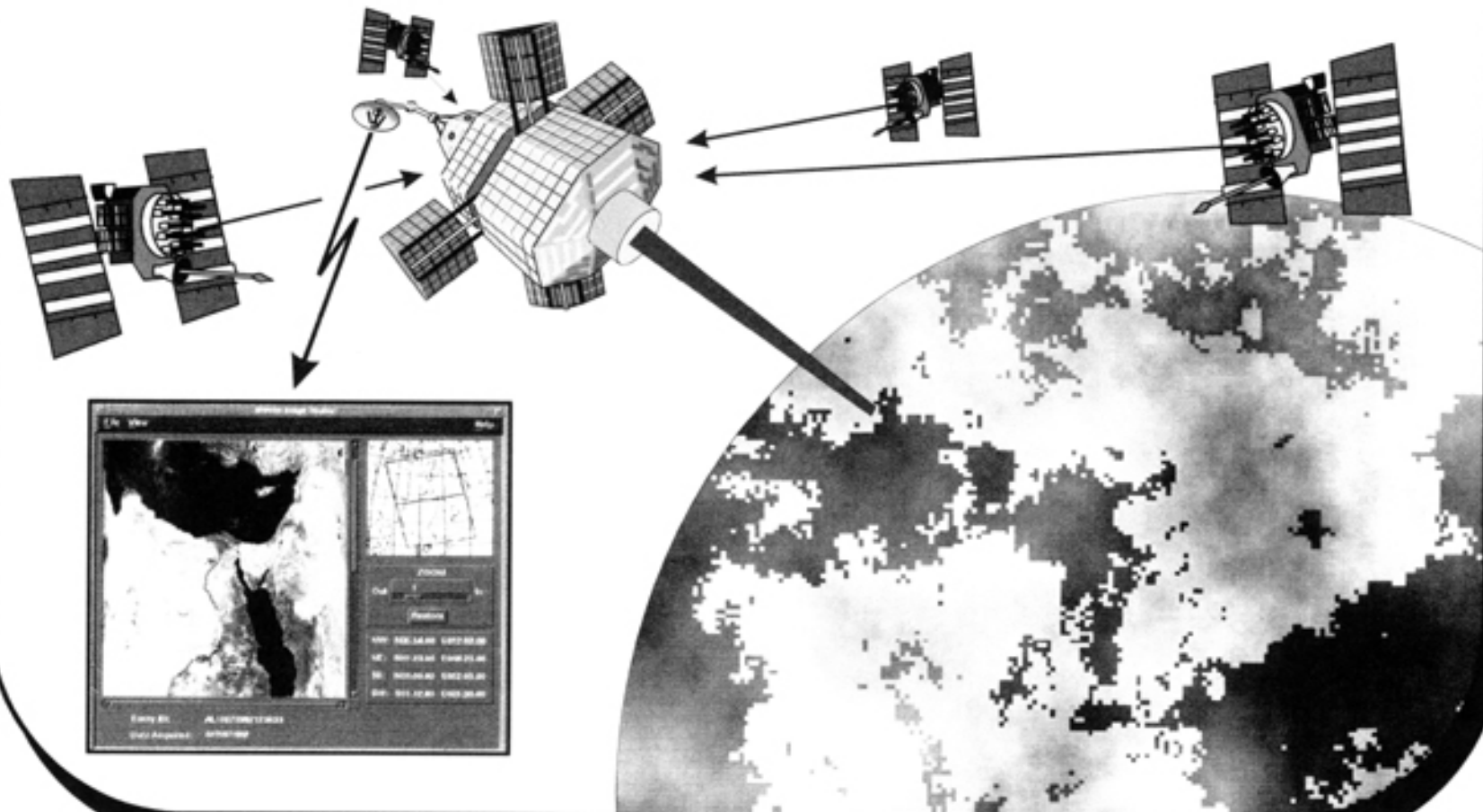
Geographic Information Systems acquire their information from a wide variety of sources. Major sources include ground surveys, pre-existing maps, and aerial photos. Ground surveys and maps have, almost by definition, some form of geographic reference assigned to them (there are some non-geographically controlled maps in existence-though they are generally of little use in a practical sense, and of no use at all to a GIS).

Aerial photos, on the other hand, while providing a virtually infinite amount of visual information about the area covered, have no intrinsic reference to the real world. In a sense, one street looks like another. The difference is *where* each street is located. That's usually referenced by some form of geodetic or geographic system such as latitude and longitude. But, in spite of the enormous amount of information available from a photographic image, *that* critically important piece of information is simply not there.

There are several ways to get an aerial photo geographically referenced so that a GIS can extract useful information from it. One way is to identify an equivalent point, such as a street corner, on a pre-existing map of the area. The problem is that any errors in the map are transferred to the photo. And maps are notoriously inaccurate. Another way is to send a surveyor to that same street corner and perform a traditional survey. This works great but costs a bundle since transferring a coordinate value to that (representative) street corner from a known bench mark is skill- and labor-intensive and, therefore, quite costly.

GPS solves these problems by being fast, inexpensive and accurate. The GPS receiver can simply be placed over the actual street corner seen in the image for a few moments and acquire the necessary position data for "anchoring" the photo to the real world. Any desired geographic information can then be accurately collected by the GIS from the now controlled, or "anchored" photo.

Satellite Imagery



Satellite Imagery, GPS, and GIS

The number of non-GPS orbiting satellites that are looking down on us continues to grow almost daily. There are already dozens of them taking pictures of the Earth in almost every part of the electromagnetic spectrum. Some of these include the weather satellites that we see daily images from on virtually every television weathercast, and radar imagers that can actually look beneath the surface to see, for example, the buried beds of ancient huge, Amazon-like rivers that once flowed in what is now the Sahara Desert of North Africa.

Space-based orbiting infrared imagers can “see” from space the difference between a broad-leaf and needle-leaf tree and determine its health. Satellites are being widely used for such things as observing the changes in the rainforests of the world to help monitor possible global climate change, and to “keep an eye on” the ever-changing ozone holes over the north and south poles. They’ve been used to record and map oil spills and volcanic clouds. Who *knows* what the military is doing with theirs. Like low-level aerial photography, high-altitude, space-based satellite imagery is finding an almost unlimited range of GIS and mapping applications.

GPS is being applied to satellite imagery at two levels. The most obvious is that, just like aerial photography, the satellite images must somehow be registered to the Earth to give it geographic meaning. GPS can be used to control, or geo-reference most satellite imagery in much the same way that it’s used for aerial photography. Once properly registered, the GIS information can accurately be extracted from the image.

The other level where GPS is being applied to satellites is in space itself. More and more, GPS is being used to position the satellites in their orbits. In fact, many new satellites are equipped with on-board GPS receivers as part of their own internal navigation and position holding systems. Even the space shuttle now flies by GPS positioning.

Features

Anything That Exists in the World Can Be Called A
Feature

Any Feature Will Have *Attributes*
Any Attribute Will Have a *Value*



Feature

Tree

Attribute

Species

Age

Health

Value

Oak

Pine

Ash

Maple

This is the **FAV**orite Way To Record GIS Data

Geographic Features

The previously discussed relationship between GPS and GIS involved applying GPS geo-coordinate control to a data source from which GIS data can then be extracted. The other method where GPS works closely with GIS is in acquiring data about “stuff” in the real world directly, thereby eliminating the middle step.

The “proper” term for “stuff” in the world is *Feature*. This term is every bit as widely encompassing as is “stuff;” it just sounds more professional. A tree is a feature. So is a road, as is a fire hydrant, and so on.

All features have *Attributes*, such as how big, how long, serial number, etc. There are always *Values* applied to those attributes, such as the actual measurements for how big or how long, and some digits for a serial number. This *Feature, Attribute, Value* set of definitions can be thought of as a GIS’s *FA Vorite* way of recording data.

An example of a feature might be a tree. Some attributes for that tree might be species, age, diameter, height, and health. Values, then, for species (for example), might be Elm, Oak, Pine, Maple, or whatever. When all of this is put together in a GIS, an analyst can then easily find out (again for example) how many Oak trees there are that are healthy, are taller than 20 meters, and are at least 75 centimeters around, as well as where each one is located.

GPS GIS Point Data Capture



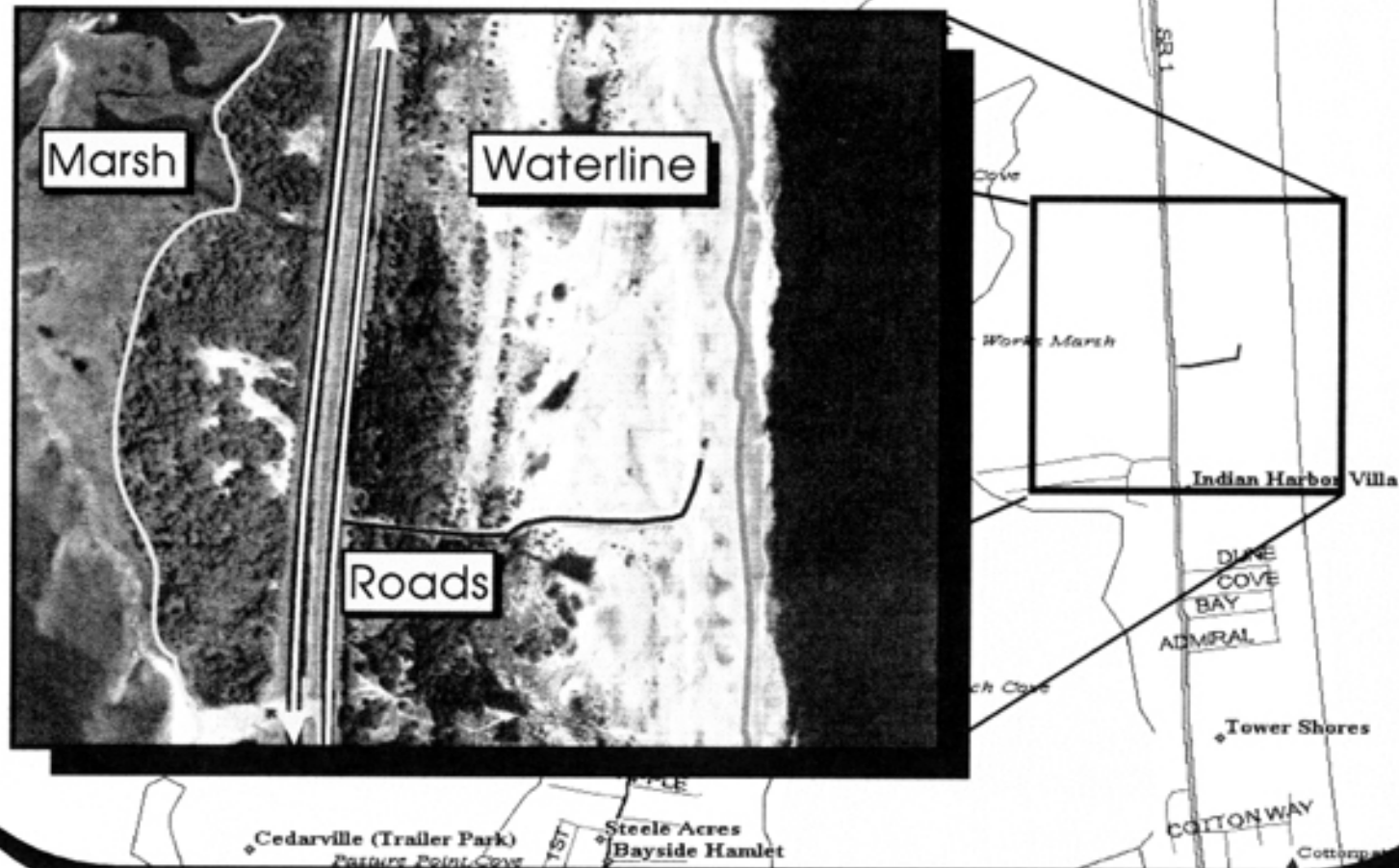
GPS/GIS Point Data Capture

The features described here, such as trees, telephone poles and fire hydrants, are very often not clearly visible on aerial photographs or maps. Surveyors *could* go out and survey each one, but it would take forever at phenomenal cost.

Here GPS truly comes into its own. Receivers with software designed for GIS data collection can be taken to that representative telephone pole, tire hydrant, or whatever, and, with only a few keystrokes, record all of the necessary information about them such as number of cross-pieces, number of transformers, condition, number of spigots, serial number and, of course, their precise geographic locations.

That information can then simply be downloaded into the GIS, usually with little to no manipulation, where it can be analyzed and/or displayed in map form. With this technique, hundreds of such features can be recorded in a single day, impossible with traditional survey methods.

GPS GIS Line Data Capture



GPS/GIS Line Data Capture

The differences between traditional cartographic “mapping” and Geographic Information Systems are increasingly becoming blurred. Conventional mapping as we generally think of it is quickly becoming absorbed into the purely digital world of GIS. And the Global Positioning System is helping to fuel this digital revolution. No longer is it necessary for cartographers and technicians to spend endless hours huddled over a scribing table to create a road map. Instead, with GPS, a user can simply turn it on in kinematic/dynamic line-feature mode and take off down the street recording it as it goes, basically creating “instant” maps as fast as the road can be driven over.

The roads thus recorded are also features, just as are the telephone poles and fire hydrants. And, just like the poles and hydrants, they too have attributes and values. For linear features, such as roads, those attributes might be condition, width, and construction material. Again, such data collection is a “piece of cake” for the Global Positioning System.

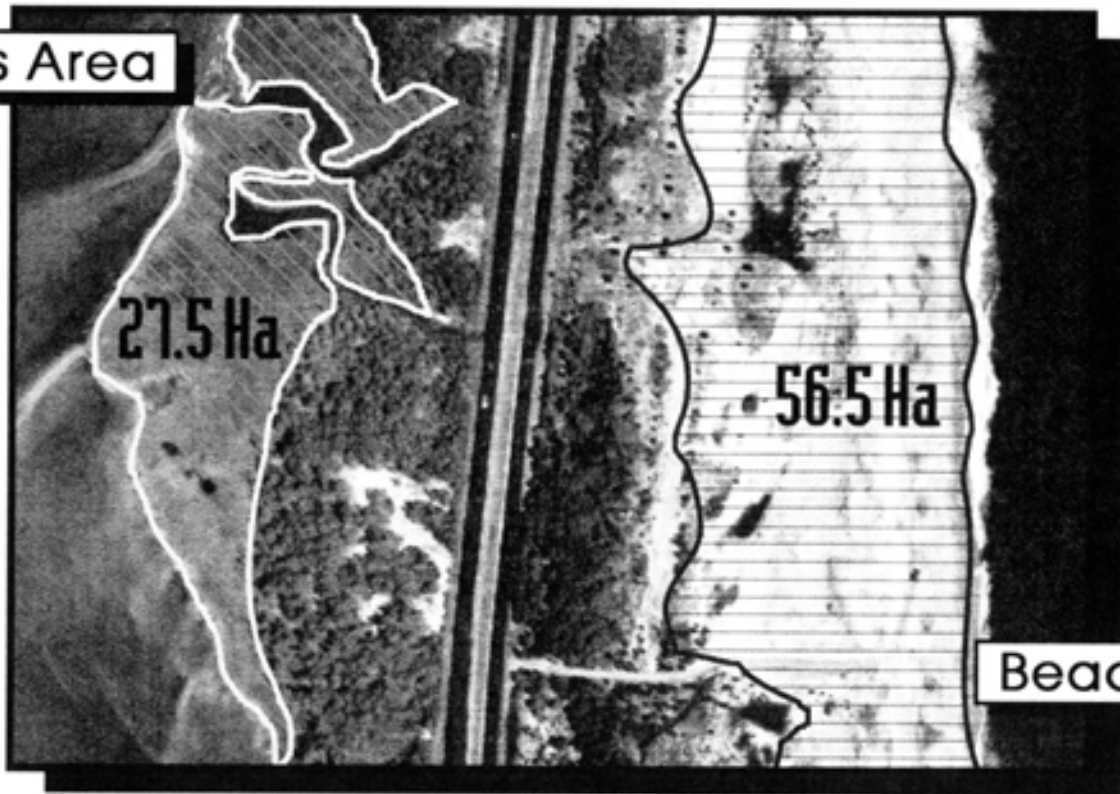
Areas From Points

Wetlands Area

27.5 Ha

56.5 Ha

Beach Area



Areas From Points

Frequently, GIS systems are used to map areas such as bird habitats, wetlands, sandy beach areas and more. This is another application where GPS excels. The acquisition of such information is simply an extension of the collection of linear data. The user only has to traverse the boundary of the area of interest, and the receiver and software combination does the rest. Areas can be calculated directly by the receiver's internal software or post-processing software (depending on manufacturer and model) or downloaded to a GIS which can then do the calculations. In any event, any difficulty involved in acquiring such information is more related to simply getting around the study area, rather than in collecting, calculating or managing the data.

External Data Sources

**DATA
OUTPUT**

**SENSOR
INPUT**

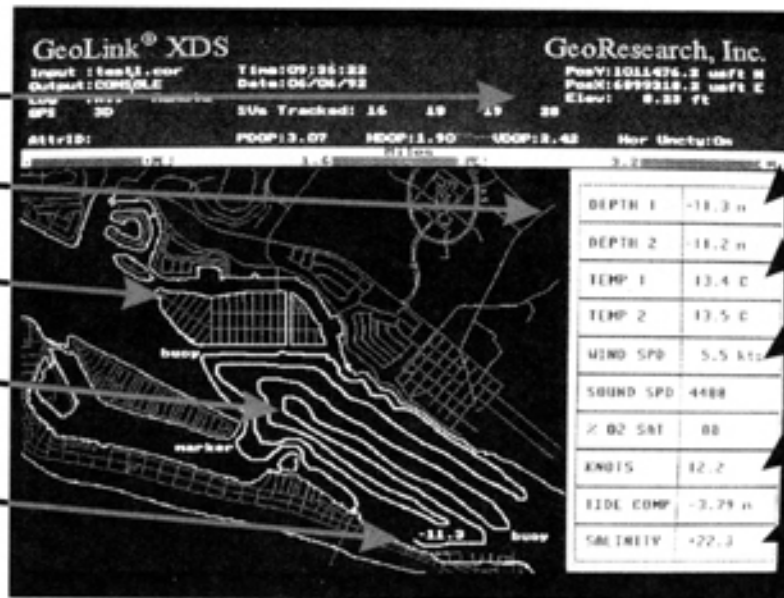
GPS Position

Sensor Data

Background Map

Bottom Contours

Sensor Location



Depth

Temperature

Wind Velocity

Speed

Salinity

All Sensor and Geo-Coordinate Data Recorded Simultaneously

External Data Sources

Geographic Information Systems depend on feature attribute values as well as geo-coordinate positions for them to be of use. This is typically a two-step process; first to get a position for a feature, then to somehow record information (attribute values) about that feature. If that attribute information happens to be, for example, tree species, fire hydrant color, or utility pole identification number, then, as we've seen, it's pretty straight-forward, the data collector simply hits the "GO" button on their receiver and then inputs the feature and its attribute values, usually through an attached keypad. The software then "attaches" the feature data to the position data for later use in a GIS.

There are, however, other types of GIS data that might be of interest to a researcher. Such information might be wind speed, or temperature, or radioactivity, or even crop yield, all of which would be positionally variable. Such data cannot easily be input into a GPS/GIS receiver because of the continuously variable nature of the data.

To deal with this difficulty, some receiver/software combinations permit the direct input of external sensor data which are instantaneously "attached" to position data on a continuous basis. The diagram at left illustrates such a package. In this example, oceanographic data are automatically recorded along with GPS position data. The right side of the screen shows that sensor data of depth, temperature, wind speed, salinity, and more, are being recorded at the location of the vessel. As the vessel moves along the area, this information is continuously updated and presented geographically accurately on a pre-existing background map. In addition, the position of the vessel is continuously given at the top right of the screen.

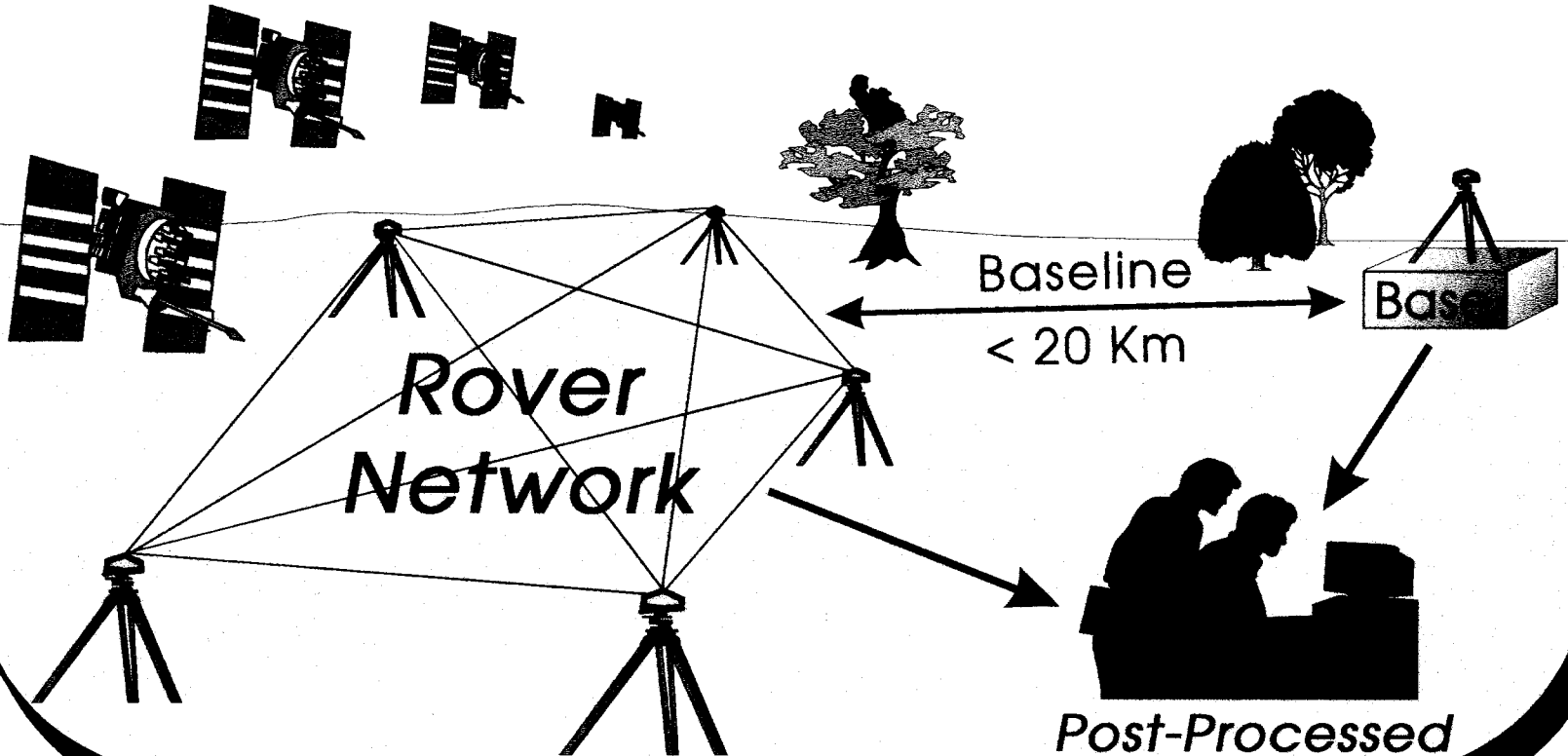
GPS Surveying

Points

Lines

Areas

Surveying Methods Go The Extra Mile



GPS Surveying

GPS *Surveying* has a special meaning. By implication, surveying means precision positioning' performed by highly trained professionals who are certified to produce legally defensible data. This can include both GPS and traditional surveying techniques. Examples of this kind of data collection might include property lines or construction surveying. The key here is that their positional data must be *legally defensible*. Surveyors are, therefore, extremely concerned with reliability and accuracy. Consequently, surveyors "go the extra mile" when acquiring positional data by whatever method they choose to use, be it by traditional transit-and-rod (or, perhaps more accurately, theodolite and total station), or by GPS surveying.

Surveying, like mapping, is also undergoing a revolution because of GPS. GPS can survey points in minutes that conventional surveys might require hours or even days to perform. High-end GPS receivers can result in positions within millimeters of truth. But such accuracy comes at a price. Survey-grade receivers are usually dual-frequency to eliminate ionospheric refraction. Base lines are kept under 20 kilometers or so to minimize any variability in what the rover and base receivers "see." To guarantee the highest levels of accuracy, multiple points are usually tied together in a network. This allows any errors to be either averaged out or made obvious so that the surveyor can identify and correct the problem.

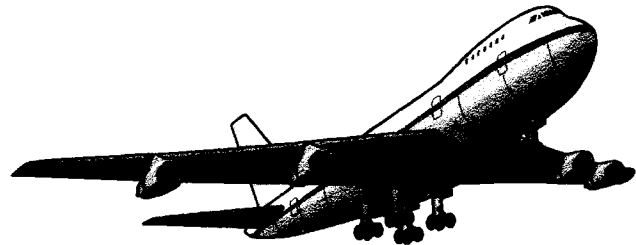
Even the satellite-transmitted ephemeris data are thrown out. This is because the ephemerides, or positions of the satellites in space that are transmitted down to the receiver are *predicted* positions. They are usually very accurate but there is always the possibility of some error. "*Precise Ephemerides*" are the actual, measured positions in space that the satellites were at for the time period of interest. Obviously, this information can only be acquired after the fact, and is available from several sources such as the National Geodetic Survey bulletin board, usually three to five days later.

GPS Navigation



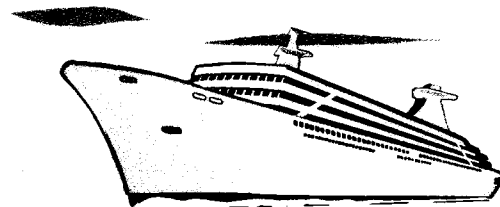
AIR

WAAS/ILS



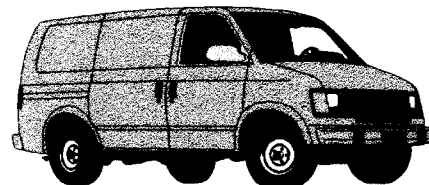
SEA

Beacons



LAND

IVHS



GPS Navigation

Unlike the applications already discussed, **Navigation** is concerned with *instantaneous position, velocity, and heading*. GPS is as well suited to these tasks as for point, line, and area determination.

Real-time navigation applications involving air travel, using augmentation systems such as the WAAS and ILS already discussed, are gaining wide acceptance and large systems are currently either being planned or installed. Precise sea navigation through the use of differential broadcast beacons already exist for most of the U.S. coastal waters. Precision farming is now regularly using GPS to guide tractors in real-time at sub-meter levels of precision, accurate enough to be able to keep aligned with individual furrows. Experiments are even currently ongoing with fully automated farming vehicles that essentially drive themselves around and through the fields.

One of the real “up-and-coming” application for GPS, and the area perhaps most likely to directly impact the “average” person, is in land navigation. This includes rail, trucking, emergency (police, fire, ambulance, etc.), and private vehicles. Collectively, these land GPS navigation applications are broadly referred to as IVS, or *Intelligent Vehicle Systems*, or more specifically **IVHS**: *Intelligent Vehicle Highway Systems*. Another term that is beginning to find increasingly wide use is **AVL**, or *Automated Vehicle Location*. There is no substantial difference between IVS and AVL and they are essentially synonymous. IVHS, on the other hand, tends to be more specific to road vehicles, and is the term that will be used throughout this book. You will no doubt come across these and other, even newer terms as the technology evolves.

IVHS

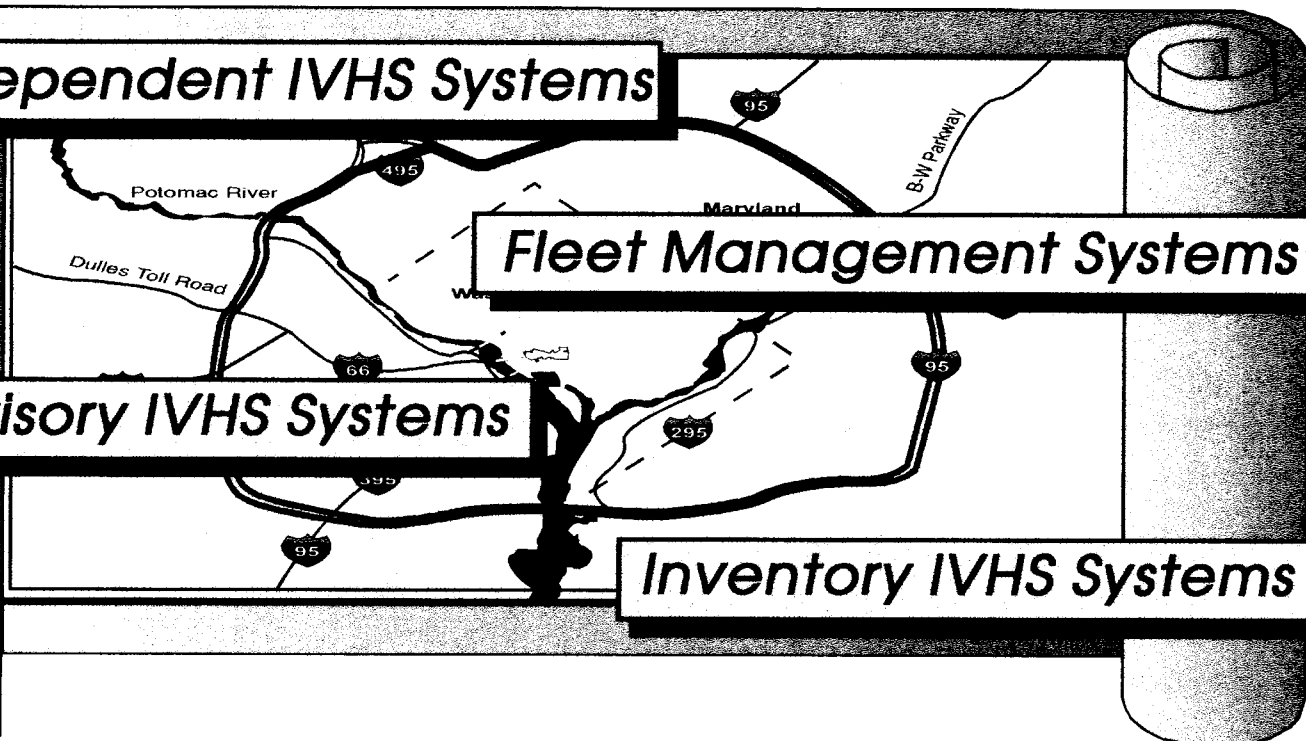
Intelligent Vehicle Highway Systems

Independent IVHS Systems

Fleet Management Systems

Advisory IVHS Systems

Inventory IVHS Systems

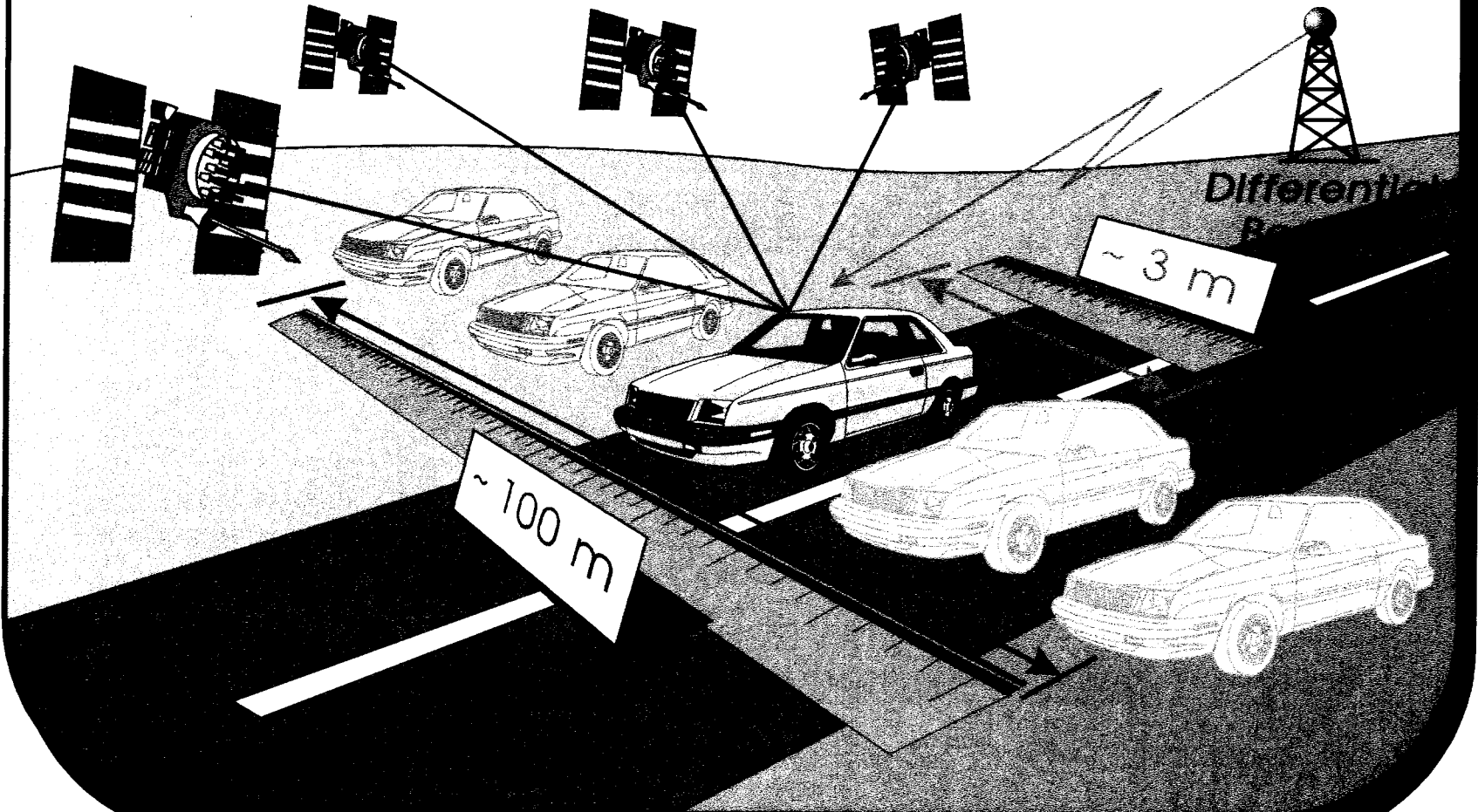


IVHS

Intelligent Vehicle Highway Systems, or simply IVHS, is a very broad term that can encompass almost any form of land road transportation that uses some form of guidance system. The term has, however, generally come to mean some form of Global Positioning System guidance. This can range from simply placing an inexpensive hand-held receiver on the dash board of the family station wagon, to the installation of a complex combination of GPS receivers, inertial guidance systems, and two-way communication to a central control facility.

In general, IVHS can be broken into four broad categories: Autonomous or *Independent IVHS Systems*, *Fleet Management Systems*, *Advisory IVHS Systems*, and *Inventory IVHS Systems* (obviously, use of the term “system” following “IVHS” is redundant, although such usage has become standard).

Independent IVHS



Independent IVHS

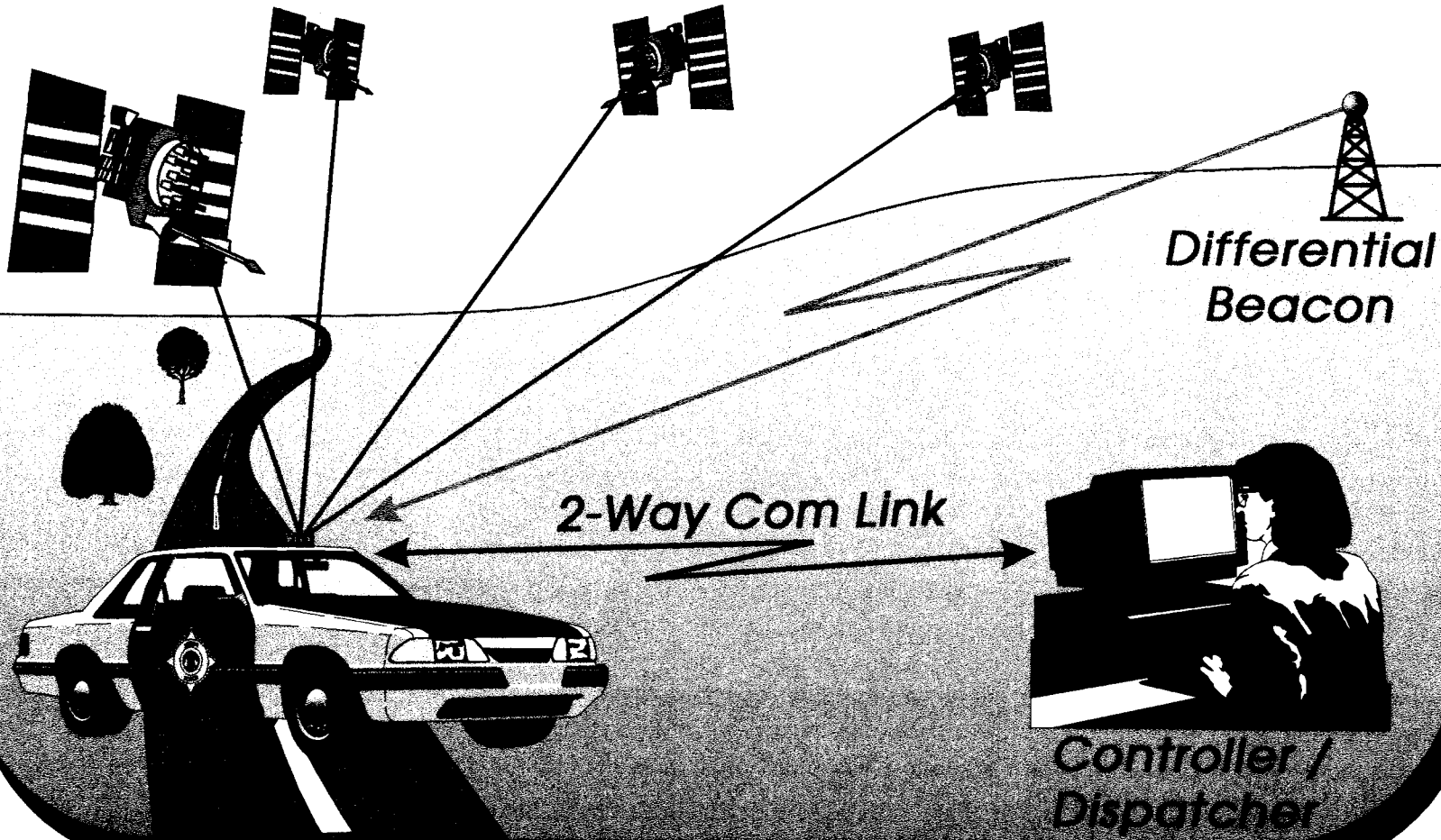
Independent IVHS is the simplest form. Independent means just that: independent. This is the same thing as “autonomous” positioning discussed earlier. A user who simply places a \$200 receiver on the dashboard of their family car is using an independent IVHS in that they are navigating without transmitting any information back out.

This form of IVHS is already available as optional built-in installations in some Oldsmobiles and Lincolns. Avis Rent a Car offers systems in their rental cars in many areas. These systems usually include a moving map display that shows the vehicle’s position on a television-like screen that moves and rotates as the vehicle travels along a road and makes turns, keeping the vehicle’s position centered on the display. Some of the more sophisticated systems also include a voice component that “speaks” to the driver, offering directions and warning of upcoming turns.

When these systems depend only on GPS, the positional error is still in that -100 meter range (assuming SA is turned on), meaning that the system doesn’t really know, for example, which lane it’s in. Based on its location and calculated direction of travel, the map display software can guess which lane it’s in and can assume that (possibly contrary to the actual received GPS position) the vehicle isn’t driving through some building, automatically adjusting the display to the nearest road in its database.

Some systems include differential correction capabilities such as FM sub-carrier reception now available in most cities. This gets the error down to about three meters which can easily place the vehicle in its actual lane. Other systems even include some form of inertial guidance system to keep track of the vehicle’s position when satellite reception is temporarily lost, such as when driving through tunnels. Usually this is done by keeping track of the vehicle’s speed with sensors and direction with accurate compasses.

Fleet Management IVHS



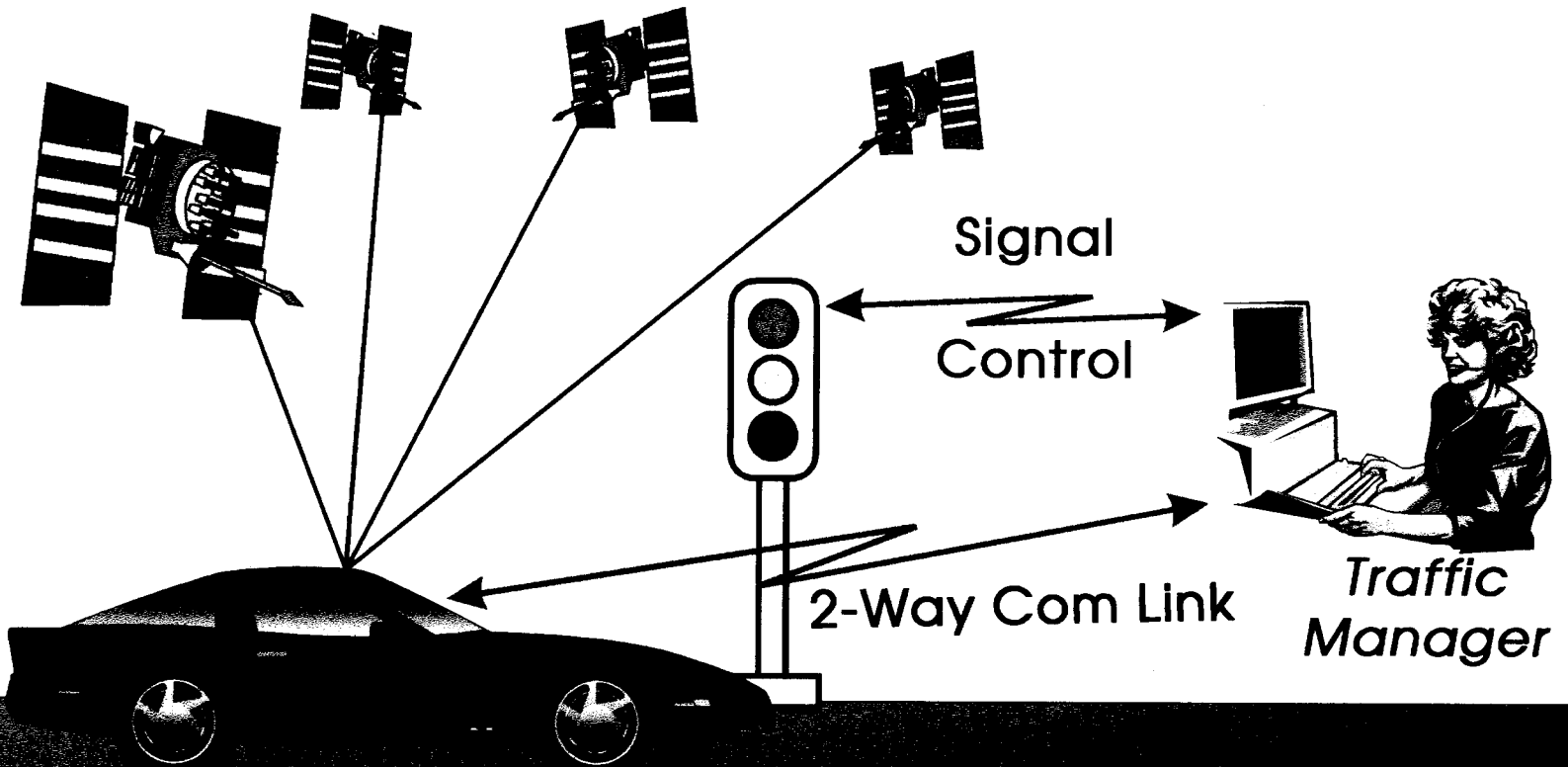
Fleet Management IVHS

Fleet Management IVHS systems allow for the rapid location and dispatching of vehicles. In these systems, the individual vehicles have accurate, usually differentially corrected, GPS receivers connected to radio transmitters that automatically send the vehicle's position back to a central processing facility. By this method, controllers and dispatchers can keep track of where their vehicles are at all times.

This has a wide range of applications. When, for example, a police emergency arises, a controller/dispatcher will immediately know which police car is closest to the scene and can get there the fastest. Ambulances can be directed straight to the emergency in the shortest possible time without ever getting lost. Such systems are already in place in many municipalities and reliance on them is quickly growing. Because of their relatively low cost, it is expected that most communities, even small ones, will implement such systems within the next few years.

Fleet management IVHS doesn't stop at emergency applications. Trucking and shipping companies, for example, are embracing this technology as fast as they can get it installed. By knowing exactly where every one of their trucks is at all times, efficiency can be improved substantially, saving the companies many times the cost of the systems.

Advisory IVHS

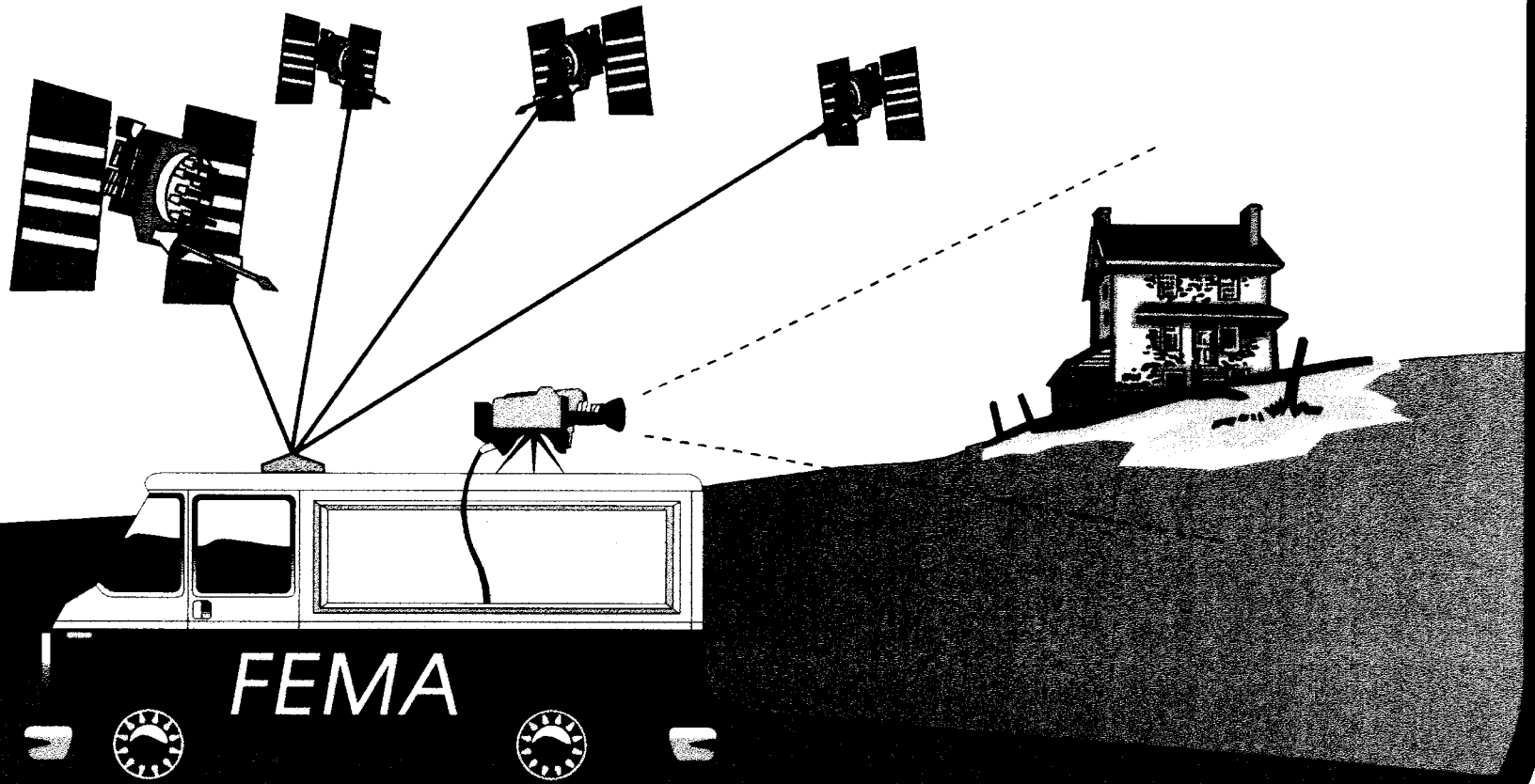


Advisory IVHS Systems

Advisory IVHS systems are an extension of Fleet Management systems. These systems are designed to smooth the flow of traffic, particularly in densely populated urban areas. In this case, users can subscribe to an advisory service that will monitor their and others' vehicles during their commutes. By tracking traffic conditions, traffic managers can signal subscribers when they're about to run into traffic problems and offer alternatives. Traffic managers can also adjust the flow of traffic by adjusting the traffic signal changes to help reduce backups.

Currently, as of this writing, there are no such systems yet fully implemented and no such subscription service for commuters is available. However, detailed plans have already been drawn up for many major urban areas. With traffic conditions in these areas continuing to deteriorate, actual implementation is expected within the next few years as traffic managers desperately search for solutions to gridlock.

Inventory IVHS



Inventory IVHS Systems

Inventory IVHS systems are designed to permit the rapid collection of data without ever having to leave the highway or road. Typically, such a system would include a GPS receiver interfaced with a recording device such as a camcorder or a digital frame camera which takes a computer-compatible digital image of the feature of interest and automatically records its geographic coordinates.

An example of such a system is already in use by the Federal Emergency Management Agency (FEMA) which is using it to record information about structures that have either been damaged by natural disasters or are at risk of damage from some future event.

Another application for this form of IVHS is in real estate where agents are able to quickly record data about the houses that they are putting on the market.

Because these systems are digital, they can quickly and easily be downloaded into Geographic Systems where risk (in the case of FEMA), or neighborhood values (in the case of real estate) can quickly be assessed.

Receiver Types

Coarse Navigation/Positioning

- Single Channel (Fast-Multiplexing) C/A < 100 m <math>< \\$200-500</math>

Mapping/High Resolution Navigation

- Single Freq, Multi-Ch. (4-12) C/A Code <math>< 100\text{m}/3-5\text{m}^*</math> ~\$400-\$2,000
- Single Freq, Multi-Ch (4-12) C/A, Smoothed <math>< 100/< 1\text{m}^*</math> ~\$1,500-\$8k

Surveying/High-Resolution Mapping

- Single Freq, Multi-Ch (6-12) Carrier Phase <math>< 30\text{ cm}^*</math> ~\$6,000-\$15k
- Dual Freq, Multi-Ch (8-12) Carrier L1/L2 millimeter* ~\$12k-\$50k

*Differentially Corrected

Receiver Types

Any discussion of specific GPS receiver types is almost self-defeating because the technology is advancing so rapidly. Nevertheless, receivers can broadly be categorized as either *Coarse Positioning*, *Mapping*, and *Survey grade*. Coarse positioning receivers are at the lowest end of cost and capability. Typical costs range from around \$200 to \$500. They are usually single or dual channel, able to collect data from only one or two satellites at a time. Since at least four satellites are required for a 3D fix, the receiver rapidly switches among the visible satellites. This is called *Fast-Multiplexing*. These units frequently don't have differential capability and don't utilize carrier-smoothing. As a result, SA accuracy is in the 100 meter range. Still, the low cost and ease-of-use make them useful for a wide range of "consumer" applications.

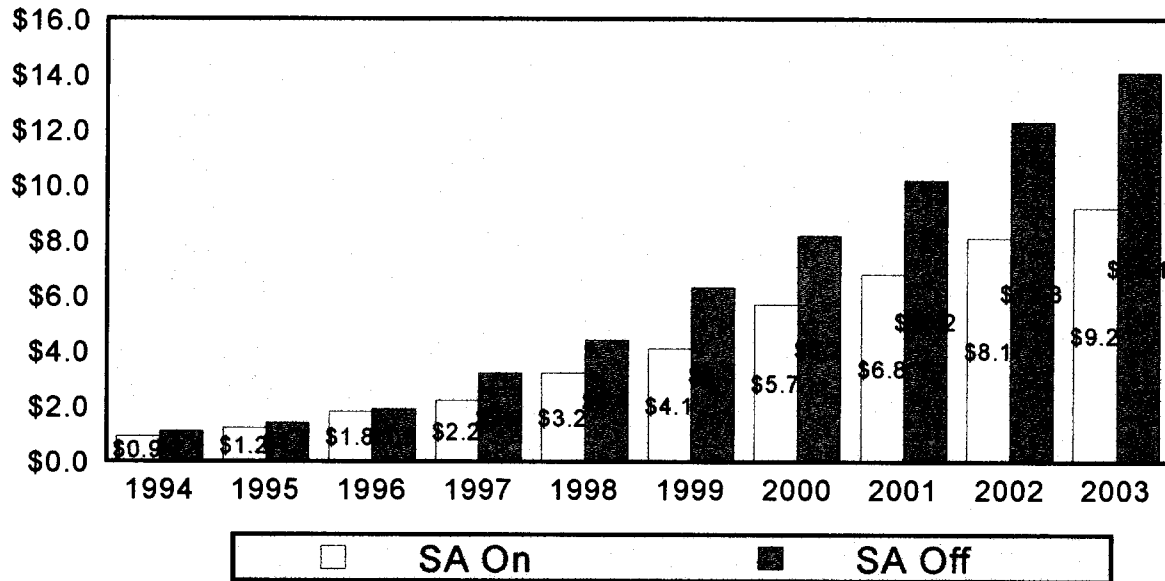
Mapping and high resolution navigation systems have a wider range of application and they're much more accurate, principally because they are differential-capable. In addition, mapping grade receivers have some form of Feature/Attribute/Value recording capability for GIS applications. Without carrier smoothing, accuracy can be expected to be in the three to five meter range. These units range from as low as \$400 to around \$2,000. If they have smoothing capability, they can achieve sub-meter accuracy. These units range in cost from around \$1,500 to \$8,000, depending on features.

Survey grade receivers can cost as little as \$6,000 for single frequency units capable of decimeter accuracy, to more than \$50,000 for the highest end dual-frequency, geodetic grade units. Clearly, there is a close correlation between how much one pays and how much accuracy one gets. Note that these are only representative of current costs in a dynamically changing field of electronics and are likely to change rapidly and continuously.

The Future Of GPS

Estimated Total N. Am. GPS Sales
1994 - 2003

US \$ Billions



Source: Booz Allen & Hamilton

The Future of GPS

Up until 1996, U.S. sales of GPS units has run around \$1 billion annually, with Selective Availability making little difference. Booze, Allen, and Hamilton, a high-tech think tank and industry analysis outfit predicts that future sales will be hampered by the degrading of the signal. By the year 2003, they predict annual sales to approach \$10 billion if S/A is left on. They believe, as do others, that that figure may exceed \$14 billion or more if S/A were to be turned off now, rather than four years from now as instructed by President Clinton's March, 1996 Policy Statement, creating thousands of jobs in this country alone.

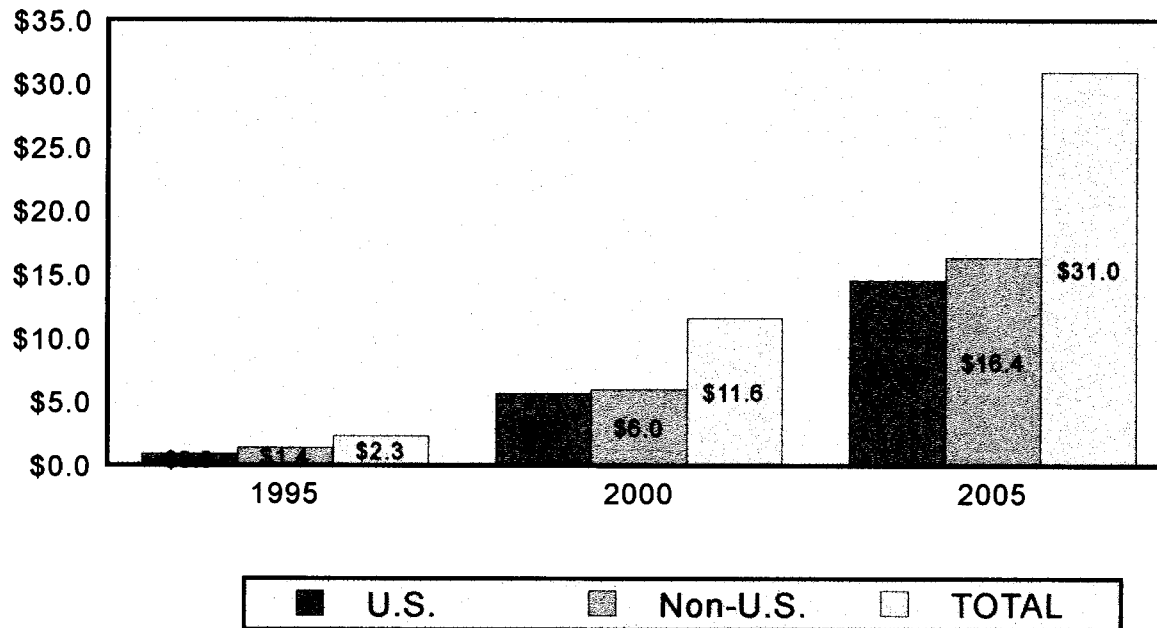
No one really knows where GPS is going in the future. Ten years ago, no one had a clue that GPS would take off as it has. Right now, the real issue of contention between the civilian and military users is still Selective Availability. In 1995, a National Research Council (NRC) report strongly recommended that the Selective Availability be immediately turned off and kept off except in times of national emergency. The military is concerned that if this is done, then during peacetime the public, in particular public transportation such as the airlines, will become so dependent on the higher, non-Selective Availability accuracy, that turning Selective Availability off, even in a national or military emergency, would impose such a safety threat that they would not be able to do so.

On the other hand, the NRC points out the absurdity of having some branches of the Federal government (i.e., the NGS and U.S. Coast Guard) spending millions of dollars to compensate for other Federal agencies (i.e., the military) that are spending millions of dollars to degrade the system. Further muddying the waters is a 1996 Rand report suggesting that the military *should* continue to employ Selective Availability until some other form of military protection is developed. The military's position is that, in the event of a national emergency, the augmentation/differential systems can simply be turned off. Where it will go remains to be seen. In any case, S/A *will* go away eventually. The issue now is how much in terms of jobs and technology-application will be lost if it takes four years.

The Future Of GPS

World GPS Market
1995, 2000, and 2005

U S \$ Billions



Source: NAPA 1995 Industry Survey

The Future of GPS

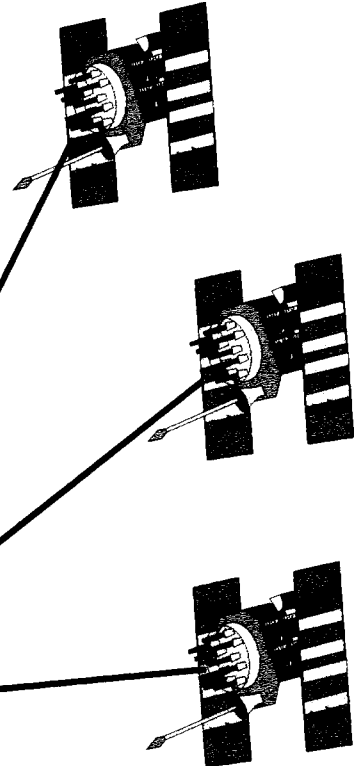
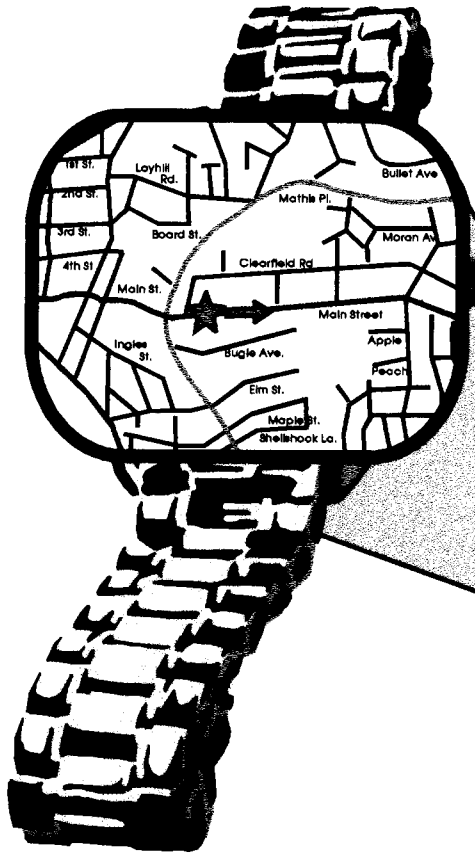
According to a 1995 NAPA Industry Survey, worldwide sales of GPS receivers in 1995 was just over \$2.3 billion, with just under \$1 billion occurring in the United States. However, ten years down the road, that figure is expected to exceed \$3 1 billion, with the majority of sales occurring overseas.

How long Selective Availability stays with us continues to be an issue. Because of SA signal degradation, the European Community has threatened to build and launch their own GPS system. Since their system would not include military “hardening,” it will be considerably cheaper than the U.S. system yet provide much higher accuracy (now loosely referred to as the European “EconoSat”).

One might respond to this with a “so what-go right ahead” attitude. After all, the Navstar GPS system is free to the user, U.S. and European alike, so why should they complain? *They’re* not paying for it. The problem is that, if this happens, there is concern that, like televisions, VCRs, and other consumer electronics, the United States will lose control of an extremely lucrative industry, along with billions of dollars in U.S. jobs, since it is expected that users will flock to the higher accuracy system. As it stands, only time will tell.

The Future?

Direction: East Speed: 2.6 MPH
Distance To Destination: 0.8 Miles
Time To Destination: 18.5 Minutes



The Future?

Politics aside, where might GPS technology take us in the near future? It is impossible to tell, but it seems that even the most outrageous guesses might not be far from fact. Up until just a few years ago, nearly all GPS receivers were produced in the United States. Now that the Japanese have started mass producing receivers, systems have abruptly begun getting cheaper and smaller, and continue to do so almost every day. How long before the South Koreans and Taiwanese begin production? Probably not long. This is going to force U.S. production up and prices down in order to compete.

Where will it lead? How about cars that tell us where we've been, where we're at, and how to get to where we want to go. Airplanes that can virtually fly themselves from takeoff to landing. Addresses given as coordinate values instead of street numbers. There is even every reason to believe that a wrist-receiver is only a few years away. After all, current GPS receiver chips, the heart of receiver systems, are already smaller than credit cards. It's all of the peripheral components that are keeping the current small units as "large" as cigarette packs. In the very near future the idea of being "lost" will become something of an anachronism and the Star Trek futurist view of taking for granted the knowing of our location anywhere on a "Class-M Planet" will become a reality. This author, for one, will be first in line.

APPENDIX I

Glossary of Terms

Appendix I

Accuracy The degree of conformance between the measured position and its true position. GPS accuracy is usually given as a statistical measure of system error. Not to be confused with **precision**. Something can be measured very precisely, but still be inaccurate. Radionavigation system accuracy can be characterized as:

Predictable Accuracy of a position solution with respect to the charted solution, where both are based on the same geodetic datum.

Repeatable The accuracy with which a user can return to a position whose coordinates have been measured at a previous time with the same system.

Relative The accuracy with which a user can measure a position relative to that of another user of the same system at the same time.

Ambiguity The uncertainty in the measurement of the number of wave cycles of the GPS carrier between the satellite and the receiver.

Argos A French-American system for positioning that uses a small transmitter which sends a signal from the surface to U.S. weather satellites, then back down to ground-based processing centers in France and the U.S. Doppler-shift distortion of the signal is used to determine position which is then transmitted back to the user via telephone, microwave links or other conventional communication channels.

Automatic Vehicle Location (See also IVHS) Any system using any sort of technology to locate or track a vehicle.

Base In GPS, refers to a GPS receiver stationed over a known, accurately surveyed point, used to apply **differential corrections** to **Rover** receivers.

Appendix I

Block I/II/IIR/IIIF The four generations of GPS satellites. Block I's were the prototypes first launched in 1978. Block II satellites are the fully operational production space vehicles. Block IIR are the Replacement, or Replenishment satellites. Block IIIF will be the Follow-on satellites.

C/A-Code The coarse acquisition code (sometimes called the Civilian Acquisition Code) is a sequence of 1,023 pseudo-random binary bi-phase modulations of the GPS carrier wave at a Chipping rate (See **Chip**) of 1.023 MHz, and a code repetition period of 1 millisecond (see **Pseudo-Random Code**).

Carrier Phase Positioning GPS measurements based on measuring the number of wavelengths of the L1 or L2 GPS carrier signal between a satellite and receiver.

Carrier Smoothing A signal processing method by which the **carrier wave** is used to increase the accuracy of **C/A-Code** positioning by smoothing the variability in individual positions in a sequence of positions.

Carrier Wave A simple radio sine wave characterized by frequency, amplitude, and wavelength, any of which can be modulated from a known reference value.

Carrier Frequency The frequency of an unmodulated output of a radio transmitter. The GPS L1 carrier frequency is 1575.42 MHz, and the GPS L2 carrier frequency is 1227.6 MHz.

CEP See Circular Error Probable.

Chip In GPS, the switch, or opportunity to switch, between a digital "0" or a "1" in a binary pulse code. Also more generally used to define most any type of electronic integrated circuit.

Appendix I

Chip Rate The number of chips per second. For example, the C/A GPS code has a chipping rate of 1.023 MHz.

Circular Error Probable (CEP) A statistical measure of GPS accuracy. In a circular normal distribution the radius of a circle containing 50% of the individual measurements being made, or the radius of a circle inside of which there is a 50% probability of being located.

Civilian Code See **C/A-code**

Clock Bias The difference between a clock's indicated time and true universal time.

Clock Offset Difference in the time reading between two clocks

Coastal Confluence Zone (CCZ) Harbor entrance to 93 kilometers (50 nautical miles) offshore or the edge of the continental shelf (at the 100 fathom curve), whichever is greater.

Code-Correlation The method by which a satellite-receiver range is determined by comparing the time difference between the satellite's and receiver's copies of a unique code.

Code-Phase GPS GPS measurements based on measuring the time delay of the **C/A-code** between a group of satellites and a receiver.

Common-Use Systems Systems used by both military and civil sectors.

Control Segment The ground-based segment of the **GPS triad**, consisting of a worldwide network of **GPS monitor stations**, uplink antennas, and **master control stations** that ensure the accuracy of the satellites' positions and their clocks.

Appendix I

Coverage The surface area or space volume covered by a radionavigation system in which the signals are adequate to permit the user to determine position to a specified level of accuracy. Such coverage is influenced by system geometry, signal power levels, receiver sensitivity, atmospheric noise, and other factors which affect signal availability.

Cycle Slip A discontinuity, or skip, in the counting of carrier sine waves between the satellites and receiver

Data Message See **Nav-msg**

Datum In **geodesy**, a fundamental measurement base fixed to the physical Earth to which a three-dimensional Cartesian coordinate system is associated which has a specified origin and orientation (usually based on an **ellipsoid**).

Delta-Range The difference in range between two different **epochs**.

DGPS Differential GPS (See **Differential Correction**)

DifferentialCorrection A technique used to improve the accuracy of a radionavigation system by determining position error at a known location and applying that determined error-correction value to the data of a user of the same system at another, unknown position. This method can be applied to post-processed (see **Post-Processing**) and **real-time** data correction.

Dilution of Precision (DOP) A measure of the contribution of satellite geometry to the uncertainty in a position fix. Values for the various DOPs typically range from 1 to 10, with lower values representing higher quality satellite geometry. Values of 1 to 3 are considered excellent; 4 to 5 good; 6 marginal; and over 6 poor to unacceptable. (See **Geometric Dilution of Precision (GDOP)**, **Position Dilution of Precision (PDOP)**, **Horizontal Dilution of Pre-**

Appendix I

cision (HDOP), Vertical Dilution of Precision (VDOP), Time Dilution of Precision (TDOP).)

Distance Root Mean Square (DRMS) A statistical measure of GPS accuracy. The root mean square value of the distances from the true location of the position fixes in a collection of measurements. 1DRMS is the radius of a circle that contains at least 68% of the fixes in a given collection at any one place. 2DRMS is the radius of a circle that contains at least 95% of the fixes in a given collection at any one place and is the most common GPS error term used. (See **CEP, SEP**.)

Dithering The deliberate introduction of clock timing error in the GPS signal in the implementation of **selective availability**.

DOP See Dilution of Precision.

Doppler Shift The predictable apparent change in frequency of a signal caused by the relative motion of the transmitter and receiver. As the transmitter and receiver move toward each other the signal is compressed and therefore the frequency becomes higher. As the transmitter and receiver move away from each other the signal is expanded, or “stretched” and the frequency becomes lower.

DRMS See Distance Root Mean Square.

Dual-Frequency Receiver A receiver capable of simultaneously receiving both the **L1** and **L2** frequencies transmitted by the GPS satellites. Such receivers have the advantage of being able to compensate for most **ionospheric refraction**.

Dynamic Positioning The determination of position while in motion, as in real-time mapping. This has come to mean by use of **C/A-code**, rather than with **carrier-phase**. (See **Kinematic Positioning**.)

Appendix I

Electromagnetic Spectrum The organization and categorization of electromagnetic waves according to their wavelength along a continuum.

Elevation Height above mean sea level. The vertical distance above the **geoid**. Not to be confused with **ellipsoidal height**.

Ellipse A two-dimensional curve where the sum of the distances from any point of the curve to two fixed points, called foci, is constant. Rotating an ellipse around its minor axis results in an ellipsoid of revolution, or, simply, an **ellipsoid**.

Ellipsoid (See Geodesy) A mathematical figure formed by revolving an **ellipse** around its minor axis. Often used interchangeably with spheroid. Two quantities define a ellipsoid: The length of the semimajor axis (a) and the flattening ($f=(a-b)/a$, where (b) is the length of the semiminor axis). Frequently defined by the semi-major axis (a) and the eccentricity (e) where the eccentricity is defined by:

$$e = \sqrt{\frac{a^2 - b^2}{a^2}}$$

Ellipsoidal Height The measure of vertical distance above the ellipsoid. Not to be confused with height above sea level or the **Geoid**. Ellipsoidal height is the measurement used by the GPS and is based on the **WGS 84**. Usually given as HAE, or Height Above the Ellipsoid.

Ephemeris A list of accurate positions or locations of a celestial object as a function of time. In GPS, refers to the positions of the satellites as they move through their orbits. Available from the satellites as “broadcast ephemeris” (also called “predicted ephemeris”), or from other sources as “precise ephemeris” for post-processing. Precise ephemeris data are the actual, measured positions of the

Appendix I

satellites for a particular point in time, whereas the broadcast ephemeris data are predicted and are, therefore, somewhat less accurate. Plural form: Ephemerides.

Epoch An instantaneous event of measurement that is sometimes, incorrectly, referred to as the *interval* between the measurements. For example, “ 10 second epochs” (incorrect), rather than “ 10 second epoch intervals” (correct).

Fast Multiplexing In GPS, a single channel receiver that samples a number of satellite ranges in rapid sequence. Cheaper to manufacture than multi-channel receivers that are capable of receiving multiple satellite signals simultaneously. Typical of low-end “consumer” or “sport” receivers.

Frequency In electromagnetic radiation, the number of sine waves passing a fixed point per unit of time.

Full Operational Capability (FOC) The level of GPS system capability achieved when a full complement of Block II satellites were placed in operational orbit. This occurred in July, 1995. (See **Initial Operational Capability (IOC)**.)

Geocentric Relative to the Earth as a center, measured from the center of mass of the Earth.

Geodesy The science related to the determination of the size and shape of the Earth by such measurements as triangulation, leveling, and gravimetric observations.

Geodetic Datum A mathematical model designed to best fit part or all of the **geoid**. Defined by **an ellipsoid** and the relationship between the ellipsoid and a point on the topographic surface established as the origin of a **datum**. (See **Geodesy**.)

Appendix I

Geoid The equipotential (equal-gravity) surface that coincides with mean sea level that may be imagined to extend through the continents. This surface is everywhere perpendicular to the force of gravity.

Geoidal Height The height above the **geoid**. Usually referred to as the elevation above mean sea level (MSL). (See **Ellipsoidal Height**.)

Geometric Dilution of Precision (GDOP) The collection of geometric factors that degrade the accuracy of position fixes derived from externally-referenced navigation systems. In GPS, related to the spatial orientation and distribution of the satellites above and around the user.

GPS The U.S. Department of Defense, satellite-based Global Positioning System comprising 24 satellites in 6 orbits spaced 60 degrees apart with 4 satellites each, at an altitude of 20,200 kilometers (10,900 nautical miles). GPS satellites transmit signals that permit the accurate determination of position anywhere on the Earth.

GPS Triad The three principal components of the Global Positioning System, consisting of the **space segment**, the **control segment**, and the **user segment**.

HAE Height Above the **ellipsoid**.

Hertz (Hz) Cycles per second.

Horizontal Dilution of Precision (HDOP) The Dilution of Precision in two dimensions horizontally. This value is often lower (i.e., “better”) than some other DOP values because it ignores the vertical dimension. (See **Dilution of Precision**.)

ILS Instrument Landing System. (See **Precision Approach**.)

Appendix I

Inclination One of the orbital parameters that specifies the elevation of an orbit relative to a reference plane. In the GPS, that plane is the Earth's equator.

Initial Operational Capability (IOC) The level of GPS system capability achieved when a minimum of 24 GPS satellites (Block I and II) first became operational in their orbits and were available for navigation. This occurred in December, 1993. (See **Full Operational Capability (FOC)**.)

Integrity The ability of a system to provide timely warnings to users when the system should not be used for navigation.

Intelligent Vehicle Highway Systems (IVHS) Also, IVS, or Intelligent Vehicle Systems. Any of a number of systems by which road vehicles can be accurately tracked. More and more this is coming to mean vehicle monitoring, tracking, and positioning by some method that includes the GPS as at least one of its components. IVS, or Intelligent Vehicle System is more general in that it includes non-highway land transportation such as rail.

Ionosphere A band of electrically charged particles between 50 and 200 kilometers (25 to 120 miles) altitude (variable) which acts as a refractive medium for radio signals. (See **Refraction**.)

IVHS See **Intelligent Vehicle Highway Systems**

IVS Intelligent Vehicle Systems. (See **Intelligent Vehicle Highway Systems**.)

Kinematic Positioning The determination of position while in motion. This has come to mean by use of **carrier-phase** techniques, rather than **C/A-code**. (See **Dynamic Positioning**.)

Appendix I

L-Band The group of radio frequencies ranging from 1.2 GHz to 1.6 GHz. The GPS carrier frequencies (1575.42 MHz and 1227.6 MHz) are within the L-Band.

L1 Signal (or L1 Frequency) The primary L-Band frequency used by GPS satellites centered at 1572.42 MHz. The L1 signal is modulated with the **C/A-Code**, **P-Code**, and **Nav-msg** data strings.

L2 Signal (or L2 Frequency) The second L-Band frequency used by GPS satellites centered at 1227.6 MHz. The L2 signal carries the **P-Code** and **Nav-msg** data.

Loran A long-range, ground-based navigation system where position is determined by measuring the time intervals between pulsed radio signals from two or more pairs of ground stations at known locations.

Mask Angle The (usually) user-definable angle above the horizon below which satellites will not be tracked. Normally set to between 15 and 20 degrees to avoid interference caused by buildings, trees, **multipath**, and to minimize atmospheric **refraction**.

Master Control Station (MCS) One component of the ground-based control segment of the **GPS triad**. The main Master Control Station is located at Falcon Air Force Base in Colorado Springs, Colorado, with two back-up Master Control Stations located in Sunnyvale, California, and Rockville, Maryland. Unmanned **monitor stations** passively track all GPS satellites visible to them at any given moment, collecting signal (ranging) data from each. This information is then passed on to the Master Control Station where the satellite position (**ephemeris**), and clock-timing data are estimated and predicted.

MHz Megahertz. Millions of cycles per second

Appendix I

Microsecond One-millionth of a second

Millisecond One-thousandth of a second

Monitor Stations One component of the ground-based control segment of the **GPS triad** used to track satellite clock and orbital parameters. Data collected at monitor stations are linked to the **master control station** where corrections are calculated and from which correction data are uploaded to the satellites as needed.

MSL Mean Sea Level. (See **Geoidal Height**.)

Multichannel Receiver A receiver capable of continuously receiving several signals from different satellites simultaneously, each on a different channel. This allows the derivation of position solutions from the simultaneous calculations of pseudo ranges.

Multipath Interference caused by reflected GPS signals arriving at a receiver as a result of nearby reflective surfaces such as buildings and bodies of water. Multipath signals have traveled longer distances and so produce erroneous **pseudo-range** estimates and consequent positioning errors.

Multiplexing See **Fast-Multiplexing**.

AND-83 North American Datum of 1983. (See **Datum**.)

Nanosecond (ns) One billionth of a second.

Nautical Mile (nm) 1.156 statute miles; 6,076 feet.; 1,852 meters

Navigation The process of planning, recording, and controlling the movement of a craft or vehicle from one place to another.

Appendix I

Nav-msg A modulation superimposed on both the C/A- and **P-Codes** with a data rate of 50 bits per second. The Nav-msg contains GPS system time of transmission, a hand-over word (HOW) for the transition from C/A- to P-Code tracking, **ephemeris** and clock data for the particular satellite being tracked, and **almanac** data for all the satellite vehicles in the constellation. Additionally, it contains information such as satellite health, coefficients for ionospheric delay models for C/A-Code users, and coefficients to calculate **Universal Coordinated Time (UTC)**.

Omega A ground-based radionavigation system that uses extremely low frequency radio waves. The system determines position by measuring phase differences between two or more transmitters. Coverage is global using only 8 transmitters. Accuracy is low; on the order of 3-6 kilometers (1.6-3.2 nautical miles).

P-Code The precise, protected, or precision code of the GPS signal, typically used without **differential correction** by U.S. military and other authorized users. It consists of a 267-day long sequence of **pseudo-random** binary biphasic modulations of the GPS carrier waves with a **chip** rate of 10.23 MHz. Each satellite is randomly assigned a one-week segment of the code which is changed weekly. The P-Code becomes the **Y-Code** when encrypted, necessitating a special decryption key available only to authorized users.

PDOP See Position Dilution of Precision

Position Dilution of Precision (PDOP) The dilution of Precision in three dimensions. Sometimes called the Spherical Dilution of Precision. Probably the most commonly used DOP value. (See **Dilution of Precision**.)

Post-Processing A technique of **differential correction** where data from **rover** receivers are corrected for numerous sources of error

Appendix I

by some form of computer processing with **base** receiver data some time after the data from all receivers have been collected.

PPS See **Precise Positioning Service**.

Precise Positioning Service (PPS) The highest level of military dynamic positioning provided by GPS using the **dual-frequency P-Code**. Specified to provide **6.0 meters 1DRMS** horizontal, or 16 meter **SEP** accuracy.

Precise Time A time requirement accurate to within 100 **milliseconds**.

Precision In GPS, the degree, or size of unit, to which a position can be measured. Not to-be confused with **accuracy**. It is possible to be very precise but inaccurate.

Precision Approach A standard aviation instrument approach procedure in which an electronic glide slope is provided, as in the **Instrument Landing System (ILS)**. There are three main ILS categories:

Category I (CAT I) Conditions must be such that visibility height above touchdown is 61 meters (200 feet) or greater, and runway visual range is 550 meters (1,800 feet) or greater. Instrument navigation positioning accuracy must be 17.1 meters (56 feet) horizontal, 4.1 meters (13.5 feet) vertical.

Category II (CAT II) Conditions must be such that visibility height above touchdown is 30.5 meters (100 feet) or greater, and runway visual range is 366 meters (1,200 feet) or greater. Instrument navigation positioning accuracy must be 5.2 meters (17 feet) horizontal, 1.7 meters (5.6 feet) vertical.

Appendix I

Category III (CAT III) Conditions must be such that runway visual range is 213 meters (700 feet) or greater. There is no height visibility limitation. Instrument navigation positioning accuracy must be 4.1 meters (13.5 feet) horizontal, 0.6 meters (2 feet) vertical. (This is for **CAT IIIA** Instrument approaches. There are, in addition, **CAT IIIB** and **CAT IIIC** classifications that have runway visibility limitations of 46 meters (150 feet), and no limitations, respectively. Instrument position specifications remain the same.)

PRN Pseudo-Random Noise. (See **Pseudo-Random Code**.)

Pseudolite Pseudo-Satellite. A ground-based transmitter that simulates a GPS satellite that can be used for ranging. The **NAV-msg** portion of the signal may contain **differential correction** data. Because the transmitter is ground-based, it can significantly improve the **DOP** values for receivers located above it, such as aircraft.

Pseudo-Random Code Also called Pseudo-Random Noise (PRN). A signal so complexly modulated with I's and O's that it appears to be random noise, even though it is not.

Pseudo-Random Noise (PRN) See **Pseudo-Random Code**

Pseudorange A range (distance) measurement based on the correlation of a satellite-transmitted code and a receiver's internally generated code that has not been corrected for clock synchronization errors between the satellite and receiver's clocks.

Radionavigation The determination of position, or the obtaining of information relating to position, for the purposes of navigation by means of the propagation properties of radio waves.

Appendix I

Real-Time Differential Correction A method by which a **differential correction** signal can be instantaneously applied in real-time to a **rover** receiver. (See **RTCM**.)

Refraction In GPS, the bending, or slight redirection, of the incoming signal satellite signal. Most commonly due to the **ionosphere** and the **troposphere**.

Relative Positioning See **Differential Correction**.

RINEX (Receiver Independent Exchange) A data format based on a set of standard definitions for GPS time, phase, and range observables. GPS manufacturers use their own proprietary formats for GPS data collected with their equipment. Consequently, one system cannot read the data of another. RINEX permits appropriate software to process GPS data even if such data are collected using receivers of differing manufacturers.

Rover In GPS, those receivers used to determine the positions of unknown points. Data from rover receivers are generally **differentially corrected** with data collected by **base** receivers.

RTCM (also **RTCM-SC104**) A common language for the **real-time differential** signal transmission developed by the Radio Technical Commission for Maritime Services Special Committee 104 (currently Version 2.1).

SA See **Selective Availability**

Satellite Constellation In GPS, the collection of orbiting satellites, located in 6 orbital planes elevated 55 degrees to the equator at an altitude of 20,200 kilometers (10,900 **nautical miles**), with three or four satellites per orbit for a total of 24 (21 functioning, 3 operational backups).

Appendix I

Selective Availability (SA) A Department of Defense method by which the accuracy of the **C/A-Code** is degraded to 100 meters **2DRMS** by **dithering** the satellite clocks and **ephemeris** data sent on the **Nav-msg**.

SEP See **Spherical Error Probable**.

Sigma See **Standard Deviation**

Space Segment The space segment of the **GPS triad**, consisting of the constellation of orbiting GPS satellites (see **Satellite Constellation**), and any ancillary spacecraft that provide GPS augmentation information such as **differential correction**.

Spherical Error Probable (SEP) A statistical measure of GPS accuracy. In a spherical normal distribution the radius of a sphere containing 50% of the individual measurements being made, or the radius of a sphere inside of which there is a 50% probability of being located. (See **DRMS**, **CEP**.)

Spheroid See **Ellipsoid**.

SPS See **Standard Positioning Service**.

Standard Deviation (Sigma) The statistical measure of the dispersion of random errors about the mean. If a large number of measurements or observations are made, the standard deviation is the square root of the sum of the squares of deviations from the mean value divided by the number of observations minus one.

Standard Positioning Service (SPS) The civilian positioning service level of accuracy achieved with a single frequency receiver using the **C/A-code**. With **selective availability** implemented, that accuracy is specified to be no worse than 100 meters 95% of the time (**2DRMS**).

Appendix I

Static Positioning In GPS, location determination using a stationary receiver, allowing the use of multiple-position averaging over time. The term has come to mean **carrier phase positioning** exclusively.

Survey The science and technology of making measurements to determine the relative position of points on, above, or beneath the Earth's surface.

Surveying The branch of applied mathematics used in the art of accurately determining the area of any part of the Earth's surface, the lengths and directions of the bounding lines, or the contour of the surface, and accurately delineating the whole on a map or chart for a specified datum. In modern time, it has come to be synonymous with legal certification, denoting that work performed by a legally certified surveyor is legally defensible in a court of law.

SV Space Vehicle, or Satellite Vehicle.

Time Dilution of Precision (TDOP) Dilution of precision with respect to time (clock offset).

Troposphere The lowest layer of the atmosphere. The zone where weather phenomena and atmospheric turbulence are most marked. Contains 75% of the total molecular or gaseous mass of the atmosphere and virtually all of the water vapor. Extends from the surface to the floor of the tropopause at between 8 and 16 kilometers (5 to 10 miles) altitude.

Universal Time Coordinated (UTC) Previously GMT. An international, highly accurate and stable uniform atomic time scale. The basis for civil time. It is occasionally adjusted by one-second increments to ensure that the difference between the uniform time scale, defined by atomic clocks, does not differ from the Earth's rotation by more than 0.9 seconds. Maintained by the U.S. Naval Observatory.

Appendix I

Universal Transverse Mercator (UTM) Grid A military grid system based on the Transverse Mercator projection applied to maps of the Earth's surface extending to 84 degrees North and South latitudes.

User Segment One segment in the **GPS triad**. Includes all end-user receivers that determine position and velocity with the GPS.

Vehicle Location Monitoring A service provided to maintain the orderly and safe movement of platforms or vehicles, encompassing the systematic observation of airspace, surface, and subsurface areas by electronic, visual or other means to locate, identify, and control the movement of vehicles. (See **IVHS**.)

Vertical Dilution of Precision (VDOP) The **dilution of precision** in the vertical dimension alone.

Wavelength In electromagnetic radiation, the distance between two equivalent points on two successive sine waves.

World Geodetic System (WGS) A consistent set of parameters describing the size and shape of the Earth and the positions of a network of points with respect to the center of mass of the Earth. The most current is the WGS-84.

Y-Code The encrypted version of the **P-Code**.

APPENDIX II

*Suggested Readings
on the GPS*

A Basic Library

Appendix II

The following is a list of suggested readings that will give further insight into the workings of the GPS. The list is by no means comprehensive as the wealth of information continues to grow daily. The list only includes books and booklets (both hardback and soft-cover) and does not include any of the long list of technical journal papers available through any engineering library. Because these books deal with a rather esoteric topic they tend to be expensive. In addition, some are considerably more technically advanced than others. Therefore, the reader may want to seek some of these volumes out at a library before purchasing them. Most books are available through Navtech Seminars' GPS book store (1-800-628-0885, 703-931-0500) in Virginia, though this is not meant to be an endorsement.

A Technical Report to the Secretary of Transportation on a National Approach to Augmented GPS Services U.S. Department of Commerce, National Telecommunications and Information Administration, U.S. Department of Transportation. 1994. Softcover book. Detailed study report on the various types of GPS augmentation systems currently being considered by the U.S. Government. Straightforward and informative. Around \$25.00. May be available through U.S. Gov. Printing Office.

A Comprehensive Guide to Land Navigation With GPS by Noel J. Hotchkiss. 1994. Softcover. Basic. An easy to use, graphically based guide to navigating with the GPS. Lots of color illustrations, user friendly. Bases its examples on several brands of inexpensive hand-held receivers. Heavy treatment of orienteering in general. Around \$30.00.

Basic Geodesy (NOAA Reprint) by the Defense Mapping Agency. 1977. Softcover booklet. Excellent introduction to the basics of geodesy. Clear and concise with many useful diagrams. Only about 40 pages but a wealth of information. Available through NOAA-NGS at (301) 713-3242. A steal at around \$4.00.

Appendix II

Differential GPS Explained by Jeff Hum. 1993. Trimble Navigation. Softcover booklet. Very basic but clear and quite easy to understand. Great for beginners. Around \$10.00.

Geodesy for the Layman (NOAA Reprint) by the Defense Mapping Agency. 1977. Softcover booklet. Excellent basic introduction to the basics of geodesy. More technically advanced than NOAA's reprint of **Basic Geodesy**. Some use of mathematics. Still, clear and concise with many useful diagrams. Only about 100 pages but a wealth of information. Available through NOAA-NGS at (301) 713-3242. A steal at around \$7.00.

Geodetic and Charting Publications. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, National Geodetic Survey. Continually updated. A listing of incredibly useful publications on geodesy, earth sciences and GPS available for pennies each. A real gold mine. Catalog is free from NOAA-NGS at (301) 713-3242.

GIS GPS Sources by Ralph Heatly. 1996. Hardcover. A directory of GPS and GIS sources along with articles on history and future of GPS and GIS. Available from Avanstar Marketing 1-800-598-2839. Fairly expensive at \$85 (+S&H) for 281 pages, but a useful reference.

Global Positioning System by the Institute of Navigation. Four Volumes I, 1980 (\$20), Volume II, 1984 (\$20), Volume III, 1986 (\$20), and Volume IV, 1993 (\$30) (or \$85 for entire set). The "bible" of GPS. Highly technical. The "red books." Available through Navtech.

Global Positioning System: Management and Operation of A Dual Use System by the Joint DoD/DOT Task Force. 1993. Softcover. Government report on the civilian and military interests in policy management of the GPS and recommendations for resolution of differences between them. Around \$15.00.

Appendix II

GPS Bibliography Subscription Service. Canadian GPS Associates. Updated twice a year. Comprehensive collections of GPS titles. Very useful resource. Available in hardcopy or computer disk for both PC or MAC. Expensive, but worth it for the professional. Ranges from \$150 for a single computer copy, up to \$500 per year for annual subscription for both hard and disk copy.

GPS Navstar Global Positioning System, User's Overview. ARINC Research Corporation, for the Navstar Joint Program Office. Document YEE-82-09D. 1991. Softcover. Overview of the User Segment of the GPS. Reads well for a government document. May be available through U.S. Coast Guard GPS Information Service or the USAF GPS Technical Library (3 1 0-363-3596) for free. Commercially available for around \$25.00.

GPS Surveyor's Field Guide; A Field Guidebook for Dynamic Surveying by Trimble Navigation. 1991. Softcover booklet. Useful, but more appropriate for surveyors. Around \$10.00.

GPS Surveyor's Field Guide; A Field Guidebook for Static Surveying by Trimble Navigation. 1991. Softcover booklet. Useful, but more appropriate for surveyors. Around \$10.00.

GPS The Guide to the Next Utility by Jeff Hum. 1989. Trimble Navigation. Softcover booklet. Very basic but clear and quite easy to understand. An excellent beginner's book. Around \$10.00.

GPS Theory and Practice by B. Hoffmann-Wellenhof, H. Lichtenegger, and J. Collins. 1994. Softcover book. Highly technical, engineer-oriented. Extensive mathematics. Can provide extremely detailed answers to most any question. Around **\$65.00.**

Appendix II

Guide to GPS Positioning by David Wells. 1987. Canadian GPS Associates. Softcover. Comprehensive, but fairly technical. Somewhat dated in that it doesn't recognize the newest developments. Around \$40.00.

Interface Control Document GPS (200) ARINC Research Corporation, for the Navstar Joint Program Office. 1993. Softcover. Military/Government document and reads like it. Only get it if you need to impress your friends with things like “. . . bits 23 and 24 of word three in subframe 1 shall be the two MSBs of the ten-bit IODC term. . . ” Does contain useful information but other sources present it better. May be available through U.S. Coast Guard GPS Information Service or the USAF GPS Technical Library (3 1 O-363-3596) for free. Commercially available for around \$25 .00.

Navstar GPS User Equipment - Introduction, Public Release Version by the NATO Team at USAF Space Systems Division, GPS Joint Program Office. 1991. Softcover. Another government document but reads very well and contains much useful information. A “must-have.” May be available through U.S. Coast Guard GPS Information Service or the USAF GPS Technical Library (310-363-3596) for free. Call it the “yellow-cover” and they'll know. Commercially available for around \$25.00 and worth it.

Technical Characteristics of the Navstar GPS. NATO Navstar GPS Support Group (public release version). 1991. Softcover. Government document and reads like it. Useful information on GPS system characteristics. Softcover. May be available through U.S. Coast Guard GPS Information Service or the USAF GPS Technical Library (3 10-363-3596) for free. Commercially available for around \$25.00.

Appendix II

The Global Positioning System and GIS, An Introduction by Michael Kennedy. 1996. Hardcover. A fine book but it's really a hands-on, step-by-step how-to book very specifically written for Trimble receivers and Arc-Info. If you don't use these systems, then this book won't be of much use. Comes with a CD of practice data. Available directly from Ann Arbor Press, MI, (currently the only source) at \$54.95 + \$3.50 S&H via fax 1-8 10-362-4932 (credit card) or mail order to KLC, BNP, PO Box 7069, Troy, MI 48007-7069.

Understanding the NAVSTAR GPS, GIS, and IVHS by Tom Logsdon. 1995. Hardback book. Fairly comprehensive. Not too deep. Tends to lack some detail but informative. Around \$50.00.

About the Author

Gregory T. French is a Senior GPS/GIS Project Manager at GeoResearch, Inc., a company specializing in integrated GPS / GIS applications and field data collection. He holds Bachelor of Science and Master's degrees in Physical Geography with minors in Geology and Computer Applications in Geography from the University of Maryland. He received his basic GPS training from *Ashtech, Inc.*, in 1992, and advanced Master GPS training from *Corvalis Microtechnology, Inc. (CMT)*, where, in 1996, he received his GPS Instructor's certification.

Mr. French has a long and varied background in earth sciences technology. He has served the EPA's *Environmental Photographic Interpretation Center (EPIC)* as a remote sensing imagery analyst and as Chief of Remote Sensing Imagery Research and Acquisition. As a research scientist at the University of Maryland's *Laboratory for Coastal Research*, he was deeply involved in research in state-of-the art digital shoreline mapping technology. He has taught Remote Sensing Laboratories, as well as undergraduate courses in Geomorphology and Coastal Environments at the University of Maryland, and is currently adjunct Professor of General Sciences at University College, University of Maryland, where he instructs courses in Global Environmental Change and Oceanography.

He has authored over two dozen technical reports on environmental, earth science, and GPS topics for Federal and State agencies as well as for open publication, and continues to explore the ever-expanding horizons of GPS applications in modern geography.