

Global Positioning System

ir. G.J. Husti

CIP-gegevens Koninklijke Bibliotheek, Den Haag

Husti, G.J.

Global Positioning System / G.J. Husti - Delft : Delft University Press – Ill.

Uitg. in opdracht van : Vereniging voor Studie- en Studentenbelangen te Delft. –
1e dr. Delft : 2000.

Met lit. opg., reg.

ISBN 90-407-1977-2

NUGI 816, 841

Trefw.: global positioning system, GPS

© VSSD

Eerste druk 2000

Uitgegeven door:

Delft University Press

Postbus 98, 2600 MG Delft

tel. 015 278 3254, fax 015 278 1661,

e-mail dup@dup.tudelft.nl; website <http://www.library.tudelft.nl/dup>

In opdracht van:

Vereniging voor Studie- en Studentenbelangen te Delft

Poortlandplein 6, 2628 BM Delft

tel. 015 - 2782124, telefax 015 - 2787585, e-mail: vssd@tudelft.nl

internet: <http://www.oli.tudelft.nl/vssd/hlf/handleidingen/html>

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of op enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

ISBN 90-407-1977-2

Voorwoord

Het boek “Global Positioning System” heb ik samengesteld ten behoeve van de basisstudie van de opleiding geodetisch ingenieur.

In 1989 was ik redacteur van “Navigatie en geodetische puntsbepaling met het Global Positioning System”. Dit boek is inmiddels verouderd. Begrijpelijk, want de ontwikkeling van GPS heeft in de laatste 10 jaar een enorme vlucht genomen. Bij de moderne meetmethoden van geodetische puntsbepaling (grondslag- en detailmetingen), navigatie, geodynamica, en Geografische Informatie Systemen, is GPS niet meer weg te denken.

Men vindt deze ontwikkeling ook terug in het onderzoek en onderwijs bij de Afdeling Geodesie. Het GPS-onderzoek is sterk uitgebreid en het aantal college-uren met GPS als onderwerp is aanzienlijk opgevoerd. In het onderwijs maken de studenten tegenwoordig reeds vanaf de basisstudie kennis met GPS. Er is dus behoefte aan een leerboek dat de elementaire begrippen van GPS behandelt. Daarnaast verwijs ik naar diverse uitstekende Engelstalige boeken, o.a. “GPS for Geodesy” van P.J.G. Teunissen en A. Kleusberg (Eds) uit 1998 (zie Referenties). Deze boeken zijn meer geschikt voor specialisatie in de eindstudie.

In Hoofdstuk 1 wordt GPS met de terrestrische radioplaatsbepalings-systemen vergeleken. Een overgang van 2-D naar 3-D. Hoofdstuk 2 geeft een overzicht van de basisbegrippen. De hoofdstukken 3 en 4 heb ik gedeeltelijk overgenomen van “Navigatie en geodetische puntsbepaling met het Global Positioning System”. De coördinatenstelsels in § 3.1 de en de satellietbanen in § 3.3 worden naar de beschrijving van B.H.W. van Gelder behandeld. Hoofdstuk 4, de invloed van de atmosfeer op het GPS-sigitaal, is naar de beschrijving van J.C. de Munck bewerkt en uitgebreid. Hierbij heb ik dankbaar gebruik gemaakt van de kennis van Frank Kleijer en Dennis Odijk, beiden werkzaam als promovendus bij de sectie Mathematische Geodesie en Puntsbepaling (MGP). Hoofdstuk 5 en Hoofdstuk 6 geven een overzicht van de toepassing van GPS, respectievelijk met code-afstanden (bijv. voor de navigatie) en fase-afstanden (voor precieze geodetische puntsbepaling). In Hoofdstuk 7 wordt de coördinatentransformatie behandeld, speciaal gericht op toepassing in Nederland. Sommige figuren en tabellen heb ik hiervoor met toestemming uit de HTW 1996 overgenomen. In Hoofdstuk 8 worden de permanente referentiestations t.b.v. GPS-metingen behandeld. In Hoofdstuk 9 tenslotte wordt GLONASS besproken, de Russische tegenhanger van GPS. Hierbij heb ik gebruik gemaakt van het (bewerkte) artikel van N.F. Jonkman: “Een inleiding tot GLONASS” (Geodesia, 1998/12).

Tenslotte bedank ik een aantal collegae voor hun bijdrage aan dit boek. Mijn bijzondere dank aan Piet Sluiter en de docenten van de sectie Mathematische Geodesie en Puntsbepaling: Kees de Jong, Frank Kenselaar, Hans van der Marel en Christiaan Tiberius voor het doornemen van concept-teksten en het doen van suggesties.

Inhoud

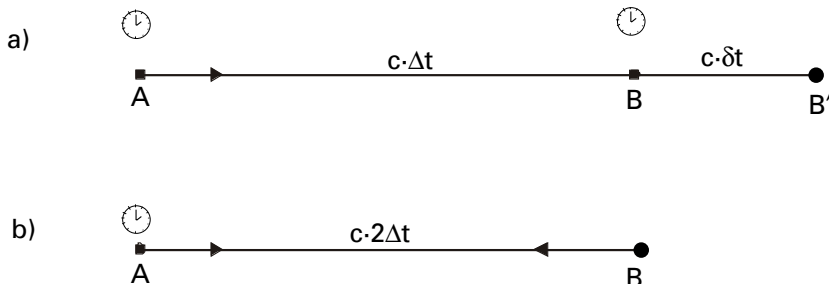
1	INLEIDING	9
2	EEN OVERZICHT VAN GPS	14
	2.1. Algemene kenmerken.	14
	2.2. GPS satelliet signaal. De uitgezonden datablokken	18
	2.3. Decodering van de data	21
	2.4. Overzicht van enkele basisbegrippen	22
	Fysische begrippen 22 / De meetbare grootheden 22 / Meting van code-afstanden 23 / Fasemeting 24 / Frequentieverschuiving (dopplereffect) 24 / Selective Availability (SA) 24 / Anti-Spoofing (A-S) 24 / GPS-ontvangers 25 / Invloed van de atmosfeer 27 / Meetmethoden van GPS 27 / Multipad 28 / Software 29 / Foutenbronnen 29 / Satellietconstellatie 30	
3	COÖRDINATENSTELSELS, TIJDSYSTEMEN EN SATELLIETBANEN	33
	3.1. Cartesische en ellipsoïdische coördinaten	33
	WGS-84 en ETRS coördinaatstelsels 35 / Topocentrisch coördinaatsysteem 36 / Inertiaal coördinaatstelsel 36	
	3.2. Tijdsystemen	37
	Universele tijd en sterrentijd 37 / GPS-tijd 39	
	3.3. Satellietbaan	40
	Baanvlak van de satelliet 41 / Oriëntering van de baanellips 42	
	3.4. Broadcast Ephemeris	44
	Berekening van klokfout voor de satellieten 45	
	3.5. Berekening van de satellietcoördinaten uit Broadcast Ephemeris	45
	Data format 46	
	3.6. Precise Ephemeris	47
	3.7. Yuma GPS Almanac	47
	3.8. Omschrijving van enkele begrippen	47
	Precessie 47 / Nutatie 47 / Aardrotatie 48 / Poolbeweging 48	
	3.9. Voorbeeld van RINEX file	49
4	INVLOED VAN DE ATMOSFEER	52
	4.1. Inleiding	52
	4.2. Geometrische optika	52
	4.3. De troposfeer	54
	De brekingsindex in de troposfeer 54 / De troposferische vertraging naar het zenit 55 / De troposferische vertraging in schuine richting 56 / Standaardatmosfeer 58	
	4.4. De ionosfeer	58
	Inleiding 58 / De brekingsindex in de ionosfeer 58 / Groepslooptijd en	

	faselooptijd 59 / De ionosferische vertraging 60 / Consequenties van dispersie 62	
	4.5. Meting van korte basislijnen	63
	Schaafeffect 63 / Hellingseffect 64	
5	PLAATSBEPALING UIT CODE-AFSTANDEN	65
	5.1. Inleiding	65
	5.2. Het mathematische model	66
	5.2.1. GPS-tijd, satelliet-tijd en ontvanger-tijd 66 / 5.2.2. Pseudoafstand uit code-meting 66 / 5.2.3. Pseudoafstand uit fasemeting 67	
	5.3. Plaatsbepaling uit code-afstanden	68
	5.4. Invoering van geografische coördinaten als onbekenden	71
	5.5. Toepassing van constraints voor de onbekenden	73
	5.6. Invloed van de satellietconfiguratie	74
	5.7. Kinematische toepassingen	75
	5.8. Relatieve plaatsbepaling: Differentiële GPS (DGPS)	76
	5.9. Kalman filter	82
6	RELATIEVE POSITIE UIT FASEWAARNEMINGEN	84
	6.1. Inleiding	84
	6.2. Model van fase-afstanden	84
	6.3. Covariantiematrix van de verschil-afstanden	87
	Covariantiematrix van de enkel-verschillen 87 / Covariantiematrix van de dubbel-verschillen 88 / Covariantiematrix van de drievoudig-verschillen 88	
	6.4. Cycle slips; detectie en reparatie	88
	6.5. Lineaire combinatie van L1 en L2 fasewaarnemingen	90
	6.6. Vereffening van fasewaarnemingen	92
	6.6.1. Inleiding 92 / 6.6.2. Relatieve positie uit dubbel-verschillen 94	
	6.7. Oplossen van meerduidigheden	96
	6.7.1. Inleiding 96 / 6.7.2. Oplossing van meerduidigheden met de LAMBDA methode 97	
	6.8. Meting van korte basislijnen (< 30 km)	100
	6.8.1. Meetmethoden 100 / 6.8.2. Meetkundige grondslag met GPS 102 / 6.8.3. Detailmeting met GPS 105 / 6.8.4. Continu kinematische meting 110	
7	DATUMTRANSFORMATIE	114
	7.1. Ruimtelijke transformatie met zeven parameters	114
	7.2. Transformatie tussen ETRS89 (= WGS-84) en RD-stelsel	117
	7.3. Getallenvoorbeeld voor transformatie	120
8	REFERENTIESTATIONS	122
	8.1. Inleiding; meetkundige grondslag	122
	ETRS89-RD 124 / Naar ETRS89-RD ? 124	
	8.2. Referentiestations voor DGPS metingen	125
	Radio Data System (RDS) 126 / Amplitude Modulated Data System (AMDS) 127 / Radiobeacons 127 / International Maritime Satellites	

	(Inmarsat) 127	
	8.3. Permanente referentiestations voor fasemeting	128
	Actief GPS Referentie Systeem (AGRS) 128 / LNR Globalcom 130	
9	OVERZICHT VAN GLONASS	131
	9.1. GLONASS-GPS integratie	132
	9.2. GLONASS ruimtesegment, besturingssegment en gebruikerssegment	134
	9.2.1. GLONASS ruimtesegment 135 / 9.2.2. GLONASS besturingssegment 137 / 9.2.3. GLONASS gebruikerssegment 139	
	9.3. GLONASS tijdsignaal en navigatieboodschap	140
	9.3.1. Structuur GLONASS tijdsignaal 141 / 9.3.2. Inhoud GLONASS navigatieboodschap 143	
	9.4. Relatieve plaatsbepaling	147
	LITERATUUR	148
	INDEX	150

1 Inleiding

De ontwikkeling van radioplaatsbepaling begint na de tweede wereldoorlog met de terrestrische radioplaatsbepalingssystemen. Diverse terrestrische systemen werden ontwikkeld: Decca, Hifix, Loran-C, Pulse-8, Syledis, Omega etc. Elk systeem heeft specifieke eigenschappen met betrekking tot de gekozen frequentie van het radiosignaal. Het kenmerkende hierbij is de samenhang tussen frequentie, precisie en reikwijdte. Met een radiogolf van hoge frequentie kan men een hoge precisie verkrijgen, maar de reikwijdte is beperkt. Hoogfrequente radiosignalen hebben nu eenmaal een rechtlijnige voortplanting en daardoor, in verband met de kromming van de aarde, een kleine reikwijdte. Een voorbeeld hiervan is het systeem van Syledis (ca. 400 MHz, ofwel $\lambda = 0,75$ m) met een reikwijdte van 100 km en met een precisie op meterniveau. In tegenstelling tot Syledis heeft het systeem van Omega een zeer grote reikwijdte, want de laagfrequente signalen van dit systeem (10 kHz, ofwel $\lambda = 30$ km) buigen met de aardkromming mee. Zo heeft men met acht grondstations voldoende voor een wereldwijde bedekking. Maar de precisie ligt op kilometerniveau! Met de keuze van de frequentie moet men dus bij de terrestrische plaatsbepalingssystemen een compromis sluiten tussen precisie en reikwijdte.



Figuur 1.1. Meting van de looptijden tussen station A en B op twee manieren.

Radioplaatsbepalingssystemen zijn op de meting van looptijden (of looptijdverschillen) van radiogolven gebaseerd. De looptijd kan men op twee manieren meten: a) enkele looptijd met twee klokken, of b) dubbele looptijd met één klok. Beschouw de punten A en B volgens figuur 1.1a, waarbij de looptijd van een puls wordt gemeten met twee klokken. Vroeger werd hierbij een analoge puls toegepast, later werd de analoge puls door digitale code vervangen. Met een digitale code kan men de tijdmeting met een hogere precisie uitvoeren. De looptijd wordt gedefinieerd als aankomsttijd minus vertrektijd. De afstand tussen A en B kan men berekenen door de gemeten looptijd met de snelheid van de radiogolf (c) te vermenigvuldigen:

$$\overline{AB} = c \cdot \Delta t \quad (1.1a)$$

Het blijkt echter praktisch onmogelijk om de twee klokken precies gelijk te laten lopen

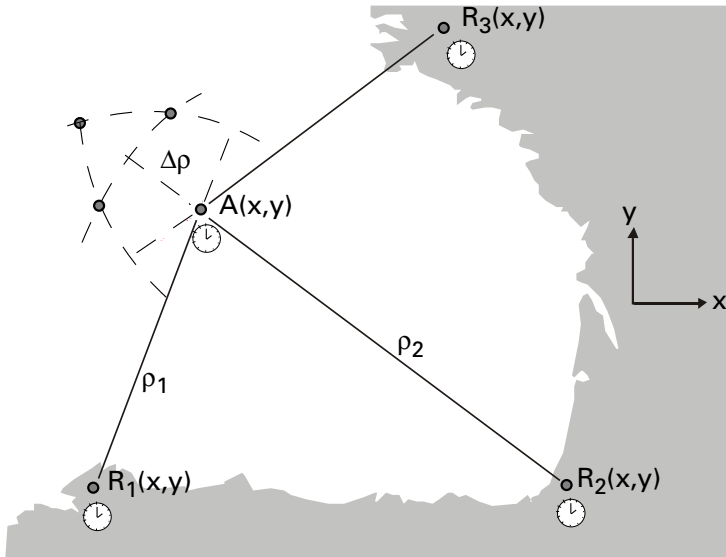
(synchroniseren). Door het verschil tussen twee klokken, de zogenaamde klokfout (δt), krijgt men in (1.1a) een fout in de afstand. Deze foutieve afstand wordt pseudo-afstand genoemd. De pseudo-afstand is in formule uitgedrukt:

$$\overline{AB'} = c \Delta t + c \delta t \quad (1.1b)$$

Bij de methode in figuur 1b wordt het signaal na aankomst in punt B door een elektronische *interrogator* direct teruggestuurd, waarna de aankomsttijd in punt A wordt gemeten. Het resultaat is de dubbele looptijd. De geometrische afstand daaruit is:

$$\overline{AB} = c \cdot \frac{1}{2} \Delta t \quad (1.2)$$

Deze meettechniek heeft het voordeel dat de klokfout geëlimineerd wordt, maar het nadeel dat het aantal gebruikers beperkt is. De interrogator kan slechts een kleine tental gebruikers tegelijkertijd bedienen. Daarom meet men de looptijd bij voorkeur met twee klokken omdat daarbij het aantal gebruikers onbeperkt is.



Figuur 1.2. Terrestrische radioplatsbepaling in de Golf van Biskaje.

In figuur 1.2 wordt een simpel voorbeeld gegeven van een gefingeerd terrestrisch radioplatsbepalingssysteem. Hierbij zijn drie referentiestations op vaste punten opgesteld langs de kust (met bekende coördinaten in een twee-dimensionaal referentiesysteem) en er is een punt A met onbekende coördinaten, bijvoorbeeld een schip. De coördinaten van punt A kan men met een insnijding bepalen uit afstanden die berekend worden uit gemeten looptijden. In de navigatie spreekt men van het snijden van positielijnen. Positielijnen zijn verzameling van punten met gelijke afstanden, in dit geval cirkels.

Stel dat de afstanden uit enkele looptijden worden bepaald waarbij alle klokken perfect gelijk lopen. In dit geval kan men het volgende model van (niet-lineaire) waarnemingsvergelijkingen opstellen:

$$\begin{aligned}
 \rho_1 &= \left[(x_A - x_1)^2 + (y_A - y_1)^2 \right]^{\frac{1}{2}} \\
 \rho_2 &= \left[(x_A - x_2)^2 + (y_A - y_2)^2 \right]^{\frac{1}{2}} \\
 \rho_3 &= \left[(x_A - x_3)^2 + (y_A - y_3)^2 \right]^{\frac{1}{2}}
 \end{aligned} \tag{1.3}$$

Aan de linkerzijde hebben we drie geometrische afstanden. In dit model hebben we drie waarnemingen (ρ_i) en twee onbekenden (x_A, y_A); de overtaligheid is dus $m - n = 3 - 2 = 1$.

In de praktijk is het zeer goed mogelijk de klokken van de referentiestations onderling te synchroniseren, bijvoorbeeld door gebruik te maken van een kabelverbinding of radioverbinding. Echter, de klok van de waarnemer in punt A zal in de regel een afwijking (klokfout) hebben t.o.v. de klokken van de referentiestations. Het gevolg is dat alle op een gegeven moment gemeten pseudo-afstanden dezelfde klokfout zullen hebben en elke combinatie van twee positiecirkels een andere positie zal opleveren (zie figuur 1.2). We krijgen in dit geval drie verschillende posities. Het model van waarnemingsvergelijkingen (1.3) is hierdoor onvolledig. Men spreekt van een modelfout. We stellen een nieuw model op:

$$\begin{aligned}
 P_1 &= \left[(x_A - x_1)^2 + (y_A - y_1)^2 \right]^{\frac{1}{2}} + c \cdot \delta t \\
 P_2 &= \left[(x_A - x_2)^2 + (y_A - y_2)^2 \right]^{\frac{1}{2}} + c \cdot \delta t \\
 P_3 &= \left[(x_A - x_3)^2 + (y_A - y_3)^2 \right]^{\frac{1}{2}} + c \cdot \delta t
 \end{aligned} \tag{1.4}$$

In model (1.4) zijn P_i de gemeten pseudo-afstanden en $c \cdot \delta t$ is de onbekende klokfout, omgerekend in een afstandsmaat. We hebben nu drie waarnemingsvergelijkingen met drie onbekenden.

Zoals eerder gezegd, een hoogfrequent terrestrisch radioplaatsbepalingssysteem heeft een beperkte reikwijdte. Met de komst van de satellieten werd het nadeel van deze beperking drastisch verminderd. Met toepassing van satellieten kan men hoogfrequente radio-signalen toepassen zolang de satellieten “radio” zichtbaar zijn voor de waarnemer, dus als er zich geen obstakels bevinden tussen satellieten en waarnemer, die het signaal afschermen. Het eerste satelliet-systeem, TRANSIT was door de U.S. Marine opgezet en was reeds in 1964 operationeel met 6 satellieten. De ontwikkeling van het Global Positioning System (GPS) is vanaf 1973 van start gegaan en in 1994 volledig operationeel geworden met 24 satellieten. Bijna parallel met GPS werd in de toenmalige Sovjetunie het GLONASS (GLObal NAVigation Satellite System) ontwikkeld. Met de komst van de satellietplaatsbepaling is de noodzaak van terrestrische systemen grotendeels vervallen. Veel terrestrische systemen, o.a. Syledis, zijn niet meer in gebruik.

In figuur 1.2 werd een twee-dimensionaal terrestrisch systeem als voorbeeld gegeven. Dit zou bijvoorbeeld Syledis kunnen zijn die qua frequentie en meettechniek met GPS kan

worden vergeleken. Bij GPS zijn echter de referentiestations in de ruimte geplaatst en worden door satellieten gedragen. De satellieten vliegen met een snelheid van 3,8 km/s rondom de aarde en op een hoogte van 20.000 km boven het aardoppervlak. In totaal zijn er 24 satellieten nodig voor een wereldwijde bedekking. De satellieten zijn geen vaste punten meer, maar dat is geen bezwaar. Want de referentiecoördinaten kan men voor elk tijdstip voldoende nauwkeurig bepalen uit de baanparameters van de satellieten. Deze coördinaten zijn nodig voor het moment waarop het meetsignaal (de digitale code) vanuit de satelliet vertrekt. Hiervoor gebruikt men een drie-dimensionaal coördinaatsysteem. De klokken van de satellieten worden via een telemetrie verbinding door grondstations gesynchroniseerd. Ook hier hebben we te maken met een klokfout tussen de klok van de waarnemer en de klokken in de satellieten. Voor GPS wordt het vergelijkingssysteem (1.3) uitgebreid tot drie dimensies:

$$\begin{aligned}
 P_1 &= \left[(x_A - x_1)^2 + (y_A - y_1)^2 + (z_A - z_1)^2 \right]^{\frac{1}{2}} + c \cdot \delta t \\
 P_2 &= \left[(x_A - x_2)^2 + (y_A - y_2)^2 + (z_A - z_2)^2 \right]^{\frac{1}{2}} + c \cdot \delta t \\
 P_3 &= \left[(x_A - x_3)^2 + (y_A - y_3)^2 + (z_A - z_3)^2 \right]^{\frac{1}{2}} + c \cdot \delta t \\
 P_4 &= \left[(x_A - x_4)^2 + (y_A - y_4)^2 + (z_A - z_4)^2 \right]^{\frac{1}{2}} + c \cdot \delta t
 \end{aligned}
 \tag{1.5}$$

In de waarnemingsvergelijkingen (1.5) worden de pseudo-afstanden (P) gemeten en de coördinaten van de satellieten (x_i, y_i, z_i) als bekend verondersteld. Men gebruikt hiervoor meestal de cartesische x, y, z coördinaten in een *Earth Centered Earth Fixed (ECEF)* systeem. We hebben dan vier onbekenden, de positie van de waarnemer (x_A, y_A, z_A) en de klokfout. Deze onbekenden kunnen we uit minimaal vier satellieten bepalen. Opgemerkt kan tenslotte worden dat in vergelijking (1.5) alle parameters in principe tijdafankelijk zijn. Ook de klokfout, want de klokken hebben een drift, waardoor de stand van de klok verandert. De enige constante term is de snelheid van de radiogolf, hoewel deze (in geringe mate) ook afhankelijk is van een aantal factoren en zo de nauwkeurigheid nadelig beïnvloedt.

Enige eigenschappen van GPS kunnen we, vooruitlopend op de volgende hoofdstukken, als volgt samenvatten:

- bepaling van enkele (absolute) en relatieve 3D posities;
- positiebepaling van een stilstaand punt (statisch) of van een punt in beweging (kinematisch);
- wereldwijd en continu (24 h per dag) beschikbaar;
- passief en vrijwel storingsvrij;
- werkt onder alle weeromstandigheden;
- ongelimiteerd aantal gebruikers;
- militair systeem, met beperkingen voor civiele gebruikers;
- breed toepassingsgebied (o.a. navigatie en geodetische puntsbepaling);
- precisie: absolute positie met pseudo-afstanden: 100m;
- precisie van relatieve positie met code-afstanden: 1m;

- precisie van relatieve positie met fasemeting op centimeterniveau.

In Hoofdstuk 2 wordt een overzicht gegeven van het Global Positioning System.

2 Een overzicht van GPS

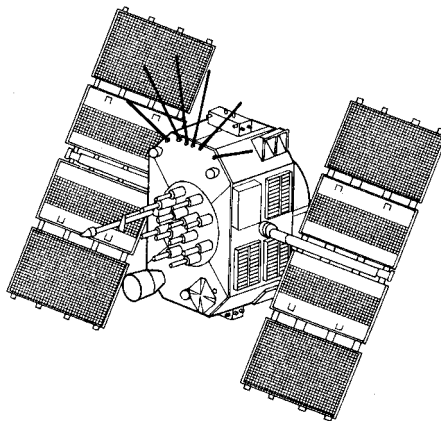
2.1. Algemene kenmerken.

GPS is een plaatsbepalingssysteem dat op afstandsmeting naar satellieten is gebaseerd. De afstanden worden uit gemeten looptijden van radiogolven afgeleid. Het systeem is continu 24 uur per dag in bedrijf, overal ter wereld bruikbaar en werkt onder alle weersomstandigheden. Het is een militair systeem dat door de overheid van de Verenigde Staten kosteloos aan civiele gebruikers ter beschikking wordt gesteld; ook buiten de Verenigde Staten. Voor civiele gebruikers zijn echter twee beperkingen opgelegd die bekend staan onder de naam Selective Availability (SA) en Anti-Spoofing (A-S). Deze beperkingen hebben consequenties voor de nauwkeurigheid van het systeem. De voornaamste componenten van GPS zijn: het ruimtesegment, het besturingssegment en het gebruikerssegment.

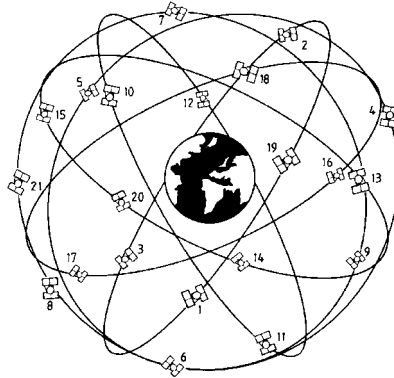
Het *ruimtesegment* bevat nominaal 24 satellieten, waarvan 21 actieve en 3 actieve reserve satellieten. Een GPS satelliet wordt in figuur 2.1 afgebeeld. Een configuratie van 21 satellieten volgens een oude opzet wordt in figuur 2.2a afgebeeld en de verdeling van 27 satellieten in zes baanvlakken (configuratie in 1999) wordt in figuur 2.2b gegeven. De satellieten worden vaak 'Space Vehicles' (SV's) genoemd.

Typische baangegevens zijn:

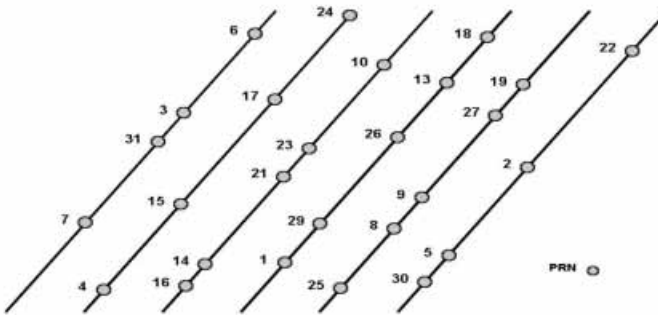
- hoogte van de baan (boven de aardoppervlakte): 20.240 km;
- vorm van de baan: bijna cirkelvormig;
- omlooperperiode (in sterrentijd): 12 uur;
- inclinatie van het baanvlak t.o.v. de equator: $i = 55^\circ$.



Figuur 2.1. GPS satelliet met zonnepanelen en gerichte antenne.



Figuur 2.2a. Configuratie van 21 GPS satellieten volgens een oude opzet.



Figuur 2.2b. Verdeling van 27 GPS satellieten (1999) in zes baanvlakken.

Het *besturingssegment* bevat vijf volgstations: Hawaii, Colorado Springs (master en volgstation), Ascension, Diego Garcia, Kwajalein (figuur 2.3). Deze zijn bij benadering gelijkmatig verdeeld over de aarde dichtbij de evenaar. De volgstations houden het systeem in bedrijf door de klok- en de baanparameters van de satellieten uit eigen metingen te bepalen en deze gegevens geregeld naar de satellieten te zenden.

Met het *gebruikerssegment* wordt de GPS-ontvanger bedoeld. Meer dan honderd typen ontvangers zijn op de markt te koop voor civiele toepassingen. De meest bekende merken voor geodetische doeleinde zijn: Ashtech, Leica, Rogue, Rockwell, Sercel, Trimble.

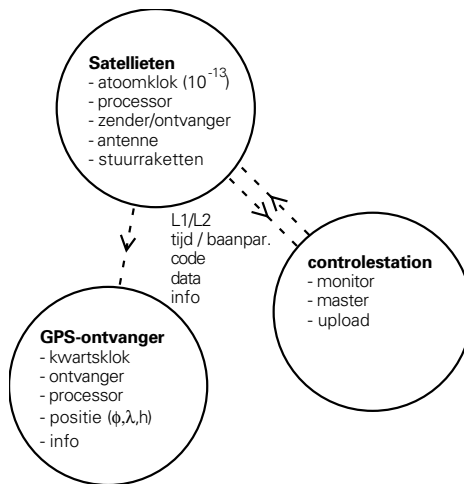
Het aantal GPS satellieten is overigens vastgesteld na een afweging van de kosten tegen de beschikbaarheid. De tegenwoordige configuratie van satellieten is ontstaan door bestudering van een aantal computermodellen. Oorspronkelijk had men voor een gunstige constellatie van 24 satellieten gekozen, maar uit kostenoverwegingen dacht men dit aantal tot 18 terug te kunnen brengen. Bij deze opzet zouden echter te veel hiaten in de positiebepaling optreden. Uiteindelijk werd in 1989 een compromis gesloten door voor een constellatie met 21+3 satellieten te besluiten, d.w.z. 21 satellieten voor normaal gebruik en 3 reserve, maar wel actieve satellieten.

Deze 21+3 constellatie is verdeeld over zes baanvlakken, waarbij in elk baanvlak vier satellieten zijn geplaatst. De zes baanvlakken snijden de equator met gelijke equatoriale

hoekverschillen van 60°. Sinds GPS opgestart is, zijn verschillende typen satellieten in gebruik geweest¹.



Figuur 2.3. Besturingssegment.



Figuur 2.4. Systeemopzet van GPS.

Block-I. Er zijn tien "Navigation Development Satellites" gelanceerd in de jaren 1978-1985. De laatste satelliet van deze serie functioneerde tot 1995.

Block-II (1989-1997). Eind 1998 heeft men meer dan 24 van dit type in gebruik. Bij deze satellieten wordt het luchtweerstand gecompenseerd, ze zijn dus zogenaamd "drag free". Hierbij werd de Selective Availability (1991) en Anti Spoofing (1994) geïntroduceerd.

Block-IIR (1997-2002). Volgens de plannen zullen 21 "cross-link ranging" type satellieten gelanceerd worden (R = Replacement). Dit betekent afstandsmeting en communicatie tussen de satellieten onderling. Hierdoor zijn vanuit de grondstations veel minder metingen nodig en de baanparameters kunnen bovendien nauwkeuriger worden bepaald.

Block-IIF Vanaf 2002 zal men beginnen met lancering van een nieuw type (F = Follow), volledig gemoderniseerde satelliet. In totaal zullen 33 satellieten worden gelanceerd. De

¹ De beschikbare satellieten worden op het Internet vermeld (zie: Internet bij Referenties).

klokstabiliteit wordt bij deze satellieten verbeterd tot 10^{-15} per dag². Dit betekent een afwijking van 0,09 ns per dag ($86400 \text{ seconden} \times 10^{-15} = 0,000\,000\,000\,09 \text{ s}$).

Het principe van plaatsbepaling berust op de meting van enkele looptijden van radiogolven van satelliet naar ontvanger. Deze meettechniek is de standaardtoepassing van GPS en wordt meestal codemeting genoemd. Hierbij worden pseudo random noise (PRN) signalen toegepast. Dit zijn bifase gemoduleerde radiogolven met digitale code, die over een grotere bandbreedte worden uitgesmeerd (bijv. 20 MHz). Hoe groter de bandbreedte des te hogere nauwkeurigheid behaald kan worden. Bij een constellatie van 21+3 satellieten zal een waarnemer overal op aarde meestal 4 tot 8 satellieten in zicht hebben boven een elevatiehoek van 15° . Er worden simultaan (of bij eenvoudige GPS-ontvangers sequentieel, kort achter elkaar) looptijden gemeten naar de beschikbare satellieten.

Voor de meting van enkele looptijden met “onafhankelijke” klokken zijn, vooral tussen de satellieten onderling, nauwkeurig gesynchroniseerde klokken vereist. Hierbij wordt voor de stabiliteit van de klok minstens 10^{-13} over 24 h geëist. Dit betekent een afwijking van 9 ns (0,000000009 sec) per dag, in afstandsmaat bijna 3m. Een dergelijke nauwkeurigheid is alleen haalbaar met atoomklokken. Bij de ontvanger kan men meestal volstaan met een kwartsklok, omdat de hoge stabiliteit daar slechts over een korte periode is vereist. Deze looptijden worden echter systematisch beïnvloed door een klokfout Δt tengevolge van de afwijking (synchronisatie-fout) tussen de klok van de ontvanger en het GPS tijdsysteem. Omdat alle klokken een drift hebben, heeft ook de klokfout een drift. Onder klokdrift wordt verstaan het verlopen van de klokfout in tijd. De klokdrift wordt voornamelijk veroorzaakt door de minder nauwkeurige klok in de ontvanger.

De gemeten looptijden worden vervolgens omgerekend in afstanden door vermenigvuldiging met de voortplantingsnelheid van radiogolven. Door de klokfout hebben alle afstanden op een gegeven moment ook een zelfde afwijking, vandaar de naam “pseudo”-afstand. Het zal duidelijk zijn dat ook de afwijking in de afstanden onderhevig is aan drift. Uit de pseudo-afstanden en de corresponderende satellietcoördinaten (berekend uit de door de satellieten uitgezonden baanparameters) kan men vier onbekenden bepalen, nl. de driedimensionale positie van de waarnemer (X, Y, Z of φ, λ, h), en de onbekende klokfout ΔT tussen de onderling gesynchroniseerde klokken van de satellieten en de klok van de ontvanger. De oplossing is het snijpunt van drie bollen, nadat de stralen met dezelfde lengte ΔR zijn gecorrigeerd. Het mathematisch model wordt in Hoofdstuk 5 in detail beschreven.

² De hierna gegeven details over de signalen en frequenties hebben betrekking op de satellieten die momenteel (1999) operationeel zijn. Voor Block-IIIF gelden andere details, hoewel voor “backward compatibility” ook de oude signaalstructuur nog lange tijd wordt uitgezonden, naast de nieuwe.

2.2. GPS satelliet signaal. De uitgezonden datablokken

De basisfrequentie van de atoomklok in de satelliet bedraagt 10,23 MHz. Hieruit worden twee draaggolven afgeleid door de basisfrequentie met een constante factor te vermenigvuldigen. Zo ontstaan de L1 en L2 banden met twee verschillende frequenties:

$$L1: f_1 = 154 \times 10,23 \text{ MHz} = 1575,42 \text{ MHz}$$

$$L2: f_2 = 120 \times 10,23 \text{ MHz} = 1227,60 \text{ MHz}$$

Deze frequenties komen overeen met golflengten van resp. 19,05 cm en 24,45 cm. Men heeft bewust voor deze hoge frequentie L-banden gekozen omdat deze frequenties goede eigenschappen hebben ten aanzien van het doordringen door de ionosfeer en de troposfeer. De twee draaggolven worden op diverse manieren gemoduleerd. De modulatie wordt als volgt voorgesteld:

$$\begin{aligned} L1(t) &= a_1 P(t) W(t) D(t) \cos(f_1 t) + a_1 C(t) D(t) \sin(f_1 t) \\ L2(t) &= a_2 P(t) W(t) D(t) \cos(f_2 t) \end{aligned} \quad (2.1)$$

waarin:

a = amplitude

P = P-code (Precision-code)

W = W-code om de P-code te versleutelen: $P + W = Y$ -code (Anti-Spoofing)

C = C/A-code (= Clear Acquisition, ook wel eens Civil Access code genoemd)

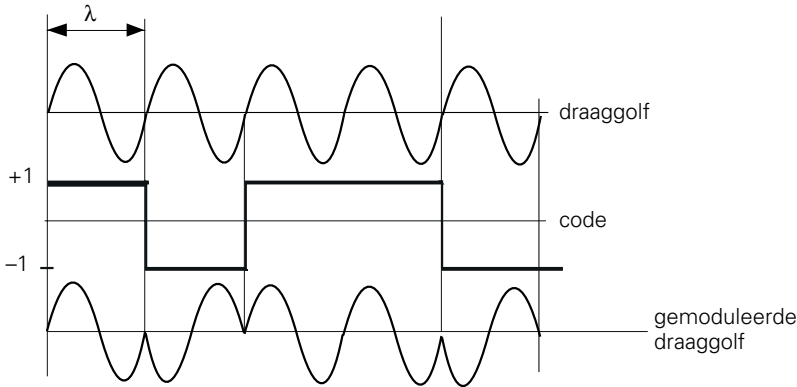
D = Data (modulatie t.b.v. gegevensoverdracht met een snelheid van 50 bits/s)

f = frequentie

t = tijdschaal.

Ten behoeve van de meting van de looptijden worden twee digitale codes uitgezonden: de C/A-code en de P-code. Beide codes bevatten een reeks nullen en enen met specifieke eigenschappen die m.b.v. een codegenerator worden opgewekt. De nullen en enen worden door middel van bifasemodulatie van de draaggolf doorgegeven. Bifasemodulatie betekent het omschakelen van de fase met 180° in een bepaald basisritme. Het basisritme heet *chiprate* en de kleinste eenheid hiervan is de *chiplengte*. Op deze wijze worden de code en de data op de draaggolf meegestuurd. Het principe van bifasemodulatie wordt in figuur 2.5 voorgesteld.

In tabel 2.1 worden enige specificaties van de codes vermeld die met bifasemodulatie worden doorgegeven. De C/A-code (of C1-code) wordt alleen op de L1-band uitgezonden, waarbij elke satelliet zijn eigen unieke code heeft. (Deze code is vergelijkbaar met de “ouderwetse” analoge puls die vroeger bij terrestrische radioplotsbepaling werd toegepast.) De codelengte is 1 ms, wat equivalent is met een afstand van 300 km. Deze code wordt iedere msec herhaald. Bij een afstand van 24.000 km tot een satelliet zijn er 80 gelijke digitale stukjes op een gegeven moment onderweg. Met behulp van deze digitale code wordt de looptijd als verschil tussen de aankomsttijd en vertrektijd bepaald.



Figuur 2.5. Bifasemodulatie van het GPS signaal.

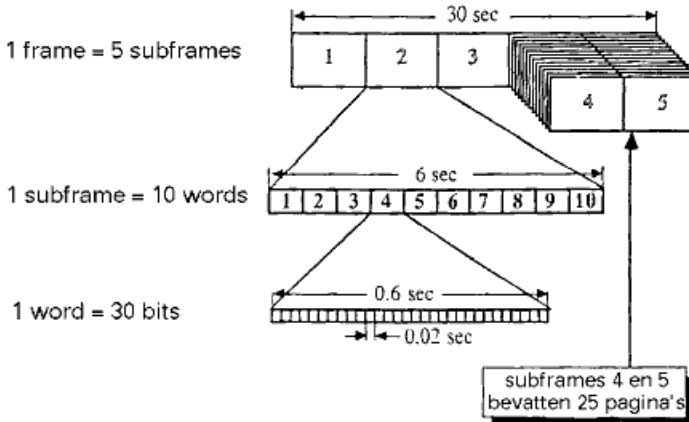
Tabel 2.1. Specificatie van het GPS signaal.

	C/A code (op L1)	P-code (op L1 en L2)
Bandbreedte (MHz)	2	20
Chirate (MHz)	1.023	10.23 (= basis frequentie)
Chiplengte (T_p)	1000 ns = 300 m	100 ns = 30 m
Resolutie < 1% van T_p	3 m	0.3 m
Codelengte	1023 bits = 1 ms	38 weken = 267 dagen

Deze looptijd is eenduidig want elk stukje code bevat informatie waaruit de vertrektijd af te leiden is. De C/A-code is vooral bedoeld voor plaatsbepaling t.b.v. civiele toepassingen. Hierbij bedraagt de resolutie (meetnauwkeurigheid) in de afstandsmeting 3 m.

De P-code (Precision-code) wordt op de L1 en L2 frequentie op identieke wijze uitgezonden. Ter onderscheiding spreekt men over P1-code en P2-code. De P-code bevat $2.36 \cdot 10^{14}$ bits, met een totale lengte van 38 weken (267 dagen). Elke satelliet heeft daaruit een eigen stuk van een week, men kan aan 37 satellieten een eigen unieke code toewijzen. De P-code start telkens opnieuw zondags om 0:00 h GPS-tijd. Het kraken van deze code is praktisch onmogelijk. Met deze code kan men een hogere nauwkeurigheid dan met de C/A-code bereiken op twee redenen: a. door de hogere resolutie (< 0.3m); b. doordat hij op beide frequenties wordt uitgezonden, waardoor de invloed van de ionosfeer geëlimineerd kan worden. De P-code is in principe voor alle gebruikers beschikbaar, maar kan omgeschakeld worden naar een gelijksoortige Y-code, die alleen voor geselecteerde (militaire) gebruikers beschikbaar is. (Ook hier ontstaan twee identieke codes, de Y1-code en de Y2-code, respectievelijk op de L1 en L2 frequenties). Deze bescherming wordt Anti-Spoofing (A-S) genoemd. Omdat A-S sinds 1994 is aangezet, kunnen civiele GPS-ontvangers alleen met C/A-code werken. We komen hierop nog terug in § 2.4.

De zogenaamde navigation message wordt door iedere GPS satelliet in perioden van 30 seconden (frames) op beide frequenties uitgezonden. Zo'n periode wordt in 5 gelijke datablokken (subframes) van 6 seconden onderverdeeld (zie figuur 2.6).



Figuur 2.6. Structuur van doorgeven van boodschappen (Navigation message).

Tabel 2.2. De uitgezonden 5 datablokken in 30 sec (= 5 x 6 sec).

Subframe	word #1	word #2	Word #3 - #10	Aantal bits
1	TLM	HOW	klokparameters	300
2	TLM	HOW	ephemeris(1)	300
3	TLM	HOW	ephemeris(2)	300
4	TLM	HOW	message,UTC/ionosfeer parameters	300
5	TLM	HOW	almanak	300
totaal:				1500

De vijf datablokken, die door elke satelliet worden uitgezonden, zijn in tabel 2.2 schematisch aangegeven. Ieder subframe bevat 10 woorden van 30 bits, totaal 300 bits. In een periode van 30 sec worden dus totaal 1500 bits uitgezonden. Het eerste woord is een TeLeMetrie (= TLM) woord t.b.v. synchronisatie gevolgd door een Hand-Over-Word (= HOW). Uit de HOW is af te leiden de verstreken tijd sinds de laatste zaterdag/zondag nacht 0^h UTC tot het moment waarop het eerstvolgende datablok begint. Deze tijd wordt Time-Of-Week (TOW) genoemd die dikwijls als GPS-tijd wordt aangeduid.

De *klokparameters* in subframe 1 worden gebruikt voor de onderlinge synchronisatie van de satellietklokken. Elke satellietklok heeft namelijk zijn eigen tijd die naar GPS-tijd dient te worden gecorrigeerd m.b.v. drie klokparameters (a_0, a_1, a_2), die gelden op een epoche t_{0c} . De klokfout kan hieruit voor een gegeven epoche (t) als volgt worden berekend:

$$\delta t^s(t) = a_0 + a_1(t - t_{0c}) + a_2(t - t_{0c})^2$$

De *ephemeris* in subframe 2 en 3 bevat de Keplerse baanparameters (efemeriden) en een aantal stoorparameters. De satellieten zenden ieder uur nieuwe baanparameters uit. Deze parameters worden gebruikt voor de berekening van de satellietposities voor elk gewenst tijdstip. De formules voor de berekening worden in Hoofdstuk 3 gegeven. Voor de actuele

baanparameters heeft de satelliet een geheugencapaciteit van 28 uur. Als binnen deze periode geen nieuwe gegevens worden door de grondstations toegezonden dan worden “oude” parameters gebruikt. In dit geval krijgt men steeds minder nauwkeurige satellietposities.

De *message* in subframe 4 bevat 25 pagina's met veel nuttige informatie, waarvan lang niet alles beschikbaar wordt gesteld aan civiele gebruikers. De belangrijkste informatie wordt verstrekt in pagina 17: special message en in pagina 18: ionosfeermodel en UTC data.

De *almanac* in subframe 5 geeft de benaderde baanparameters van alle satellieten (in 24 pagina's) en de toestand (health) van deze satellieten in de 25e pagina. De almanak wordt in de eerste plaats door de ontvanger gebruikt om de satellieten bij een koude start op de juiste frequentie te kunnen zoeken. De frequentie van de satellieten is namelijk bij ontvangst verschillend i.v.m. het dopplereffect. Voor de bepaling van deze frequentie heeft men de benaderde positie van de satellieten nodig. De almanak kan ook (door middel van postprocessing) voor planning van metingen worden gebruikt.

2.3. Decodering van de data

De door de satelliet uitgezonden data, een lange reeks van bits, wordt door de GPS-ontvanger verwerkt en gebruikt voor de positie berekening, maar wordt ook via een RS232 poort aan de gebruiker doorgegeven. Aan deze poort kan een personal computer worden aangesloten, waarmee de data m.b.v. een log-programma wordt uitgelezen t.b.v. een directe verwerking (in real-time), of opgeslagen voor latere verwerking (post-processing). Bij vele type ontvangers wordt ook de mogelijkheid geboden de data eerst in het interne geheugen op te slaan en later door middel van een PC uit te lezen (“downloaden”).

Deze data bevat o.a. de volgende gegevens:

- a. baanparameters (navigation message);
- b. waarnemingen (tijd, code-afstanden en fase-afstanden);
- c. posities (berekend door de ontvanger in real-time).

De navigation message, zoals eerder besproken, is afkomstig uit de satellieten en moet dus bij iedere type ontvanger dezelfde zijn. De overige blokken zijn echter ontvanger-afhankelijk, want de wijze waarop een ontvanger de metingen uitvoert en berekent, kan verschillend zijn.

Het oudste type GPS-ontvanger TI-4100 (van Texas Instrument uit 1984) leverde de data in de meest ruwe vorm, namelijk in bit format, die de gebruiker zelf moest decoderen. Dit komt er op neer dat er een aantal bytes, in de vorm van 8 bits getallen (in ASCII tussen 0 en 255), volgens bepaalde regels samengesteld moet worden, bijvoorbeeld twee of meer bytes aan elkaar gekoppeld en vervolgens vermenigvuldigd met een gegeven factor. Zo krijgt men zinvolle getallen. De tegenwoordige ontvangers voeren de decodering meteen uit en leveren de datablokken in een compacte vorm. Dit bespaart een hoop rekenwerk en er is veel minder opslagruimte nodig voor de data.

2.4. Overzicht van enkele basisbegrippen

Fysische begrippen

Het meetsysteem van GPS is gebaseerd op de voortplanting van radiogolven van satelliet naar ontvanger, waarvan de voortplantingssnelheid in vacuüm is aangenomen als: $c = 299792458 \text{ m/s} \approx 3 \cdot 10^8 \text{ m/s}$. Deze voortplantingssnelheid heeft de volgende relatie met de frequentie (f), de golflengte (λ) en de periode (T):

$$c = f \cdot \lambda = \frac{2\pi}{T} \lambda \quad (2.2)$$

Het is mogelijk om zowel code-modulatie als de fase van de draaggolf te gebruiken voor de looptijdmeting (en daaruit de afstand te bepalen). Voor wat betreft de fase geldt het volgende. In een locatie neemt in een periode T toe van 0 tot 2π , dus ook een tijdsinterval t verandert de fase met $\varphi = \frac{t}{T} 2\pi = t f$, waarin f is het aantal perioden van 2π per seconde. Men kan dus schrijven:

$$\frac{d\varphi}{dt} = f \quad (2.3)$$

Een signaal dat de satelliet met $\varphi = 0$ verlaat komt bij de ontvanger op een afstand ρ aan (eveneens met $\varphi = 0$) op $t_\rho = \frac{\rho}{c}$. Op een willekeurig later tijdstip t is de fase:

$$\varphi_t = \int_{t_\rho}^t f \cdot dt = (t - t_\rho) f = \left(t - \frac{\rho}{c}\right) f \quad (2.4)$$

Dit is de fasevergelijking van de radiogolven tussen een zender en een ontvanger. Een probleem bij fasemeting is dat alle perioden identiek zijn. We kunnen wel de fracties van de perioden meten, maar niet het aantal volledige perioden. Deze zg. meerduidigheid moet door een berekeningsmethode worden opgelost.

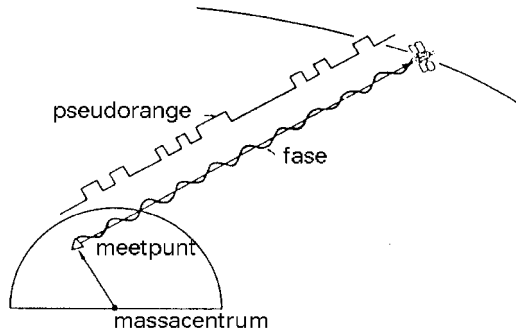
De meetbare grootheden

In het voorgaande hebben we de drie vitale componenten van GPS behandeld: a) het ruimte segment, b) het besturingssegment en c) het gebruikers segment (zie: figuur 2.4). De gebruikers van het systeem hebben geen invloed op a) en b). Ze mogen alleen maar hopen dat het systeem goed wordt onderhouden en dat de uitgezonden signalen voldoen aan de verwachtingen. De civiele gebruiker schaft een GPS-ontvanger aan (er is voldoende keuze voor alle doeleinden) en maakt kosteloos gebruik van het systeem.

Een GPS-ontvanger kan in principe met behulp van de satelliet signalen verschillende soorten grootheden meten, zoals:

- code-afstand (C/A of C1-code, P1-code en P2-code)
- fase-afstand (op L1 en L2)
- frequentieverschuiving (dopplerverschuiving).

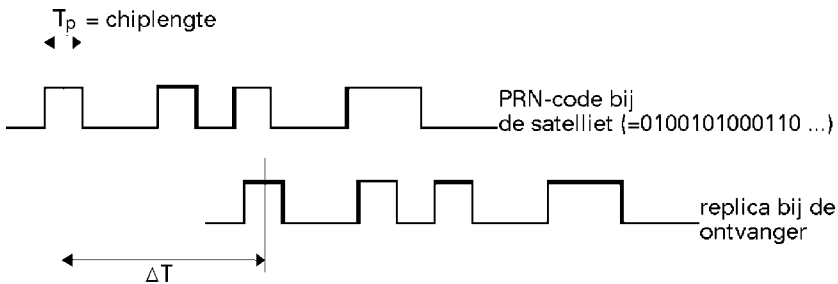
Het kan in het algemeen worden gesteld dat de pseudo-afstanden vooral bij navigatie worden gebruikt, waarbij het meterniveau voor de positiebepaling voldoende is, en de fase-afstanden bij nauwkeurige geodetische toepassingen, waarbij het centimeterniveau vereist is. In figuur 2.7 worden deze twee meetmethoden geschetst.



Figuur 2.7. Code-afstand en fasewaarneming.

Meting van code-afstanden

De looptijd van de digitale code van satelliet tot ontvanger wordt met correlatietechnieken bepaald. Hiervoor moet de ontvanger een replica van de binaire PRN-code genereren die door de satellieten wordt uitgezonden (het algoritme van de PRN-code moet dus in de ontvanger bekend zijn). Het principe van de meting wordt in figuur 2.8 voorgesteld.



Figuur 2.8. Meting van de aankomsttijd van het satelliet signaal.

De structuur van de toegepaste PRN-codes (voor C/A de zogenaamde Gold-code) is zodanig dat er een maximale correlatie wordt verkregen als de twee codes precies op elkaar passen en dat de correlatie tussen de verschillende PRN-codes erg klein is. Door een stapsgewijze verschuiving van de codes, totdat de maximale correlatie is gevonden, wordt de aankomsttijd van het signaal op de klok van de ontvanger gemeten. De looptijd wordt als aankomsttijd (op de ontvangerklok) minus vertrektijd (op de satellietklok) bepaald. De nauwkeurigheid van de looptijd hangt hoofdzakelijk van twee factoren af: de chip lengte en de precisie waarmee de maximale correlatie wordt gevonden. Voor de laatste wordt meestal uitgegaan van een precisie van 1% van de chip lengte. We krijgen hiermee voor de gemeten pseudo-afstand ruwweg de volgende interne precisie:

C/A-code waarnemingen: 10 ns = 3 m

P-code waarnemingen: 1 ns = 30 cm

Het blijkt dat een interne precisie van de code-afstand bij nieuwe types geodetische ontvangers van 0.1% van de chip lengte mogelijk is. Dit geeft voor de C/A-code een *interne precisie* van enkele decimeters! Maar het is een interne precisie, want de werkelijke precisie is veel slechter i.v.m. de Selective Availability. Een groot voordeel van de code-afstand is, in tegenstelling tot de fase-afstand, dat het eenduidig is.

Fasemeting

De fasemeting op de L1 en L2 draaggolven wordt bij de geodetische ontvangers toegepast. Voordat de fasemeting in de ontvanger kan plaats vinden dient de draaggolf te worden hersteld. Herstel betekent het verwijderen van de code en de data (na de verwerking) om zo een sinus-vormig signaal terug te krijgen. Door de gescheiden modulatie van de C/A-code en de P-code op de L1 draaggolf, kan de L1 draaggolf in de ontvanger wat gemakkelijker worden hersteld, ook zonder kennis van de P-code (of Y-code). Herstel van de L2 draaggolf is daarentegen veel lastiger als de Y-code (de versleutelde P-code) niet bekend is.

De fase-afstand kan met een hoge precisie van ongeveer 3 mm worden gemeten, veel nauwkeuriger dus dan de code-afstand. De fase-afstand bevat evenwel bij elke satelliet een onbekende meerduidigheid (d.w.z. een aantal onbekende golflengten) die in de data-verwerking moet worden bepaald (zie § 6.6).

Frequentieverschuiving (dopplereffect)

De GPS satellieten zenden twee draaggolven uit van constante frequentie. Het ontvangen ("received") signaal ondergaat een frequentieverschuiving als gevolg van het dopplereffect. Voor de frequentieverschuiving (Δf) geldt (bij benadering):

$$\Delta f = (f_r - f_s) = -\frac{1}{c} f_s v_p \quad (2.5)$$

waarin:

f_s = de uitgezonden (constante) frequentie,

f_r = de ontvangen (variabele) frequentie, en

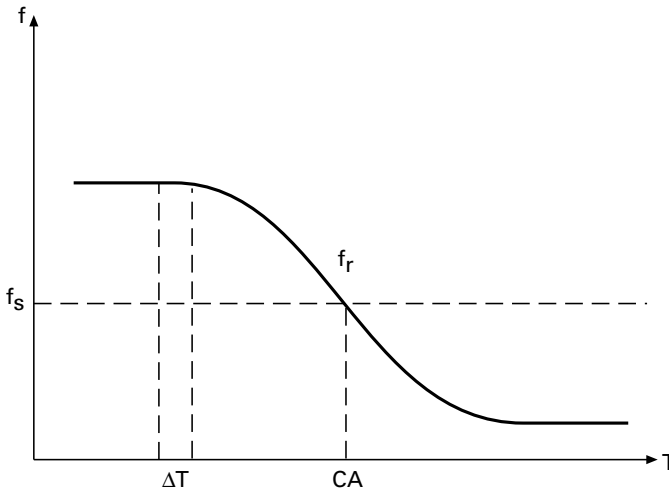
v_p = de snelheidscomponent in radiale richting, d.w.z de afstandsverandering in de afstand tussen de satelliet en de waarnemer.

De frequentie van het ontvangen signaal is dus hoger dan de uitgezonden frequentie als de satelliet nadert, en lager als de satelliet zich verwijdert. Het omslagpunt ligt bij de kortste afstand (c.a. = closest approach), waarbij de radiale snelheid en dus ook de frequentieverschuiving nul zijn (zie figuur 2.9).

Selective Availability (SA)

Bij alle Block-II satellieten vindt een degradatie van de meetprecisie plaats die speciaal voor niet-militaire gebruikers van GPS is bedoeld. Deze degradatie wordt Selective Availability (SA) genoemd. De degradatie wordt toegepast op de precisie van de satellietklok (van 10^{-13} naar 10^{-9}) en mogelijk ook op de precisie van de baanparameters.

SA is zodanig opgezet dat bij een enkele positiebepaling (*single point positioning*) met C/A-code, 95% van de horizontale positiefouten kleiner zijn dan 100 meter. Voor de verticale component van de positie (de hoogte boven de ellipsoïde) geldt 156 meter. De hoogtecomponent is dus altijd slechter. De meetprecisie van een enkele code-afstand is ongeveer 25 meter.



Figuur 2.9. Voorstelling van het dopplereffect. (CA = closest approach, het omslagpunt bij de kortste afstand).

Sinds 1991 is SA door het besturingssegment daadwerkelijk aangezet. Echter, met toepassing van differentiële technieken kan het SA-effect geëlimineerd worden. Hierbij wordt gebruik gemaakt van een referentieontvanger die op een bekend punt wordt opgesteld en in feite de invloed van SA kan bepalen (zie § 5.8). Vermoedelijk zal de toepassing van SA omstreeks het jaar 2006 verdwijnen.

Anti-Spoofing (A-S)

Hieronder verstaat met het omzetten van de bekende P-code (P1 en P2) in een geheime Y-code (Y1 en Y2). Het voornaamste doel hiervan is de immuniteit van het systeem voor moedwillige radiostoringen (“jamming”) te verhogen. Het heeft enige nadelige gevolgen voor civiele gebruikers, vooral omdat de P-code niet meer beschikbaar is en daardoor het meten van code-afstanden op twee frequenties onmogelijk wordt.

GPS-ontvangers

Er zijn meer dan honderd verschillende types GPS-ontvangers op de markt in alle prijsklassen en met verschillende mogelijkheden en doelstellingen (navigatie, geodetische toepassingen, tijdsoverdracht, etc). De belangrijkste kenmerken van een GPS-ontvanger zijn het aantal frequenties (L1 of L1 en L2), het aantal kanalen (geeft aan hoeveel satellieten er gelijktijdig kunnen worden ontvangen) en de typen waarnemingsgrootheden. De meest verkochte geodetische ontvangers hebben tegenwoordig twee frequenties (L1 en L2) en 9 of 12 kanalen. Voorbeelden: Ashtech (12 kan.), Leica (9 kan), TurboRogue (9 of

12 kan.), Trimble SSI (9 of 12 kan.). Deze ontvangers kunnen in principe 5 verschillende grootheden meten en registreren, nl. drie pseudo-afstanden: C1, P1, P2 en twee fasewaarnemingen op L1 en L2. (Sommige ontvangers meten ook de dopplerverschuiving op L1 en L2). Als echter Anti-Spoofing aangezet is en dus de P-code vervangen wordt door een geheime Y-code, dan kunnen geen pseudo-afstanden met P1 en P2-code worden gemeten. Geodetische ontvangers hebben ook nog extra opties voor speciale toepassingen, bijv. voor Real-Time Kinematic (RTK).

De ontwikkeling van de ontvangers is op het gebied van fasemeting nog gaande en de fabrikanten passen verschillende methoden toe om de consequenties van A-S te omzeilen. Er zijn op dit moment vier principiële oplossingen voor de code en de fasemeting op L2 bekend:

1. kwadrateren van het signaal (deze techniek werd vroeger bij de Trimble SSE toegepast);
2. code-kwadrateren: eerst code-correlatie en dan kwadrateren (Leica);
3. kruis-correlatie (TurboRogue, Trimble SSI);
4. Z-tracking (Ashtech Z-12).

De vier methoden worden met elkaar in tabel 2.3 vergeleken. De methode van kwadrateren is inmiddels door de technische ontwikkeling achterhaald. Nadelen waren de halve golflengte (i.v.m. de verdubbeling van de frequentie) en het aanzienlijke verlies aan signaalsterkte (-30 dB).³ Dit gaf vooral veel problemen bij de meting van satellieten met een lage elevatiehoek. Bij code-kwadrateren is het verlies aan signaalsterkte kleiner, maar ook hier wordt de frequentie verdubbeld en dus de golflengte gehalveerd. De halve golflengte gaf problemen bij het oplossen van meerduidigheden⁴.

Tabel 2.3. Code en fasemeting op L2 (Y-code onder A-S).

parameters:	C/A-code	Y-code	golflengte	SNR
kwadrateren	Nee	Y2	half (9,5cm)	-30 dB
code-kwadrateren	Ja	Y2	half	-17 dB
kruis-correlatie	Ja	Y2-Y1	vol (19cm)	-27 dB
Ashtech	Ja	Y1&Y2	vol	-14 dB

De kruis-correlatiemethode bepaalt het verschil in aankomsttijd van twee identieke signalen. Men maakt gebruik van het feit dat de P-codes en ook de Y-codes bij een satelliet op beide frequenties exact gelijk zijn. Op deze wijze kan het verschil in aankomsttijd worden gemeten, ongeacht of de code bekend is of niet. De Y-code wordt namelijk uit de bekende P-code afgeleid door een onbekende W-code toe te voegen: $Y1 = P1 + W$ en $Y2 = P2 + W$. De verschiltijd is dus:

$$X = P2 - P1 = Y2 - Y1 \quad (2.7)$$

Men kan uit deze grootheid een soort fictieve pseudo-afstand voor de C/A-L2-code berekenen:

³ De kolom SRN geeft het verlies in Signal to Noise Ratio aan t.o.v. het zonder A-S beschikbare signaal.

⁴ Zie Hoofdstuk 6 voor een nadere uitleg over het oplossen van meerduidigheden.

$$C2 = C1 + X \quad (2.8)$$

Het ruisniveau van C2 wordt echter aanzienlijk hoger, omdat de kruis-correlatie een vrij hoog ruisniveau heeft. Uit de kruis-correlatie kan men in principe de ionosferische correctie berekenen omdat de grootte X het looptijdverschil is tussen de signalen op L1 en L2 frequentie.

De Ashtech Z-12 ontvanger werkt met een zg. Z-tracking, waarbij de A-S weinig invloed heeft op de waarnemingsgrootheden. Volgens Ashtech werkt de Z-12 ontvanger onder alle omstandigheden als een P-code ontvanger, afgezien van een gering verlies aan signaalsterkte.

Invloed van de atmosfeer

Bij GPS meet men de looptijd van radiogolven, zowel bij codemeting als bij fasemeting. Om de looptijd in een afstand om te zetten moet men de voortplantingssnelheid van de radiogolf weten. In vacuüm is deze goed bekend ($c = 299792458$ m/s) maar op weg naar de ontvanger passeren de signalen de atmosfeer waarbij de voortplantingsnelheid vertraagd wordt. Twee lagen van de atmosfeer veroorzaken een significante vertraging: de ionosfeer (op de hoogte tussen 70 - 1000 km) en de troposfeer (0 - 40 km). De effecten van deze twee lagen zijn verschillend.

Ionosfeer. Het GPS signaal bereikt eerst de ionosfeer en komt in botsing met de elektrisch geladen deeltjes daarin. Gevolg is dat het signaal wordt vertraagd. De mate van vertraging hangt van een aantal factoren af:

- a. de hoeveelheid aanwezige elektronen (hangt weer af van de stand van de zon die de deeltjes in de ionosfeer ioniseert);
- b. de elevatiehoek van de satelliet;
- c. de frequentie van het radiosignaal. Dit geeft een goede mogelijkheid om de ionosferische vertraging te schatten door gebruik van twee verschillende frequenties. De ionosferische vertraging is vooral overdag significant en 's nachts vrijwel nihil. De orde van grootte bedraagt 0 - 20 meter.

Troposfeer. De vertraging door de troposfeer hangt af van de luchtdruk, temperatuur en de vochtigheid van de luchtlagen die door het signaal worden gepasseerd. Het probleem is dat de meteorologische waarden alleen bij de ontvanger eenvoudig kunnen worden gemeten. De luchtdruk is hierbij niet zo'n probleem omdat de op de grond gemeten luchtdruk een maat is voor de hoeveelheid lucht moleculen die het signaal passeert. Maar de vochtigheid kan erg variëren als bijvoorbeeld het signaal een wolk of een gebied boven watermassa's passeert. Er zijn diverse modellen ontwikkeld om de troposferische vertraging te bepalen. De frequentie van het signaal heeft geen invloed op de troposferische vertraging, zodat het meten op twee frequenties geen oplossing biedt.

De atmosferische effecten worden in Hoofdstuk 4 in meer detail behandeld.

Meetmethoden van GPS

Afhankelijk van de doelstelling van de plaatsbepaling met GPS zijn diverse meetmethoden bekend. Men kan de volgende indeling maken:

- codemeting versus fasemeting;
- Absolute plaatsbepaling (enkel punt) versus relatieve plaatsbepaling (twee of meer punten);
- Statische meting versus kinematische meting;
- Postprocessing (nabewerking) versus real-time processing (momentane bewerking).

Codemeting

Enkel puntsbepaling of plaatsbepaling (in absolute zin) wordt meestal t.b.v. navigatie met één ontvanger uitgevoerd. Hierbij wordt codemeting toegepast of soms codemeting met fasemeting versterkt. De nauwkeurigheid is 100 m horizontaal en 156 m vertikaal op basis van 95% kansniveau, voornamelijk als gevolg van de Selective Availability.

Indien een hogere nauwkeurigheid wordt verlangd kan men pseudo-afstandcorrecties toepassen die op een referentiestation worden bepaald. Dit is een relatieve plaatsbepaling en men spreekt in dit geval van Differentiële GPS of kortweg DGPS. Diverse systematische fouten (o.a. de invloed van SA maar ook gedeeltelijk de atmosferische vertraging) kunnen op deze wijze worden geëlimineerd waardoor de nauwkeurigheid tot meterniveau kan worden verbeterd.

Codemeting kan statisch (stilstaand) of kinematisch (in beweging) toegepast worden en de berekening met postprocessing of in real-time uitgevoerd worden. De codemeting wordt in hoofdstuk 5 behandeld.

Fasemeting

Fasemeting wordt t.b.v. geodetische puntsbepaling toegepast waarbij een hoge nauwkeurigheid op centimeterniveau wordt gevraagd. Deze methode wordt alleen relatief, dus met gebruik van een referentiestation toegepast. Geodetische puntsbepaling met fasemeting kan statisch of kinematisch worden toegepast.

De statische methode heeft diverse verbeterde vormen zoals snelstatisch (rapid static of fast static) en semi-kinematic (stop and go). Een belangrijke aspect is hierbij de snelle oplossing van de meerduidigheden in de fase-afstanden. Hiervoor is de methode “On The Fly” (OTF) of “On The Move” ontwikkeld die de meerduidigheden snel kan oplossen, zelfs wanneer de ontvanger in beweging is.

De kinematische methode in real-time noemt men Real-Time Kinematic (RTK). Hierbij worden alle waarnemingen van het referentiestation via een radioverbinding (of telefoonverbinding) naar de mobiele ontvanger doorgegeven en de relatieve positie wordt door de mobiele ontvanger berekend. Kenmerkend zijn de efficiëntie en de directe beschikbaarheid van de coördinaten.

De continue kinematische (true kinematic) methode is vooral van belang voor de bepaling van nauwkeurige posities van schepen en van vliegtuigen die remote sensing meetinstrumenten meevoeren.

De mathematische modellen van de fasemeting worden in Hoofdstuk 6 behandeld.

Multipad

In de plaatsbepaling wordt aangenomen dat het signaal tussen de satelliet en de ontvanger de kortste weg aflegt. Er kunnen echter reflecties optreden die door gebouwen en andere

objecten in de buurt van de ontvanger worden veroorzaakt. Hierdoor zullen een direct signaal en een gereflecteerd signaal de ontvanger bereiken. Dit heet het multipadefect; het leidt altijd tot nauwkeurighedsverlies. Het multipadefect kan men voorkomen door de antenne in een reflectievrije omgeving op te stellen. Sommige ontvangers proberen het effect in real-time te detecteren en onschadelijk te maken.

Software

Naast de keuze voor een bepaalde type GPS-ontvanger, heeft de software waarmee de postprocessing wordt uitgevoerd een belangrijke invloed op de precisie en betrouwbaarheid van de plaatsbepaling. De ontwikkeling van GPS-software is een lange weg die vermoedelijk nooit zal ophouden. De fabrikanten van de ontvangers bieden meestal commerciële software aan waarmee vele soorten metingen kunnen worden verwerkt. Deze software voldoet in de praktijk in het algemeen goed. Voor wetenschappelijke doeleinden wordt vaak van speciale software gebruik gemaakt met uitgebreidere mogelijkheden.

Foutenbronnen

De nauwkeurigheid van de GPS positie uit code-afstanden en fase-afstanden wordt door een aantal factoren beïnvloed. De voornaamste foutenbronnen zijn:

- satellietbanen;
- klokparameters;
- atmosfeer (ionosfeer en troposfeer);
- antenne (fasecentrum en multipad);
- configuratie van de satellieten t.o.v. de geografische positie van de ontvanger;
- fasemeerduidigheden;
- Selective Availability.

In tabel 2.4 worden enige a priori gegevens verstrekt van de bijdragen van verschillende foutenbronnen, zowel voor C/A-code als voor P-code. Deze gegevens zijn ontleend aan “NAVSTAR GPS User Equipment, Public Release, February 1991”. Het totale effect van respectievelijk 10,8-13,9m en 6,6m werd verkregen door kwadratische sommatie van alle invloeden (voortplantingswet van varianties), zonder rekening te houden met een eventuele correlatie. Deze gegevens beschouwen we als a-priori standaardafwijkingen van de pseudo-afstand waarmee de precisie van de positiebepaling geschat kan worden.

Deze “officiële” getallen dient men voorzichtig te behandelen. In de eerste plaats, omdat de bijdrage van de meetruis bij nieuwe type ontvangers intussen tot enkele decimeters is verbeterd (C/A-code benadert dus de nauwkeurigheid van de P-code!). In de tweede plaats, omdat een degradatie van de meetprecisie plaatsvindt (speciaal bedoeld voor niet-militaire gebruikers van GPS) door toepassing van de Selective Availability (SA)⁵. De meetprecisie in de code-afstanden loopt hierdoor tot ongeveer 25 meter op. Omdat het SA-effect vrij groot is, ligt de totale nauwkeurigheid aanzienlijk lager dan de gegevens in tabel 2.4 aangeven.

⁵ SA heeft in korte tijdintervallen een systematisch karakter.

Tabel 2.4. Foutenbronnen in de meting van code-afstanden zonder SA (in meters).

foutenbron	soort	C/A-code	P-code
satelliet	klok stabiliteit	3,0	3,0
	baanverstoring	1,0	1,0
	diversen	0,5	0,5
besturingssegment	ephemeriden	4,2	4,2
	diversen	0,9	0,9
propagatie	ionosfeer	5,0-10,0	2,3
	troposfeer	2,0	2,0
	multipad	1,2	1,2
ontvanger	meetruis	7,5	1,5
	diversen	0,5	0,5
totaal (σ)		10,8-13,9	6,6

Echter, met toepassing van differentiële plaatsbepaling kunnen een aantal systematische fouten grotendeels worden geëlimineerd. Hierbij wordt gebruik gemaakt van een referentieontvanger die gelijktijdig op een bekend punt wordt opgesteld. De invloeden van baanfouten, klokfouten, atmosferische effecten en SA-effect kunnen hierdoor aanzienlijk worden gereduceerd. Differentiële (relatieve) plaatsbepaling wordt zowel met codemeting als met fasemeting toegepast.

Satellietconstellatie

Ondanks het feit dat GPS thans volledig operationeel is, verandert de beschikbaarheid van het aantal satellieten voortdurend. Daarmee verandert ook de sterkte van de geometrische oplossing waarmee de (absolute) positie op een gegeven moment bepaald kan worden. De geometrische sterkte voor de positiebepaling wordt door een getal gedefinieerd. Dit getal heet PDOP (Position Dilution Of Precision) waarvan de beschrijving later, in § 5.6, wordt gegeven. Een slechte PDOP waarde wordt veroorzaakt door de combinatie van te weinig beschikbare satellieten en/of door een ongunstige verdeling van de satellieten voor een bepaalde meetlocatie. Men spreekt van een slechte PDOP-waarde als het getal groter is dan 6. Het is vaak zinvol om de satellietconstellatie (configuratie) te beoordelen, voordat men een meting gaat uitvoeren. Men kan hiervoor berekeningen uitvoeren met behulp van beschikbare programma's.

In figuur 2.10 wordt de constellatie over een periode van 24 uur in een tabel weergegeven. In het hoofd van deze tabel vindt men informatie over o.a. de datum, locatie (Delft: φ , λ), de gewenste minimum elevatiehoek van de satellieten (15°) en het aantal beschikbare 27 (!) satellieten. In de eerste kolom worden de satellieten gegeven in tijdperioden tussen opkomst en ondergang (Rise/Set), de duur (dT) en de corresponderende PDOP-waarden. Het blijkt dat de constellatie per etmaal ruim 70 maal verandert. Er is volgens deze tabel één periode waarin slechts vier satellieten beschikbaar zijn, met een duur van 6 minuten en een hoge PDOP-waarde. In figuur 2.11 wordt dezelfde tabel grafisch gegeven: boven het aantal satellieten (Nsats) en onder de PDOP-waarden. In figuur 2.12 wordt een *Sky plot* voorgesteld. De constellatie wordt voor een volgende dag praktisch dezelfde, met het verschil dat alle satellieten bijna 4 minuten (3^m56^s) vroeger komen.

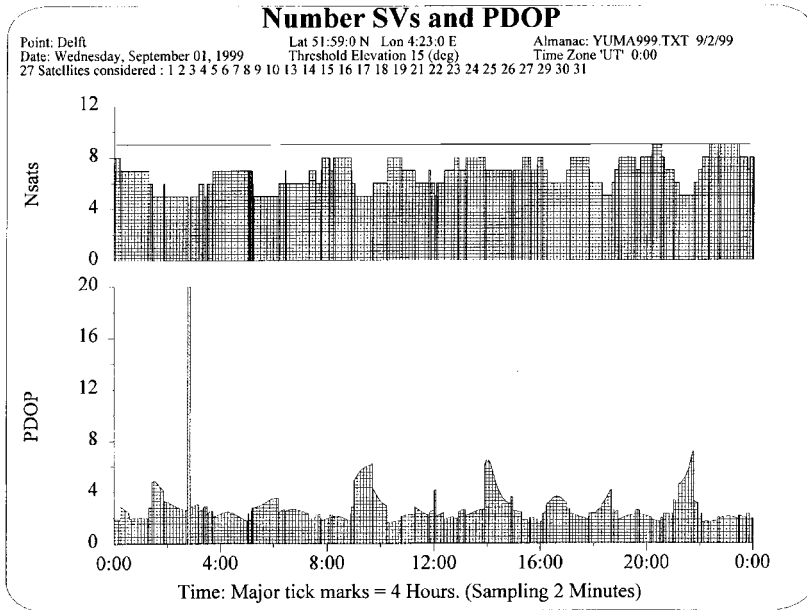
SV Constellations

Point: Delft
 Date: Wednesday, September 01, 1999
 27 Satellites considered : 1 2 3 4 5 6 7 8 9 10 13 14 15 16 17 18 19 21 22 23 24 25 26 27 29 30 31
 Sampling Rate: 2 Minutes

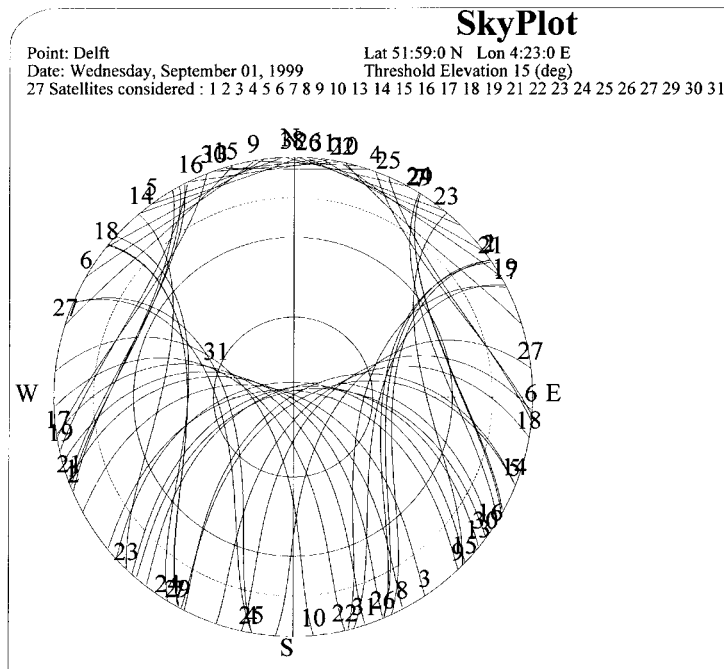
Lat 51:59:0 N Lon 4:23:0 E
 Threshold Elevation 15 (deg)
 Almanac: YUMA999.TXT 9/2/99
 Time Zone 'UT' 0:00

Constellation	T Rise	T Set	dT	PDOP Rise	PDOP Set
1 5 6 8 22 24 25 30	0:00	0:14	0:14		1.9
1 5 6 22 24 25 30	0:14	0:36	0:22	2.9	2.4
1 6 17 22 24 25 30	0:36	1:18	0:42	1.9	2.0
6 17 22 24 25 30	1:18	1:26	0:08	2.8	2.8
6 17 22 25 30	1:26	1:52	0:26	4.8	4.2
3 6 17 22 25 30	1:52	1:54	0:02	2.4	2.4
3 6 17 22 25	1:54	2:46	0:52	3.3	2.6
3 6 17 22	2:46	2:52	0:06	21.8	924.5
3 6 17 22 23	2:52	3:10	0:18	2.9	3.1
3 6 17 19 22 23	3:10	3:22	0:12	2.5	2.7
3 17 19 22 23	3:22	3:30	0:08	2.9	2.9
3 17 19 22 23 31	3:30	3:42	0:12	2.4	2.6
3 17 19 21 22 23 31	3:42	5:02	1:20	2.1	1.8
3 17 21 22 23 31	5:02	5:04	0:02	2.4	2.4
3 15 17 21 22 23 31	5:04	5:10	0:06	1.9	1.8
3 15 21 22 23 31	5:10	5:12	0:02	2.5	2.5
3 15 21 23 31	5:12	6:10	0:58	2.8	3.5
3 15 21 23 29 31	6:10	6:24	0:14	2.4	2.7
2 3 15 21 23 29 31	6:24	6:26	0:02	2.4	2.4
2 3 15 21 29 31	6:26	7:18	0:52	2.5	2.3
2 3 14 15 21 29 31	7:18	7:32	0:14	1.9	1.9
2 14 15 21 29 31	7:32	7:44	0:12	2.2	2.3
1 2 14 15 21 29 31	7:44	7:46	0:02	2.0	2.0
1 2 7 14 15 21 29 31	7:46	8:04	0:18	1.8	1.9
1 2 7 14 15 29 31	8:04	8:12	0:08	2.2	2.2
1 2 7 14 15 16 29 31	8:12	8:52	0:40	2.1	1.8
1 2 7 14 15 16 29	8:52	8:54	0:02	2.3	2.3
1 7 14 15 16 29	8:54	9:00	0:06	2.8	2.9
1 7 14 15 16	9:00	9:42	0:42	4.9	6.2
1 7 14 15 16 25	9:42	10:14	0:32	4.2	2.9
1 4 7 14 15 16 18 25	10:14	10:46	0:32	1.6	1.8
1 4 7 14 16 18 25	10:46	11:16	0:30	2.1	2.2
1 4 7 14 16 18	11:16	11:46	0:30	2.8	2.3
1 4 7 14 16 18 19	11:46	11:50	0:04	2.0	2.0
1 4 14 16 18 19	11:50	12:00	0:10	2.5	2.5
4 14 16 18 19	12:00	12:04	0:04	4.1	4.1
4 14 16 18 19 24	12:04	12:22	0:18	2.1	2.4
4 13 14 16 18 19 24	12:22	12:44	0:22	1.9	2.1
4 10 13 14 16 18 19 24	12:44	12:54	0:10	1.9	1.9
4 10 13 16 18 19 24	12:54	13:10	0:16	2.4	2.6
4 10 13 16 18 19 24 27	13:10	13:54	0:44	2.2	2.7
4 10 13 18 19 24 27	13:54	14:54	1:00	6.1	3.1
10 13 18 19 24 27	14:54	14:58	0:04	3.6	3.6
10 13 18 19 24 26 27	14:58	15:18	0:20	2.5	2.4
2 10 13 18 19 24 26 27	15:18	15:36	0:18	1.8	1.8
2 10 13 19 24 26 27	15:36	15:52	0:16	2.0	2.0
2 10 13 17 19 24 26 27	15:52	16:04	0:12	1.7	1.7
2 10 13 17 19 26 27	16:04	16:14	0:10	2.3	2.4
2 10 13 17 26 27	16:14	16:58	0:44	3.0	3.2
2 9 10 13 17 26 27	16:58	17:06	0:08	2.7	2.7
2 9 10 13 17 23 26 27	17:06	17:48	0:42	2.3	1.9
2 9 10 13 23 26 27	17:48	17:50	0:02	2.1	2.1
2 9 13 23 26 27	17:50	18:08	0:18	2.4	2.3
2 7 9 23 26 27	18:08	18:18	0:10	2.6	2.8
2 7 9 23 26	18:18	18:40	0:22	3.0	4.1
2 7 8 9 23 26	18:40	18:46	0:06	2.5	2.5
2 7 8 9 21 23 26	18:46	18:56	0:10	2.3	2.5
2 5 7 8 9 21 23 26	18:56	19:34	0:38	1.8	2.2
5 7 8 9 21 23 26	19:34	19:44	0:10	2.6	2.6
5 7 8 9 21 23 26 29	19:44	20:10	0:26	2.0	2.0
5 7 8 9 21 23 26 29 30	20:10	20:32	0:22	1.8	1.7
5 7 8 9 21 23 29 30	20:32	20:38	0:06	2.0	1.9
5 7 8 9 21 29 30	20:38	21:00	0:22	2.3	2.3
5 8 9 21 29 30	21:00	21:12	0:12	3.3	3.3
5 8 9 29 30	21:12	21:46	0:34	4.5	7.1
4 5 8 9 29 30	21:46	21:56	0:10	3.3	3.1
4 5 6 8 9 29 30	21:56	22:04	0:08	2.3	2.3
4 5 6 8 9 25 29 30	22:04	22:20	0:16	1.6	1.7
4 5 6 8 9 24 25 29 30	22:20	22:40	0:20	1.6	1.7
4 5 6 8 24 25 29 30	22:40	22:44	0:04	2.0	2.0
1 4 5 6 8 24 25 29 30	22:44	23:28	0:44	1.9	1.9
1 5 6 8 24 25 29 30	23:28	23:46	0:18	2.1	2.0
1 5 6 8 24 25 30	23:46	23:52	0:06	2.3	2.2
1 5 6 8 22 24 25 30	23:52	0:00	0:08	1.9	1.9

Figuur 2.10. Satellietconfiguratie in tabelvorm.



Figuur 2.11. Satelliet configuratie in grafische vorm.



Figuur 2.12. Sky plot: de satellietbanen in een periode van 24 uur (met station Delft in het midden).

5 Plaatsbepaling uit code-afstanden

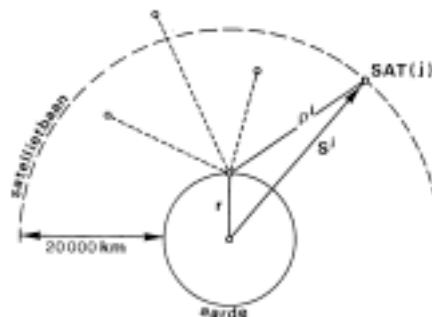
5.1. Inleiding

De code-afstandsmeting (of pseudo-afstandsmeting) uit PRN-code wordt beschouwd als de normale opzet van GPS en is vooral bedoeld voor plaatsbepaling in real-time. De belangrijkste toepassingen hiervan zijn de navigatie te land, ter zee en in de lucht.

In § 2.4 werd de meting van de looptijden met digitale code besproken. Uit de gemeten enkele looptijd van een satelliet naar de waarnemer wordt een afstand berekend door deze looptijd te vermenigvuldigen met de voortplantingsnelheid van de radiogolf. De looptijd wordt met twee, in feite onafhankelijke klokken gemeten die niet exact gesynchroniseerd kunnen worden. Men mag aannemen dat alle satelliet-klokken (na het aanbrengen van een klokcorrectie) in voldoende mate gelijk lopen in GPS-tijd terwijl de klok in de ontvanger een significante afwijking heeft. Deze afwijking wordt de klokfout van de ontvanger genoemd. Door deze klokfout hebben ook alle afstanden een afwijking. Vandaar de ook veel gebruikte naam pseudo-afstand (zie Hoofdstuk 1).

Zonder deze klokfout zouden drie satellieten voldoende zijn om de positie van de waarnemer vanuit de bekende satelliet posities te bepalen met een ruimtelijke insnijding. De oplossing is het snijpunt van drie bollen met de satellieten in het middelpunt. De meting van een vierde satelliet is nodig om ook de klokfout als een vierde onbekende te bepalen.

Een simpele opzet van enkelpuntsbepaling (*single point positioning*) uit gemeten code-afstanden naar vier satellieten, wordt in figuur 5.1 voorgesteld.



Figuur 5.1. Positiebepaling uit vier afstanden.

In figuur 5.1 zijn: r de onbekende positievector van de waarnemer (X_r, Y_r, Z_r) , $S(t) = [X^s(t), Y^s(t), Z^s(t)]$ de positievector van de satelliet (deze kan uit de uitgezonden Keplerse baanparameters worden berekend) en $\rho(t)$ de geometrische afstand (range) tussen satelliet en waarnemer.

5.2. Het mathematische model

5.2.1. GPS-tijd, satelliet-tijd en ontvanger-tijd

Het is zeer belangrijk om de verschillende tijden die bij GPS voorkomen exact te definiëren en de onderlinge relaties vast te leggen. Elke satelliet en elke ontvanger heeft een eigen klok (oscillator) en zo doende zijn eigen tijd. Omdat het onmogelijk is om alle klokken met een gewenste precisie gelijk te laten lopen, heeft men een referentietijd geïntroduceerd. Deze referentietijd heet GPS-tijd. GPS-tijd is gekoppeld aan de internationale atoomtijd (IAT) zodanig dat het verschil IAT-GPStijd altijd 19s is (zie § 3.2). De satelliet en de ontvanger (receiver) klok tijden worden als volgt gedefinieerd t.o.v. de GPS-tijd ($t = t_{\text{GPS}}$):

$$\begin{aligned} t^s(t) &= t + \delta t^s(t) \\ t_r(t) &= t + \delta t_r(t) \end{aligned} \tag{5.1}$$

waarin $\delta t^s(t)$ en $\delta t_r(t)$ de klokfouten zijn, respectievelijk van de satellietklokken en van de ontvangerklokken¹. (De klokken worden dus gecorrigeerd door de fouten er van af te trekken). Hier blijkt dat de klokfouten slechts op een gegeven tijdstip gelden, want alle klokken hebben een drift, dus ook de klokfouten. Daar de afstandsmetingen op tijdmetingen zijn gebaseerd, dient men bij het opstellen van de waarnemingsvergelijkingen rekening te houden met deze klokfouten en klokdriften. De klokfout $\delta t^s(t)$ wordt voor elke satellietklok in de Broadcast Ephemeris met behulp van drie klokparameters (§ 2.2) vrij nauwkeurig gegeven, terwijl de klokfout $\delta t_r(t)$ van de ontvanger gemodelleerd (en dus bepaald) moet worden.

5.2.2. Pseudoafstand uit code-meting

De bepaling van code-afstanden is gebaseerd op de meting van de enkele looptijd van de code van satelliet tot ontvanger. De code kan zijn C/A, P1 of P2 code. Voor de gemeten “pseudo-looptijd” $t_r^s(t)$ op het tijdstip van aankomst bij de ontvanger in GPS-tijd (t) kunnen we schrijven:

$$t_r^s(t) = t_r(t) - t^s(t - \tau_r^s(t)) \tag{5.2}$$

waarin:

- $t^s(t - \tau_r^s(t))$ de tijd van vertrek (in satelliettijd)
- $t_r(t)$ de tijd van aankomst (in ontvangertijd)
- $\tau_r^s(t)$ de looptijd van het signaal (in GPS - tijdschaal)
- t epoche, het moment van aankomst in GPS tijd.

Hierbij worden voor de vertrektijd en voor de aankomsttijd twee verschillende tijdsystemen gehanteerd. Om alles in GPS-tijd uit te drukken, substitueren we (5.1) in (5.2). Na substitutie volgt:

¹ Voor de satellieten wordt een bovenindex s gebruikt (de satellieten bevinden zich boven) en voor de ontvanger een onderindex r .

$$t_r^s(t) = \tau_r^s(t) + \delta t_r(t) - \delta t^s(t - \tau_r^s(t)) \quad (5.3)$$

De satelliet klokkfout is ook niet constant; deze klokkfout dient voor het tijdstip van het vertrek van het signaal (aankomsttijd minus looptijd) te worden bepaald. Vermenigvuldiging van (5.3) met de snelheid c van de radiogolven, substitutie van $\rho_r^s(t) = c \tau_r^s(t)$ en toevoeging van de ionosferische (I_r^s) en de troposferische (T_r^s) vertraging (zie: Hoofdstuk 4) geeft de waarnemingsvergelijking van de pseudoafstand $P_r^s(t)$:

$$P_r^s(t) = \rho_r^s(t) + c \delta t_r(t) - c \delta t^s(t - \tau_r^s(t)) + I_r^s + T_r^s \quad (5.4)$$

waarin:

I_r^s = ionosferische vertraging van het signaal

T_r^s = troposferische vertraging van het signaal

De ionosferische en troposferische vertragingen worden bij code-metingen meestal niet gemodelleerd, maar gecorrigeerd, geëlimineerd of verwaarloosd.

5.2.3. Pseudoafstand uit fasemeting

De meting van fase-afstanden is gebaseerd op de meting van de enkele looptijd van de draaggolf. Met een twee-frequentie ontvanger kan men op de L1 en L2 banden behalve twee code-afstanden ook twee onafhankelijke fase-afstanden meten. De fase-afstand wordt als volgt geschreven:

$$\varphi_r^s(t) = \varphi_r(t) - \varphi^s(t - \tau_r^s(t)) + N_r^s \quad (5.5)$$

waarin:

$\varphi^s(t - \tau_r^s(t))$ de fase van draaggolf vanuit de satelliet, op het tijdstip van uitzenden;

$\varphi_r(t)$ de fase van draaggolf van de ontvanger op het tijdstip van aankomst;

$\tau_r^s(t)$ de looptijd van het signaal van satelliet tot ontvanger;

N_r^s het aantal gehele golflengten (integer getal).

Merk op dat de faseafstand uit twee componenten wordt samengesteld:

a) het faseverschil: $\varphi_r - \varphi^s$; dit is een fractie van de golflengte;

b) het aantal gehele golflengten of cycles.

Voor de fase van de satelliet en de ontvanger kunnen we schrijven:

$$\begin{aligned} \varphi^s(t - \tau_r^s(t)) &= \varphi^s(t_0) + f(t - \tau_r^s(t) - t_0) + f(\delta t^s(t - \tau_r^s(t)) - \delta t^s(t_0)) \\ \varphi_r(t) &= \varphi_r(t_0) + f(t - t_0) + f(\delta t_r(t) - \delta t_r(t_0)) \end{aligned} \quad (5.6)$$

waarin $\varphi^s(t_0)$ en $\varphi_r(t_0)$ de initiële fasen en f de frequentie van een van de GPS draaggolven zijn. Substitutie van (5.6) in (5.5) geeft de gemeten fase in cycles:

$$\varphi_r^s(t) = f \tau_r^s(t) + f \delta t_r(t) - f \delta t^s(t - \tau_r^s(t)) + A_r^s \quad (5.7)$$

waarin:

$$A_r^s = N_r^s + \varphi_r(t_0) - f \delta t_r(t_0) - \varphi^s(t_0) + f \delta t^s(t_0) \quad (5.8)$$

De grootheid A_r^s is de zg. *ambiguity term*. Deze term is geen integer want met uitzondering van N_r^s hebben alle termen een reële waarde. Vergelijking (5.7) kan in afstand eenheid (meters) worden omgezet door vermenigvuldigen met $\lambda = c / f$. Met toevoeging van de atmosferische vertraging voor de ionosfeer en de troposfeer volgt de waarnemingsvergelijking van de fasemeting:

$$\boxed{p_r^s(t) = \rho_r^s(t) + c \delta t_r(t) - c \delta t^s(t - \tau_r^s(t)) + \lambda A_r^s - I_r^s + T_r^s} \quad (5.9)$$

Hierbij kan worden opgemerkt dat

1. De fase-afstand wordt met een klein letter p (phase) genoteerd;
2. De vergelijkingen (5.4) en (5.9) lijken veel op elkaar, echter,
 - de ionosferische parameters hebben tegengestelde tekens. Het tekenverschil ligt aan het feit dat de code bij het passeren van de ionosfeer wordt vertraagd terwijl de draaggolf juist versneld. De grootte van de vertraging en de versnelling is vrijwel gelijk.
 - de vergelijking (5.9) voor de fase heeft voor elke satelliet een onbekende meerduidigheid.

Het model van fasemeting volgens vergelijking (5.9) wordt verder in Hoofdstuk 6 behandeld.

5.3. Plaatsbepaling uit code-afstanden

We gaan uit van een simpel *functie model* voor de code-afstand volgens (5.4), met verwaarlozing van de ionosferische en de troposferische vertraging²:

$$P_r^s(t) + c \delta t^s(t - \tau_r^s(t)) = \rho_r^s(t) + c \delta t_r(t) \quad (5.10)$$

waarin de geometrische afstand is:

$$\rho_r^s(t) = \sqrt{(X^s(t) - X_r)^2 + (Y^s(t) - Y_r)^2 + (Z^s(t) - Z_r)^2} \quad (5.11)$$

met de coördinaten $X^s(t), Y^s(t), Z^s(t)$ van de satelliet op het moment van *vertrek* van het signaal, en de coördinaten X_r, Y_r, Z_r van de waarnemer (antenne van de ontvanger) bij aankomst van het signaal. Als de waarnemer in beweging is, zoals dat het geval is bij kinematische metingen dan dient men bij de coördinaten van de waarnemer ook de epoche (t) toe te voegen.

In vergelijkingen (5.10) en (5.11) hebben we een niet-lineaire relatie tussen de waargenomen code-afstand en een aantal onbekenden. De onbekenden zijn: de coördinaten X_r, Y_r, Z_r van de ontvanger en de klokcorrectie $c \delta t_r$. Om deze onbekenden volgens de kleinste kwadraten methode op te lossen, dienen we alleen de geometrische afstand ρ_r^s te lineariseren. Volgens de Taylor-reeks, met verwaarlozing van de hogere orde termen, geldt:

² In de ruwe data worden bij de meeste ontvangers de ongecorrigeerde code-afstanden geven, zodat men de klokfout $\delta t^s(t)$ zelf moet aanbrengen. Hiervoor gebruikt men de uitgezonden klokparameters (§ 2.2).

$$f(x, y, z) = f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) = \quad (5.12)$$

$$f(x_0, y_0, z_0) + \frac{\partial f(x_0, y_0, z_0)}{\partial x_0} \Delta x + \frac{\partial f(x_0, y_0, z_0)}{\partial y_0} \Delta y + \frac{\partial f(x_0, y_0, z_0)}{\partial z_0} \Delta z$$

Met de partiële afgeleiden van (5.11) krijgen we:

$$\rho_r^s = \rho_{r0}^s - \frac{X^s - X_{r0}}{\rho_{r0}^s} \Delta X_r - \frac{Y^s - Y_{r0}}{\rho_{r0}^s} \Delta Y_r - \frac{Z^s - Z_{r0}}{\rho_{r0}^s} \Delta Z_r \quad (5.13)$$

waarin ρ_{r0}^s uit de benaderde positie van de ontvanger wordt berekend:

$$\rho_{r0}^s = \sqrt{(X^s - X_{r0})^2 + (Y^s - Y_{r0})^2 + (Z^s - Z_{r0})^2} \quad (5.14)$$

Gemakshalve laten we de epoche (t) bij alle tijdafhankelijke termen vanaf (5.13) achterwege. Substitutie van (5.13) in (5.10) geeft een stelsel lineaire waarnemingsvergelijkingen voor een aantal satellieten ($s = 1, 2, \dots, m$):

$$P_r^s + c \delta t^s - \rho_{r0}^s = - \frac{X^s - X_{r0}}{\rho_{r0}^s} \Delta X_r - \frac{Y^s - Y_{r0}}{\rho_{r0}^s} \Delta Y_r - \frac{Z^s - Z_{r0}}{\rho_{r0}^s} \Delta Z_r + c \delta t_r \quad (5.15)$$

De coëfficiënten van ΔX , ΔY en ΔZ zijn de zg. richtingscosinussen van de ontvanger naar de satellieten. Met de gebruikelijke notatie van de kleinste kwadraten vereffening geldt:

$$E\{\Delta \underline{y}\} = A \cdot \Delta \underline{x} \Rightarrow \quad (5.16)$$

$$E\left\{ \begin{bmatrix} \Delta \underline{y}^1 \\ \Delta \underline{y}^2 \\ \vdots \\ \Delta \underline{y}^m \end{bmatrix} \right\} = \begin{bmatrix} -\frac{X^1 - X_{r0}}{\rho_{r0}^1} & -\frac{Y^1 - Y_{r0}}{\rho_{r0}^1} & -\frac{Z^1 - Z_{r0}}{\rho_{r0}^1} & +1 \\ -\frac{X^2 - X_{r0}}{\rho_{r0}^2} & -\frac{Y^2 - Y_{r0}}{\rho_{r0}^2} & -\frac{Z^2 - Z_{r0}}{\rho_{r0}^2} & +1 \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{X^m - X_{r0}}{\rho_{r0}^m} & -\frac{Y^m - Y_{r0}}{\rho_{r0}^m} & -\frac{Z^m - Z_{r0}}{\rho_{r0}^m} & +1 \end{bmatrix} \begin{bmatrix} \Delta X_r \\ \Delta Y_r \\ \Delta Z_r \\ c \delta t_r \end{bmatrix}$$

waarin $\Delta \underline{y}^s = P_r^s + c \delta t^s - \rho_{r0}^s$ de gelineariseerde waarnemingsvector is, het verschil tussen de waargenomen en de benaderde afstand. De waargenomen afstand wordt hierbij gecorrigeerd met de klokfout δt^s van de satellietklokken die aan de navigatie boodschap wordt ontleend.

Uit minimaal vier satellieten kan men de vier onbekenden uit (5.16) oplossen. Als $m > 4$ dan voert men een kleinste kwadraten vereffening uit:

$$\Delta \hat{\underline{x}} = (A^* Q_y^{-1} A)^{-1} A^* Q_y^{-1} \Delta \underline{y} = Q_{\hat{x}} A^* Q_y^{-1} \Delta \underline{y} \quad (5.17)$$

waarin Q_y de variantie-covariantiematrix is van de waargenomen code-afstanden. In de praktijk wordt hiervoor vaak een diagonaal matrix gebruikt:

$$Q_y = \sigma_y^2 I_m \quad (5.18)$$

waarin σ_y^2 de a-priori variantie is van de waarnemingen. Bij gebruik van een geodetische ontvanger kan hiervoor een waarde van $\sigma = 1 - 15$ meter worden genomen. De hoogste waarde geldt als SA aangezet is. Voor de oplossing van de onbekenden uit (5.17) zijn altijd een aantal iteraties nodig. In deze iteraties heeft de klokfout van de ontvanger een significante invloed op het tijdstip waarvoor de satellietcoördinaten moeten worden berekend, nl. het vertrek van het signaal. Dit tijdstip wordt uit de aankomsttijd en de onbekende looptijd bepaald. De looptijd zal na elke iteratie stap iets verbeterd worden volgens:

$$t' = t - \tau_r^s(t) = t - \frac{\hat{P}_r^s + c\hat{\delta t}_r}{c}$$

In de volgende stap worden hierdoor de satelliet coördinaten beïnvloedt. Na de oplossing van de normaalvergelijkingen (5.17) dient men de positie te bepalen door de schattingen bij de benaderde waarden op te tellen:

$$\begin{aligned} \hat{X} &= X_0 + \Delta\hat{X} \\ \hat{Y} &= Y_0 + \Delta\hat{Y} \\ \hat{Z} &= Z_0 + \Delta\hat{Z} \end{aligned} \quad (5.19)$$

Verbetering van de code-afstanden met fasewaarnemingen

De code-afstanden worden dikwijls gefilterd m.b.v. de simultaan gemeten fase, waardoor de positie met een hogere precisie kan worden bepaald. Deze methode maakt gebruik van het feit dat veranderingen in de afstand tussen waarnemer en satelliet veel nauwkeuriger uit fase-waarnemingen dan uit code-waarnemingen kunnen worden bepaald. Een simpel filter geeft de zogenaamde fase-ondersteunde code-afstand (*phase-aided pseudorange*) per satelliet:

$$\bar{P}_k = \frac{P_k + (k-1)(\bar{P}_{k-1} + \Delta p_{k-1,k})}{k} \quad k = 1, 2, \dots, k_{\max} \quad (5.20)$$

waarin k het aantal waarnemingen voor het filter is. Voor het afstandsverschil in (5.20) geldt:

$$\Delta p_{k-1,k} = p_k - p_{k-1} \approx P_k - P_{k-1}$$

Door toepassing van het filter (5.20) wordt uit de voorgaande code-afstand en de gemeten afstandsverandering telkens de actuele code-afstand geschat en het gewogen gemiddelde berekend met de gemeten code-afstand. Het gladstrijken van de code-afstanden kan in principe tot in het oneindige doorgaan waarbij de precisie van de code-afstand bij iedere stap iets verbeterd wordt. Dit betekent dat men geleidelijk overgaat van code-afstand naar fase-afstand.

Wil men echter een zeker evenwicht bewaren tussen de code-afstand en de fasemeting, dan dient men aan k een maximale waarde te verbinden. Deze waarde bedraagt in de praktijk ca. 100. Als er in de gemeten afstandsverandering nu en dan een kleine fout (enkele cycle slips) optreedt, dan zal deze fout slechts een geringe invloed hebben. Bij een

grove fout moet echter het filter opnieuw gestart worden. Het nadeel van dit filter is dat het starten (of het stoppen) van het filter bij een satelliet een ongewenste sprong in de positie veroorzaakt. De precisieverbetering in de positiebepaling blijft trouwens beperkt, want de invloeden van andere systematische effecten, met name baanfouten en atmosferische refractie, kunnen hiermee niet worden verminderd. Later zal blijken dat deze invloeden veel effectiever met differentiële techniek kunnen worden geëlimineerd.

5.4. Invoering van geografische coördinaten als onbekenden

Bij navigatie is het praktischer geografische coördinaten (φ , λ) en hoogte (h) boven de ellipsoïde als onbekenden te bepalen. De geocentrische en de geografische coördinaten zijn gelijkwaardige WGS84 coördinaten. De transformatie is gebaseerd op de bekende relatie (zie § 3.1):

$$\begin{aligned} X &= (N + h) \cos\varphi \cos\lambda \\ Y &= (N + h) \cos\varphi \sin\lambda \\ Z &= ((1 - e^2)N + h) \sin\varphi \end{aligned} \quad (5.21)$$

waarin N de radius van de kromming is in de eerste vertikaal:

$$N = a / (1 - e^2 \sin^2\varphi)^{1/2} \quad (5.22)$$

Men kan nu de positie eerst in het geocentrische stelsel volgens model (5.16) oplossen en daarna de XYZ- coördinaten naar het geografische stelsel transformeren door middel van een inverse transformatie van (5.21). Het heeft echter voordelen om het model in het geografische stelsel op te stellen en de positie direct in dat stelsel op te lossen. Hiervoor schrijven we de partiële afgeleiden van (5.21) als volgt:

$$\begin{aligned} \frac{\partial X}{\partial \varphi} &= -(M + h) \sin\varphi \cos\lambda & \frac{\partial X}{\partial \lambda} &= -(N + h) \cos\varphi \sin\lambda & \frac{\partial X}{\partial h} &= \cos\varphi \cos\lambda \\ \frac{\partial Y}{\partial \varphi} &= -(M + h) \sin\varphi \sin\lambda & \frac{\partial Y}{\partial \lambda} &= (N + h) \cos\varphi \cos\lambda & \frac{\partial Y}{\partial h} &= \cos\varphi \sin\lambda \\ \frac{\partial Z}{\partial \varphi} &= (M + h) \cos\varphi & \frac{\partial Z}{\partial \lambda} &= 0 & \frac{\partial Z}{\partial h} &= \sin\varphi \end{aligned} \quad (5.23)$$

waarin M de kromtestraal is in de meridiaan volgens:

$$M = \frac{a(1 - e^2)}{(1 - e^2 \sin^2\varphi)^{3/2}} \quad (5.24)$$

Voor de nieuwe waarnemingsvergelijkingen met geografische coördinaten als onbekenden hebben we de partiële afgeleiden $\partial\rho/\partial\varphi$, $\partial\rho/\partial\lambda$ en $\partial\rho/\partial h$ nodig. Deze kunnen we als volgt schrijven:

$$\frac{\partial\rho}{\partial\varphi} = \frac{\partial\rho}{\partial X} \frac{\partial X}{\partial\varphi} + \frac{\partial\rho}{\partial Y} \frac{\partial Y}{\partial\varphi} + \frac{\partial\rho}{\partial Z} \frac{\partial Z}{\partial\varphi}$$

$$\frac{\partial\rho}{\partial\lambda} = \frac{\partial\rho}{\partial X} \frac{\partial X}{\partial\lambda} + \frac{\partial\rho}{\partial Y} \frac{\partial Y}{\partial\lambda} + \frac{\partial\rho}{\partial Z} \frac{\partial Z}{\partial\lambda} \quad (5.25)$$

$$\frac{\partial\rho}{\partial h} = \frac{\partial\rho}{\partial X} \frac{\partial X}{\partial h} + \frac{\partial\rho}{\partial Y} \frac{\partial Y}{\partial h} + \frac{\partial\rho}{\partial Z} \frac{\partial Z}{\partial h}$$

Substitutie van (5.23) in (5.25) geeft:

$$\frac{\partial\rho}{\partial\varphi} = -\frac{\partial\rho}{\partial X}(M+h)\sin\varphi\cos\lambda - \frac{\partial\rho}{\partial Y}(M+h)\sin\varphi\sin\lambda + \frac{\partial\rho}{\partial Z}(M+h)\cos\varphi$$

$$\frac{\partial\rho}{\partial\lambda} = -\frac{\partial\rho}{\partial X}(N+h)\cos\varphi\sin\lambda + \frac{\partial\rho}{\partial Y}(N+h)\cos\varphi\cos\lambda \quad (5.26)$$

$$\frac{\partial\rho}{\partial h} = \frac{\partial\rho}{\partial X}\cos\varphi\cos\lambda + \frac{\partial\rho}{\partial Y}\cos\varphi\sin\lambda + \frac{\partial\rho}{\partial Z}\sin\varphi$$

Deze vergelijking, met toevoeging van de klokfout, kunnen we in matrix vorm schrijven:

$$\Delta\rho_r^s = \begin{bmatrix} \frac{\partial\rho}{\partial X} & \frac{\partial\rho}{\partial Y} & \frac{\partial\rho}{\partial Z} & 1 \end{bmatrix} G \begin{bmatrix} (M+h)\Delta\varphi \\ (N+h)\Delta\lambda\cos\varphi_0 \\ \Delta h \\ c\delta t_r \end{bmatrix} \quad (5.27)$$

waarin G de transformatie matrix is naar geografische coördinaten:

$$G = \begin{bmatrix} -\sin\varphi_0\cos\lambda_0 & -\sin\lambda_0 & \cos\varphi_0\cos\lambda_0 & 0 \\ -\sin\varphi_0\sin\lambda_0 & \cos\lambda_0 & \cos\varphi_0\sin\lambda_0 & 0 \\ \cos\varphi_0 & 0 & \sin\varphi_0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.28)$$

Met de matrix G wordt dus de coëfficiëntmatrix A volgens (5.16) naar een nieuwe coëfficiëntmatrix getransformeerd. Hierbij zijn benaderde waarden voor de geografische breedte en lengte voldoende. We kunnen nu de nieuwe waarnemingsvergelijkingen als volgt schrijven:

$$P_r^s + \delta t^s - \rho_{r0}^s = A G \begin{bmatrix} \Delta N \\ \Delta E \\ \Delta h \\ c \delta t_r \end{bmatrix} \quad (5.29)$$

waarin ΔN and ΔE de noord en oost componenten zijn van de onbekenden, uitgedrukt in meters ten opzichte van de ingevoerde benaderde positie φ_0 en λ_0 :

$$\begin{aligned} \Delta N &= (M + h)\Delta\varphi \cong (R + h)\Delta\varphi \\ \Delta E &= (N + h)\Delta\lambda \cos \varphi_0 \cong (R + h)\Delta\lambda \cos \varphi_0 \end{aligned} \quad (5.30)$$

De nieuwe coëfficiënten matrix in (5.29) noteren we als $\bar{A} = AG$. Zo hebben we de waarnemingsvergelijkingen in de gebruikelijke notatie:

$$E\{\Delta \underline{y}\} = A \cdot \Delta x \Rightarrow E\left\{ \begin{bmatrix} \Delta \underline{y}^1 \\ \Delta \underline{y}^2 \\ \vdots \\ \Delta \underline{y}^m \end{bmatrix} \right\} = \bar{A} \begin{bmatrix} \Delta N \\ \Delta E \\ \Delta h \\ c \delta t_r \end{bmatrix} \quad (5.31)$$

5.5. Toepassing van constraints voor de onbekenden

Het bovenstaand model volgens (5.31) wordt vaak met invoering van constraints opgelost. Beschouw het gepartitioneerd model³

$$E\left\{ \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} \right\} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x ; \quad Q_y = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \quad (5.32)$$

Dit model kan opgelost worden uit een stelsel normaalvergelijkingen:

$$\hat{x} = (A_1^* Q_1^{-1} A_1 + A_2^* Q_2^{-1} A_2)^{-1} (A_1^* Q_1^{-1} \underline{y}_1 + A_2^* Q_2^{-1} \underline{y}_2) \quad (5.33)$$

Substitutie van $A_2 = I$; $Q_2 = P_x^{-1}$; en $\underline{y}_2 = 0$ in het model (5.32) geeft de oplossing:

$$\hat{x} = (A_1^* Q_1^{-1} A_1 + P_x)^{-1} A_1^* Q_1^{-1} \underline{y}_1 \quad (5.34)$$

Naar analogie van (5.17) kan men het model (5.34) in gelineariseerde vorm schrijven:

$$\Delta \hat{x} = (A_1^* Q_y^{-1} A_1 + P_x)^{-1} A_1^* Q_y^{-1} \Delta \underline{y}_1 \quad (5.35)$$

³ Prof.dr.ir. P.J.G. Teunissen: Adjustment Theory, hoofdstuk 6.

waarin P_x een gewichtsmatrix van de onbekenden is. Deze gewichtsmatrix heeft de volgende vorm:

$$P_x = \begin{bmatrix} \sigma_{\Delta N}^2 & & & \\ & \sigma_{\Delta E}^2 & & \\ & & \sigma_{\Delta h}^2 & \\ & & & \sigma_{c\delta t_r}^2 \end{bmatrix}^{-1} \quad (5.36)$$

Het voordeel van deze methode is dat men de P_x matrix kan manipuleren en daarmee constraints opleggen. Met een bewuste keuze van P_x kan men de positiebepaling versterken. Dit kan nuttig zijn bij navigatie op zee als de onbekenden in het geografische stelsel worden bepaald. Zo kan men aan de hoogte een "soft-constraint" toekennen door voor zijn standaardafwijking een kleine waarde te kiezen, bijvoorbeeld $\sigma_{\Delta h} = 1\text{m}$, of zelfs "vasthouden" met een zeer kleine waarde, bijvoorbeeld $\sigma_{\Delta h} = 0.001\text{m}$. Echter, met een keuze van nul kan men de P_x matrix niet berekenen! Wil men een vrije oplossing, dan kiest men voor de standaardafwijking een grote waarde, bijvoorbeeld $\sigma_{\Delta h} = 1000$ meter.

Op een vergelijkbare wijze kan men (eventueel) ook aan de klok-onbekende δt_r een soft-constraint opleggen. Bij een stabiele klok in de ontvanger of bij de aansluiting van een externe atoomklok zal namelijk deze klok-onbekende slechts met een kleine drift in de tijd veranderen. In dit geval kan men de drift van de klok als een extra onbekende modelleren. Opleggen van constraints werd in de tijd toegepast toen de satelliet configuratie nog niet volledig was. Zo heeft men op zee ook met drie-, en zelfs met twee-, satellieten kunnen navigeren.

5.6. Invloed van de satellietconfiguratie

Het aantal beschikbare satellieten en de relatieve positie van de satellieten t.o.v. de waarnemer heeft een aanzienlijke invloed op de precisie van de positiebepaling. Om de sterkte van de geometrische oplossing uit te drukken heeft men het begrip Dilution of Precision (DOP) ingevoerd. Diverse DOP-waarden kunnen worden afgeleid uit de gewichtscoefficientenmatrix: $Q_x = (A^* A)^{-1}$. Hierbij wordt dus een covariantiematrix $Q_y = I_m$ aangenomen. De Geometric Dilution of Precision (GDOP) wordt hiermee als volgt gedefinieerd:

$$\text{GDOP} = \sqrt{\text{spoor}(Q_x)} = \sqrt{Q_{xx} + Q_{yy} + Q_{zz} + Q_{tt}} = \sqrt{Q_{NN} + Q_{EE} + Q_{hh} + Q_{tt}} \quad (5.37)$$

Men onderscheidt tevens de volgende andere DOP waarden:

$$\text{Positie:} \quad \text{PDOP} = \sqrt{Q_{xx} + Q_{yy} + Q_{zz}} = \sqrt{Q_{NN} + Q_{EE} + Q_{hh}}$$

$$\text{Horizontaal:} \quad \text{HDOP} = \sqrt{Q_{NN} + Q_{EE}} \quad (5.38)$$

$$\text{Vertikaal:} \quad \text{VDOP} = \sqrt{Q_{hh}}$$

$$\text{Tijd: } \text{TDOP} = \sqrt{Q_{tt}}$$

Bij een gelijk aantal satellieten geldt bovendien:

$$\text{GDOP}^2 = \text{PDOP}^2 + \text{TDOP}^2 \quad \text{en} \quad \text{PDOP}^2 = \text{HDOP}^2 + \text{VDOP}^2$$

Volgens (5.37) en (5.35) blijken GDOP, PDOP en TDOP invariant te zijn voor het gekozen coördinaatstelsel. Het maakt dus niet uit of de positie in X, Y, Z of N, E, h wordt uitgedrukt. Bovendien zijn alle DOP-waarden een weerspiegeling van de geometrische sterkte en hebben niets met de precisie van de waarnemingen te maken. PDOP wordt gebruikt als een maat voor de nauwkeurigheid van de positie (hoewel de covariantiematrix $Q_{\hat{x}}$ een completer informatie geeft). Hoe kleiner de PDOP waarde, hoe beter de geometrische oplossing. In het algemeen wordt $\text{PDOP} = 2$ als een uitstekende waarde ervaren, terwijl $\text{PDOP} > 7$ als een slechte waarde geldt.

De gemiddelde waarde van PDOP bedraagt bij een volledig operationeel systeem ongeveer 2.5. De a priori standaardafwijking van een enkele positie bepaling wordt hiermee als volgt geschat:

a) met P-code: $\sigma_{\text{pos}} = \sigma_0 \cdot \text{PDOP} = 6.6 \text{ m} \times 2.5 = 17 \text{ m}$

b) met C/A-code: $\sigma_{\text{pos}} = \sigma_0 \cdot \text{PDOP} = 12.4 \text{ m} \times 2.5 = 31 \text{ m}$.

De waarden van σ_0 werden aan tabel § 2.4 ontleend. Dit is slechts een ruwe schatting van de precisie in de enkele positiebepaling zonder de invloed van de Selective Availability. Zoals eerder vermeldt verlaagt SA de nauwkeurigheid met C/A code tot de orde van 100 meter.

5.7. Kinematische toepassingen

Het mathematische model, beschreven in § 5.2, geldt zowel voor de positiebepaling van een stilstaand punt (stationair), als voor een punt in beweging (kinematisch). Het kenmerkende van een stationaire toepassing is dat men de metingen naar eigen behoefte kan herhalen en er geen behoefte is om de berekening in real-time uit te voeren. Met postprocessing kan men dan het best mogelijke resultaat verkrijgen, waarbij meer onbekenden (bijvoorbeeld ook de klokdrift) kunnen worden gemodelleerd.

Bij een kinematische toepassing is er sprake van een bewegend punt (schip, auto of vliegtuig), waarbij op elk tijdstip op een andere punt wordt gemeten. Er is dus telkens slechts één positiebepaling mogelijk die niet herhaald kan worden. Bij navigatie wil men natuurlijk de berekening van de positie in real-time bepalen. Met toepassing van formule (5.31) krijgt men de vier onbekenden (φ , λ , h en $c\delta t_r$) in bepaalde tijdsintervallen. Het tijdsinterval is bij de meeste ontvangers instelbaar; bijvoorbeeld vanaf 0.5 seconde tot een aantal minuten. Korte waarnemingstijd is een belangrijke optie bij kinematische metingen waarbij de vier onbekenden in korte tijdsintervallen kunnen worden opgelost. Men heeft dan praktisch gesproken een continue plaatsbepaling.

De precisie van zo'n kinematische plaatsbepaling kan verbeterd worden door toepassing van een Kalman filter. Men maakt dan gebruik van de informatie van voorgaande posities om de actuele positie te verbeteren. "Ik weet waar ik ben en ik weet ook waar ik vandaan kom". Dit zal tevens de betrouwbaarheid van de plaatsbepaling verhogen en informatie

geven over de snelheid van het vaartuig. Het principe van het Kalman filter wordt in § 5.9 kort toegelicht.

5.8. Relatieve plaatsbepaling: Differentiële GPS (DGPS)

De plaatsbepaling met code-afstanden kan, zowel bij statische als bij kinematische toepassingen, in belangrijke mate verbeterd worden door toepassing van relatieve plaatsbepaling, de zogenaamde differentiële GPS, of DGPS. Wanneer namelijk op twee (of meer) stations gelijktijdig wordt gemeten naar dezelfde satellieten, en deze stations niet al te ver van elkaar liggen, dan zal de invloed van bepaalde systematische fouten nagenoeg gelijk zijn op de berekende posities van beide punten. De verschilvector, d.w.z. de differentiële positie zal dan bijna vrij zijn van die systematische invloeden.

Dergelijke systematische fouten zijn vooral de fouten in de satellietbanen en de fouten veroorzaakt door de atmosferische refractie, met name de ionosfeer en de troposfeer. Ook de invloed van SA kan beschouwd worden als een systematische fout. Gezien de grote hoogte van de GPS satellieten, kan men zelfs bij afstanden tot 1000 km een gunstig resultaat verkrijgen door het gebruik van DGPS.

Als voorbeeld beschouwen we de invloed van de baanfouten. Het blijkt dat de invloed van de radiale component $\Delta\rho$ van de baanfout op een basislijn \overline{AB} ruwweg berekend kan worden met een verhoudingsformule:

$$\frac{\Delta\overline{AB}}{\overline{AB}} = \frac{\Delta\rho}{\rho} = \frac{20m}{20.000km} = 10^{-6} \quad \Rightarrow \quad \Delta\overline{AB} \approx 10^{-6} \overline{AB}$$

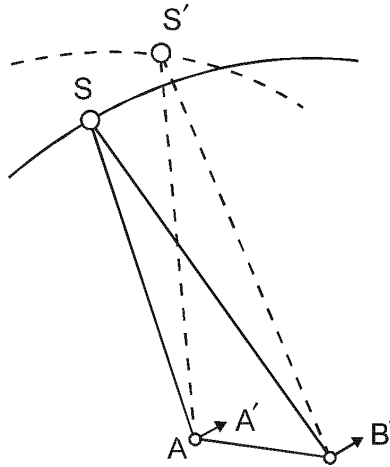
Dit betekent dat (kleine) baanfouten zeer gunstig worden verkleind, in dit voorbeeld met een factor 10^{-6} . Dit betekent dat de invloed van een baanfout van 20 m op een basislijn van 100 km slechts 10cm bedraagt.

In figuur 5.2 wordt een simpel voorbeeld gegeven. Stel dat in de punten A en B simultane GPS waarnemingen worden uitgevoerd. De simultaneïteit houdt in dat dezelfde satellieten gelijktijdig worden waargenomen. De invloed van systematische effecten zal in A en B nagenoeg gelijk zijn. Noem A het referentiestation met bekende WGS-84 coördinaten en B het mobiele station. Er zijn nu diverse methoden mogelijk.

Methode 1: positiecorrectie. De positie van A en B wordt simultaan bepaald met twee gelijke type ontvangers. De positie van het referentiestation A zal meestal een afwijking hebben ten opzichte van de bekende positie. Deze afwijking heeft twee componenten: een systematisch en een stochastisch gedeelte (ruis). De positie van punt B kan gecorrigeerd worden met de positie-afwijking bij station A. De systematische fout wordt hierdoor vrijwel geëlimineerd, maar er wordt wel een extra ruis aan de positie van B toegevoegd. Deze methode blijkt in de praktijk niet erg betrouwbaar te zijn omdat in A en B vaak niet dezelfde satellieten worden waargenomen.

Methode 2: correctie aan de gemeten afstanden. Men krijgt betere resultaten als men correcties aan de code-afstanden aanbrengt. In dit geval berekent men code-afstand correcties in punt A voor alle satellieten. De correctie is gelijk aan gemeten code-afstand

minus berekende code-afstand. De laatste wordt berekend uit de bekende coördinaten van punt A en de satelliet coördinaten. Vervolgens corrigeert men hiermee de gemeten code-afstanden op het onbekende punt B.



Figuur 5.2. De invloed van een baanfout op de positie van stations A en B is vrijwel gelijk.

Deze methode zou in principe hetzelfde resultaat moeten opleveren als de methode met een positie correctie, maar heeft een aantal voordelen: men mag in A en B verschillende type instrumenten gebruiken, men hoeft niet dezelfde software te gebruiken, men heeft een directe controle voor wat de gemeten satellieten betreft en men mag zelfs het aantal satellieten naar eigen behoefte selecteren. Wel dient men altijd dezelfde (recente) ephemeris van de satellieten te gebruiken.

Methode 3: postprocessing. Indien men de DGPS posities niet direct nodig heeft dan kan men de berekening later uitvoeren door middel van postprocessing. Het voordeel van postprocessing is dat men de positie van een of meer stations uit een vereffening kan bepalen. Zo doende heeft men beter vat op de systematische fouten. De meeste gangbare softwarepakketten hebben een optie voor de differentiële berekening uit code-afstanden. De oplossing met postprocessing kan op een vergelijkbare wijze uitgevoerd als de methode van enkele verschillen (single differences) met fase-afstanden. Zie Hoofdstuk 6.

Correcties aan de gemeten code-afstanden

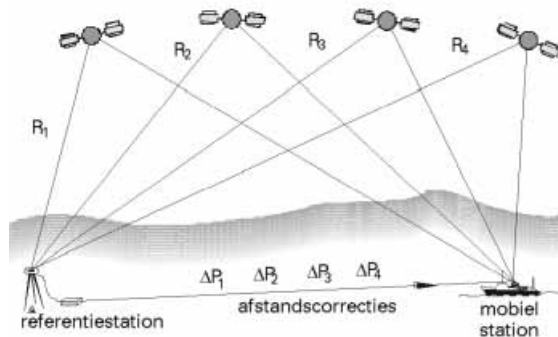
In figuur 5.3 worden simultane waarnemingen van een aantal satellieten voorgesteld op een referentiestation (met bekende coördinaten) en op een onbekende mobiel station, in dit geval een schip.

Het model van code-afstanden op het referentiestation schrijven we afgekort als volgt (zie: vergelijking (5.4)):

$$P_r^s = \rho_r^s + c \delta t_r - c \delta t^s + I_r^s + T_r^s \quad (5.39)$$

waarin de geometrische afstand

$$\rho_r^s = \sqrt{(X^s - X_r)^2 + (Y^s - Y_r)^2 + (Z^s - Z_r)^2} \quad (5.40)$$



Figuur 5.3. Voorstelling van DGPS.

berekend kan worden uit de *bekende* coördinaten van het referentiestation. Hiermee zijn we in staat om voor elke waargenomen satelliet een correctie te berekenen:

$$\delta P_r^s = \rho_r^s - P_r^s = -c \delta t_r + c \delta t^s - I_r^s - T_r^s \quad (5.41)$$

Het model van de code-afstand op een onbekende mobiel station i is:

$$P_i^s = \rho_i^s + c \delta t_i - c \delta t^s + I_i^s + T_i^s \quad (5.42)$$

Uit (5.41) en (5.42) volgt de *gecorrigeerde* code-afstand als nieuwe waarnemings-grootheid:

$$\bar{P}_i^s = P_i^s + \delta P_r^s = \rho_i^s + c \delta t_{ri} + \delta I_{ri}^s + \delta T_{ri}^s \quad (5.43)$$

waarin: $\delta t_{ri} = \delta t_i - \delta t_r$ enz.

Na het aanbrengen van de correctie kan men de positieberekening van het mobiele station met deze gecorrigeerde code-afstanden uitvoeren volgens formule (5.16). De gecorrigeerde code-afstand heeft de volgende eigenschappen:

- er is een gemeenschappelijke (onbekende) klokfout geïntroduceerd;
- de fout in de satelliet klok is volledig geëlimineerd;
- de invloed van SA wordt geëlimineerd;
- de invloed van ionosfeer en troposfeer (en de invloed van de baanfouten) worden gereduceerd naarmate de afstand tot het referentiestation kleiner is;
- de waarnemingsruis neemt toe, omdat in (5.43) twee waarnemingen worden opgeteld.

RTCM formaat

De *Radio Technical Commission for Maritime Services* (RTCM) is opgericht in 1947 te Washington als een adviserende organisatie en houdt zich o.a. bezig met de standaardisatie van data formaten. Deze service heeft onder de titel “RTCM SC-104 DGPS Standard” (SC = Special Committee) een publicatie uitgegeven die speciaal op DGPS is gericht. De eerste versie is reeds verschenen in 1985, daarna volgden een aantal verbeterde versies. Men streeft naar:

- uniform formaat van DGPS correcties;

- een minimaal tijdsinterval tussen correcties;
- eenheid, groot bereik en hoge resolutie van de correcties.

De data typen worden genummerd van type nr. 1 t/m 63. In de volgende tabel worden enkele data typen vermeld.

Type	C/Uitgezonden data
1	C/A-code correcties
2	Delta differentiële correcties
3	Positie referentiestation
7	Almanac informatie
10	P-code correcties
16	Speciale boodschap
18	Fasewaarnemingen
19	Codewaarnemingen
20	Correcties aan fasewaarnemingen
21	Correcties aan codewaarnemingen

Type 1 bevat correcties aan de C/A-code afstanden in binaire formaat. De meeste GPS ontvangers hebben een RTCM aansluiting waarop een radio-ontvanger kan worden aangesloten en de ingevoerde correcties in real-time worden verwerkt.

Tabel 5.1. Voorbeeld van DGPS correcties.

RTCM ASCII formaat (type = 1)					
Tijd [sec]	SV #	UDRE (code)	AODE(baan)	PRC [m]	RRC [m/s]
616.2	19	0	209	18.18	0.192
616.2	17	0	37	15.06	0.016
616.2	23	0	226	-5.62	-0.208
616.2	28	0	149	-17.40	0.128
616.2	26	0	27	-55.88	-0.032
616.2	22	0	68	-16.32	-0.384
616.2	21	0	31	-25.28	0.288
616.2	6	0	76	-17.60	-0.576
616.2	31	0	719	-12.48	0.576
618.0	19	0	209	18.52	0.192
618.0	17	0	37	15.08	0.016
618.0	23	0	226	-5.98	-0.208
618.0	28	0	149	-17.16	0.128
618.0	26	0	27	-55.86	-0.032
618.0	22	0	68	-16.64	-0.352
618.0	21	0	31	-24.64	0.288
618.0	6	0	76	-18.56	-0.576
618.0	31	0	179	-11.52	0.576
621.0	19	0	209	19.10	0.192
621.0	17	0	37	15.02	-0.016
621.0	23	0	226	-6.62	-0.208
...

In tabel 5.1 wordt een voorbeeld gegeven van de correcties aan de code-afstanden in ASCII formaat. Per blok worden hier de correcties voor 9 satellieten uitgezonden. In de eerste kolom wordt de tijdstip in seconden gegeven (elke uur begint met 0), Sv verwijst naar satelliet nummer, UDRE is een kwaliteitscode (0 = OK), AODE (Age Of Data Ephemeris) is een willekeurig nummer dat aan een set baanparameters wordt toegekend, PRC = Pseudo Range Correction en RRC = Range Rate Correction. Men krijgt hier een idee van de grootte van de correcties (bijvoorbeeld bij satelliet #26 is de correctie vrij groot: -55.8 meter!) en van de verandering van de correcties in tijd (bijvoorbeeld bij satelliet #31 is de verandering 0.576 m/s).

De methode van correctie aan de code-afstanden leent zich uitstekend voor real-time toepassingen. De correcties bevatten weinig data: voornamelijk tijdstippen, afstands-correcties en de verandering van de correcties voor een aantal beschikbare satellieten. Deze worden doorgegeven door middel van een telemetrie verbinding. Men dient hier rekening te houden met enige vertraging in de verwerking (latency). Voor het doorgeven van DGPS correcties hanteert men het RTCM internationaal standaard formaat.

De meeste systematische fouten veranderen langzaam, behalve de invloed van de Selective Availability (SA). Als SA geactiveerd is, dient men de correcties in korte tijdsintervallen (liefst iedere 2 sec) toe te passen, omdat SA een hoogfrequente variatie vertoont. Als SA uitgeschakeld zou worden (en er zijn hiervoor plannen) dan zal de toepassing van DGPS correcties veel minder noodzakelijk zijn.

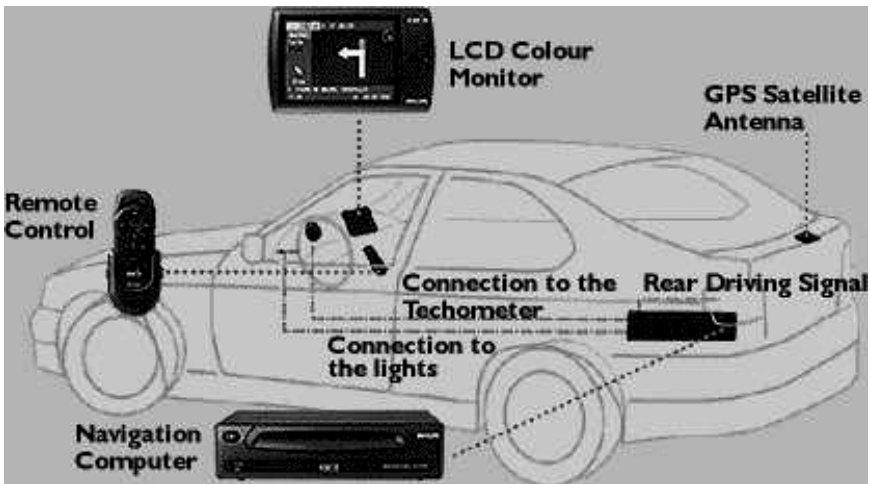
DGPS wordt ook toegepast over grotere gebieden van honderden tot duizenden kilometers. In dit geval spreekt men van Wide Area Differential GPS (WADGPS), waarbij een netwerk van referentiestations wordt opgezet. De precisie en de betrouwbaarheid kunnen worden verbeterd door de correcties van meerdere stations te ontvangen en het gewogen gemiddelde van deze correcties te gebruiken. Bij de toekenning van de gewichten wordt de o.a. de afstand tot de referentiestations in aanmerking genomen.

Relatieve plaatsbepaling met code-afstanden wordt bij statische of kinematische plaatsbepaling toegepast waarbij een nauwkeurigheid op meterniveau voldoende is. Voorbeelden hiervan zijn navigatie, offshore metingen en bepaling van de positie van het vliegtuig ten behoeve van diverse meettechnieken, zoals fotogrammetrie, radaropname en radaraltimetrie.

Diverse soorten navigatiemethoden worden bij voertuigen met GPS (of met DGPS) geïntegreerd. Een typisch voorbeeld uit de praktijk is het CARIN (CAR Information and Navigation) auto-navigatiesysteem.

Het Carin systeem, zoals dat sinds de jaren '80 door Philips Car Systems (sinds begin 1998: VDO Car Communication B.V. (VCC)) is ontwikkeld, maakt gebruik van GPS, gyro (oorspronkelijk kompas) en tachometer (snelheidssensor), zie figuur 5.4. Met behulp van GPS worden de gyro en de tachometer gecalibreerd. Voor de gyro betekent dit, dat het azimut van de rijrichting (*heading*), zoals berekend door de GPS ontvanger, wordt vergeleken met de uitvoer van de gyro. Het verschil tussen de twee geeft de gyro-offset, het verloop in de gyro-uitvoer tussen twee opeenvolgende vergelijkingen van GPS azimut en gyro-uitvoer geeft de gyro-drift. De pulsen van de tachometer, geïntegreerd over een bepaald tijdsinterval en gedeeld door de lengte van dit interval, is een maat voor de snelheid van de auto. Vergelijking van deze grootte met de horizontale snelheid, berekend door de GPS ontvanger, resulteert in de afstand per puls, de schaalfactor voor de

conversie van pulsen of pulsen per seconde naar meters of meters per seconde. Behalve voor calibratie van gyro en tachometer, wordt GPS ook gebruikt voor de bepaling van de absolute positie van de auto. De positie wordt tenslotte gecorrigeerd door de positie te vergelijken met de digitale kaart (*map matching*), het inpassen van de berekende positie van de auto op de meest waarschijnlijke positie op de digitale wegenkaart, die ook onderdeel uitmaakt van het Carin systeem. De bestuurder wordt geïnformeerd via een monitor (zijn positie wordt op een kaart aangegeven) en via een luidspreker, waarbij een stem aanwijzingen geeft voor het bereiken van de gewenste bestemming. Diverse opties, o.a. de bestemming (plaatsnaam en straat), worden m.b.v. de afstandsbediening ingevoerd (zie figuur 5.5).



Figuur 5.4. Het CARIN systeem met drie sensoren: GPS, gyro en tachometer.



Figuur 5.5. Onderdelen van CARIN: navigatie systeem (inclusief GPS-ontvanger, CD-speler en gyro), monitor, luidspreker, afstandbediening en GPS antenne. De bestuurder wordt via de monitor en luidspreker geïnformeerd.

5.9. Kalman filter

De precisie en betrouwbaarheid van de kinematische plaatsbepaling kan verbeterd worden door toepassing van een Kalman filter. Hierbij wordt gebruik gemaakt van de voorgaande positie en snelheid om de actuele positie en snelheid te berekenen. Dit leidt tot een betere plaatsbepaling mede door het feit dat elke nieuwe positie ook met betrekking van de positie getoetst kan worden. Deze methode wordt vooral bij navigatie toegepast. We zullen hier het principe behandelen.

De state vector wordt gedefinieerd door de positie (p), snelheid (v) en versnelling (a) op een epoce t :

$$x(t) = \begin{bmatrix} p(t) \\ v(t) \\ a(t) \end{bmatrix} \quad (5.44)$$

Deze state vector kan worden voorspeld uit de state vector van een voorgaand tijdstip ($t - \Delta t$). We beschouwen een dynamisch model met een constante versnelling:

$$\begin{aligned} p(t) &= p(t - \Delta t) + \Delta t v(t - \Delta t) + \frac{1}{2} \Delta t^2 a(t - \Delta t) \\ v(t) &= v(t - \Delta t) + \Delta t a(t - \Delta t) \\ a(t) &= a(t - \Delta t) \end{aligned} \quad (5.45)$$

ofwel:

$$x'(t) = \begin{bmatrix} p(t) \\ v(t) \\ a(t) \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & \frac{1}{2} \Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p(t - \Delta t) \\ v(t - \Delta t) \\ a(t - \Delta t) \end{bmatrix} = T x(t - \Delta t) \quad (5.46)$$

Deze voorspelling heet "*transition*" van de state vector, en de matrix T "*transition*" matrix. Toepassing van de voortplantingswet der varianties geeft de covariantie matrix van de voorspelde state vector:

$$Q_{x'}(t) = T Q_x T^* + Q_{Trans} \quad (5.47)$$

waarin Q_{Trans} een additionele covariantiematrix is, die de onzekerheid in het dynamisch model (modelfout) in rekening brengt. We hebben nu de voorspelde en de actuele state vector op een tijdstip t , waarvoor de waarnemingsvergelijkingen luiden:

$$E\{x'\} = x \quad ; \quad Q_{x'} \quad (\text{voorspeld volgens 5.46 en 5.47}) \quad (5.48)$$

$$E\{y\} = Ax \quad ; \quad Q_y \quad (\text{waargenomen})$$

We kunnen hieruit de actuele state vector x berekenen door toepassing van een recursieve vereffening. Het model hiervoor luidt:

$$E\left\{\begin{bmatrix} \underline{x}' \\ \underline{y} \end{bmatrix}\right\} = \begin{bmatrix} I \\ A \end{bmatrix} x ; \quad Q = \begin{bmatrix} Q_{x'} & 0 \\ 0 & Q_y \end{bmatrix} \quad (5.49)$$

De oplossing van een recursief model wordt voor een algemeen geval in tabel 5.2 gegeven. Naar analogie van dit model krijgen we de bekende formules voor het Kalman filter uit het model (5.49). De oplossing van het eerste gedeelte voor een tijdstip t luidt:

$$\begin{aligned} \hat{\underline{x}}_{(1)} &= Q_{x'} I Q_{x'}^{-1} x' = x' \\ Q_{\hat{\underline{x}}_{(1)}} &= Q_{x'} \end{aligned}$$

en de update:

$$\hat{\underline{x}} = x' + Q_{x'} A^* (Q_y + A Q_{x'} A^*)^{-1} (y - Ax') \quad (5.50)$$

ofwel:

$$\hat{\underline{x}} = x' + K (y - Ax') \quad (5.51)$$

waarin de matrix K "gain matrix" wordt genoemd. Deze matrix wordt dus berekend volgens:

$$K = Q_{x'} A^* (Q_y + A Q_{x'} A^*)^{-1}. \quad (5.52)$$

De covariantiematrix van $\hat{\underline{x}}$ volgt uit tabel 5.2:

$$Q_{\hat{\underline{x}}} = (I - KA) Q_{x'}. \quad (5.53)$$

Tabel 5.2. Formule systeem voor het recursieve model⁴.

Het gepartitioneerde model:

$$E\left\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}\right\} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x ; \quad Q_y = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}$$

De oplossing van het eerste gedeelte:

$$\begin{aligned} \hat{\underline{x}}_{(1)} &= (A_1^* Q_1^{-1} A_1)^{-1} A_1^* Q_1^{-1} \underline{y}_1 \\ Q_{\hat{\underline{x}}_{(1)}} &= (A_1^* Q_1^{-1} A_1)^{-1} \end{aligned}$$

De update:

$$\begin{aligned} \hat{\underline{x}} &= \hat{\underline{x}}_{(1)} + Q_{\hat{\underline{x}}_{(1)}} A_2^* (Q_2 + A_2 Q_{\hat{\underline{x}}_{(1)}} A_2^*)^{-1} (\underline{y}_2 - A_2 \hat{\underline{x}}_{(1)}) \\ Q_{\hat{\underline{x}}} &= Q_{\hat{\underline{x}}_{(1)}} - Q_{\hat{\underline{x}}_{(1)}} A_2^* (Q_2 + A_2 Q_{\hat{\underline{x}}_{(1)}} A_2^*)^{-1} A_2 Q_{\hat{\underline{x}}_{(1)}} \end{aligned}$$

⁴ Prof.dr.ir. P.J.G.Teuissen: Adjustment Theory, hoofdstuk 6.