GPS OP AFSTAND BEKEKEN

Prof.dr.ir. P.J.G. Teunissen

Het toekomstige NAVSTAR/GPS systeem, dat naar verwachting rond 1992 operationeel zal zijn, zal tot ingrijpende veranderingen leiden op het gebied van de geodetische puntsbepaling. Voor navigatie toepassingen voorziet het systeem via de looptijdmetingen in de mogelijkheid van wereldwijde, continue, snelle drie-dimensionale enkelpuntsbepaling van hoge precisie, 20 m of beter. Voor landmeetkundige toepassingen voorziet het systeem, via de fasemetingsmethode met meer ontvangers, in de mogelijkheid van relatieve puntsbepaling met een precisie in de orde van 10^{-6} of beter. Hierbij komt de min of meer scherpe scheiding tussen situatie en hoogte, gebruikelijk in de traditionele geodesie, te vervallen en behoren de hiermee gepaard gaande tijdvergende verkenningen tot het verleden.

Dit is samengevat, de strekking van veel inleidende artikelen over GPS*. GPS lijkt dus hèt puntsbepalingssysteem van de jaren '90 te worden, met ongekende mogelijkheden voor zowel stationaire als kinematische toepassingen. Bij het bereiken van de globale en continue 2D- en 3D-puntsbepalingscapaciteit, in de eerste helft van respectievelijk 1991 en 1992, kan GPS op operationele basis in

---

*Een helder overzicht van principes en configuratie van het GPS-systeem vindt u in de inleidende tekstboeken:
- GPS: Navigatie en Geodetische puntsbepaling met het Global Positioning System, F.J.J. Brouwer et al., DUP 1989.
- Guide to GPS Potioning, D.E. Wells (ed.), Canadian Associates, 1986.
- Surveying with GPS, R.W. King et al., School of Surveying, University of New South Wales, Australia.

In: Een halve eeuw in de goede richting. Lustrumboek Snellius 1985-1990. pp. 215-233.

de puntsbepaling worden ingevoerd. Tabel 1 geeft een overzicht van enkele stationaire en kinematische toepassingen (met een ruwe afschatting van de precisie-eisen) die dan binnen handbereik komen. Een vijftal van deze toepassingen wordt in dit lustrumboek nader besproken. Zo gaat *ir. D. Boswinkel* in op de mogelijkheden die het GPS-systeem biedt voor de pas-puntbesparing in de luchtfotogrammetrie. Het belangrijke concept van de "Differentiële GPS" en de mogelijkheden die dit concept biedt voor toepassingen op zee, wordt besproken door *ir. P. Sluiter*. De toepassing van GPS in de autonavigatie en de daarmee gepaard gaande problemen in met name een stedelijke omgeving, wordt behandeld door *prof.dr.ir. D. van Willigen*. De laatste twee bijdragen richten zich op het m.b.v. GPS bepalen van verticale en horizontale deformaties. *dr.ir. F.J.J. Brouwer* bespreekt naast de oorzaken en gevolgen van de zeespiegelrijzing, de toepassing van GPS voor het monitoren van de stand van de zeespiegel. *ir. G.J. Husti* gaat tenslotte in op de planningsprocedure van GPS-projecten en bespreekt kort het door de faculteit der Geodesie in 1987 in Griekenland uitgevoerde GPS-project ter bepaling van de lokale stabiliteit van laserstations.

LAND	SEA	AIR
• Geodynamics (plate tectonics, sealevel rise; 0.01 - 0.1 ppm over intercontinental distances)	• Hydrographic surveying (0.1 - 10 m)	• Aero-triangulation (depending on mapping scale: submetre to several meters)
• Continental 3D reference frame (0.1 - 1 ppm)	• Marine 3D seismic surveys (1 - 5 m)	• Airborne laserprofiling (hor.: 1 - 10 m, vert.: 0.5 - 1 m)
• National hor. and vert. control networks (1 - 10 ppm)	• Marine gravity surveys (< 10 cm/s for Eötvös correction < 1 mgal)	• Airborne gravimetry (hor.: 50 m, vert.: 2 m, vel.: 10 cm/s)
• Surface and platform subsidence monitoring (1 ppm)	• Navigation in coastal area (50 - 100 m)	• Airborne laser bathymetry (hor.: 15 m)
• Large scale topography (10 - 100 ppm)	• Navigation in open waters (1 - several km)	• Airtransport terminal approach (hor.: 0.1 - 0.5 km)
• Land navigation (10 - 50 m)		• Airtransport terminal area (hor.: 0.5 - 1 km)
		• Airtransport en route (1 - several km)

tabel 1

Met de ongekende mogelijkheden van GPS voor zowel stationaire als kinematische toepassingen in het verschiet, is het niet ver wonderlijk dat in menig onderzoeksinstituut, waaronder het Laboratorium voor Geodetische Rekentechniek (LGR) van de faculteit der Geodesie, koortsachtig gewerkt wordt aan de verdere modelontwikkeling voor het GPS-systeem. Op het gebied van de theorievorming met haar essentiële praktische invulling hebben we echter nog een lange weg te gaan; adequate en theoretisch onderbouwde verkenningsregels ontbreken nog, de ontwikkeling van een referentiekader voor de verwerking en aansluiting van GPS-metingen staat nog in de kinderschoenen, en een op GPS toegesneden toetsingstheorie kan nog niet systematisch toegepast worden vanwege het gammele kansmodel en het gebrek aan inzicht in de te kiezen verzameling van alternatieve hypothesen. Behandeling van deze onderzoeksthema's, welke deel uitmaken van het 3D-puntsbepalingsonderzoek van het LGR, vergt uiteraard meer ruimte dan mij voor deze lustrumbijdrage is toegewezen. Ik zal me daarom beperken tot een inleidende en elementaire discussie van een aantal in de GPS-literatuur gepropageerde *vereffeningsconcepten.*

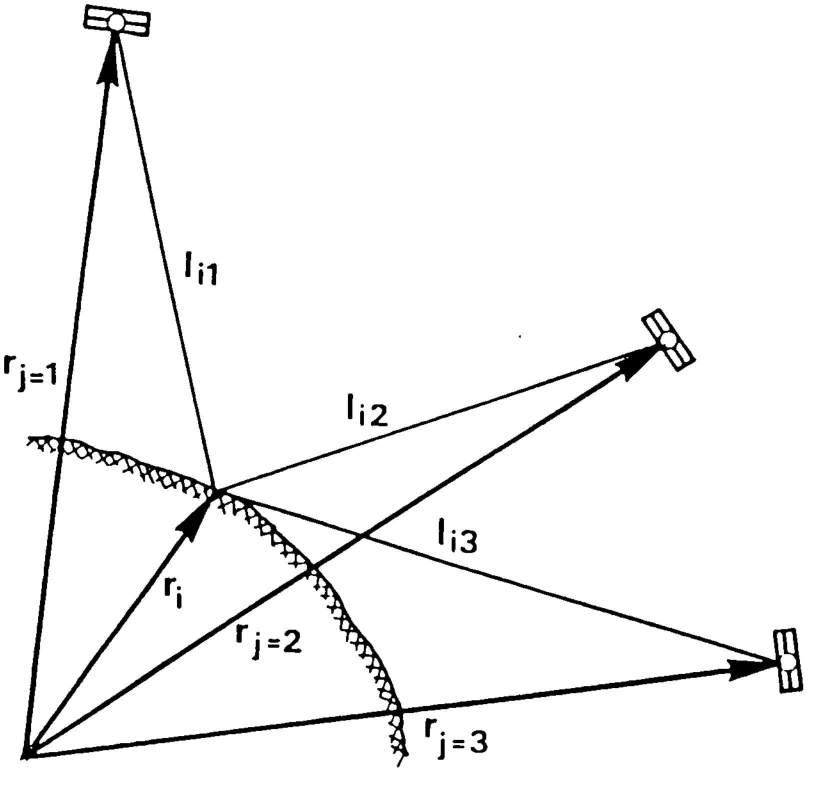
Afstandmeting

De enkelpuntsbepaling met behulp van GPS is gebaseerd op looptijdmeting van de door de GPS-satellieten uitgezonden signalen. In de GPS-ontvanger wordt, op ingenieuze wijze, de tijd τ gemeten die het GPS-signaal nodig heeft om de afstand satelliet - ontvanger te overbruggen. Onder ideale omstandigheden kan dan de afstand ℓ_{ij} tussen ontvanger *i* en satelliet *j* berekend worden als het produkt van de gemeten looptijd τ en de bekend verondersteldde voortplantingssnelheid van licht c: $c\tau=\ell_{ij}$. Lineariseren van deze vergelijking naar de ontvangerpositie r_i en satellietpositie r_j geeft dan de bij de looptijdmeting behorende waarnemingsvergelijking: $c\Delta\tau=\Delta\ell_{ij}=c_{ij}^{*}\Delta r_{ij}$ (* staat voor "getransponeerd"). Hierbij is c_{ij} gelijk aan de genormeerde richtingsvector $r_{ij}/|r_{ij}|$ en r_{ij} gelijk aan het verschil r_j-r_i.

Voor de puntsbepaling zijn we voornamelijk geïnteresseerd in de bepaling van Δr_i, de onbekende positie van de ontvanger. Dit wordt echter problematisch als naast Δr_i ook nog de Δr_j als onbekende parameters zouden optreden. Gelukkig bevat het door de GPS-satellieten uitgezonden signaal tevens informatie over r_j, de positie van de satelliet. Deze informatie is van dien aard dat voor veel toepassingen de satellietpositie bekend verondersteld mag worden. Met $\Delta r_j=0$, reduceert de bij de looptijdmeting behorende waarnemingsvergelijking tot:

$$\Delta \ell_{ij} = -e_{ij}^{*} \Delta r_i \,. \tag{1}$$

Om de drie onbekende ontvanger coördinaten, welke vervat zijn in de vector Δr_i, eenduidig te kunnen berekenen, hebben we drie vergelijkingen van het type (1) nodig. Met andere woorden, meting vanuit ontvanger i naar drie satellieten $j=1$, $j=2$, $j=3$ is nodig en voldoende om de positie van de ontvanger te kunnen bepalen (zie figuur 1). Iedere extra gemeten afstand levert een additionele overtalligheid en daardoor de noodzaak om tot een vereffening over te gaan.



Figuur 1: Enkelpuntsbepaling.

Uitgaande van het principe van de kleinste-kwadraten vereffening volgt, onder de aanname dat de waarnemingen ongecorreleerd zijn en alle een gelijke variantie σ^2 hebben, dat de formele precisie van de berekende ontvangercoördinaten beschreven wordt door de drie-bij-drie variantie matrix:

$$Q_{ri} = \sigma^2 \left[\sum_{j=1}^{n} e_{ij}\, e_{ij}^{*} \right]^{-1}. \tag{2}$$

Hierbij is n gelijk aan het aantal aangemeten satellieten. De precisie van de berekende ontvangercoördinaten wordt voor een belangrijk deel bepaald door de geometrie van de satelliet-configuratie. De stand en relatieve oriëntering van de n eenheidsvectoren e_{ij} speelt hierbij een beslissende rol. Indien er bijvoorbeeld een vector a zou bestaan zodanig dat $e_{ij}^{\bullet}.a \approx 0$ voor $j = 1,\ldots,n$, dan liggen alle vectoren e_{ij} (bijna) in een vlak en heeft de bij de waarnemingsvergelijkingen (1) behorende ontwerpmatrix een (bijna) singulariteit in de richting a. De ontvangerpositie is dan slecht of helemaal niet bepaald in de richting van de vector a.

Bijvoorbeeld, als alle vectoren e_{ij} ongeveer evenwijdig met de lokale horizon liggen dan richt de vector a zich naar het lokale zenit en is de hoogtecomponent van de ontvangerpositie slecht of helemaal niet bepaald. (Zie ook de bijdrage van *dr. F. Brouwer* over het meten van de zeespiegelrijzing met behulp van GPS). Voor een goede precisie is dus een voldoende spreiding van de eenheidsvectoren e_{ij} over de eenheidsbol, geslagen om de ontvangerpositie, noodzakelijk.

Een in de navigatie literatuur veel gebruikte maat voor de spreiding van de eenheidsvectoren e_{ij} wordt gegeven door:

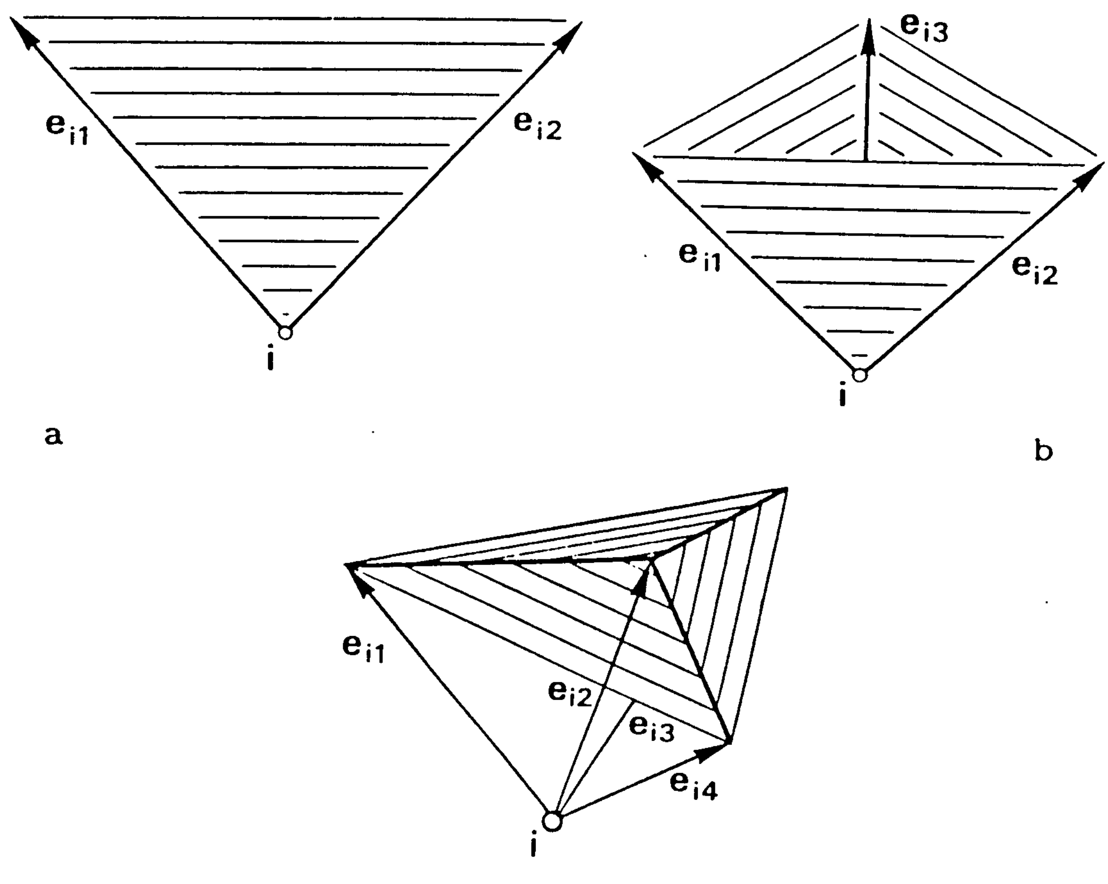
$$\text{PDOP} = (\sigma^{-2} \text{ spoor } Q_{ri})^{1/2} , \tag{3}$$

waarbij de afkorting PDOP staat voor <u>P</u>osition <u>D</u>ilution <u>O</u>f <u>P</u>recision. Voor het geval er van overtalligheid geen sprake is en de ontvangerpositie dus eenduidig bepaald is, is PDOP eenvoudig geometrisch te interpreteren. We maken hiervoor gebruik van het feit, dat de inverse van een matrix gelijk is aan de matrix van cofactoren gedeeld door de determinant en dat de determinant van een matrixprodukt gelijk is aan het produkt van de determinanten. Met de commutativiteitseigenschap van het spoor volgt dan, dat het kwadraat van PDOP geschreven kan worden als:

$$\text{PDOP}^2 = \frac{\text{spoor}(\text{cofactor}[e_{ij}^{\bullet}\, e_{ik}])}{(\det[e_{ik}])^2} \tag{4}$$

waarbij $e_{ij}^{\bullet}.e_{ik} = 1$ als $j = k$ en $e_{ij}^{\bullet}.e_{ik} = \cos(e_{ij}, e_{ik})$ als $j \neq k$. Voor het tweedimensionele geval is de teller van (4) gelijk aan 2 en is $\det[e_{ik}]$ gelijk aan de sinus van de door de twee eenheidsvectoren e_{i1} en e_{i2} ingesloten hoek. Hieruit volgt dat voor het twee-dimensionale geval PDOP gelijk is aan de reciproque waarde van het oppervlak opgespannen door de twee eenheidsvectoren e_{i1} en e_{i2}: $\text{PDOP} = 1/\text{oppervlak}(e_{i1}, e_{i2})$ (zie figuur 2a).

Dit eenvoudige en inzichtelijke resultaat doet vermoeden dat PDOP voor het meer realistische drie-dimensionale geval wel eens gelijk kan zijn aan de reciproque waarde van het door de drie eenheidsvectoren e_{i1}, e_{i2}, e_{i3} opgespannen volume. In artikelen en tekstboeken handelend over GPS komt men dan ook wel eens de bewering tegen dat $PDOP=1/volume(e_{i1},e_{i2},e_{i3})$. Dit resultaat is echter helaas onjuist. Het juiste antwoord is vrij gemakkelijk uit (4) af te leiden en luidt: $PDOP=1/3[\sqrt{\{\Sigma(O_j)^2\}}]/[volume(e_{i1},\ e_{i2},\ e_{i3})]$, waarbij O_j, $j=1,2,3$, de oppervlakken zijn van de drie door de eenheidsvectoren e_{ij}, $j=1,2,3$, opgespannen zijvlakken (zie figuur 2b).



Figuur 2: PDOP

Wanneer we PDOP vermenigvuldigen met σ en delen door 3 krijgen we een maat voor de gemiddelde standaardafwijking. Het PDOP-concept komt dan ook overeen met het door de geodeet F.R. Helmert [1843-1917] geïntroduceerde, en nog steeds in de Duitse geodetische literatuur veel gebruikte, concept van de "Mittler Punktfehler". Het voordeel van het PDOP-concept is dat het een eenvoudig te berekenen scalaire maat betreft. Dit is echter tegelijkertijd ook haar grootste nadeel. Hoewel het mogelijk is om met PDOP eenvoudig defecten in de satelliet-configuratie aan te wijzen (immers, als de satellieten in een vlak liggen dan geldt dat het volume opgespannen door de vectoren e_{ij} nul en de PDOP dus oneindig is) moet er met het PDOP-concept toch met de nodige voorzichtigheid omgesprongen worden.

Voor een precisiebeschrijving van de berekende ontvangerpositie is het PDOP-concept ontoereikend. Immers, de PDOP-waarde veronachtzaamt de correlatie tussen de berekende coördinaten van de ontvanger. Ook voor verkenningsberekeningen dient men het PDOP-concept voorzichtig te hanteren. Een in de literatuur veel gelezen argument voor het gebruik van PDOP luidt, dat bij een goede satelliet-configuratie, namelijk een waarbij de satellieten met een niet te grote elevatiehoek regelmatig over de horizon zijn verspreid, een kleine PDOP-waarde hoort. Dit is juist. Het omgekeerde is echter niet per se juist. Dat wil zeggen, dat bij een kleine PDOP-waarde geen gunstige satelliet-configuratie hoeft te horen, tenminste, als men waarde hecht aan een zekere homogeniteit in de precisie van de ontvangercoördinaten.

Het is niet moeilijk om in te zien dat de PDOP invariant is onder een willekeurige rotatie van de satelliet-configuratie. Dat dit geldt voor de eenduidige enkelpuntsbepaling volgt direct uit de gegeven geometrische interpretatie. Het geldt echter ook voor de overtallige enkelpuntsbepaling. Daar het spoor commutatief is in zijn argument en de getransponeerde van een rotatiematrix gelijk is aan zijn inverse, volgt voor een willekeurige rotatiematrix R namelijk dat:

$$\text{spoor}[\Sigma\, R\, e_{ij}\, e_{ij}^{*}\, R^{*}]^{-1} = \text{spoor}[\Sigma\, e_{ij}\, e_{ij}^{*}]^{-1}\,. \tag{5}$$

PDOP is dus een maat voor de uitgestrektheid van de satelliet-configuratie en niet een maat voor de stand van deze configuratie ten op zichte van het zenith ter plekke van de ontvanger. De stand van de satelliet-configuratie heeft echter wel invloed op de homogeniteit in de precisie van de ontvangercoördinaten.

Pseudo afstandmeting

Waarnemingsvergelijking (1) is geldig als we ervan uit mogen gaan dat we over een voldoende precieze klok in de ontvanger beschikken. Wanneer dit niet het geval is dienen we (1) te vervangen door:

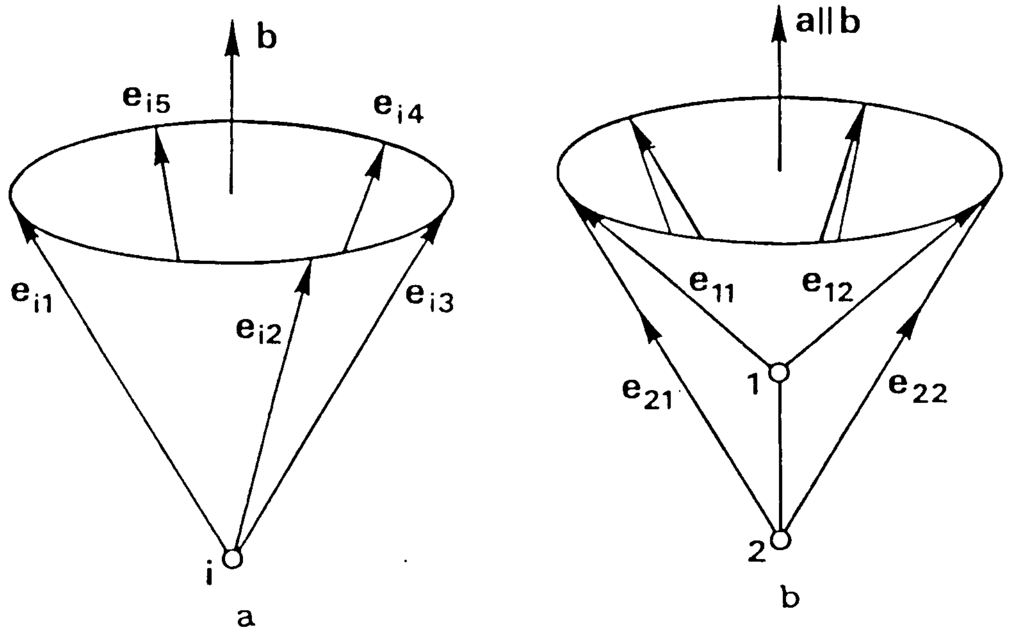
$$\Delta \ell_{ij} = -c_{ij}^{*} \Delta r_i + c \Delta t , \tag{6}$$

waarbij Δt het tijdsverschil tussen de ontvangertijd en de satellietentijd modelleert. Als (6) geldig is spreekt men in de navigatie literatuur over pseudo-afstandmeting. Merk op, daar de afstanden ℓ_{ij} tussen ontvanger i en GPS-satellieten j qua orde van grootte constant zijn ($=2*10^4$ km), dat de term $(c/\ell_{ij})\Delta t$ bij benadering als een constante schaalfactor geïnterpreteerd kan worden. De pseudo-afstandmeting uit de navigatie literatuur komt dan ook mooi overeen met de pseudo-afstandmeting zoals we die in de landmeetkunde kennen.

Bij simultaan uitgevoerde metingen en stabiele satellietklokken, mogen we Δt constant veronderstellen. In dat geval hebben we vier onbekende parameters en dus minimaal vier metingen nodig om de ontvangerpositie en de klokterm Δt te bepalen. Iedere extra meting levert weer een extra overtalligheid. De bij de kleinste-kwadraten vereffening behorende variantiematrix van de berekende ontvangerpositie wordt nu verkregen door de eenheidsvectoren e_{ij} in (2) te vervangen door de vectoren $\bar{e}_{ij}$, waarbij de $\bar{e}_{ij}$ gelijk zijn aan het verschil van de e_{ij} met de gemiddelde richting van alle n eenheidsvectoren e_{ij}, $j=1,..,n$.

De precisie van de berekende ontvangerpositie in het geval van pseudoafstandmeting is over het algemeen wat slechter dan in het geval van afstandmeting. De variantiematrix van pseudoafstandmeting is alleen identiek aan die van (2) als de som van alle n eenheidsvectoren e_{ij} gelijk is aan de nulvector. Dit kan, daar de satellieten onder de horizon niet aangemeten kunnen worden, echter alleen voorkomen bij eenheidsvectoren e_{ij} waarvan de elevatiehoek nul graden is, met als gevolg dat de hoogtecomponent van de ontvangerpositie onbepaalbaar wordt. In alle andere gevallen zal de som van de eenheidsvectoren e_{ij} altijd een component bevatten in de richting van het lokale zenith.
In het geval van afstandmeting hebben we gezien dat voor een goede precisie een voldoende spreiding van de eenheidsvectoren e_{ij} over de eenheidsbol noodzakelijk is. Geldt dit nu ook voor de pseudo-afstandmeting? Sommige tekstboeken beweren van wel. Dit antwoord is echter op zijn minst onvolledig.

Bij de afstandmeting hebben we gezien dat er een singulariteit optreedt als de eenheidsvectoren alle in een vlak liggen. Dit geldt onverminderd ook voor de pseudo-afstandmeting. Bij de pseudo-afstandmeting kan echter nog een ander type singulariteit optreden. De bij de waarnemingsvergelijkingen (6) behorende ontwerpmatrix is namelijk ook singulier als er een vector b zou bestaan zodanig dat $e_{ij}^{*}.b=1$ voor $j=1,...,n$. Deze vergelijking kunnen we ook schrijven als $|b|\cos(e_{ij},b)=1$. We zien dan onmiddellijk dat er sprake is van een singulariteit, als de vector b een willekeurige, maar constante, hoek maakt met alle eenheidsvectoren e_{ij} (zie figuur 3a). De ontvangerpositie is dan onbepaald in de richting van de vector b. Voor een satelliet-configuratie met een (bijna) gelijke elevatie voor iedere satelliet betekent dit, dat de hoogte van de ontvangerpositie slecht of helemaal niet bepaalbaar is. In het geval van pseudo-afstandmeting mogen dus zowel de vectoren e_{ij} zelf, als hun eindpunten niet in een vlak liggen. Voor de eenduidige enkelpuntsbepaling volgt dit ook direct uit de geometrische interpretatie die aan de PDOP gegeven kan worden.



Figuur 3: Positie onbepaald in richting b.

Volkomen analoog aan de gevolgde afleiding bij de afstandmeting kan men eenvoudig laten zien, dat de PDOP van de pseudo-afstandmeting gelijk is aan de verhouding van eenderde maal de wortel van de kwadratische som van de vier oppervlakken van het door de eindpunten van e_{i1}, e_{i2}, e_{i3}, e_{i4}, bepaalde viervlak en het bij dit viervlak behorende volume (zie figuur 2c). Dus PDOP-$=\infty$ als de vectoren e_{i1}, e_{i2}, e_{i3}, e_{i4} in een vlak liggen of als hun eindpunten in een vlak liggen.

Tot nu toe hebben we stilzwijgend aangenomen dat de variantiematrices van de enkelpuntsbepaling een realistische precisiebeschrijving van de ontvangerpositie geven. Wil dit echter het geval zijn, dan dienen we de onvermijdelijke stochastiek in de "vastgehouden" grootheden in rekening te brengen. Dit is in essentie volkomen analoog aan de methode van precisiebeschrijving welke, in de mathematische geodesie, gevolgd wordt bij het aansluiten van landmeetkundige netwerken aan netwerken van hogere orde (bijvoorbeeld het RD-netwerk).

Bij de landmeetkundige aansluiting wordt op grond van praktische overwegingen het netwerk van hogere orde vrijwel altijd onveranderd gelaten. Met andere woorden, het netwerk van hogere orde wordt in de vereffening "vastgehouden". We spreken dan over een "pseudo kleinste-kwadraten" vereffening. De term "pseudo" wordt gebruikt om aan te geven dat er van een strenge vereffening geen sprake is. Immers, in de vereffening wordt de onvermijdelijke stochastiek in de coördinaten van het hogere orde netwerk verwaarloosd.

Hoewel deze stochastiek in de vereffening verwaarloosd wordt, betekent het natuurlijk niet, dat dit ook is toegestaan bij het toepassen van de, voor de toetsingstheorie en kwaliteitsbeschrijving zo belangrijke, voortplantingswet der varianties en covarianties. Bij de precisie-beschrijving van de landmeetkundige pseudo kleinste-kwadraten aansluiting wordt dan ook de stochastiek van het hogere orde netwerk via een geschikt gekozen en uit de criterium-theorie afgeleide, vervangingsmatrix in rekening gebracht. Vertaald naar het probleem van de precisie-beschrijving van de ontvangerpositie betekent dit, dat we in staat dienen te zijn een adequate vervangingsmatrix op te stellen voor alle in de vereffening "vastgehouden" grootheden. We kunnen hierbij denken aan bijvoorbeeld de satellietbanen en de ionosferische- en troposferische refractie. Helaas ontbreekt het ons nog aan een dergelijke vervangingsmatrix.

Om toch een idee te krijgen van de orde van grootte van de haalbare precisie van de looptijdmetingen, nemen we voor de eenvoud aan dat de vervangingsmatrices van de "vastgehouden" grootheden door geschaalde eenheidsmatrices beschreven mogen worden. Dit betekent, dat we de correlatie verwaarlozen en een homogene precisie in de verschillende coördinaatrichtingen veronderstellen. Toepassing van de voortplantingswet laat dan zien dat de structuur van de variantiematrix van de ontvangerpositie ongewijzigd blijft en alleen de variantiefactor σ^2 vervangen dient te worden. Dat wil zeggen, σ^2 bevat nu niet alleen de variantie van de metingen maar ook de varianties van de in de vereffening "vastgehouden" grootheden. De totale variantie factor wordt nu verkregen als:

$$\sigma^2 = \sigma^2_{(s)} + \sigma^2_{(p)} + \sigma^2_{(o)} \, . \tag{7}$$

waarin, σ_s = standaardafwijking satelliet-effecten, σ_p = standaardafwijking propagatie-effecten en σ_o = standaardafwijking ontvanger-effecten. Een ruwe afschatting, voor civiele toepassingen, van de verschillende elementen van σ^2 is $\sigma_s=2$ m, $\sigma_p=5$-16 m en $\sigma_o=11$ m. Hiermee volgt voor de enkelpuntsbepaling, dat σ = 12-20 m.

Dit ongekend goede precisieniveau voor de real-time enkelpuntsbepaling met behulp van looptijdmetingen laat zien, dat het GPS-systeem de potentie heeft zich te ontwikkelen tot hèt globale puntsbepalingssysteem voor (bijna) alle toepassingen die liggen op het gebied van de land-, zee- en lucht*navigatie*. Helaas heeft het U.S. Department of Defense echter besloten om voor de nog te lanceren GPS-satellieten het concept van "*Selective Availability*" toe te passen. Hierbij wordt de precisie van zowel de uitgezonden baan-ephemeriden als de satellietklokken opzettelijk gedegradeerd. De mate van degradatie is afhankelijk van de door het U.S. Department of Defense gemaakte militaire overwegingen, maar zal zich vooralsnog richten op een precisie van 100 m in de (horizontale) enkelpuntsbepaling voor 95% van de tijd. Alléén de (bevriende) partijen, welke over de juiste code beschikken, zullen in staat zijn om de effecten van de kunstmatige degradatie te niet te doen. Voor een bespreking van enkele consequenties van "Selective Availability" wordt verwezen naar de lustrumbijdragen van *prof.dr.ir. D. van Willigen* en *ir. P. Sluiter.*

Naast precisie speelt uiteraard ook betrouwbaarheid een belangrijke rol in de land-, zee- en luchtnavigatie. (Betrouwbaarheid van een navigatiesysteem wordt in de navigatieliteratuur ook wel aangeduid als de "integriteit" van het systeem). Veiligheidsvoorschriften in de civiele luchtvaart schrijven bijvoorbeeld voor dat het navigatiesysteem uitgerust dient te zijn met een procesbewakingssysteem voor het tijdig signaleren van kwaliteitsdegradatie in de geschatte positie. Een procesbewakingssysteem ten behoeve van betrouwbaarheid kan bijvoorbeeld gebaseerd worden op de recentelijk bij het LGR ontwikkelde recursieve DIA-theorie voor de real-time *Detectie*, *Identificatie* en *Adaptatie* van in navigatiesystemen optredende modelfouten. Een dergelijke procesbewaking is echter pas mogelijk als men de beschikking heeft over verschillende elkaar aanvullende en controlerende positiesensoren. Voor de civiele luchtvaart betekent dit, dat GPS nooit als een "stand alone" of "solemeans" navigatiesysteem toegestaan zal worden. Het zal samen met bijvoorbeeld traagheidssystemen, Loran-C of het GLONASS-systeem (het "GPS-systeem" van de Sovjet-Unie) deel moeten gaan uitmaken van een *geïntegreerd navigatie systeem*. (Zie ook de lustrumbijdrage van *ir. M.A. Salzmann* over "Dynamische Gegevensverwerking" en de lustrumbijdrage van *prof.dr.ir. D. van Willigen* over het belang van de geïntegreerde navigatie voor landverkeer.)

Afstandverschilmeting

Naast looptijdmeting is het met sommige (zogenaamde geodetische) GPS-ontvangers ook mogelijk om de (fractionele) fase van de door de satelliet uitgezonden draaggolf te meten. De fase ϕ van een harmonische beweging is gedefinieerd als: $\phi = ft + \Phi$, waarbij ϕ voor frequentie, t voor tijd en Φ voor initiële fase staat. De fase-waarneming ϕ_{ij} tussen ontvanger *i* en satelliet *j* bestaat nu uit het verschil van de fase ϕ_j en ϕ_i. De fase ϕ_j behoort bij het door de satelliet *j* uitgezonden signaal en de fase ϕ_i is van het in de ontvanger *i* gegenereerde, in frequentie identieke, signaal modulo een periode. Het faseverschil wordt dus gegeven door $\phi_{ij} = \phi_j - \phi_i = f(t_j - t_i) + \Phi_j - \Phi_i + (\text{mod } 1 \text{ per})$. Er is sprake van modulo een periode, omdat de meting een fractionele fase betreft en niet een complete fase.

Noteren we $-(c/f)\phi_{ij}$ als ℓ_{ij}, de modulo één periode als N_{ij}, en het tijdsverschil tussen de ontvangertijd en satelliettijd als Δt, dan volgt de gelineari-seerde waarnemingsvergelijking voor de fase-waarneming als:

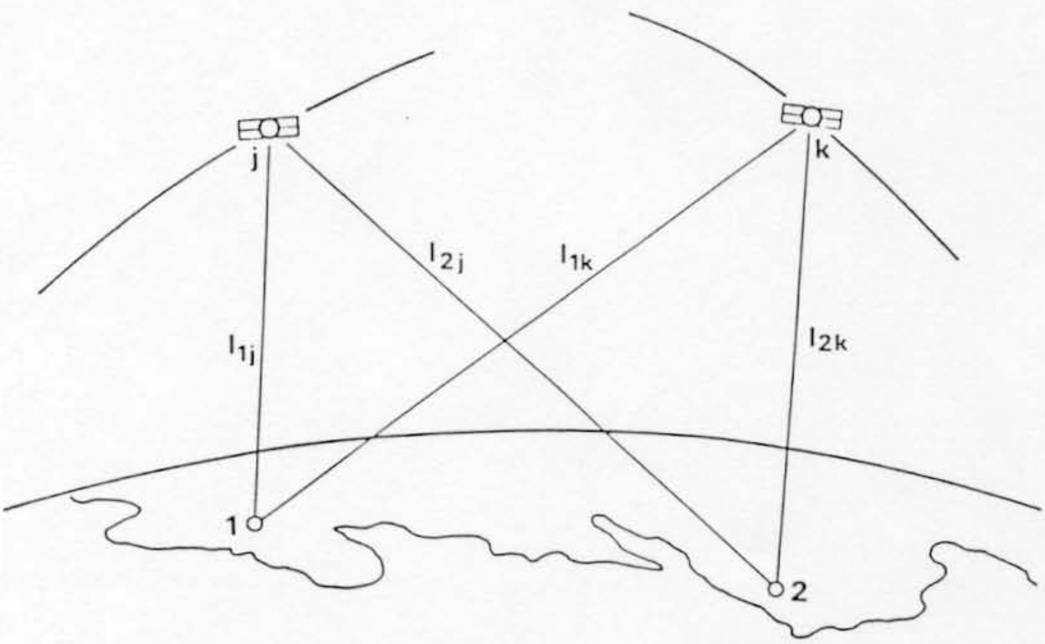
$$\Delta\ell_{ij} = e_{ij}^{\bullet} \Delta r_{ij} + c\,\Delta t + (c/f)\,(\Phi_i - \Phi_j - N_{ij}). \tag{8}$$

Substitueren we weer $\Delta r_j = 0$ en nemen we aan dat we over voldoende precieze klokken beschikken, dus $\Delta t = 0$, dan laat een vergelijking tussen (8) en (6) zien dat de fasemeting in dit geval ook als een pseudo-afstandmeting geïnterpreteerd kan worden. Immers, de laatste term in (8) is bij continue meting constant voor een vast ontvanger-satellietenpaar. Fase-waarneming biedt dan in principe ook de mogelijkheid om enkelpuntsbepaling uit te voeren.

Waarom zien we deze methode van pseudo-afstandmeting dan toch niet als zelfstandig meetprincipe in de navigatiekunde toegepast? De reden is, dat de methode een snelle puntsbepaling onmogelijk maakt. Immers, (8) met $\Delta r_j = 0$ en $\Delta t = 0$, is alleen voor een vast ontvanger-satellietenpaar als pseudo-afstand te interpreteren. Om een voor de navigatie acceptabel precisie-niveau in de berekende ontvangercoördinaten te garanderen dient dus een vast gekozen satelliet voldoende lang gevolgd te worden. Daarmee zijn we er echter nog niet. Met het volgen van een satelliet hebben we (vanwege de relatief langzame rotatie van de aarde) immers alleen nog maar de mogelijkheid om de ontvangerpositie in het baanvlak van de satelliet te bepalen. Een tweede satelliet (liefst met een baanvlak dat loodrecht staat op het baanvlak van de eerste) is dan nog nodig om tot een drie-dimensionale enkelpuntsbepaling te komen.

Voor de navigatie is de fasemeting als *zelfstandig* meetprincipe dus minder geschikt. (Fasemeting *geïntegreerd* met pseudo-afstandmeting vindt natuurlijk wel zijn toepassing in de navigatie.) Voor landmeetkundige toepassingen daarentegen zijn de fasemetingen, in tegenstelling tot de looptijdmetingen, bijzonder geschikt. De fasemetingen zijn ongeveer twee ordes preciezer dan de beste looptijdmetingen. Bovendien maakt simultane fasemeting met meer ontvangers het mogelijk om de schadelijke invloed van modelfouten in de vastgehouden satelliet-banen en atmosferische modellen voor een belangrijk deel te dempen. Een relatieve precisie in de orde van 10^{-6} lijkt voor veel landmeetkundige toepassingen alhaalbaar.

Bij de landmeetkundige toepassing van GPS is het model van waarnemingsvergelijkingen impliciet opgebouwd uit verschillen van afstanden tussen ontvanger- en satellietposities. Om dit te illustreren beschouwen we weer vergelijking (8). Veronderstel dat vanuit twee verschillende lokaties met twee ontvangers, 1 en 2, simultaan naar dezelfde satelliet j gemeten kan worden. Nemen we het verschil van beide metingen dan volgt dat in dit verschil de initiële fase Φ_j van het satellietsignaal geëlimineerd is en de klokterm alleen nog maar de instabiliteit bevat van de ontvangersklokken.



Figuur 4: Dubbel-verschilmeting.

Veronderstellen we verder, dat een dergelijke verschilmeting op hetzelfde tijdstip ook naar satelliet k wordt uitgevoerd, dan volgt uit het verschil van

$\Delta(\ell_{1j}-\ell_{2j})$ en $\Delta(\ell_{1k}-\ell_{2k})$ de zogenaamde "double difference" (zie figuur 4):

$$\Delta\{(\ell_{1j}-\ell_{2j})-(\ell_{1k}-\ell_{2k})\} = e_{1j}^{\bullet}\Delta r_{1j}-e_{2j}^{\bullet}\Delta r_{2j}-e_{1k}^{\bullet}\Delta r_{1k}+e_{2k}^{\bullet}\Delta r_{2k}+N_{12jk} \qquad (9)$$

met $N_{12jk}=-N_{1j}+N_{2j}+N_{1k}-N_{2k}$. In dit dubbele verschil zijn op de meerduidigheid N_{12jk} na, alle niet-geometrische grootheden geëlimineerd. De meerduidigheid N_{12jk}, welke constant is bij een continu volgen van het satellietenpaar (j,k), kunnen we tenslotte nog elimineren door het nemen van verschillen tussen de "double differences". De zo verkregen, van niet-geometrische grootheden gevrijwaarde, verschillen worden in de GPS-literatuur aangeduid als "triple differences".

Om enig inzicht in de betekenis van deze afstandsverschillen te verkrijgen beschouwen we voor de eenvoud het enkele afstandsverschil:

$$\Delta(\ell_{1j}-\ell_{2j}) = e_{1j}^{\bullet}\ \Delta r_{1j} - e_{2j}^{\bullet}\ \Delta r_{2j}\ . \qquad (10)$$

Eventuele singulariteiten van (10) kunnen dan vrij gemakkelijk vertaald worden naar singulariteiten behorende bij de zogenaamde dubbele en drievoudige afstandsverschillen. Nemen we weer aan dat $\Delta r_j=0$, dan zijn in principe 6 vergelijkingen van het type (10) nodig en voldoende om de twee ontvangerposities r_1 en r_2 te bepalen.

Bij bepaalde configuraties treden er echter weer singulariteiten op. De bij de waarnemingsvergelijkingen (10) behorende ontwerpmatrix is singulier als er vectoren a en b bestaan, zodanig dat $e_{1j}^{\bullet}.a=e_{2j}^{\bullet}.b$ voor $j=1,...$ Hieruit volgt direct dat er sprake is van een singulariteit als $e_{1j}^{\bullet}.a=0$ en/of $e_{2j}^{\bullet}.b=0$. Dit betekent natuurlijk weer, dat een voldoende spreiding van de satellietposities noodzakelijk is. Tevens volgt ook dat er sprake is van een singulariteit als $e_{1j}^{\bullet}.a=1$ èn $e_{2j}^{\bullet}.b=1$. Dit is het geval als de verschilvector r_{12} een willekeurige maar constante hoek maakt met de twee sets van eenheidsvectoren e_{1j} en e_{2j}. Als de ontvangerposities zich op de symmetrie-as bevinden van de twee door de eenheidsvectoren e_{1j} en e_{2j} gevormde kegels (zie figuur 3b), doet zich dus een singulariteit voor.

Naast de twee genoemde type singulariteiten, treedt er in de praktijk nog een belangrijk derde type singulariteit op. We beschouwen daartoe het verschil $e_{1j}-e_{2j}$. Realiseren we ons nu, dat de afstanden ℓ_{1j} en ℓ_{2j} qua orde van grootte nagenoeg constant zijn, namelijk $\ell_{1j}=\ell_{2j}=\ell=2*10^4$ km en benaderen we vervolgens e_{1j} met $(1/\ell)r_{1j}$ en e_{2j} met $(1/\ell)r_{2j}$, dan volgt de afschatting:

$$e_{1j} - e_{2j} = r_{12} / \ell. \qquad (11)$$

Dit resultaat laat zien dat het verschil e_{1j}-e_{2j} voor relatief korte ontvangerafstanden praktisch verwaarloosd kan worden (r_{12}/ℓ = 10^{-4} voor r_{12} = 2 km). De consequenties hiervan kunnen we het beste illustreren, als we (10) herschrijven tot:

$$\Delta(\ell_{1j} - \ell_{2j}) = e_{2j}^{\bullet}\,\Delta r_{12} + (e_{1j}^{\bullet} - e_{2j}^{\bullet})\Delta r_{1j}. \tag{12}$$

Uit het feit dat de tweede term in het rechterlid van (12) praktisch verwaarloosbaar is voor relatief korte ontvangerafstanden, kunnen we nu een tweetal belangrijke conclusies afleiden.

Ten eerste laat (12) zien dat het met afstandsverschilmeting (en dus ook met de dubbele en drievoudige varianten) *praktisch* onmogelijk is de ontvangerpositie r_1 te bepalen. Dus hoewel r_1 *theoretisch* wel te bepalen is, zal men in de praktijk bij het opstellen van de normaalvergelijkingen een numerieke singulariteit aantreffen. De drie kleinste eigenwaarden en het *collineariteitsgetal* van de normaalmatrix van de enkele afstandsverschillen zullen dan in de orde van het kwadraat van $|r_{12}|/\ell$ liggen. Dit komt overeen met een orde van 10^{-6} bij een ontvangerafstand van 20 km. Voor de dubbele en drievoudige afstandsverschillen zal deze afschatting zelfs nog aan de optimistische kant zijn. De conclusie luidt dan ook dat men voor de coördinaatberekening in de praktijk genoodzaakt zal zijn de positie van een ontvanger bekend te veronderstellen. Merk overigens op, dat met dezelfde afschatting (11), het afstandsverschil $\Delta(\ell_{1j}$-$\ell_{2j})$ gedeeld door ℓ, te interpreteren is als een afstandsverhouding. Dat wil zeggen dat bij benadering geldt: $\Delta\ln(\ell_{1j}/\ell_{2j}) = \Delta(\ell_{1j}-\ell_{2j})/\ell$. Dit laat nogmaals zien, daar de afstandsverhouding een vormgrootheid is, dat de ontvangerpositie r_1 voor relatief korte ontvangerafstanden praktisch onbepaalbaar is.

De tweede belangrijke conclusie betreft de betekenis van Δr_j in vergelijking (12). Analoog aan Δr_1 is ook Δr_j niet voor relatief korte ontvangerafstanden te bepalen. Dit betekent dat het pas zin heeft om aan baanbepaling te gaan denken als men zich van een redelijke uitgestrektheid van de ontvangerposities verzekerd heeft. Dit is de reden waarom men voor landmeetkundige toepassingen genoodzaakt is om te veronderstellen dat Δr_j=0. Omgekeerd impliceert (11) gelukkig ook, dat de doorwerking van eventuele modelfouten Δr_j in de satellietbanen voor relatief korte ontvangerafstanden voor een belangrijk deel gedempt wordt. Dit en de gegeven dynamische eigenschappen van de hoge en bijna circulaire GPS-banen, doen vermoeden dat het mogelijk moet zijn om een vervangingsmatrix van relatief eenvoudige structuur voor de stochastiek van de baanverstoringen te ontwikkelen. Of de demping van dien aard is, dat de vervangingsmatrix voor de vastgehouden baan een eenvoudig geschaalde eenheidsmatrix mag zijn, moet echter nog blijken.

Bij enkele toepassingen in de navigatie kan ook dankbaar gebruik gemaakt worden van genoemde dempingseigenschap. Hierop is het principe van de zogenaamde "differentiële GPS" gebaseerd. Veronderstel dat twee GPS-ontvangers, 1 en 2, simultaan de looptijden naar een en dezelfde satelliet j meten. Veronderstel verder, dat de positie van ontvanger 1 bekend is en dat de door deze ontvanger waargenomen afstand ℓ_{1j} via een telemetrieverbinding snel aan de rekenmachine van een zich verplaatsende, en niet al te ver verwijderde, ontvanger 2 bekend gemaakt kan worden. Dan beschikt men dus ter plekke van ontvanger 2 over zowel ℓ_{1j} als ℓ_{2j}, en kan de vergelijking voor het afstandsverschil ℓ_{1j}-ℓ_{2j} opgesteld worden, zodat de relatieve positie r_{12}, die voor atmosferische effecten en baanverstoringen (inclusief de effecten van "Selective Availability") praktisch gevrijwaard is, snel berekend kan worden. Zie de lustrumbijdrage van *ir. P. Sluiter* voor een verdere uiteenzetting van het concept van "Differentiële GPS".

Ter afsluiting van deze lustrumbijdrage wil ik de lezer nog kort op een belangrijk aspect van de huidige verwerkingsmethodiek van GPS wijzen; de vereffening van de bij de "double difference" behorende meerduidigheid N_{12jk}.

We weten dat de meerduidigheid N_{12jk} in vergelijking (9), in tegenstelling tot de geometrische onbekenden, alleen integer waarden kan aannemen. Het liefst zouden we deze informatie in de vereffening mee willen nemen. Immers, de stelling is dat hoe meer informatie in de vereffening wordt meegenomen, des te preciezer de onbekende parameters geschat kunnen worden. Helaas is het zo, dat de traditionele vereffeningstheorie alleen reqelwaardige grootheden aankan.

Om dit probleem op te lossen lijken mij twee wegen open te staan:

1. Men probeert een nieuwe vereffeningstheorie te ontwikkelen welke wèl in staat is integer grootheden te vereffenen en tevens, analoog aan de traditionele vereffeningstheorie, de mogelijkheid kent van een praktische kwaliteitsbeschrijving. Aan een dergelijke theorie moet overigens een ruimere betekenis worden toegekend, daar zij niet alleen van belang is voor GPS.
2. Men beschouwt N_{12jk} als een reqelwaardige grootheid en neemt daarmee genoegen met het feit dat het resultaat wat minder precies zal zijn. Deze benadering is natuurlijk geen echte oplossing van het probleem, maar wel een praktische, die misschien voor de meeste toepassingen goed genoeg blijkt te zijn.

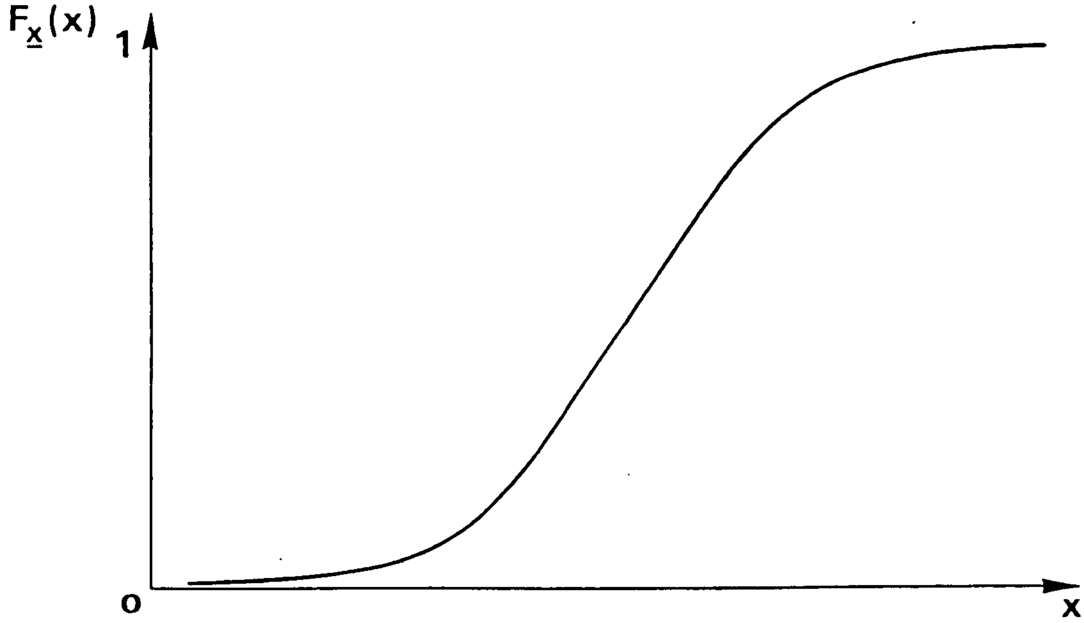
In de huidige praktijk van de GPS-gegevensverwerking wordt echter geen van beide wegen gevolgd. De gangbare methode die in de GPS-literatuur beschreven staat, komt kort op het volgende neer. Eerst wordt een traditionele vereffening uitgevoerd met behulp van de dubbele verschilgrootheden. In deze vereffening worden de meerduidigheden N_{12jk} als reqelwaardige parameters geschat. Met behulp van een afrondingsmechanisme worden dan de geschatte reqelwaardige meerduidigheden gezet op de dichtstbij liggende integer waarden. Tenslotte wordt dan een nieuwe vereffening uitgevoerd, waarbij nu de meerduidigheden niet meer als onbekende parameters worden aangemerkt, maar in plaats daarvan als bekend veronderstelde grootheden op de in de vorige stap bepaalde integer waarden worden vastgehouden.

Wat is er nu mis met deze procedure? Ik maak geen bezwaar tegen het idee van het afronden naar integer waarden. Hoewel over het in de praktijk toegepaste afrondingsmechanisme nog wel het een en ander gezegd kan worden, is het afrondingsidee niet slecht en kan het als een additionele vereffeningsstap geïnterpreteerd worden. Mijn bezwaar richt zich meer op de laatste stap in de gevolgde procedure. Het is principieel onjuist om te doen alsof de uit het waarnemingsmateriaal geschatte parameters in een volgende stap, waar het identieke waarnemingsmateriaal wordt gebruikt, als constante en bekend veronderstelde grootheden aangemerkt mogen worden. Dit heeft namelijk, geheel analoog aan hetgeen is besproken over de "vastgehouden" grootheden, tot gevolg, dat de formele variantiematrix van de berekende parameters een te optimistische precisiebeschrijving oplevert. Het in de praktijk veelvuldig gebruik van de zogenaamde "*fudge factors*" is hier dan waarschijnlijk ook niet vreemd aan.

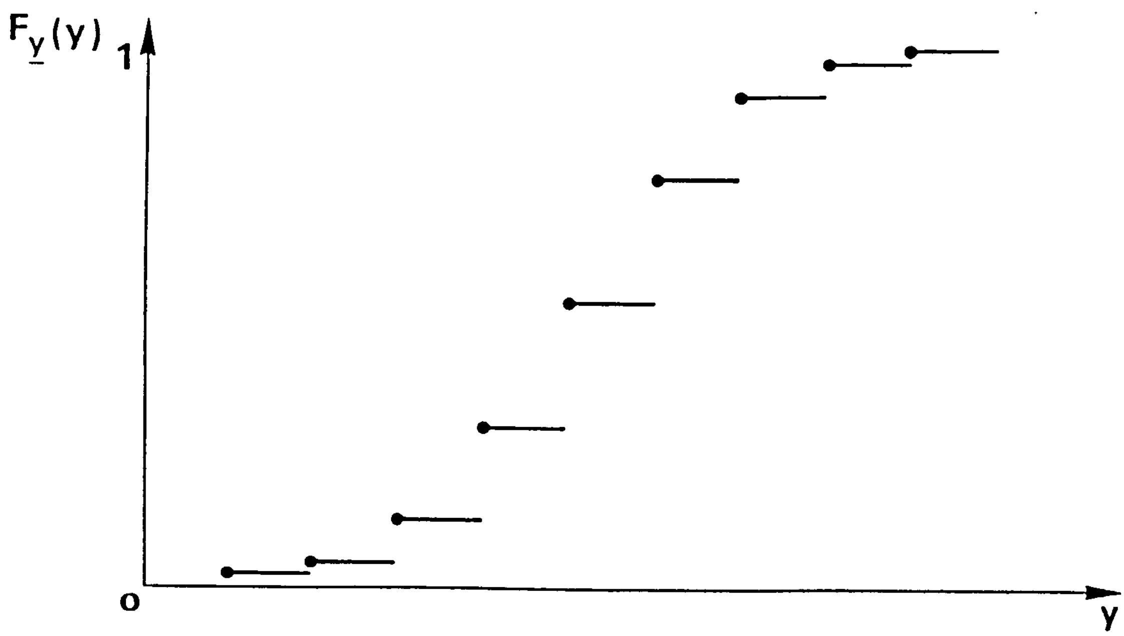
In tekstboeken handelend over GPS wordt deze laatste stap nu juist verdedigd, op grond van de bewering dat de precisie van de overgebleven geometrische parameters in vergelijking met de eerste stap beter is, omdat er sprake is van een kleiner aantal te bepalen onbekende parameters. Inderdaad geldt in het algemeen dat de (formele) precisie van een verzameling geschatte parameters beter wordt naarmate het aantal parameters in de verzameling kleiner wordt. In de huidige situatie is dit echter maar schijn. Immers, de stochastiek van de "vastgehouden" grootheden, zou met een bijbehorende kansdichtheidsfunctie beschreven dienen te worden.

Het volgende eenvoudige voorbeeld maakt dit misschien wat duidelijker. Veronderstel dat $\underline{x}$ een normaal verdeelde scalaire grootheid is met onbekende verwachting μ en bekende standaardafwijking σ, dus $\underline{x} \sim N(\mu,\sigma^2)$. De beste schatting van de onbekende verwachting wordt dan verkregen door de steekproefwaarde (dat wil zeggen de gemeten waarde) x van $\underline{x}$ toe te kennen aan μ.

In dit triviale geval verkrijgen we de zuivere schatter $\underline{\hat{\mu}}=\underline{x}$ van μ en is de verdeling van $\underline{\hat{\mu}}$ identiek aan die van $\underline{x}$ (zie figuur 5).



Figuur 5: Cumulatieve verdelingsfunctie vóór afronding



Figuur 6: Cumulatieve verdelingsfunctie ná afronding.

Veronderstel nu echter dat we ook nog weten dat de verwachting μ alleen maar integer waarden kan aannemen. In dit geval kunnen we besluiten de steekproefwaarde x van $\underline{x}$ af te ronden naar de dichtstbij gelegen integer waarde en deze afgeronde waarde toe te kennen aan μ. Dit betekent dat we een functie y van x, y(x), gedefinieerd hebben volgens de toewijzing: als $x \in [i-1/2, i+1/2)$ dan $y=i$, waarbij nu y(x) de waarde is die toegekend wordt aan de verwachting μ. De waarden die y kan aannemen zijn dus alleen geheeltallig.

Hieruit mag men echter niet concluderen, zoals in het geval van GPS wel gedaan wordt, dat $\underline{y}$ een constante is. Immers, y is een functie van een stochastische grootheid en daarmee zelf een stochastische grootheid geworden.
De schatter van μ, welke nu gegeven wordt door $\hat{\underline{\mu}}=y(\underline{x})$, heeft een kansdichtheidsfunctie en wel een van het discrete type (zie figuur 6). De kans dat $y(\underline{x}) \neq \mu$ is dus niet nul en kan afhankelijk van σ wel degelijk significant zijn.

Om de toetsingstheorie adequaat te kunnen toepassen en een behoorlijke kwaliteitsbeschrijving te kunnen leveren, dient men dus in beginsel de gehele kansdichtheidsfunctie in rekening te brengen. Voor GPS betekent dit, dat het van belang is inzicht te krijgen in het gedrag van de *gemengde* simultane verdeling van de vereffende geometrische grootheden samen met de vereffende en afgeronde integer meerduidigheden. Pas dan wordt het mogelijk om een met argumenten onderbouwde strategie te ontwikkelen voor het verwerken van de geheeltallige meerduidigheden.