

Geodesy and Satellite Navigation: A Lighthearted Look at a Strange Relationship

Vidal Ashkenazi

Nottingham Scientific Ltd, United Kingdom

There has always been a love-hate relationship between Geodesy and Satellite Navigation. Indeed, Satellite Positioning started life as an extension of Terrestrial Geodesy. When the first satellite, Sputnik 1, started orbiting the Earth in 1957, geodesists in several countries realised that satellites offered substantial potential as a geodetic positioning and navigation tool. The basic technologies of Terrestrial Geodesy of the day, notably Triangulation, Traversing and Precise Levelling were slow and cumbersome, mainly because of the effect of the curvature of the surface of the Earth, which limited the range of measurements to theodolite observations between points situated on hilltops, observation towers and triangulation masts. The advent of EDM (Electronic Distance Measurement) in the 1960's helped Terrestrial Geodesy, but it too was affected by the same limitation, namely the shortness of observable EDM ranges, due to the Earth's curvature.

Earth orbiting satellites did not suffer from this drawback. They could be viewed simultaneously from several points on Earth, and therefore direction and range measurements made, provided they were not obscured by high natural features or tall man-made structures. This led to several new Satellite Geodesy positioning methodologies. The first of these was Satellite Triangulation, which was used initially to supplement and strengthen terrestrial triangulation networks. Satellite Triangulation consisted of geodetic direction measurements derived from high power photographs of satellite orbits made against a stellar background of stars, with known right ascension and declination. A few years later, this was followed by range measurements to satellites, made from Earth bound EDM equipment to corner cube reflectors placed on the early satellites. The methodology used so far was an extension of Geodetic Astronomy, with little reference to Physical Geodesy.

This changed significantly when geodesists realised that they could use the Doppler shift on the signal broadcast from a satellite to obtain differential range measurements which, together with the known Keplerian orbit of the satellite, could lead to a relatively fast positioning, or navigation, method. The Keplerian orbital motion of satellites is primarily based on the Earth's gravity field, a subject of expertise by practitioners of Physical Geodesy. This gave birth to Transit-Doppler, the first satellite navigation technology. Transit-Doppler was used in the late 1970's and early 1980's not only for the positioning of naval ships and of submarines surfacing in the polar regions, but also for the strengthening and scaling of national and continental terrestrial triangulation networks. However, it did not take long to realise that positioning by Transit-Doppler to a reasonable degree of accuracy took several minutes, and therefore it could not be used as a full navigation methodology, which requires quasi-instantaneous positioning.

These were the early days of a new global satellite positioning, navigation and timing system, first called NAVSTAR Global Positioning System, a name later shortened to just GPS. The rest is history. The early decision to base GPS on a constellation of 24 medium-Earth orbit satellites was taken on the advice, as you would expect, of geodesists at the US Naval Surface Weapons Centre in Dalgren, Virginia. The close relationship between the early GPS and Geodesy was further demonstrated by the adoption of WGS84, the World Geodetic System

1984, as the basis of the 3-d coordinate system of GPS. As GPS was born during the Cold War, it was declared a US military navigation system, with full access to NATO, but only restricted access and down-graded positioning accuracies for civilian users. This so-called Selective Availability (SA) gave the green light to the civilian geodetic community to come up with new methodologies which would counter the effects of SA. As always, human ingenuity did not disappoint, and two new differential techniques were developed.

The first was the Differential GPS (DGPS) technique, which improved relative positioning accuracies of GPS by at one order of magnitude, down to a few metres. As a result, DGPS soon became the standard methodology for the offshore positioning of oil platforms, pipelines, etc. The next advance in improving the accuracy of satellite positioning was made on the advice of radio-astronomers, who proposed replacing the standard GPS pseudorange measurements, which are based on timing the modulated signal from satellite to receiver. Instead, they suggested making measurements on the basic carrier frequencies of these signals, just as they did with extra-galactic signals arriving at, say two, widely spaced radio telescopes in so-called Very Long Baseline Interferometry (VLBI), leading as a by-product to the Cartesian coordinate differences between the two telescopes.

This was the beginning of centimetric Positioning by the Carrier Phase GPS method, which was later developed further by geodesists into Kinematic GPS and centimetric Navigation. GPS had now become the universal high precision quasi-instantaneous positioning and navigation tool, leading the way for hundreds of new applications. Again, geodesists led the way, concentrating on high precision scientific and engineering applications. These included surveying and mapping, positioning in offshore engineering, the monitoring of local crustal dynamics and plate tectonics, the relative vertical movements of tide gauges, and the continuous 3-d movements of critical engineering structures, like tall buildings, dams, reservoirs, long suspension bridges, etc. All of these applications required very high relative positioning accuracies, but not quasi-instantaneously as in the safety-critical navigation and landing of civilian aircraft. This came much later.

Initially, GPS was considered as a standard navigation tool for military vehicles on land, sea and air, but not for safety-critical civilian transportation. This was because, unlike military positioning and navigation, safety-critical civilian transportation not only requires quasi-instantaneous and accurate positioning, but also so-called “high integrity and good coverage”. Geodesists will immediately realise that “integrity” stands for the geodetic concept of “reliability”, whereas “coverage” refers to the availability of a sufficient number of satellites which can be sighted by a receiver continuously, and are not obscured by natural or man-made obstructions, like high mountains, tall buildings, and the wings of an aircraft. On its own, GPS cannot meet these requirements to the level required in safety-critical civilian transportation. Military transportation, on the other hand, has relatively modest requirements which can be met by GPS. Indeed, you do not become a NATO Air Force pilot, if you want a safe life. Flying as a passenger in a commercial airline is something else all together.

The penetration of satellite navigation, and primarily GPS, into Civil Aviation involved yet again, as you would expect, geodesists. They had to develop jointly with the Civil Aviation community the necessary theoretical and practical tools, which could be used to establish and quantify their requirements of accuracy, integrity and coverage. This involved the use of existing geodetic tools, like the Covariance Matrix, the analysis of Least Squares residuals, and the well established geodetic Reliability measures. New tools were also developed, such as the concept of RAIM or Receiver Autonomous Integrity Monitoring, based on the analysis of the Least Squares residuals. These geodetic tools, which were highly beneficial to the Civil

Aviation community, started a long term fruitful collaboration between the two communities. However, this has not always been a straightforward and smooth relationship, and it involved especially at the beginning a deep suspicion of these “academic” geo-scientists. Here are a few notable examples of this love-hate relationship.

As a general rule, the existing civil aviation horizontal coordinates were based on latitudes and longitudes, with no particular reference to a Reference Datum. Heights in Civil Aviation were and still are based on barometric altimetry, on the assumption that all that matters is “the relative heighting between airplanes” which is not affected significantly by a change in barometric pressure. This disregards, of course, the fact that the heights of natural features on the ground, like mountains, do not change with changing barometric pressure. The first challenge was to convince the international Civil Aviation community that their horizontal coordinates, ie latitudes and longitudes, required a proper Geodetic Datum and, as GPS was being contemplated as a future navigation tool, it made sense to adopt the same reference datum, namely WGS84. It took a while to convince the community to accept that.

The adoption of WGS84 led to the re-surveying of most airports, runways, and various en route and landing navigation aids, in order to bring them into WGS84, in preparation for the introduction of GPS. This led to the discovery of some large discrepancies in airports and navaids in many countries, between the existing horizontal coordinate values and their new WGS84 equivalents. Geodesists will be familiar with such occurrences, whenever they start dealing with a new community, whether they are civil or offshore engineers, oceanographers or meteorologists.

Geodesists of a certain age will also remember some of the earliest GPS receivers, such as the TI 4100 receivers, made by Texas Instruments. These early receivers operated by measuring sequentially 4 pseudoranges to 4 different satellites, and therefore they were programmed to first check the geometry of the satellites in view and decide on the best 4, in terms of geometrical configuration. However, later on, with the emergence better receivers, which could measure all the available pseudoranges quasi-simultaneously, there was no need to carry on with measurements only to the best 4 satellites. One could process all of these measurements by Least Squares and reject those with relatively large residuals, if any. This standard processing of observations is bread-and-butter stuff to surveyors and geodesists. However, this was not the case with a number of navigation experts, who persisted on recommending the use of only the “best 4” satellites for quite sometime, before they finally abandoned the practice.

Satellite navigation and positioning has changed substantially and significantly over the last 5 to 10 years. With Galileo in its Development and Validation Phase, the future developments in GPS IIF and GPS III, the renewed interest in GLONASS, and the satellite navigation initiatives in Japan, China, India, Australia and several other countries, GNSS or Global Navigation Satellite System, is moving from being a concept, largely based on GPS alone, to a full global reality. There is now a comprehensive programme of GPS modernisation, which is currently underway, and aims to deliver significant improvements to both military and civil users.

The earliest mass-market applications of GPS involved road vehicles and mobile phones. In both cases, the twin aims are Navigation (*where am I, and how do I go to my destination?*) and Tracking (*where is he, she or it?*). In the case of Road Vehicle Tracking, successful applications include Fleet Monitoring (taxis or road transport companies), Theft Recovery of private cars, Black Box Incident Recorder, and the Transport of Hazardous or Valuable

Cargoes. Typically, most of these applications share 3 common features, namely “prior knowledge of the proposed route, the continuous tracking of position and velocity by GPS, and the trigger of an alarm by a significant deviation”.

Similarly, there have been a number of GPS tracking applications using mobile phone technology (GSM or GPRS), but these are not as developed and widespread as vehicle tracking. These involve vulnerable people, like young children, the elderly, key workers in some risky environments (like railways), individuals with a chronic or contagious disease, and even VIPs. Person tracking with GPS+Telematics could also involve judicial cases (ordered by a court of law), of suspected criminals or anti-social elements. Other proposed applications include Environmental Information, Location-based-Security, and Marketing.

On its own, a GPS-enabled phone offers location and communication. This may answer the question, “Where is she or he?” and “Where am I?”, but no more. However, when this is combined with an appropriate GIS database and a direction sensor, the combined system could answer two other very important questions, namely “*What’s around me?*” and “*What’s that building over there?*” This could be achieved by a GPS+compass device, providing positional and directional data, which the mobile phone or the PDA transmits to a remote server. The server calculates the user’s position and identifies the building along the measured azimuth, gets the relevant information from the database, and sends it back to the client. This is clearly valuable for the public utilities (water, gas, electricity, TV), shopping and leisure (restaurant menus, theatre tickets), house hunting (details of the property advertised *for sale*), and of course, for visitors and tourists (museums, notable buildings, archaeological sites).

Leaving mobile phones aside, satellite navigation can also be used for Location-Based-Security. For example, a briefcase or a portable PC can be programmed to unlock safely only in a specified location and nowhere else. This would minimise the risk of (military or commercial) high security material falling into the wrong hands. There are already some working prototype systems, which prevent access to the files of a personal computer, except when the PC is at some pre-specified locations. Other location-and-context-based applications which are being considered include the marketing and selling of goods, the reception of pay-TV, credit card security, spectator sports, and many others.

Indeed, the qualification of “critical application” is no longer restricted to safety critical transportation, but it also applies now to financial-critical, legal-critical, security-critical and business-critical applications as well. This creates a problem with standard off-the-shelf autonomous GPS receivers, which cannot operate indoors, because of signal attenuation and multipath. Over the last few years, GPS chip and receiver manufacturers have tried, with some success, the development of High Sensitivity GPS (or HS-GPS). The latest HS-GPS receivers, which incorporate up to 200,000 correlators operating in parallel, make it relatively easy to identify true pseudoranges from among the many signal and multipath reflections. There are now several manufacturers in the USA, Japan, Korea and Europe, advertising HS-GPS chips, and many other companies using such chipsets in their receivers.

The development of satellite navigation and positioning applications has not always been one of a continuous and inevitable success story. There have also been several abject public failures. A recent example of such a failure was the British government sponsored Pilot Scheme of tracking dangerous individuals, such as convicted criminals and paedophiles, by tagging them with “standard” off-the-shelf GPS-and-telematics sensors. As you would expect these trials failed, giving both the popular and the so-called serious newspapers a field day. Some newspapers dismissed satellite navigation altogether.

In conclusion, like nearly all the technologies that preceded it, satellite navigation and positioning is going through the standard stages of development from birth to maturity. Older Surveyors and Geodesists may well remember the advent of Electronic Distance Measurement (EDM), using microwaves or lightwaves in the late 1960's and the 1970's. When the first EDM instruments were introduced, the distances measured were also measured with tapes, just in case. Then came the second phase, when Surveyors became fully confident about EDM, and used it routinely for fast and precise range measurements. It took a few years and several critical mistakes in local mapping and national triangulation, to realise that EDM instruments could go wrong, and that they had to be calibrated regularly in order to determine their accuracy and systematic biases.

The development of satellite navigation and positioning is following practically the same stages as EDM did 40 years ago. Only now we can formalise these successive stages of development of a technology and give them names, by using Gartner's famous "Hype Cycle Curve", which was invented about 10 years ago, in conjunction with IT. Using a simplified version, which appeared recently in GPS World, these successive stages of Development of a Technology are now formally called "Technology Trigger", followed by "Peak of Inflated Expectation", leading to "Trough of Disillusionment", happily followed by the "Slope of Enlightenment", and hopefully leading to the "Plateau of Productivity".

As I write this, the first Galileo satellite, GIOVE, has been launched successfully, opening a new era in Satellite Navigation. Hopefully, this will lead to the development of a large number of new critical applications, and involve close collaboration with Geodesy and several other related disciplines, for the benefit of business, government and society.

Here is one last example about the strange relationship between Geodesy and GPS. The US Delegation to the International Telecommunications Union (ITU) recently proposed to abolish leap seconds, and thus cut the link between Solar Time and Coordinated Universal Time (UTC), and ipso facto GPS Time. At present, whenever the difference between UTC and Solar Time gets close to 0.7 of a second, a leap second correction is made in order to keep the difference between them under 0.9 seconds. This is done every few years, on the recommendation of the International Earth Rotation and Reference Systems Service, which monitors continuously the difference between Solar Time and UTC. This leap second correction, which has to be applied every few years to GPS Time, apparently causes software problems, because it has to be programmed in manually. However, considering the difficulties that this change would cause to other scientific communities, such as astronomers, and even to users of GPS time itself for some critical applications, the US proposal has now been postponed for the time being.

I would like to finish this article, which has been written as a tribute to the retirement of an outstanding Geodesist, with a personal remark. Among the many achievements of Professor Guenter Seeber, who authored the now famous book of "Satellite Geodesy", is also another little known talent. Professor Seeber travelled quite frequently to Brazil, to lecture and to help young Geodesy researchers at the Federal University of Parana in Curitiba. I also visited the University of Parana, some years before him. In one of our meetings in Europe, Professor Seeber asked me in what language I was lecturing there, to which I replied "very basic Portuguese, which is all I can manage". When I next met him a year or so later in a conference in South America, he delivered a paper in absolutely perfect Portuguese, and made me feel very inadequate. Clearly, Guenter Seeber is a man of many talents.