Analytical GPS Navigation Solution

Alfred Kleusberg

Abstract

The GPS navigation solution determines the coordinates $\boldsymbol{x} = (x, y, z)$ of the GPS receiver and the receiver clock offset cdT from measurements of at least four pseudo-ranges. We derive a direct solution of these observation equations without linearization and discuss the occurrence of unique solutions, double solutions, and infinitely many solutions, and the geometric conditions leading to these cases.

1. Introduction

The determination of the coordinates of a receiver position from measurement of pseudo-ranges to satellites is the standard mode of positioning for users of the Global Positioning System and similar systems; a minimum of four pseudo-ranges is necessary for three-dimensional positioning. Certain geometric constellations between the satellites and the receiver do not allow the determination of a unique position; we shall refer to these cases by using the term 'singularity'.

The equations linking the pseudo-ranges and the receiver coordinates are non-linear. The direct solution of these non-linear equations is possible, and several different solutions have been described in the literature. The widely used alternative is to linearize the pseudo-range equations and to use the tool of linear algebra in the position determination calculations.

When comparing results obtained from the solution of the non-linear and the linearized equation it was found, that differences occur for certain geometric constellations. In particular, the non-linear equations are solvable in some cases when the linearized equations lead to a singularity. To understand the reasons behind this behavior, we shall investigate the geometry leading to the above mentioned singularities.

2. The solution of the non-linear pseudo-range equations

Neglecting refraction effects, satellite clock offsets and measurement errors, the pseudo-range measured with a GPS receiver, p_i, is the sum of the satellite-to-receiver distance, s_i, and the receiver clock offset, dT, multiplied by the speed of light, c (Milliken and Zoller, 1980). The subscript i identifies the satellite.

The GPS navigation solution determines the coordinates (x, y, z) and the clock offset dT of a GPS receiver from pseudo-ranges p_i, i=0, 3 measured to four GPS satellites, and the coordinates (x_i, y_i, z_i), i=0, 3 of these satellites. These quantities are interrelated through the observation equations

$$p_i = [(x_i-x)^2 + (y_i-y)^2 + (z_i-z)^2]^{1/2} + c\cdot dT \qquad (1)$$

where we have used c as an abbreviation for the speed of light. The receiver clock offset can be eliminated from the observation equations (1) by subtracting p_0 from p_1, p_2, p_3. This yields equations for three range differences $d_i = p_i - p_0$, i=1,3 represented in terms of satellite and receiver coordinates according to

$$d_i = [(x_i-x)^2 + (y_i-y)^2 + (z_i-z)^2]^{1/2} - [(x_0-x)^2 + (y_0-y)^2 + (z_0-z)^2]^{1/2} \qquad (2)$$

From a geometric point of view, each of these three equations describes a hyperbolic surface of position. These surfaces intersect in the possible locations of the GPS receiver.

The non-linear eqns. (2) can also be solved directly without the process of linearization, thereby not requiring the availability of initial approximate values for the receiver position and being non-iterative as a consequence. Bancroft (1985) derived a rather elegant algebraic solution procedure for eqns. (2) and noted that his procedure "performs better than an iterative solution in regions of poor GDOP" (ibid.). His algorithm involves the inversion of a (4 x 4) matrix and the solution of a scalar equation of second order. Bancroft's method was further discusses and analyzed by Abel and Chaffee (1991) and by Chaffee and Abel (1994).

Krause (1987) published a two step algorithm for the direct solution of the eqns. (2). After the receiver clock offset $c \cdot dT$ is determined in a first step involving the inversion of a (2 x 2) matrix, the vectors from the satellites to the receiver can be evaluated and the receiver position is calculated through vector addition. Krause (ibid.) notes that "simulations under usual and extreme user and constellation situations showed absolute stability and precision for the algorithm".

The solution presented by Grafarend and Shan (1996) involves squaring eqn. (2) (to remove the square root), and then algebraically reducing the equations in order to provide the explicit solution for the receiver coordinates. This procedure includes the inversion of a (3 x 3) matrix.

The non-linear hyperbolic eqns. (3) were solved by Kleusberg (1994) using vector algebra. The algorithm is shown below in a modified and simplified version.

The distances b_i and the unit vectors $\boldsymbol{e}_i$ between the satellite S_0 and the satellites S_i are computed from the satellite coordinates (vectors are indicated by bold letters). These quantities completely describe the intrinsic geometry of the satellite configuration. The position of the receiver P is described by the (unknown) unit vector $\boldsymbol{e}$ pointing from S_0 to P, and the corresponding unknown distance s_0.

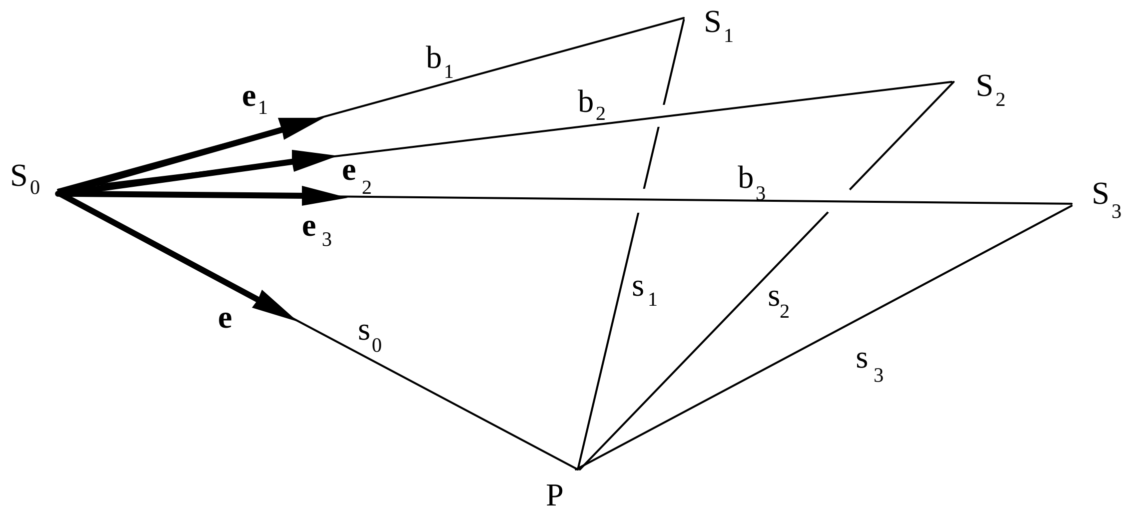

Figure 1: Geometry of the three-dimensional hyperbolic intersection

Noting that the cosine of the angle between the unit vectors $\boldsymbol{e}$ and $\boldsymbol{e}_i$ is equal to their scalar product $\boldsymbol{e} \cdot \boldsymbol{e}_i$, we can represent the geometry in each of the three triangles S_i - S_0 - P by the cosine rule according to

$$s_i^2 = b_i^2 + s_0^2 - 2 b_i s_0 (\mathbf{e} \cdot \mathbf{e}_i) \tag{3}$$

Further noting that the observation equation (2) can be rewritten as $s_i = d_i + s_0$, we also obtain in each of these three triangles a relation between the measurements of pseudo-range differences, and the distances between the receiver and the satellites

$$s_i^2 = d_i^2 + s_0^2 + 2 d_i s_0 \tag{4}$$

Equating equations (3) and (4) yields, after some basic algebraic manipulation

$$2 s_0 = \frac{b_i^2 - d_i^2}{d_i + b_i (\mathbf{e} \cdot \mathbf{e}_i)} \text{ , i = 1,3} \tag{5}$$

There are three unknowns in these three equations: the two independent components of the unit vector $\boldsymbol{e}$ and the distance s_0, all other terms are known.

In order to reduce the number of unknowns further, we equate the right hand sides of the first and the second of the eqns. (5), and similarly the second and the third, thereby eliminating the distance s_0.

$$\frac{b_i^2 - d_i^2}{d_i + b_i (\mathbf{e} \cdot \mathbf{e}_i)} = \frac{b_{i+1}^2 - d_{i+1}^2}{d_{i+1} + b_{i+1} (\mathbf{e} \cdot \mathbf{e}_{i+1})} \text{ , i = 1,2} \tag{6}$$

These two equations can be rearranged by utilizing the distributive law of vector algebra to yield

$$\left[\frac{b_i}{b_i^2 - d_i^2} \mathbf{e}_i - \frac{b_{i+1}}{b_{i+1}^2 - d_{i+1}^2} \mathbf{e}_{i+1} \right] \cdot \mathbf{e} = \left[\frac{d_{i+1}}{b_{i+1}^2 - d_{i+1}^2} - \frac{d_i}{b_i^2 - d_i^2} \right] \text{ , i=1,2} \tag{7}$$

which reads in short form by using obvious abbreviations for the terms in square brackets

$$\begin{aligned} \mathbf{F}_1 \cdot \mathbf{e} &= U_1 \\ \mathbf{F}_2 \cdot \mathbf{e} &= U_2 \end{aligned} \tag{8}$$

These are two scalar equations for the components of the unit vector $\boldsymbol{e}$. In general, there will be two solutions for $\boldsymbol{e}$ satisfying eqn. (8). For the special case that $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are parallel, the solution is undefined.

The algebraic solution of equations (8) can be derived by applying the vector triple product identity to the product of $\boldsymbol{e}$, $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$

$$\boldsymbol{e} \times (\boldsymbol{F}_1 \times \boldsymbol{F}_2) = \boldsymbol{F}_1\, \boldsymbol{e} \cdot \boldsymbol{F}_2 - \boldsymbol{F}_2\, \boldsymbol{e} \cdot \boldsymbol{F}_1. \tag{9}$$

Replacing the scalar products on the right hand side of (9) with equation (8) we obtain

$$\boldsymbol{e} \times (\boldsymbol{F}_1 \times \boldsymbol{F}_2) = U_2\, \boldsymbol{F}_1 - U_1\, \boldsymbol{F}_2. \tag{10}$$

With the abbreviations

$$\boldsymbol{G} = \boldsymbol{F}_1 \times \boldsymbol{F}_2, \;\; \boldsymbol{H} = U_2\, \boldsymbol{F}_1 - U_1\, \boldsymbol{F}_2 \tag{11}$$

we can rewrite equation (10) in a shorter form as

$$\boldsymbol{e} \times \boldsymbol{G} = \boldsymbol{H}. \tag{12}$$

Muliplying both sides of this equation by $\boldsymbol{G}$ from the left, and applying the triple vector product identity again to the left hand side of the resulting equation, we obtain

$$\boldsymbol{e}\, \boldsymbol{G} \cdot \boldsymbol{G} - \boldsymbol{G}\, \boldsymbol{G} \cdot \boldsymbol{e} = \boldsymbol{G} \times \boldsymbol{H}. \tag{13}$$

The scalar product in the second term of the left-hand side can be written in terms of the length of the vectors involved, and the angle b between them

$$\boldsymbol{G} \cdot \boldsymbol{e} = (\boldsymbol{G} \cdot \boldsymbol{G})^{1/2} \cos b\, . \tag{14}$$

The angle b also appears if we compute the length of the vector $\boldsymbol{H}$ from equation (12)

$$(\boldsymbol{H} \cdot \boldsymbol{H})^{1/2} = [(\boldsymbol{e} \times \boldsymbol{G}) \cdot (\boldsymbol{e} \times \boldsymbol{G})]^{1/2} = (\boldsymbol{G} \cdot \boldsymbol{G})^{1/2} \sin b\, . \tag{15}$$

Comparing eqns. (14) and (15) we get

$$\boldsymbol{G} \cdot \boldsymbol{e} = \pm (\boldsymbol{G} \cdot \boldsymbol{G})^{1/2} [1 - \boldsymbol{H} \cdot \boldsymbol{H} / \boldsymbol{G} \cdot \boldsymbol{G}]^{1/2} = \pm [\boldsymbol{G} \cdot \boldsymbol{G} - \boldsymbol{H} \cdot \boldsymbol{H}]^{1/2}. \tag{16}$$

Inserting this relation into equation (13) we obtain after some rearrangement the two solutions of equation (8)

$$\boldsymbol{e}^{1,2} = (\boldsymbol{G} \cdot \boldsymbol{G})^{-1} \{ \boldsymbol{G} \times \boldsymbol{H} \pm \boldsymbol{G}\, [(\boldsymbol{G} \cdot \boldsymbol{G}) - (\boldsymbol{H} \cdot \boldsymbol{H})]^{1/2} \}. \tag{17}$$

The distance s_0 from the satellite S_0 to the receiver P can now be determined from anyone of the three equations (5) according to

$$s_0^{1,2} = \frac{1}{2} \frac{b_i^2 - d_i^2}{d_i + b_i \left(\boldsymbol{e}^{1,2} \cdot \boldsymbol{e}_i \right)} \tag{18}$$

and the coordinates of the receiver are finally determined from (cf. Figure 1)

$$\boldsymbol{x}^{1,2} = \boldsymbol{x}_0 + s_0^{1,2} \boldsymbol{e}^{1,2}\, . \tag{19}$$

3. Discussion of results

Qualitatively, the results obtained from eqns. (18) and (19) can be classified in the following way:

a) The two unit vectors determined from eqn. (17) are different, and eqn. (18) yields two positive distances s_0. In this case, there are two intersections of the hyperbolic surfaces of position. Both solutions satisfy the observation eqns. (2). The correct solution can be identified if *a priori* information about the approximate receiver location is available.

b) The two unit vectors determined from eqn. (17) are different, and eqn. (18) yields one positive and one negative distance s_0. In this case, only the solution belonging to the positive distance satisfies the observation eqns. (2).

c) The two unit vectors determined from eqn. (17) are the same. In this case, the two intersections of the hyperbolic surfaces of position coincide. It will be shown elsewhere that in this particular case the receiver-to-satellite unit vectors are on a conic surface. It is known that in this case the linearized pseudo-range cannot be solved uniquely.

d) The receiver is located on the extension of one of the baselines b_i. In this case, $d_i = \pm b_i$ and one of the denominators in eqn. (7) is zero. This critical geometric situation is know from 2-dimensional hyperbolic positioning (e.g. LORAN). In 3-dimensional satellite positioning it does not occur if the receiver location is significantly lower than the orbits of the satellites.

e) The two vectors $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are parallel. In this case, the denominator in eqn. (17) is zero, and there are infinitely many solutions $\boldsymbol{e}$ satisfying eqns. (8). Since $\boldsymbol{F}_1$ is in the plane defined by the satellites S_0, S_1 and S_2, and $\boldsymbol{F}_2$ is in the plane defined by satellites S_0, S_2 and S_3, the four satellite positions are coplanar in this particular case. Not all coplanar satellite positions will lead to this singularity; it will be shown elsewhere that only the arrangement of the satellite positions in a conic section allows infinitely many solutions of the observation equations (2).

References

Abel, J.S. and Chaffee, J.W. (1991): Existence and Uniqueness of GPS Solutions. IEEE Transactions on Aerospace and Electronic Systems, Vol. 27, No. 6, pp. 952-956

Bancroft, S. (1985): An Algebraic Solution of the GPS Equations. IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-21, No. 6, pp. 56-59

Chaffee, J. and Abel, J. (1991): On the Exact Solutions of Pseudorange Equations. IEEE Transactions on Aerospace and Electronic Systems, Vol. 30, No. 4, pp. 1021-1030

Grafarend, E.W. and Chan, J. (1996): A Closed-form Solution of the Nonlinear Pseudo-Ranging Equations (GPS), ARTIFICIAL SATELLITES Planetary Geodesy, No. 28, pp. 133-147

Kleusberg, A. (1994): Die direkte Lösung des räumlichen Hyperbelschnitts. Zeitschrift für Vermessungswesen, Vol. 119, No. 4, pp. 188-192

Krause, L.O. (1987): A Direct Solution to GPS-Type Navigation Equations. IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-23, No. 2, pp. 225-232

Milliken, R.J, Zoller, C.J. (1980): Principle of Operation of NAVSTAR and System Characteristics. in: Global Positioning System Papers published in NAVIGATION, Vol. I, pp. 3-14, The Institute of Navigation, Washington, D.C.