Datum transformation by Gauss projection of the ellipsoid onto the sphere and bilinear function of complex variable: an application to ETRS89 – ED50

J. B. Mena

Alcala University, Polytechnic School, Mathematics Department.
Aptdo 20, Alcalá de Henares, E-28871 Madrid, Spain
Tel. +34 91 8854901
Fax +34 91 8854951
juan.mena@uah.es

Abstract.

In this paper a simple method for accurate transformation of geodetic coordinates is presented. The process uses the regional network of points allocated in both geodetic systems researching only the two dimensional problem (longitude and latitude). Basically, transformation of any point $P$ consists to determine first the three network vertexes nearest to $P$ such as this point is located inside the triangle. Computing both geodetic systems separately, these three vertexes are projected from the ellipsoid onto the Gauss sphere tangent along a parallel. Given that both geodetic variables are isometric on the sphere, and since only three network vertexes have been considered, it can find of closed way the bilinear transformation of complex variable which applies the triangle in study of a sphere in the other. Starting from here, the $P$ image is obtained projecting this point on the sphere corresponding to their geodetic system, applying it later on the other sphere, and computing finally the inverse transformation which projects this sphere on the second ellipsoid. This procedure achieves higher accuracy than matrix spatial methods of seven parameters, and it could be considered of the same order that methods based on distortion analysis. However, besides being simpler, the presented process provides better results there where regional reference network is not very dense.

Keywords: datum transformation, Gauss projection, bilinear function.

1. Introduction.

The establishment of WGS84 (World Geodetic System 1984) as world geodetic reference and GPS (Global Positioning System) were fundamental reasons for developing accurate methods to transform geodetic coordinates between WGS84 and regional datums like ED50 (European Datum 1950). Not only in Europe, regional ellipsoids are still used for different applications in other parts of the world. The study of the mathematical relation between WGS84 and each one of the main regional geodetic systems (NIMA, 2004) is a topic in continuous development in which progressively better results are achieved.

Many of datum transformation methods are based on estimation of the unknown terms of a matrix equation or a polynomial expression utilizing the least squares adjustment and starting of a network of vertexes whose geodesic coordinates are known in both systems. This Double Vertexes Network (DVN) is allocated in a limited region, reason for which the transformation equation only will be correct for interior of this zone. Furthermore, the accuracy of each mathematical model is attached closely to the quality determination and local density of the network.

In Gonzalez-Matesanz et al. (2003) many methods developed in order to carry out transformation from WGS84 to other geodetic reference systems are presented. The seven-parameter models are especially important for their simplicity, since infinitesimal character of transformation is assumed. An example well known is the Molodensky – Badekas model or Bursa – Wolf model. Yet but simple are
the five parameters methods, where parallelism between corresponding ellipsoid axes is supposed. With similar accuracy are the models based on determination by statistical regression of polynomial coefficients in real or complex variable (Draper and Smith, 1981). However, in a higher step for their accuracy order are the methods whose intermediate results are corrected by means of distortion analysis (Collier et al., 1998). These corrections are achieved by techniques like the minimum curvature surfaces (Briggs, 1974; Smith and Wessel, 1990), the rubber sheeting process (White and Griffin, 1985), the multiple statistic regression (Applebaum, 1982), or the least squares collocation (Tapley, 1976; Collier et al., 1998).

Researching the two dimensional geodetic problem, it is possible to generate a basic classification of different transformation methods by only two principal categories according to their accuracy order. The first group involves the simple methods, which could be defined as those where distortion analysis is not used. The second group includes any procedure where distortion modelling is applied. Thereby, all method of the second group is assumed as higher accuracy order than any simple method. According to this classification the process here presented is not belonging to any group, but rather it can be classified in the intersection of both groups. In fact, this is a simple method for datum transition, but at the same time it is able to achieve an accuracy order comparable to the methods of the second group.


2.1. The transformation function.

In order to generalize suppose that a DVN have been accurately determined in coordinates of geodetic systems S₁ and S₂. In this frame, assume a point \( P \) which is included into region and whose coordinates in the S₁ and S₂ systems are \( (\lambda_1, \phi_1) \) and \( (\lambda_2, \phi_2) \) respectively. The problem is to find the most accurate bijection \( f \) between two dimensional spaces as:

\[
(\lambda_2, \phi_2) = f(\lambda_1, \phi_1)
\]

For this goal, assume that \( f \) has to verify the following conditions:

1. It has to be conformal, since angles are preserved in both geodetic systems.
2. The linear deformation coefficient in \( P \) has to be the corresponding of the DVN.

These conditions show that \( f \) has to be a likeness function into a differential environment of \( P \), with the same likeness factor in this point than DVN. However, \( f \) deduction involves two questions. The first is that latitude variable is not isometric on the ellipsoid, and the second is that linear deformation factor is unknown while \( f \) is not calculated.

For solving the first question we could opt for a variable change in order to transform the geodetic latitudes into isometric latitudes by usual way (Bugayevskiy and Snyder, 1995; Lauf, 1983). However, longitude and latitude are isometric coordinates on the sphere. Then we can find the bijection of each ellipsoid on the corresponding sphere so that linear deformation between both projections of \( P \) on these spheres is the same as the linear deformation between both positions of \( P \) on both ellipsoids. To achieve this goal, which solves our two questions, Gauss theory for ellipsoid to sphere projection has been applied (Zakatov, 1953). In this projection sphere and ellipsoid are tangents along a parallel, and the sphere radius value is the mean of curvature of the ellipsoid in the selected parallel. The Gauss projection is conformal on entire sphere and practically equidistant in two degrees to both sides of tangent parallel. Denominating \( g_1 \) and \( g_2 \) these Gauss projections on respective parallels of \( P \), we write:

\[
(\lambda_1', \phi_1') = g_1(\lambda_1, \phi_1), \quad (\lambda_2', \phi_2') = g_2(\lambda_2, \phi_2)
\]
Here the first member coordinates are the corresponding $P$ images on each sphere. Concretely, for an ellipsoid where $e$ is the first eccentricity and $\phi_0$ is the geodetic latitude of tangent parallel, the corresponding $g$ function is achieved computing previously the projection constants by means of the following expressions where $\phi_0'$ is the projection of $\phi_0$:

\[
K = \left(1 + \frac{e^2 \cos^2 \phi_0}{1 - e^2}\right)^{1/2}; \quad \phi_0' = \sin \left(\frac{\sin \phi_0}{K}\right)
\]

\[
k = \left(\frac{1 - e \sin \phi_0}{1 + e \sin \phi_0}\right)^{K e/2} \tan \left(\frac{\phi_0 + \frac{\pi}{4}}{2}\right)^{K} \left[\tan \left(\frac{\phi_0 + \frac{\pi}{4}}{2}\right)\right]^{-1}
\]

Then, the mathematical relations that define this function are:

\[
\lambda' = K \lambda, \quad \tan \left(\frac{\phi' + \frac{\pi}{4}}{2}\right) = \frac{1 - e \sin \phi}{k (1 + e \sin \phi)}^{K e/2} \tan K \left(\frac{\phi + \frac{\pi}{4}}{2}\right)
\]

Since the applied Gauss projection is conformal, the values of linear deformation factor on meridian ($m$) and parallel ($n$) are the same in each point. In fact, equations [4] result by integration of following differential equations where $R$ is mean earth’s radius and $M$ and $N$ are the main radii of curvature of ellipsoid:

\[
m = R \frac{d\phi'}{d\phi}, \quad n = R \frac{\cos \phi' d\lambda'}{\cos \phi d\lambda} \quad \rightarrow \quad m = n \quad \rightarrow \quad \frac{d\lambda'}{d\lambda} = K
\]

\[
\frac{d\phi'}{d\phi} = \frac{M \cos \phi'}{N \cos \phi}
\]

Besides, on tangent parallel and their vicinity both linear factors are equal to one:

\[
m_0 = n_0 = K \frac{R \cos \phi_0'}{N \cos \phi_0} = 1
\]

This equation can be extended up to two degrees to both sides of tangent parallel given that the values of $R$, $K$ and $k$ are achieved computing also the following conditions (cf. Zakatov, 1953):

\[
\left[\frac{dm}{d\phi}\right]_0 = 0, \quad \left[\frac{d^2m}{d\phi^2}\right]_0 = 0
\]

Therefore, in the Gauss projection defined through Eqs. [2 – 4] angles and the linear deformation that exists in $P$ environment are preserved.

Once both equations [4] have been obtained, determination of a function $f'$ is reduced to find the relation between spherical surfaces in $P$ environment. This relation is now established through isometric variables. Therefore we will opt for representing it by means of complex variable in the following way:

\[
w = w(z) \quad \text{where} \quad \begin{cases} z = \lambda_1' + i \phi_1' \\ w = \lambda_2' + i \phi_2' \end{cases}
\]
For sought function \( w \) three basic conditions will be assumed:

1. Function \( w \) has to be conformal identically as a function \( f \).
2. Function \( w \) has to be an union of several geometric elementary transformations.
3. Function \( w \) has to be specific for each point \( P \) for accuracy reasons.

Analyzing several complex variable transformations, bilinear model was selected because this function is conformal and it includes rotation, translation, dilation (or contraction) and inversion:

\[
\frac{w}{z} = \frac{\alpha z + \beta}{\gamma z + 1}
\]

Furthermore, in order to verify condition 3, we will compute the complex coefficients \( \alpha, \beta \) and \( \gamma \) using only the three DVN vertexes \( A, B \) and \( C \) nearest to \( P \) such as this point is located inside the triangle. In this way the resolution of linear system:

\[
\alpha z_i + \beta - \gamma z_i w_i - w_i = 0 \quad (i = A, B, C)
\]

provides a close solution. This is more accurate for transformation of \( P \) than classic solution by the least squares adjustment using all DVN vertexes (Higham, 1996).

Equations [1] [2] and [5] give rise to the following decomposition of bijections for function \( f \) definition:

\[
(\lambda_2, \varphi_2) = \left( g_2^{-1} \cdot w \cdot g_1 \right) \left( \lambda_1, \varphi_1 \right)
\]

2.2. Algorithm for a point set transformation.

Once equation [8] has been specified, for transforming a point set from the system \( S_1 \) to \( S_2 \), the presented method involves applying for each one point \( P \) the following successive operations:

1. Determination of the minor area triangle \( ABC \) whose vertexes are DVN points and as it includes \( P \). This is a computational geometry problem whose solution can be consulted in De Berg et al. (1998).

2. Starting of \( S_1 \) coordinates of \( A, B, C \) vertexes, computation of mean latitude parallel \( (\varphi_{01}) \) and construction of the Gauss projection \( (g_1) \) of ellipsoid 1 on sphere 1 (Eqs. [2], [3] and [4]). Since the geodetic latitude of \( P \) is known, this value could be used as latitude of tangent parallel.

3. Starting of \( S_2 \) coordinates of \( A, B, C \) vertexes, computation of mean latitude parallel \( (\varphi_{02}) \) and construction of the Gauss projection \( (g_2) \) of ellipsoid 2 on sphere 2 (Eqs. [2], [3] and [4]). As latitude of tangent parallel, geodetic value of \( P \) in \( S_1 \) also could be used.

4. Projection of points \( A, B \) and \( C \) on sphere 1 \( (z_i \) values) and sphere 2 \( (w_i \) values) by means both equations [4] which have been obtained from steps 2 and 3.

5. Determination of the bilinear transformation [6] which applies \( z_i \) values on \( w_i \) values respectively. For this goal linear system [7] has to be solved.

6. Application of the transformation [8] to \( S_1 \) coordinates of \( P \). That is to say: \( g_1 \) projection, \( w \) transformation, and \( g_2 \) inverse projection. Last operation involves successive approaches in the latitude equation [4], but it converges quickly.
3. Results and conclusions.

The presented method has been applied to ETRS89 - ED50 transformation into Iberian Peninsula and Balearic Islands using as DVN the RE GENTE network (Quiros et al., 2003). This network embraces around thousand vertexes defined in two systems: ETRS89 (European Terrestrial Reference System 1989) whose ellipsoid is the same as WGS84, and ED50 where the International Hayford ellipsoid is assumed. REGENTE network involves a convenient density since the mean distance between points is only of 30 – 40 kilometres. In this frame, once carried out different tests in several points of the zone, it was obtained an estimated mean error of about 10 - 15 centimetres. This value is enough for many applications of diverse nature and it suppose an accuracy which is characteristic for some methods where distortion analysis is used. Table 1 shows some results which were achieved for peripheral points of the region, this is to say, for points where least accuracy is obtained. In this table corresponding results to a matrix method of seven parameters (simple method) and a method of distortion analysis (high accuracy) also are presented. It can observe that results of the new method (bilinear function) are many more accurate than results of the simple method. In fact residual values ‘distortion analysis – bilinear function’ are smaller than residual values ‘distortion analysis – simple method’.

<table>
<thead>
<tr>
<th>Name</th>
<th>ETRS89</th>
<th>ED50 (simple method)</th>
<th>ED50 (distortion analysis)</th>
<th>ED50 (bilinear function)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CADIZ</td>
<td>06 17 02.926 W</td>
<td>06 16 58.164 W</td>
<td>06 16 58.038 W</td>
<td>06 16 58.037 W</td>
</tr>
<tr>
<td>ALMERIA</td>
<td>02 06 26.179 W</td>
<td>02 06 21.811 W</td>
<td>02 06 21.757 W</td>
<td>02 06 21.780 W</td>
</tr>
<tr>
<td>BARCELONA</td>
<td>02 09 37.915 E</td>
<td>02 09 41.013 E</td>
<td>02 09 41.951 E</td>
<td>02 09 41.944 E</td>
</tr>
<tr>
<td>VIZCAYA</td>
<td>02 30 36.291 W</td>
<td>02 30 31.611 W</td>
<td>02 30 31.662 W</td>
<td>02 30 31.661 W</td>
</tr>
<tr>
<td>PONTEVEDRA</td>
<td>08 46 00.460 W</td>
<td>08 45 55.189 W</td>
<td>08 45 54.986 W</td>
<td>08 45 54.939 W</td>
</tr>
</tbody>
</table>

Table 1: Some results for peripheral points of the region.

By way an extreme test, the method was deformed using as reference only the three network vertexes located in the northeast, northwest and south limits of territory. In this case, it would have been more convenient applying the least squares adjustment for computing the bilinear transformation. However, errors around two meters were obtained for triangle inside (typical accuracy of the simple methods). This fact confirms that the method in this way deformed achieves a similar accuracy than some simple matrix methods. In Table 2 some results of the simple method, distortion analysis and deformed method are presented.

<table>
<thead>
<tr>
<th>Name</th>
<th>ETRS89</th>
<th>ED50 (simple method)</th>
<th>ED50 (distortion analysis)</th>
<th>ED50 (deformed method)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P001</td>
<td>01 18 12.435 W</td>
<td>01 18 07.967 W</td>
<td>01 18 07.956 W</td>
<td>01 18 07.922 W</td>
</tr>
<tr>
<td>P002</td>
<td>02 15 33.239 W</td>
<td>02 15 28.720 W</td>
<td>02 15 28.675 W</td>
<td>02 15 28.645 W</td>
</tr>
<tr>
<td>P003</td>
<td>03 12 17.211 W</td>
<td>03 12 12.611 W</td>
<td>03 12 12.555 W</td>
<td>03 12 12.513 W</td>
</tr>
<tr>
<td>P004</td>
<td>04 12 22.221 W</td>
<td>04 12 17.567 W</td>
<td>04 12 17.507 W</td>
<td>04 12 17.441 W</td>
</tr>
<tr>
<td>P005</td>
<td>06 16 10.362 W</td>
<td>06 16 05.535 W</td>
<td>06 16 05.522 W</td>
<td>06 16 05.340 W</td>
</tr>
</tbody>
</table>

Table 2: Some results obtained by deformed method.

Our main conclusion is that the presented method shows a simple and novel algorithm for datum transformation that solves this problem using classical theories. When a DVN is dense then the process is of similar accuracy order than methods based on distortion analysis. But if a DVN is poor then it is of similar accuracy order than simple methods applied with a dense DVN. Therefore high density of reference network is not required necessarily so that it can be applied for datum transformation in widespread zones with low density of observations.
4. References.


