Projections and coordinate systems made easy

by Laurette Coetzee, Open Spatial Solutions

Why do we use projections? And what exactly is a coordinate system?

If I ask you where you are, you will most probably give me an address, or indication in terms of well known features. However, you may also give me a position such as 25 degrees, 12 minutes and 23,23 seconds south; 28 degrees, 6 minutes and 9,09 seconds east; S25.84668 E28.12909. So what does this mean?

When we give a position in coordinates (S25.84668 E28.12909) it indicates a position on the earth itself in relation to the longitudes and latitudes in which the earth has been divided. The longitudes and latitudes are measured from two lines of zero on the earth, namely the equator and the Greenwich meridian. The equator is the line of zero running horizontally across the earth, while the Greenwich meridian is the line of zero running vertically across the earth.

In geography the lines running parallel to the equator are called parallels or latitudes and the lines running along the Greenwich line are called meridians or longitudes. The reason for calling them meridians is because they do not run parallel but converge to the poles.

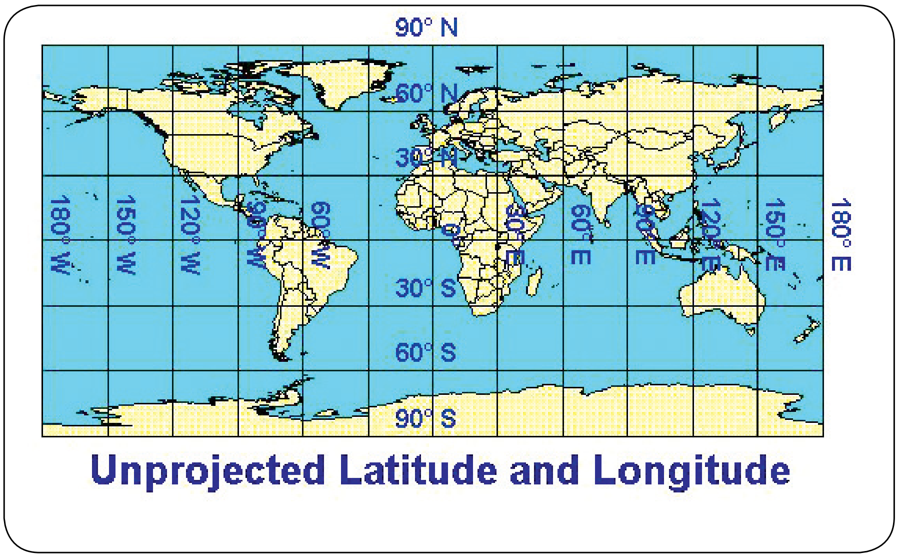

Fig. 1: World unprojected latitude and longitude. (Credit: Peter H. Dana)

So, when someone gives you a position - S25.84668 E28.12909 – it is the exact position in terms of the coordinates as they are represented on the Earth. These are the coordinates used when drawing in CAD. It is positions on the Earth, which are not projected, since it is the actual position on the Earth. Synonyms include lat/long (because it is given in terms of the latitudes and longitudes on the Earth; degrees, minutes, seconds; decimal degrees (because it is given in terms of the coordinates measured from the lines of zero); geographic coordinate system (because it is based on coordinates on the Earth itself).

The bottom line here is that geographic maps use a two-dimensional surface, which can be compared to a graph paper. Fig. 1 shows the two-dimensional surface.

Some definitions for latitudes and longitudes include:

Latitude - (shown as a horizontal line) is the angular distance, in degrees, minutes, and seconds of a point north or south of the equator. Lines of latitude are often referred to as parallels because they are parallel to each other

Longitude - (shown as a vertical line) is the angular distance, in degrees, minutes, and seconds, of a point

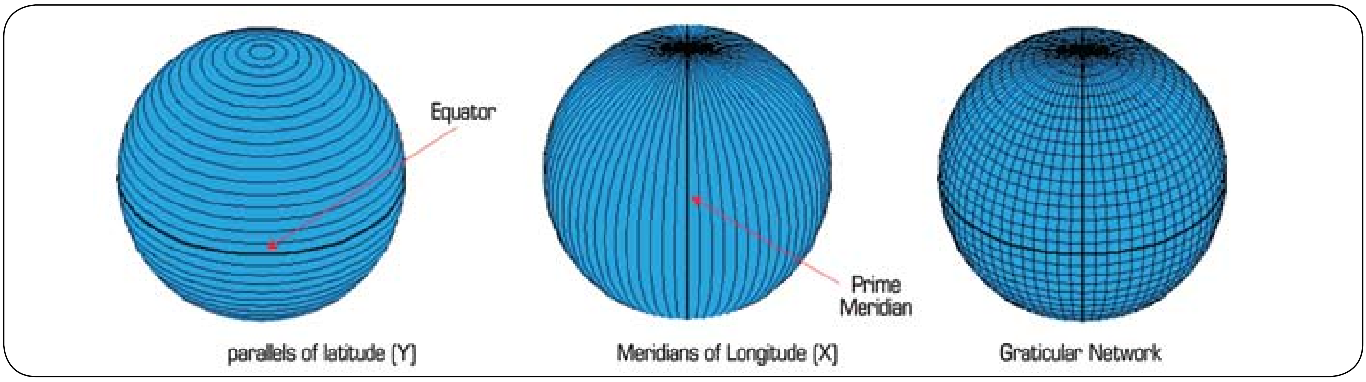

Fig. 2: Images of a globe, displaying lines of reference. (Credit: gis.washington.edu)

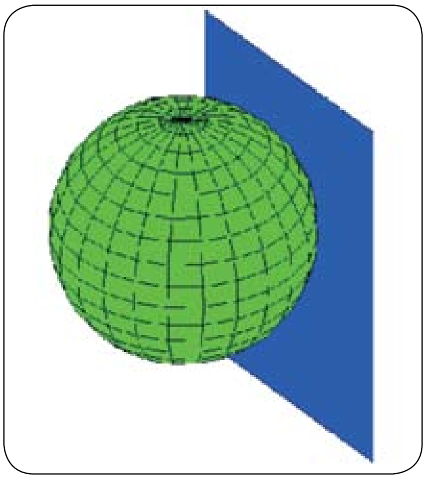

Fig. 3: Planar projection sufrace. (Credit: Peter H. Dana)

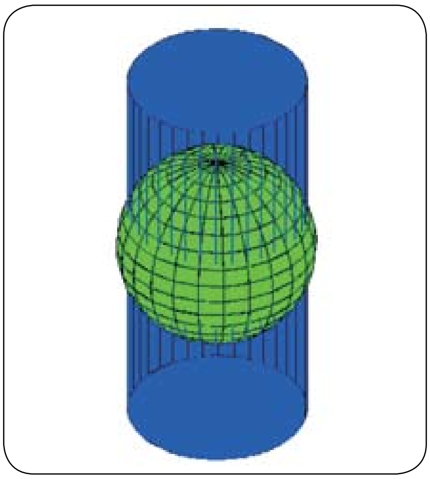

Fig. 4: Cylindrical projection surface. (Credit: Peter H. Dana)

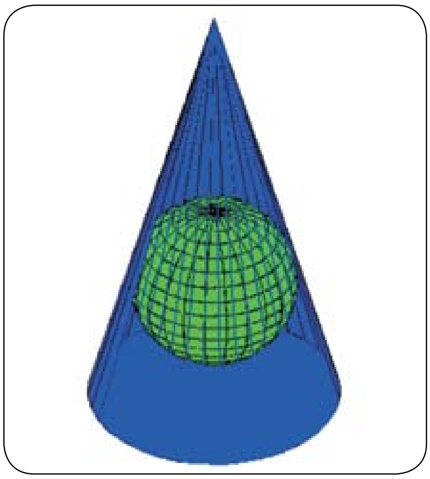

Fig. 5: Conical projection surface. (Credit: Peter H. Dana)

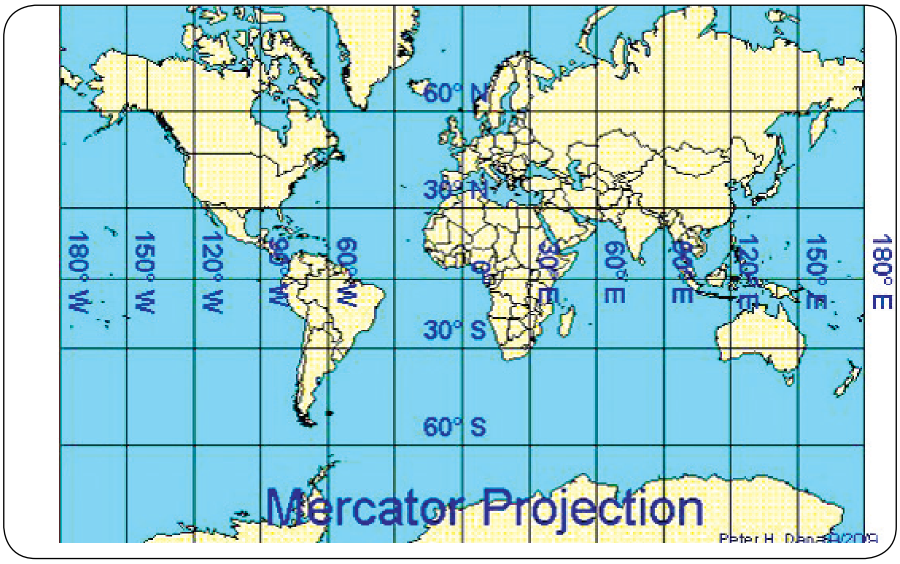

Fig. 6: Mercator projection. (Credit: Peter H. Dana)

east or west of the prime (Greenwich) meridian. Lines of longitude are often referred to as meridians because they run from the central meridian and are "bent" according to the shape of the earth.

Fig. 2 indicates the difference between lines of latitudes and longitudes and how they are measured.

But, now we – especially the GIS community – have a big problem. We need to portray the Earth, which is essentially round, onto a flat surface, a paper or computer map. How do we do that? Essentially we need to flatten the round shape to fit onto a flat surface.

In the process of flattening the Earth to fit onto the flat surface, we distort what is presented on the Earth. The poles will end up presented smaller in relation to their actual size, because of the convergence of the meridians to the poles.

In order to be able to minimise distortion, different ways of projecting the Earth are used. In essence there are three different ways to project the Earth: planar, cylindrical and conical.

A planar projection projects the Earth onto a flat surface, like a piece of paper (see Fig. 3). Basically this means that you put a light bulb in the Earth model and let it shine through onto a flat piece of paper. Distortion is limited around the equator, where the Earth touches the paper, whilst distortion increases towards the poles and moving away along the equator, as the Earth curves away from the flat piece of paper.

The second type of projection is a cylindrical projection. In this case a cylinder is put around the Earth model and the light bulb shines through reflecting on the cylinder. As you can see in Fig. 4 the distortion will be minimal along the equator, all around the Earth, where the cylinder touches the Earth. Distortion increases, the further you move to the poles, because of the Earth curving away from the cylinder.

The third way to project is the conical projection. In this case a cone is put over the Earth. The tip of the cone points upwards. As you can see in Fig. 5 distortion is minimal along the equator where the cone touches the Earth. Distortion is limited upwards, toward the North Pole, because the cone follows the curve of the Earth, although not exactly. Distortion downwards, toward the South Pole is more because the cone moves away from the curve of the Earth.

These are the three major methods of projections. But, we still have a problem with the previous projections, as distortion still happens. In order to minimise distortion, we also have to take into account what the map will be used for. Based on this, projections can also be categorised in: conformal, equal-area or equivalent and equidistant.

In the case of conformal projections, the direction is conserved, meaning the direction will always be represented correctly. This kind of map will be used for navigation by compass, where direction is important.

In the case of equal area (equivalent) projections, the area is preserved, meaning that the area will always be true. These maps will be important to surveyors, landscape architects and town and regional planners.

Equidistant projections preserve distance, meaning that distance will

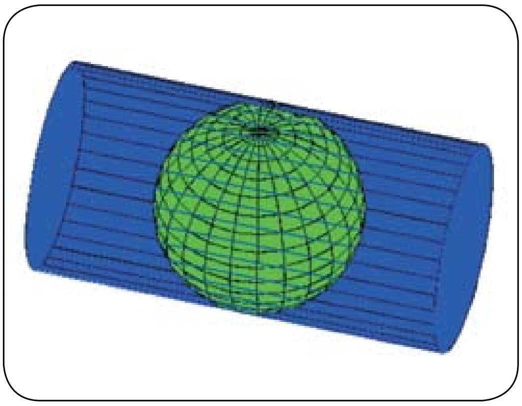

Fig. 7: Transverse cylindrical projection surface. (Credit: Peter H. Dana)

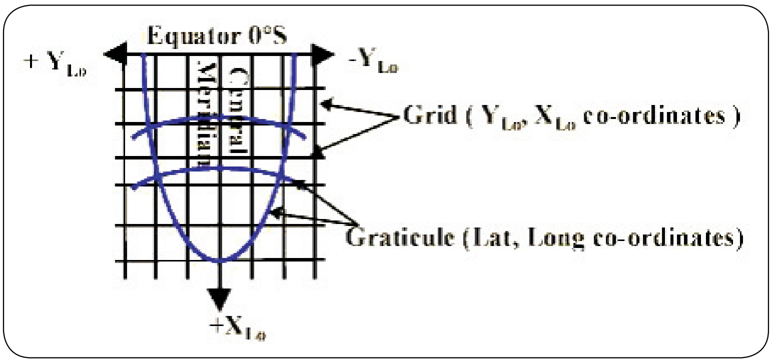

Fig. 8: The difference between the geographic coordinate system and the Mercator protection. (Credit: CD:SM)

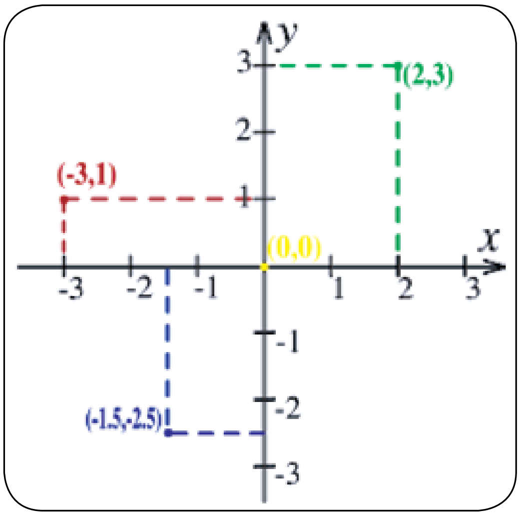

Fig. 9: The Cartesion coordinate system. Four points are marked: (2,3) in green, (-3,1) in red, (-1.5, -2.5) in blue and (0,0), the origin, in yellow. (Credit: Wikipedia)

always be true. This is very important for pilots, who have to determine fuel load and have to know exactly how far they need to fly.

Different people throughout time developed different combinations of the three major types of projections (together with the supportive types of projections) resulting in long lists of projections. Some of these include, and may sound familiar: Gauss Conformal, Lambert Conic projection, Albers Equal Area, Cylindrical Equal Area, and more.

When we look at the projections, it is clear that countries with a large east-west stretch are accommodated nicely, but countries, such as South Africa, with a large north-south stretch are not really accommodated, since the line touching the Earth runs along the equator (east-west) and not the central meridian (north-south). This logically results in projection being true along the equator and not the central meridian. To compensate for this, it was decided to create projections where the tangent line (the line touching the Earth) runs along the central meridian. These projections are referred to as transverse projections – basically meaning the projections were turned 90°.

Mercator projected the Earth based on a cylindrical projection – where a cylinder is put around the Earth and projected on the cylinder. In this case parallels and meridians are perpendicular to each other and run in straight lines as can be seen in Figure 6. This projection was used to calculate *x*-, *y*- coordinates known as the LO (local orientation) coordinate system.

As you can see, distortion increases to the poles (north-south) and South Africa stretches in a north-south direction. Therefore, to minimise distortion, in South Africa, we turn the cylinder 90°, resulting in a transverse projection. We call this the Transverse Mercator Projection, because we use Mercator's projection and make it transverse. Fig. 7 illustrates how the cylinder is turned 90°.

In South Africa we use the LO coordinate system quite extensively. Only the area within one degree of longitude on either side of the central meridian is projected. The width of each segment, often referred to as a belt, is thus two degrees of longitude and is referred to the central meridian (CM) of that belt. Each zone is named after the longitude of origin i.e. LO 17°, LO 19°, LO 21° etc. LO's are always uneven, because they start in bands - which are 1° east and 1° west - from 0°. This means that the central line will be 1°, thus resulting in LO's being uneven.

In Fig. 8 you can see the difference between the geographic coordinate system and the Mercator projection. You will see that the coordinates for the geographic coordinate system (blue) are based on the shape of the Earth itself, while the Mercator projection (black) shows the latitudes and longitudes as straight, perpendicular lines.

So where do *x*-coordinates and *y*-coordinates fit in? When do we use them and why? These are referred to as Cartesian coordinates and measure the distance from the centre of the Earth. Remember, up until now, latitudes and longitudes were lines measured from the equator and the central meridian, meaning that they were measured from lines of zero on the surface of the Earth. In the case of Cartesian coordinates, they are measured from the actual centre of the Earth – the midpoint of the Earth. Therefore they also include a *z*-value, which is height. For this same reason, Cartesian coordinates are always portrayed in metres, as measured from the midpoint of the Earth. Measurement of the *x*- and *y*-coordinates is based on the mathematical calculation of *x*- and *y*-values, as portrayed in Fig. 9.

Remember, when we discussed projections, we mentioned that in South Africa, we turn the projection 90°? When we work in Cartesian coordinates, we do the same, because we still have to project, resulting in Fig. 10, where South Africa will now have a positive *x* and a positive or negative *y* value, depending on which longitude was used as the central meridian (the tangent line).

If we work with the LO-system, we will, for instance, use longitude 27 east as our tangent line (central meridian), meaning that east of LO27 our *y*-coordinate will be negative, and west of LO27 our *y*-coordinate will be positive.

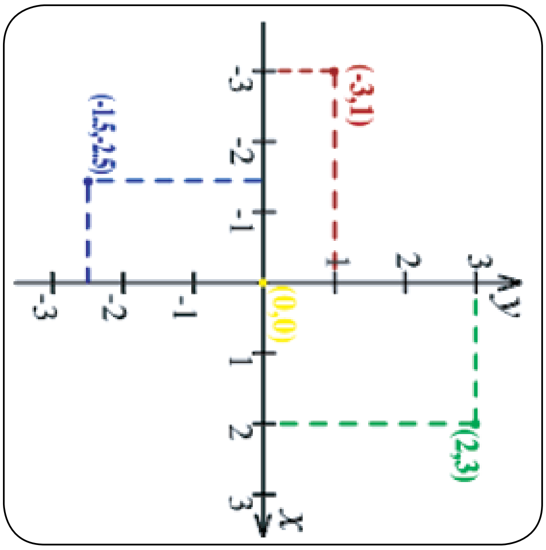

Fig. 10: Cartesian coordinates showing South Africa with a positive x and a positive y value.

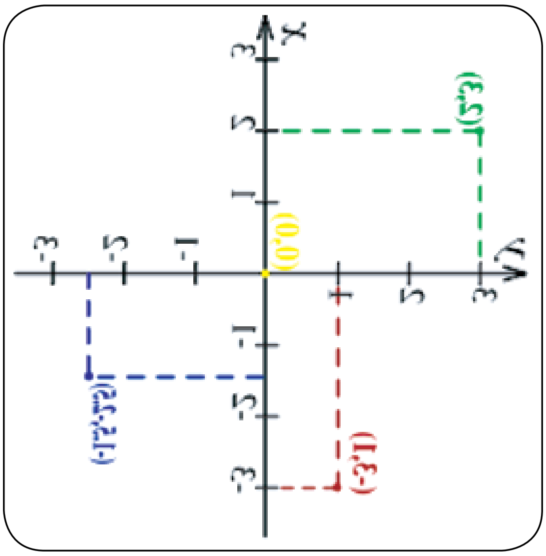

Fig. 11: Cartesian coordinates showing South Africa's y-coordinate as negative and its x-coordinate as positive.

However, most GIS software is developed in America, which is situated in the Northern Hemisphere. Because they want to be positive (necessitating that north be positive) they make south negative, resulting in us being negative. This means that we have to flip our picture over the horizontal axis, which is now the *y*-axis, because we rotated it 90°. Now our reference picture looks like that in Fig. 11, with our *y*-coordinate being negative and our *x*-coordinate being positive (South Africa is situated south of the equator and east of the Greenwich meridian, which we assume as the tangent line).

Because *x* is the distance as measured from the equator it will always be a large number. So if you receive any data, and there is no indication of which one is the *x*- and which one the *y*-coordinate, you can safely assume that the largest coordinate is your *x*-coordinate.

In some cases mapmakers make use of a constant, which is the first part of the *x*-coordinate, repeated in all the coordinates, because the *x*-coordinate becomes too long to use easily. The constant is indicated on the map, as a constant, and left out of the actual written *x*-coordinate, e.g.:

x-coord1 = -2 789 456

x-coord2 = -2 789 573

The constant is 2 789

x-coord1 = -456

x-coord2 = -573

Due to evolving knowledge regarding the shape of the earth (i.e. from a round ball to an ellipse form), the reference shape for projections changed over time. This, however, is a science on its own which will be explained in more detail in a follow-up article, where all the different factors and influences will be brought together.

Conclusion

In summary, it is important to remember:

- If you capture latitudes and longitudes in CAD, it represents an actual coordinate as on the Earth – without projection.
- In order to portray the round Earth on a flat paper surface, we need to project the Earth onto the paper.
- Projections result in distortion and different kinds of projections are used for different outcomes.
- Latitudes and longitudes are measured from the equator and the central meridian, on the surface of the Earth, based on lines of longitude and latitude, and are portrayed in degrees, minutes, seconds or decimal degrees.
- *X*- and *y*-coordinates, referred to as Cartesian coordinates, measure the distance from the centre of the Earth – the midpoint of the Earth – and are based on the mathematical calculation of your *x*-, *y*-, and *z*-axes.
- Cartesian coordinates are used when we work in projections.
- Because Cartesian coordinates measure the distance from the midpoint of the Earth, they are portrayed in metres.
- People using CAD work in latitudes and longitudes (degrees, minutes, seconds) which means not projected.
- People working in GIS work with Cartesian coordinates, which means projected.
- In South Africa we use the Transverse Mercator Projection, meaning we rotate our cylinder 90° - meaning that our *x*-axis is vertical and our *y*-axis is horizontal.
- Because most GIS software is developed in America, the Northern Hemisphere is made positive, meaning that we are in the negative hemisphere. This means that our *y*-coordinate is negative. That is also why you may sometimes get a mirror image of your map when you open it up – especially when you get it from a CAD program.
- Because our *x*-axis measures the distance from the equator, the *x*-coordinate will always be the biggest.

Acknowlegement

This article is based on a presentation given at the GISSA Gauteng general meeting, in November 2007.

Contact Laurette Coetzee, Open Spatial Solutions, Tel 012 666-7000, lcoetzee@openspatial.co.za