The Geocentric Mercator Projection: A Fast and Accurate Approximation of the Ellipsoidal Mercator Projection

Bastian Erdnüß

May 26, 2022

1 Introduction

The Mercator projection is a conformal map of the world which is globally defined (except for the polar regions) and for which grid north agrees with true north everywhere. These two properties (globally defined and properly north-orientated) make it the preferred map in many applications, despite its downside, an increasing scale of the map with increasing distance from the equator.

There are two common versions of the Mercator map, an elliptical one (EPSG:3395 World Mercator) based on the WGS 84 reference ellipsoid (EPSG:7030 WGS 84) and a spherical one (EPSG:3857 Pseudo Mercator) based on the geographic latitude. This article proposes a third version that is almost as easy to calculate as the spherical version but almost as accurate as the ellipsoidal version, the spherical Mercator projection based on the geocentric latitude.

Compared to the spherical Mercator projection based on the geographic latitude, the maximum non-conformity (maximal deviation between a measured angle in the map and in local space) reduces from 6' (arc minutes) by factor 600 to just 0.6" (arc seconds). And even in absolute coordinates, the maximum error compared to the ellipsoidal Mercator projection reduces from over 21 km by a factor of 1300 to under 16 m.

2 Formulas

Let a be the equatorial radius of the earth and f be its flattening (e.g. in WGS 84: $a = 6378137\,\text{m}$ and $f = 1/298.257223563$).

Let $\varphi \in (-\pi/2, \pi/2)$ denote the geographic latitude and $\lambda \in [-\pi, \pi]$ denote the geographic longitude of a point of interest (for simplicity of the formulas, we assume them to be already converted to radians). The geocentric latitude ϑ relates to the geographic latitude φ by

$$\tan\vartheta = (1-f)^2 \tan\varphi.$$

The Mercator coordinates are easting x and northing y, which for simplicity are also given in radians in this section. Easting

$$x = \lambda$$

simply corresponds to longitude and northing and latitude are connected by the formulas

$$\sinh y = \tan\vartheta = (1-f)^2 \tan\varphi.$$

The scale of the map is each of

$$1 : a\cos\vartheta$$
$$1 : a\,\operatorname{sech} y.$$

A common choice is to multiply x and y by a to make the scale $1 : 1$ at the equator.

3 Derivation

3.1 Spherical Mercator projection

For the spherical Mercator projection on longitude λ and (arbitrary) latitude α, the following formulas hold (among many known others [1, (2.32)]):

$$
\begin{aligned}
&x = \lambda \\
&\sinh y = \tan \alpha \\
&\tanh y = \sin \alpha \\
&\operatorname{sech} y = \cos \alpha.
\end{aligned}
$$

For a sphere of radius r the scale of the map is given by either of the following two formulas [1, (2.36)]

$$
\begin{aligned}
&1 : r \cos \alpha \\
&1 : r \operatorname{sech} y.
\end{aligned}
$$

3.2 Elliptical Mercator projection

For the elliptical Mercator projection in terms of longitude λ and geographic latitude φ it holds [1, (6.8)]

$$
\begin{aligned}
&x = \lambda \\
&y = \operatorname{arsinh}(\tan \varphi) - e \operatorname{artanh}(e \sin \varphi)
\end{aligned}
$$

for an oblate ellipsoid of revolution with numerical eccentricity e and the projection has a scale of [1, (6.17), (6.5)]

$$1 : a \cos \varphi \Big/ \sqrt{1 - e^2 \sin^2 \varphi}$$

where a is its semi-major axis. There is no closed form available to express φ in terms of y but has to be calculated by iterative methods.

3.3 Elliptical coordinates

In elliptical coordinates, Euclidean 3-space $W = (W_1, W_2, W_3)$ can be parametrized by $p = (\rho, \beta, \lambda)$ per

$$W = E\,(\cosh \rho \cos \beta \cos \lambda, \cosh \rho \cos \beta \sin \lambda, \sinh \rho \sin \beta),$$

where β is also called the parametric latitude, $E = ae$ is the linear eccentricity of an oblate ellipsoid of revolution, and e is its numerical eccentricity which is related to the flattening f by $1 - e^2 = (1 - f)^2$.

Denote by $\partial W / \partial \rho$, $\partial W / \partial \beta$, and $\partial W / \partial \lambda$ the derivative of W by the parameters p, i.e.

$$
\begin{aligned}
\partial W / \partial \rho &= E\,(\sinh \rho \cos \beta \cos \lambda, \sinh \rho \cos \beta \sin \lambda, \cosh \rho \sin \beta) \\
\partial W / \partial \beta &= E\,(-\cosh \rho \sin \beta \cos \lambda, -\cosh \rho \sin \beta \sin \lambda, \sinh \rho \cos \beta) \\
\partial W / \partial \lambda &= E \cosh \rho \cos \beta (-\sin \lambda, \cos \lambda, 0).
\end{aligned}
$$

By calculating their pairwise scalar products it turns out that these vectors are orthogonal.

For the WGS 84 reference ellipsoid it holds that

$$e = 0.0818191908426,\ e^2 = 0.00669437999014, \text{ and } E = 521854.008423\,\text{m}.$$

For each fixed ρ the parametrization describes an ellipsoid and when $a = E \cosh \rho$ it describes the surface for the reference ellipsoid which for WGS 84 is the case when

$$\rho = \rho_0 = 3.1947128245.$$

In the remainder of this article W is restricted to $\rho = \rho_0$, i.e. the reference ellipsoid. In this case $\partial W / \partial \beta$ and $\partial W / \partial \lambda$ span the tangent space to the ellipsoid and since $\partial W / \partial \rho$ is orthogonal to it, it is parallel to the local zenith direction.

3.4 Relation between the different latitudes

Since $\partial W/\partial\rho$ is the local zenith direction, for the geographic latitude φ holds that

$$\tan\varphi = (\partial W_3/\partial\rho) \Big/ \sqrt{(\partial W_1/\partial\rho)^2 + (\partial W_2/\partial\rho)^2} = \tan\beta \Big/ \tanh\rho_0,$$

where $\sqrt{(\partial W_1/\partial\rho)^2 + (\partial W_2/\partial\rho)^2} = E\sinh\rho_0\cos\beta$. On the other hand for the geocentric latitude ϑ it holds that

$$\tan\vartheta = W_3 \Big/ \sqrt{W_1^2 + W_2^2} = \tan\beta\tanh\rho_0.$$

Moreover, from $a = E\cosh\rho_0$ it follows that $\operatorname{sech}\rho_0 = E/a = e$ and therefore, $\tanh^2\rho_0 = 1 - \operatorname{sech}^2\rho_0 = 1 - e^2 = (1-f)^2$, i.e. $\tanh\rho_0 = 1 - f$, using the hyperbolic trigonometric identity $\tanh^2\alpha = 1 - \operatorname{sech}^2\alpha$. From this the relation between the different latitudes follows

$$\tan\vartheta = (1-f)\tan\beta = (1-f)^2\tan\varphi.$$

3.5 Scale of the Mercator map and deviation from conformity

The Mercator projection maps the longitude to x and the latitude to y. The map is conformal, when $\|\partial W/\partial x\| = \|\partial W/\partial y\|$, i.e. at each position the units along the x- and y-axes are of equal length in the tangent space of the ellipsoid. This is the case for the elliptical Mercator projection by construction and for its spherical approximations, the quotient $s = \|\partial W/\partial y\| \big/ \|\partial W/\partial x\|$ can be used as a measure of the "unconformity" of the projection with $s = 1$ indicating a conformal map. By the chain rule

$$s = \|\partial W/\partial\beta \;\, \mathrm{d}\beta/\mathrm{d}y\| \Big/ \|\partial W/\partial\lambda \;\, \mathrm{d}\lambda/\mathrm{d}x\| = J s_0$$

with $J = \mathrm{d}\beta/\mathrm{d}y$ and $s_0 = \|\partial W/\partial\beta\| \big/ \|\partial W/\partial\lambda\|$, since $\mathrm{d}\lambda/\mathrm{d}x = 1$ for $\lambda = x$. $\partial W/\partial\beta$ and $\partial W/\partial\lambda$ were already calculated and by taking their norms and minor simplifications it follows

$$\begin{aligned} s_0 &= \|\partial W/\partial\beta\| \Big/ \|\partial W/\partial\lambda\| = \sqrt{\tanh^2\rho_0 + \tan^2\beta} = \sqrt{1 - e^2 + \tan^2\beta} \\ &= \sqrt{1 - e^2\cos^2\beta}\sec\beta. \end{aligned}$$

3.5.1 Pseudo Mercator projection

For the spherical Mercator map based on the geographic latitude φ it holds that

$$\begin{aligned} \sinh y &= \tan\varphi = \tan\beta \Big/ (1-f) \\ \cosh y &= \sec\varphi = \sqrt{1 + \tan^2\beta \Big/ (1-f)^2} \end{aligned}$$

and therefore by differentiating the first equation

$$\cosh y\,\mathrm{d}y = \sec^2\beta\,\mathrm{d}\beta \Big/ (1-f)$$

or with some simplification involving $(1-f)^2 = 1 - e^2$ and the formula for $\cosh y$ above

$$\begin{aligned} J &= \mathrm{d}\beta/\mathrm{d}y = (1-f)\cosh y\cos^2\beta = \sqrt{1 - e^2\sin^2\beta}\cos\beta \\ s &= J s_0 = \sqrt{1 - e^2\cos^2\beta}\sqrt{1 - e^2\sin^2\beta} = \sqrt{1 - e^2 + e^4\sin^2\beta\cos^2\beta}. \end{aligned}$$

It is always $s \le 1$ and s smallest at the equator and poles where

$$s = \sqrt{1 - e^2} = 1 - f.$$

I.e. at the equator a course of 45° in the map corresponds to a course α in on the ellipsoid for which $\tan\alpha = 1 - f$ which deviates by $|\alpha - 45°| = 0.10° = 6'$ (arc minutes) from the course on the map for the WGS 84 ellipsoid.

3.5.2 Geocentric Mercator projection

For the spherical Mercator projection on the geocentric latitude ϑ it holds that

$$\sinh y = \tan\vartheta = (1-f)\tan\beta$$
$$\cosh y = \sec\vartheta = \sqrt{1+\tan^2\vartheta} = \sqrt{1+(1-f)^2\tan^2\beta}$$

and therefore analogously to the last section

$$\cosh y\,\mathrm{d}y = (1-f)\sec^2\beta\,\mathrm{d}\beta$$
$$J = \mathrm{d}\beta/\mathrm{d}y = \cosh y\cos^2\beta \,\big/\, (1-f) = \sqrt{1-e^2\sin^2\beta}\cos\beta \,\big/\, (1-f)$$
$$s = J s_0 = \sqrt{1-e^2\cos^2\beta}\sqrt{1-e^2\sin^2\beta} \,\big/\, (1-f) = \sqrt{1+e^4/(1-e^2)\sin^2\beta\cos^2\beta}.$$

It is always $s \geq 1$ and s is largest for $\beta = \pi/4$ with

$$s = \sqrt{1+e^4/4 \,\big/\, (1-e^2)} \approx 1 + e^4/8$$

I.e. at parametric latitude of $\beta \approx \pm 45°$, a course of 45° in the map corresponds to a course of α with $\tan\alpha \approx 1 + e^4/8$ on the ellipsoid which deviates by $|\alpha - 45°| \approx 0.00017° \approx 0.6''$ (arc seconds) from the course on the map for the WGS 84 ellipsoid which is by a factor of 600 smaller than for the Pseudo Mercator projection.

3.6 Absolute coordinates

The absolute values of the coordinates play only a minor role. They are useful to illustrate the accumulated deviation from conformity over the globe. The only other role that they can play is to provide redundency and improve error recognition in the form that if someone copies e.g. elliptical Mercator coordinates in a software expecting spherical Mercator coordinates or vice versa, it is advantageous if the result is either so close that it doesn't play a role (e.g. within 10 m) or it is so far appart that there is obviously something wrong (e.g. a coordinate on an entirely wrong continent).

Since for all analyzed projections it was $x = \lambda$ a deviation can only occur in the y-coordinate. For the elliptical Mercator projection it holds that

$$y = \operatorname{arsinh}(\tan\varphi) - e\operatorname{artanh}(e\sin\varphi),$$

for the spherical Mercator projection on the geographic latitude it holds that

$$y = \operatorname{arsinh}(\tan\varphi),$$

and for the spherical Mercator projection on the geocentric latitude it holds that

$$y = \operatorname{arsinh}((1-e^2)\tan\varphi).$$

3.6.1 Pseudo Mercator projection

For the spherical mercator projection on the geographic latitude obviously the error on the y-axis is

$$\Delta y = e\operatorname{artanh}(e\sin\varphi)$$

which in absolute value is largest at the poles with value $\Delta y = e\operatorname{artanh} e \approx 0.0067$ but in meters corresponds to an error of only

$$\Delta = e\operatorname{artanh}(e\sin\varphi)\; a\cos\varphi \,\big/\, \sqrt{1-e^2\sin^2\varphi}.$$

Expanding the Taylor series of Δ by e to second order using $\operatorname{artanh} x \approx x$ yields

$$\Delta \approx ae^2\ \sin\varphi\cos\varphi$$

which is largest an latitude 45° where it has a value of $\Delta > 21\,\text{km}$.

3.6.2 Geocentric Mercator projection

For the geocentric Mercator projection on the geocentric latitude, it is

$$\Delta y = \operatorname{arsinh}((1-e^2)\tan\varphi) - \operatorname{arsinh}(\tan\varphi) + e\operatorname{artanh}(e\sin\varphi).$$

Using the Taylor expansions $\operatorname{artanh}(x) \approx x+x^3/3$ and $\operatorname{arsinh}(y+x) \approx \operatorname{arsinh}(y)+x/\sqrt{1+y^2}-x^2/2\; y/\sqrt{1+y^2}^3$ together with the trigonometric identity $\cos\varphi = 1/\sqrt{1+\tan^2\varphi}$ for $|\varphi| < \pi/2$, Δy can be expanded by e to

$$\Delta y \approx -e^2\tan\varphi\cos\varphi - e^4/2\;\tan^3\varphi\cos^3\varphi + e^2\sin\varphi + e^4/3\;\sin^3\varphi,$$

where the second order terms in e cancel and the fourth order terms reduce to

$$\Delta y \approx -e^4/6\;\sin^3\varphi.$$

This leads to an deviation in meters of

$$\Delta = \Delta y\, a\cos\varphi \;\big/\; \sqrt{1-e^2\sin^2\varphi} \approx -ae^4/6\;\sin^3\varphi\cos\varphi$$

where the nominator can be ignored since it would only contribute terms in the order of e^6. The absolute deviation is largest at latitude $\varphi \approx \pm 60°$ with value $\Delta \approx \sqrt{3}/32\; ae^4 < 16\,\mathrm{m}$ which is by a factor of over 1300 smaller compared to the Pseudo Mercator projection.

3.7 Optimized calculation

In many numerical implementations, arctan and *artanh* are faster and more accurate implemented compared to their other inverse trigonometric and hyperbolic counter parts. For that reason it can be advantageous to implement the necessary formulas in the following way:

3.7.1 Geographic to map

$$s = (1-e^2)\sin\varphi \;\big/\; \sqrt{1-(2-e^2)e^2\;\sin^2\varphi}$$
$$y = \operatorname{artanh} s$$
$$\text{scale: } 1 : a\sqrt{1-s^2}$$

3.7.2 Map to geographic

$$t = \sinh y$$
$$\varphi = \arctan(1/(1-e^2)\,t)$$
$$\text{scale: } 1 : a \;\big/\; \sqrt{1+t^2}$$

4 Summary

The maximal deviation from conformity is for the spherical Mercator projection based on the geocentric latitude in the order of $e^4/8$ while it is for the spherical Mercator projection based on the geographic latitude in the order of $f = e^2/2$, which is $4/e^2 = 2/f \approx 600$ times as high for the WGS 84 ellipsoid.

Compared to the elliptical Mercator projection, each local map is squeezed slightly along the y-axis. This effect is strongest around latitude 45° where the y-axis is squeezed by about 0.006‰. This results in deviation of courses in the map around the diagonals (NE, NW, SE, and SW directions) of up to 0.6” (arc seconds, in degrees).

In comparison, for the Pseudo Mercator projection, maps at the quator are stretched along the y-axis by over 3‰ and courses around the diagonals deviations by up to about 6’ (arc minutes) which is around 600 times higher than for the Geocentric Mercator projection.

The deviation in the absolute y-values between the elliptical and the Geocentric Mercator projections increases on the ways to the poles. However, the absolute distance in meters is largest around latitude 60° and below 16 m.

On the other hand, the deviation between the y-values of the elliptical and the spherical Mercator projection on the geographic latitude is highest at about 45° latitude and goes up to over twenty kilometers which is over 1300 times as high.

5 Conclusion

The spherical Mercator projection based on the geocentric latitude increases the conformity of the map projection (in the sense of quasi-conformal maps) by orders of magnitude while keeping the necessary formulas for the calculations almost as simple as for the usual spherical Mercator projection. It is therefore suitable as replacement in almost all places where the spherical Mercator projection is favored over the elliptical Mercator projection.

References

[1] Osborne, P. (2013). The Mercator Projections. Zenodo.