THE NEW 2022 DATUMS: A BRIEF BOOK

Dr. Alan Vonderohe
Professor Emeritus
University of Wisconsin-Madison

May, 2019

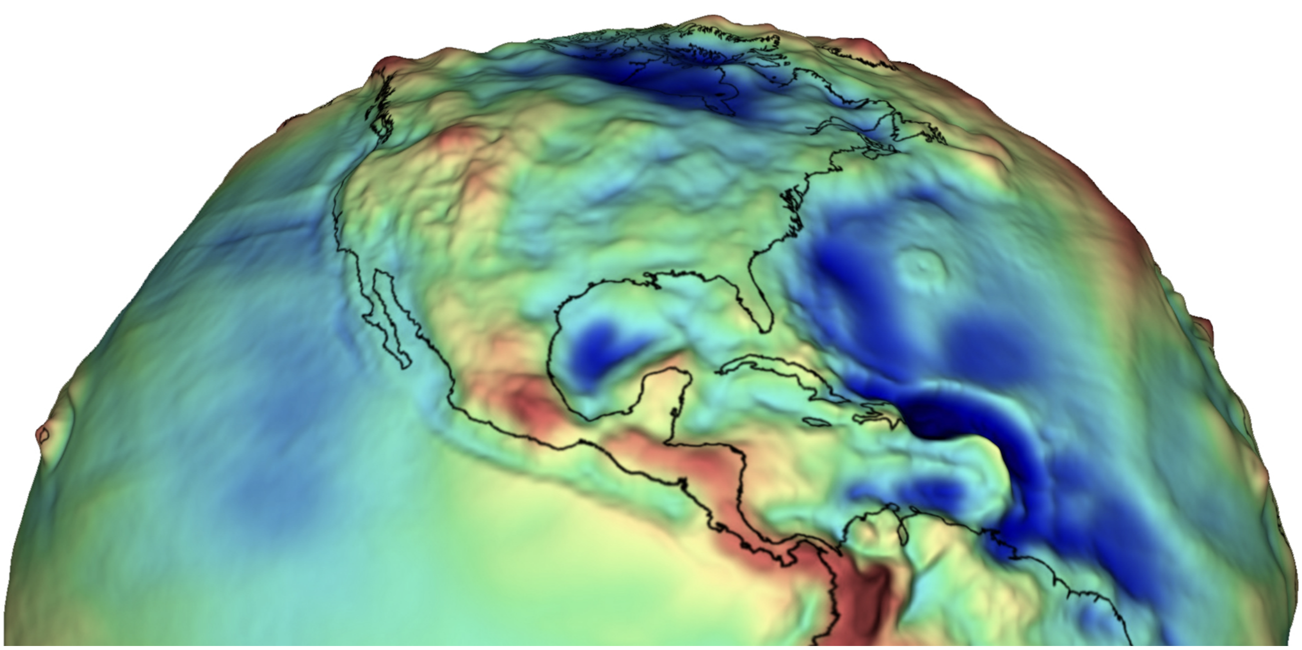


TABLE OF CONTENTS

FOREWORD

This short book was written by Alan Vonderohe in the spring of 2019, with successive chapters being released at intervals over a period of a month.

The initial audience for the book was the Wisconsin Spatial Reference System 2022 Task Force, or WSRS2022, a broad-based group organized to address how the National Geodetic Survey's plans for its new 2022 datums would impact the Wisconsin geospatial community.

The chapters are collected here in a single document available to anyone who wants to understand more about the new datums.

May, 2019

PREFACE

In 2022, the National Geodetic Survey will adopt and publish two new geodetic datums that are the underlying bases for positioning, mapping, and navigation. The new datums will impact many human activities. They will have non-negligible positional effects in most of the United States, including Wisconsin. The new datums will also account for positional changes within them over time, an aspect of reality not heretofore comprehensively addressed.

This brief book, of four chapters, is intended for a broad audience who seek background information on datums and coordinate systems to better understand the upcoming changes and their possible effects on work activities and products. The book is conversational and non-technical in nature.

The four chapters will be made available in sequence over a period of 4 to 6 weeks. Chapter One will address reference frames and how they are established. Chapter Two will build upon reference frames and will discuss the upcoming North American Terrestrial Reference Frame of 2022 (NATRF2022). Chapter Three will delve into gravity and how it behaves on Earth, forming the basis for elevations. Chapter Four will build upon concepts presented in Chapter Three to describe the North American-Pacific Geopotential Datum of 2022 (NAPGD2022).

To present nearly bare fundamentals, considerable liberty has been taken in simplifying not only concepts but also history. As examples, no mention is made of the celestial sphere, WGS 84, dynamic heights, or components of NAPGD2022 other than GEOID2022.

I would like to thank Glen Schaefer, Sam Wenz, Dan Rodman, and Brad Hollister for reviewing drafts of this book.

Enjoy.

Alan Vonderohe, April 2019

CHAPTER ONE

UNDERLYING REFERENCE FRAMES

Conversation 1

Guru: “Everything moves all the time”.

Me: “What? That rock over there isn’t moving. It’s sitting still”.

Guru: “Take two steps toward the rock”.

Me: “Ok”.

Guru: “Now, did the rock move”?

Me: “No”.

Guru: “How do you know”?

Me: “Because it’s in the same place”.

Guru: “What do you mean by ‘place’”?

Me: “I mean ‘a location on Earth’s surface’”.

Guru: “Then your stationary reference frame for ‘location’ or ‘place’ is Earth’s surface, right”?

Me: “Yes”.

Guru: “What if your stationary reference frame was inside your head, right behind your eyes? Wouldn’t it be that the steps you took caused Earth’s surface to move under your head and, therefore, move the rock toward you”?

Me: “That’s preposterous. Earth’s surface and the rock didn’t move. I did”.

Guru: “That depends on your choice of ‘reference frame’. When you fly on a clear day at 35,000 feet, doesn’t it appear that Earth and the rocks on it are moving under the aircraft”?

Me: “Yes”.

Guru: “That’s because you have chosen your reference frame to be the aircraft. If you are standing on the surface of the Moon (new reference frame) with a telescope, both Earth’s surface with its rocks and the aircraft will appear to be moving, but at different speeds and perhaps in different directions. The truth is that everything is moving all the time. The universe is expanding and taking everything with it and, as it goes, things are moving around inside of it with respect to one another…”

Me: “Alright, alright! I don’t want to know about the universe expanding. I’m just trying to make a map here on Earth”.

Guru: “Are you going to use GPS to help make your map”?

Me: “Yes”.

Guru: “Of course you don’t have to be an astrophysicist to use GPS. However, if you want to understand the underlying reference frame for GPS positioning and, therefore, your map, that reference frame does have something to do with where things are in the universe”.

Me: “How can that possibly be? And, even if it is, I don’t need to know about it. All I want to do is make my map”.

Guru: “Oh, then this stuff with new datums doesn’t matter”?

Me: “I know that new datums cause me to deform my maps or make brand new maps altogether, but I don’t understand these datums and why they keep changing. Besides, we’ve been talking about reference frames. What do reference frames have to do with these new datums”?

Guru: “A reference frame and a datum aren’t very much different from one another. Some might say they aren’t different at all”.

Me: “Oh, well, tell me more, then”.

International Celestial Reference Frame (ICRF)

Let’s say we want to establish a reference frame that is as stable as we can make it and consists of a coordinate system referenced by selected physical objects. We specify the origin to be at the most stable, physically-describable point we can find in the spatial realm of human endeavors. That would be the center of mass of the solar system, also known as its “barycenter”. This is the point around which all objects in the solar system orbit.

Now, we reference the axes of the coordinate system by objects whose directions from the origin do not change over time, at least not by any amount we can measure. Those are certain quasars, the most distant objects observed in the universe. The selected quasars are not only very far away, they are also moving farther away from us faster than any other objects. But, changes in <u>directions</u> to these quasars, if such changes exist at all, are so small we cannot observe them. Directions to the defining quasars are measured using an extremely precise radio telescope technique. Specification and publication of those directions provide us with our very stable “celestial” reference frame.

But, wait, that reference frame is <u>not perfectly</u> stable because our measurements of directions to quasars, although extremely precise, are not perfect. In addition, we keep making new measurements and we make some of them from more widely-distributed places. So, our best estimates for true values of the directions keep improving. Also, the coordinate system axes are referenced by a subset of the known quasars. If new measurements cause significant enough changes, or if we decide to add or remove quasars in the defining set, we publish a new celestial reference frame. This is similar to the horizontal geodetic datum, NAD 83, wherein new adjustments were performed as measurements continued to improve and new control points became available, such as Wisconsin's High Accuracy Reference Network (HARN). There have been three published International Celestial Reference Frames (ICRFs). The latest, ICRF3, was published on January 1, 2019.

Earth rotates inside the celestial reference frame and its rotation with respect to the frame can be measured. Those measurements tell us what time it is…well, at least according to one standard of time…there are a whole bunch of time standards, but the most familiar one has the length of a day being equal to the duration of one Earth rotation relative to the sun. Yes, but that duration is not uniform unless we let hours, minutes, and seconds continually grow and shrink…our clocks wouldn't do us much good in that case. Earth's rotation rate decreases and increases…for a number of reasons.

Conversation 2

Me: "Ok, ok…What the blankety-blank has any of that to do with the map I'm trying to make? I don't need quasars to make my map and what is this "time" stuff anyway? What has time to do with my map"?

Guru: "Have patience. At some point, I'll try to help you make connections".

Me: "Ok, ok…Let's get down to Earth. I'm getting a headache".

International Terrestrial Reference Frame (ITRF)

To do a survey or make a map of an area of Earth's surface, we need a reference frame that more suits our purpose than one we access by measurements to very distant things in the sky. We need a <u>terrestrial</u> reference frame. It makes very good sense to cause our terrestrial reference frame to be relatable, or tied, to the celestial reference frame that tells us how Earth is rotating and helps us keep time.

A Cartesian (rectangular) coordinate system is based upon specified perpendicular axes. Cartesian coordinates, which are distance offsets from the axes, are convenient because they enable simplified computations. In addition, they are similar to typical human perception of location in three spatial dimensions. So, let's establish a "geocentric," right-handed, three-dimensional Cartesian coordinate system with its origin at Earth's center of mass. That's a good place to put the origin if we are thinking "worldwide" and we might want to use artificial satellites to help us navigate. Those satellites orbit Earth's center of mass.

Now, let's orient the axes of our coordinate system by making the Z axis coincide with Earth's spin axis and making the X axis lie in the plane containing the Z axis and a very famous point at the Royal Observatory in Greenwich, England (Figure 1). Notice that "Z", in the context of geocentric coordinates, is not "up". "Up" has to do with gravity, which we will talk about later.

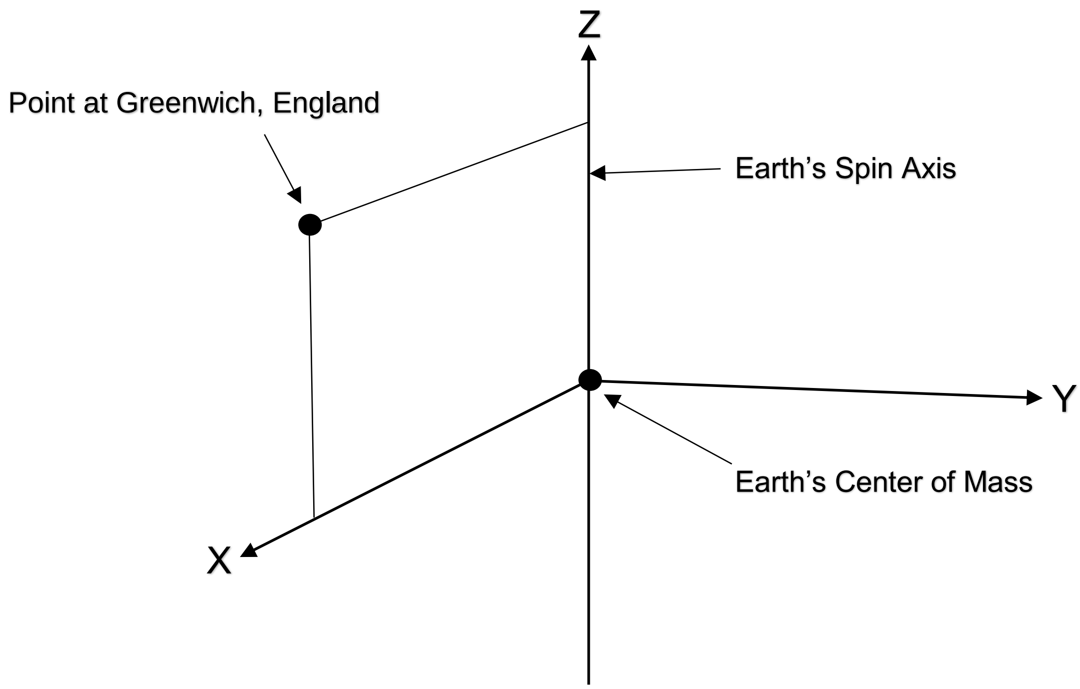


Figure 1.
Geocentric Coordinate System

A difficulty we face with straightforward concepts like this is, as mentioned earlier, all things move all the time. With respect to Earth's solid crust, moving things include Earth's center of mass and its spin axis. Why? Well, among other things, Earth is not a solid body through and through. In fact, a great deal of Earth's interior is a heterogeneous fluid that sort of "sloshes" around slowly. Even the seemingly stationary part (the crust) that we live on, and on which the oceans and atmosphere lie, is moving around because of geophysical phenomena such as plate tectonics and glacial isostatic adjustment.

We can use measurements made in the ICRF, for determining Earth's rotation rate, to also determine where the spin axis is with respect to Earth's crust at any given time (epoch). So, time-stamped spatial offsets to the spin axis constitute part of a set of "Earth orientation parameters" that relate the ICRF to the terrestrial reference frame we are building.

To complete our terrestrial reference frame, we use some additional space-based measurement techniques. GPS is a component of a larger Global Navigation Satellite System (GNSS) with satellites launched by five countries and the European Union. GNSS includes ground-based tracking stations that monitor and upload satellite orbits. Coordinates of tracking stations are very well known. They are so well known that movements and velocities of tracking stations can be determined from coordinate differences over time. The same is true for radio telescopes, used to establish the ICRF, and also for ground-based stations of some other high-precision, space-based measurement techniques, one of which is very effective for locating Earth's center of mass. When we publish these coordinates and velocities, we have an International Terrestrial Reference Frame (ITRF). Because of motions, new measurements, and measurements with less error, there have been nine ITRFs from 1992 to present, the latest being ITRF2014.

International GNSS Service Reference Frame

The International Global Navigation Satellite System Service (IGS) is a collection of government agencies and scientific organizations, under the auspices of the International Association of Geodesy, that encourages and advocates for geometrical linkages among GNSS and other precise geodetic measurement techniques, including all those used to establish ITRFs. Members of IGS own and operate additional GNSS tracking stations beyond those used for ITRF definition. They also have networks of continuously operating reference stations collecting GNSS satellite data all the time over periods of years and decades. ITRF coordinates and velocities of these stations are precisely known and when published, and so designated, constitute an IGS reference frame, considered to be an extension of its corresponding ITRF. The most recent IGS reference frame is designated "IGS14" and is an extension of ITRF2014.

Conversation 3

Guru: "Does this stuff make any sense"?

Me: "Yes, but I still don't see what any of it has to do with the map I'm trying to make of a piece of the State of Wisconsin, right here, in the USA. For one thing, I'm not going to base my map on a coordinate system with its origin at Earth's center of mass. I want my map to have an X/Y coordinate system somewhere near the ground I'm standing on and also have elevations referenced to mean sea level".

Guru: "Good, but we have a way to go before we get there from ICRFs, ITRFs, and IGS reference frames. I just wanted to give you some background on the basis for your X/Y coordinates and elevations. Hang on. Read the next document on the North American Terrestrial Reference Frame of 2022 and then the following documents on Gravity and the North American-Pacific Geopotential Datum of 2022".

CHAPTER TWO

GEODETIC COORDINATES, MAP PROJECTIONS, AND THE NORTH AMERICAN TERRESTRIAL REFERENCE FRAME OF 2022 (NATRF2022)

Conversation 4

Me: “I’m going to just go ahead and use ITRF2014 as a basis for my map. It has a nice rectangular coordinate system, referenced by plenty of recoverable points on Earth’s surface”.

Guru: “Well, ok, but your map isn’t going to make a lot of sense to most people”.

Me: “Why not”?

Guru: “ITRF2014’s X/Y plane contains Earth’s center of mass and its Z axis, which isn’t “up,” coincides with Earth’s spin axis. So, neither X nor Y on your map will have much to do with “north,” all coordinate values will be very large numbers, and your map won’t have any elevations”.

Me: “That’s true. Now I remember, in an earlier moment of clarity, I told you I wanted a coordinate system with the X/Y plane somewhere near the ground I’m standing on and I wanted elevations referenced to mean sea level. How do I make those things so”?

Guru: “First, let’s deal with your X/Y plane and a type of height that goes along with it. We’ll talk about elevations later because they have to do with gravity which is a very strange phenomenon”.

Me: “Ok”.

Guru: “To get from any ITRF or IGS to the reference frame you want, let’s first establish a “geodetic” coordinate system that has a direct mathematical relationship with the geocentric coordinate system of the ITRF or IGS”.

Geodetic Coordinate System

We need a good, three-dimensional, mathematical representation of Earth. It has been known for a long time that the general shape of Earth is not spherical, but, rather, is flattened at the poles, like a giant medicine ball. Such an object is an “oblate spheroid” and is a special case of a more general “ellipsoid.” This special-case ellipsoid is mathematically defined by its “semimajor” and “semiminor” axes, “a” and “b,” respectively, as shown in Figure 1.

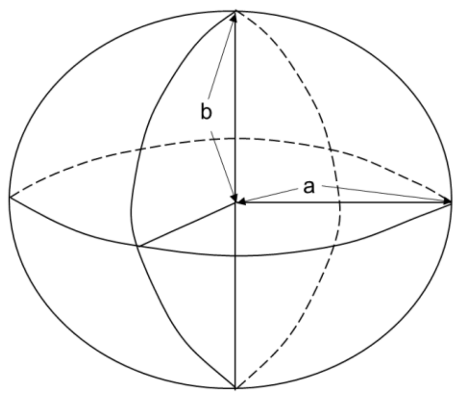


Figure 1.
Ellipsoid with Semimajor Axis (a) and Semiminor Axis (b)

For over 100 years, ellipsoids with differing values of a and b were chosen to best represent different regions of Earth because people weren’t thinking “globally.” With the emergence of space-based measurement techniques in the 1970s and 1980s, there arose clear need for a worldwide geodetic reference system. An ellipsoid, called “Geodetic Reference System of 1980”, or “GRS 80,” was selected for a new horizontal geodetic datum under development. That datum was the “North American Datum of 1983.” It was published in 1986 and then referred to as “NAD 83 (1986).” Subsequent adjustments of NAD 83 had names referring to their publication years in parentheses. All adjustments of NAD 83 used the GRS 80 ellipsoid, as will the new datum and reference frame, called the “North American Terrestrial Reference Frame of 2022” or “NATRF2022.” But, wait, we are getting ahead of ourselves. We’ve got to get back to geodetic coordinates, how they relate to geocentric coordinates, and how they will lead us to an X/Y system “near the ground we are standing on.”

We begin forcing a direct relationship between geodetic and geocentric coordinates by placing the center of the ellipsoid at the geocentric origin (Earth’s center of mass for

ITRFs and IGSs). We make the minor axis of the ellipsoid coincide with the Z axis of the geocentric system (Earth's spin axis for ITRFs and IGSs). These constraints cause the north-south reference plane, or equatorial plane, of the ellipsoid and the X/Y plane of the geocentric system to coincide. Further, we force the east-west reference plane, containing the prime meridian, of the ellipsoid to be coincident with the X/Z plane of the geocentric system. Two coordinates, latitude (Φ) and longitude (λ), in the geodetic system have angular units. The third coordinate, ellipsoid height (h), has distance units. The longitude of a point "p" (λ_p) is the angle, positive eastward and measured in the equatorial plane, between the plane of the prime meridian and the meridional plane containing p. The latitude of p (Φ_p) is the angle, positive northward and measured in the meridional plane of p, between the equatorial plane and the line, through p, that is perpendicular to the ellipsoid. That line is called the "ellipsoid normal" at p. The ellipsoid height of p (h_p), positive outward and measured along the ellipsoid normal, is the distance from the ellipsoid to p. Figure 2 depicts relationships among geocentric and geodetic coordinates. Mathematical transformations between the two coordinate systems are easily programmable and run automatically in nearly every GPS receiver.

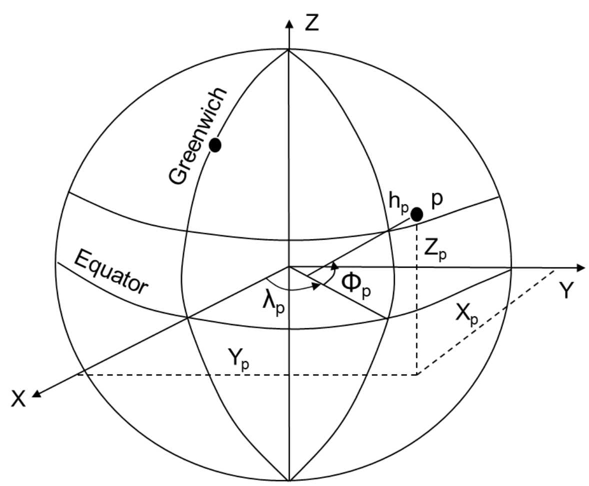


Figure 2.
Geocentric (X_p,Y_p,Z_p) and Geodetic (Φ_p,λ_p,h_p) Coordinates of Point p

Unfortunately, we cannot yet make a map because ellipsoids are not "developable." That is, no part of an ellipsoid can be laid flat without tearing. So, a seamless map cannot be made unless we "project" geodetic coordinates onto a surface that is developable.

Map Projections

Cones and cylinders are developable and are commonly used for map projections. A cone can be sliced from its base to its apex and laid flat. A cylinder can be sliced lengthwise, unrolled, and laid flat. An X/Y coordinate system can then be established on the selected flat surface. When a cone or cylinder is properly sized and oriented with respect to a reference ellipsoid, the latitude and longitude of any point can be transformed into X/Y coordinates by "projecting" the point from the ellipsoid onto the cone or cylinder (see Figures 3 and 4). Such "map projections" are mathematical and easily programmable. X and Y on a map projection are often referred to as "easting" and "northing," respectively.

All map projections contain distortions. That is, at least some things on Earth's surface appear distorted on a map projection. Figure 4 depicts "linear" distortion of distances on Earth's surface (red) as they appear on a map projection surface (yellow). A measure of linear distortion is the ratio of map projection distance to its corresponding Earth's surface distance. Linear distortion is greater than 1 if the map projection surface is above Earth's surface and less than 1 if the map projection surface is below Earth's surface. Variation of linear distortion, across an area covered by a map projection, is caused by Earth's curvature and changes in ellipsoid height of the terrain. If extents of the area and height differences of the terrain are not great, the range of linear distortion across a map projection can be held within bounds.

Optimal sizing and orientation of a map projection surface will cause it to align, as best as possible, with the mean ellipsoid height of the terrain. If the range of linear distortion is within 0.99995 to 1.00005, the National Geodetic Survey (NGS) refers to the map projection as a "low-distortion projection." All map projections of the Wisconsin Coordinate Reference Systems (WISCRS) have linear distortion ranges within 0.999967 to 1.000033. So, not only are WISCRS projections low-distortion, the linear distortions are small enough that they can usually be ignored for applications as stringent as land surveying and engineering design and construction.

Conversation 5

Guru: "So, that's how you get an X/Y plane near the ground you're standing on".

Me: "Ok, but you still have to help me understand the new datums…that's what this book is supposed to be about isn't it"?

Guru: "Ok…Here we go…"

NATRF2022

NATRF2022 will replace NAD 83 and all its adjustments. The GRS 80 ellipsoid will be used, but it will be translated and rotated from its position on NAD 83 to make it better Earth-centered and better aligned to Earth's spin axis. There will also be a scaling, between NAD 83 and NATRF2022, arising from improved measurements. Translation,

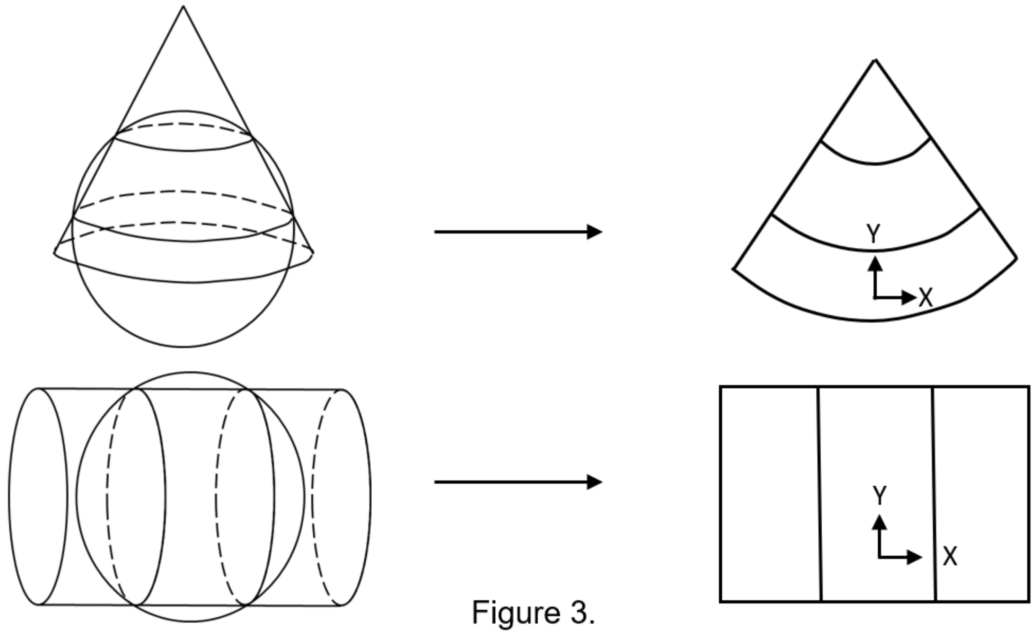


Figure 3.

Conical and Cylindrical Map Projections

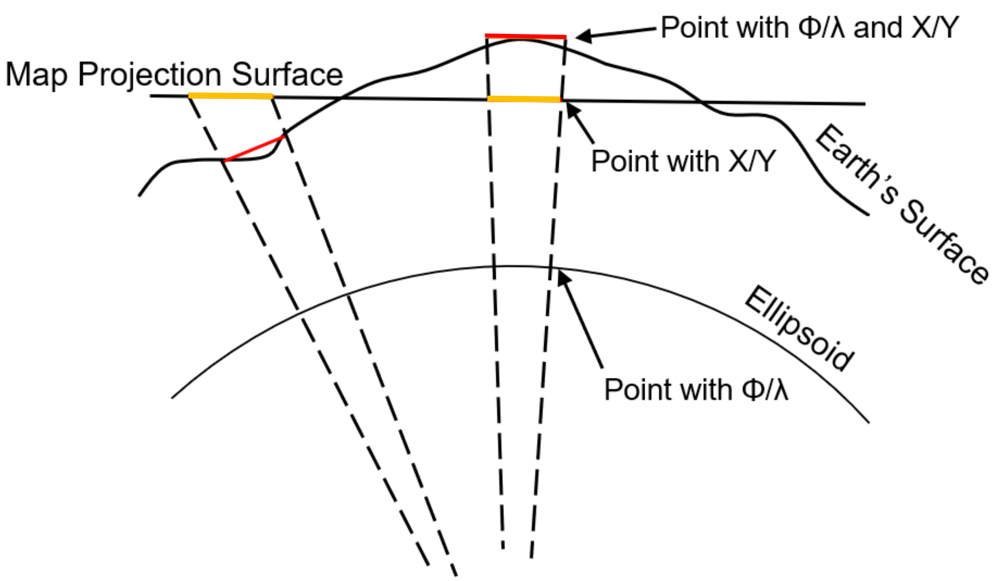
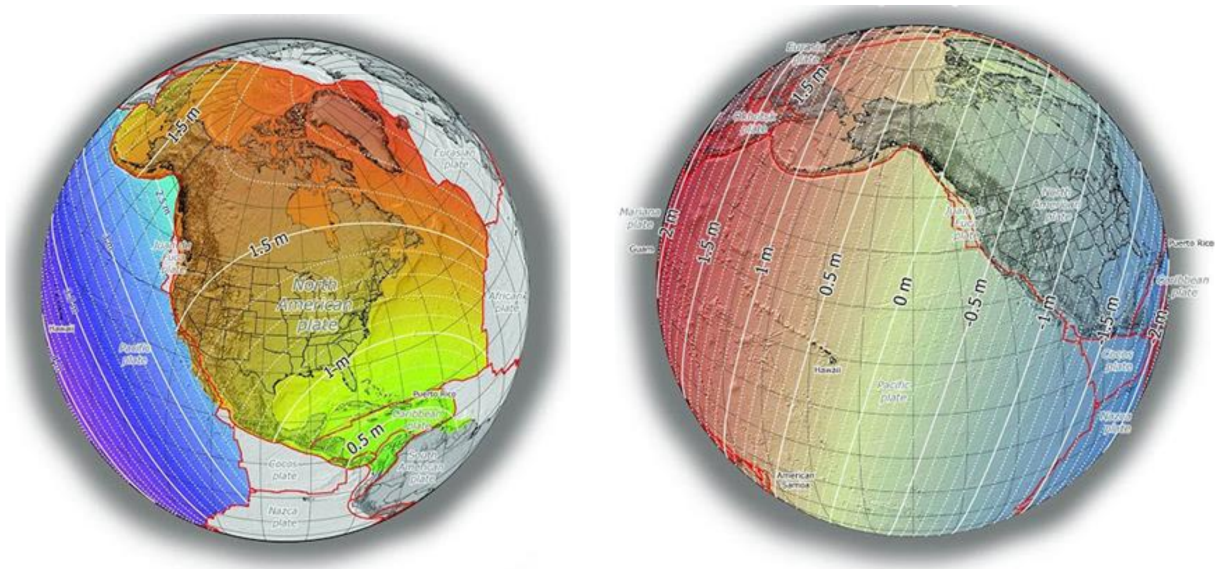


Figure 4.

Profile of Map Projection Showing Linear Distortion of Distances

rotation, and scaling will cause changes in latitude, longitude, and ellipsoid height across Wisconsin. In Figure 5, approximate horizontal position changes arise from changes in latitude and longitude. In Wisconsin, those changes range from 1.3 meters to 1.4 meters and approximate ellipsoid height changes range from -1.1 meters to -0.9



a. b.

Figure 5.
Changes in Horizontal Position (a) and Ellipsoid Height (b)
(From https://www.ngs.noaa.gov/datums/newdatums/WhatToExpect.shtml)

meters. NGS will provide software, which can be adopted by vendors, for transforming from NAD 83 to NATRF2022.

NATRF2022 will be made to coincide with the most current IGS frame at a specified epoch (t_0). Thereafter, NATRF2022 and that IGS frame will diverge because NATRF2022 will account for plate tectonics while IGS does not. That is, NATRF2022 will be "plate-fixed" and tend to hold constant over time the latitude and longitude of any point on the North American plate.

The primary motion of a tectonic plate is rotation about an "Euler axis." Figure 6 depicts horizontal velocity vectors at parts of the continuously operating reference station (CORS) network in the USA. Velocity vectors can be used to compute the position of the "Euler pole," the point where the Euler axis penetrates the reference ellipsoid. The computed position of the Euler pole is shown for the velocity vectors in Figure 6.

Figure 7 indicates how an Euler axis and pole are related to geocentric and geodetic coordinates. In the figure, λ_0 is the longitude of the plane containing the Euler axis, Θ_0 is the co-latitude (complement of the latitude) of the Euler axis, and ω_0 is the angular velocity of the rotating plate. Given λ_0, Θ_0, ω_0, the horizontal velocity at any point p with Φ_p, λ_p can be computed. Thus, for any given epoch (t), the coordinates of point p at t_0 can be computed from measured coordinates of p at time t and the velocity of point p. Therefore, the horizontal position of p can be made nearly stable over time relative to the tectonic plate. This aspect of NATRF2022 will be accessed through GPS, the CORS network, and the Online Positioning User Service (OPUS).

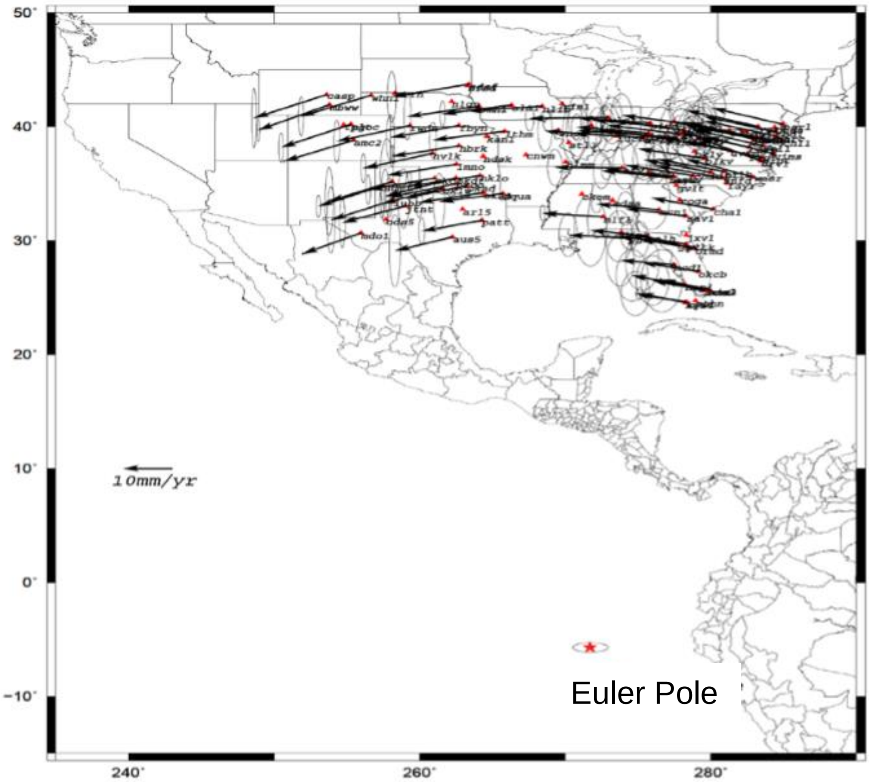


Figure 6.
Horizontal Velocity Vectors at Selected CORS and Computed Position of Euler Pole (From NOAA (2017))

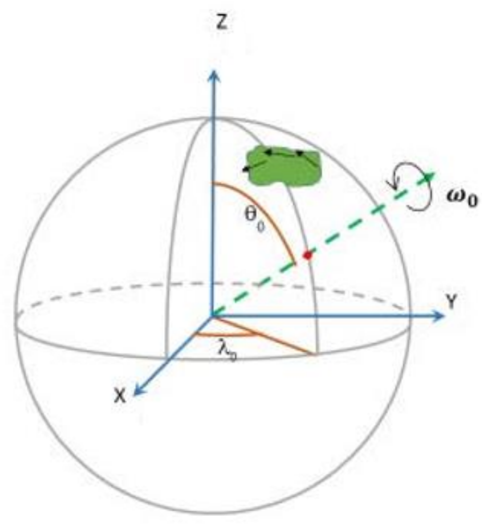


Figure 7.
Euler Axis (Dashed Green) and Pole (Red Dot) for Rotating Plate (Green Area) Located with Respect to Geocentric and Geodetic Coordinate Systems (From NOAA (2017))

Above, you will find the non-exclusive words “primary,” “tend,” and “nearly.” This is because Eulerian motion is not all that is happening with tectonic plates. The plates are not perfectly rigid but, rather, are somewhat heterogeneous in stiffness. So, not all points on a plate are rotating with exactly the same angular velocity. Also, there are additional geophysical phenomena, such as glacial isostatic adjustment, that cause non-Eulerian movements. Moreover, Eulerian motion affects only horizontal positions, but ellipsoid heights are also changing. These additional velocities will be estimated by NGS using an “intra-frame velocity model,” the details of which have yet to be decided upon. NGS will provide values for intra-frame velocities but will not include them in computation of final geodetic-quality coordinates because of uncertainties in the modeling. Therefore, NATRF2022 coordinates will be time dependent.

Another consideration is that Eulerian parameters (i.e., pole coordinates and angular velocity) are estimated from a limited number of dispersed and imperfect point velocity measurements. As new and better measurements of point velocities become available, improvements can be made in estimates for Eulerian parameters. Each time its then-current Eulerian parameters are updated, NATRF2022 will be given a new version number.

Conversation 6

Me: “Give me some time. I have to go back and re-read a bunch of this”.

Guru: “Ok. Patience I have”.

(some time passes)

Me: “Alright, now I think I get much of what is going to happen. I don’t quite understand the significance of time-dependency, though, to what I am doing”.

Guru: “Well, if your map doesn’t have to be registered to any future positional information, then time-dependency of the reference frame might not be that significant. However, most maps are time-stamped for a reason. In any case, let’s say you are establishing geodetic control for a design and construction project that has tight spatial tolerances and is expected to take a number of years to complete. Furthermore, let’s say your project is in an area where intra-frame velocities are known to be significant. Then, as your project progresses, you might see changes in your control points’ coordinates. These changes might be greater than expected differences arising from uncertainties in your measurements. The changes will be due to time dependencies of the control points’ coordinates”.

Me: “Alright. Let’s move on to elevations. Remember, I want elevations above mean sea level”.

Guru: “Ok, but we have to understand something about gravity first. Remember, I told you gravity is a very strange thing. Try reading Chapter 3”.

References

National Oceanic and Atmospheric Administration (2017), “Blueprint for 2022, Part 1: Geometric Coordinates”, NOAA Technical Report NOS NGS 62.

CHAPTER THREE
GRAVITY

Conversation 7

Me: “Last I knew, you were telling me gravity is a very strange thing. Somehow, that must have to do with elevations I want to put on my map”.

Guru: “It’s strange alright and its behavior here on Earth and in Earth’s vicinity has a <u>lot</u> to do with elevations you want to put on your map”.

Me: “Ok. Tell me about it”.

Gravity in General

Just why is it that objects are attracted to one another solely by virtue of their having masses? No one quite knows. Many folks recognize gravity as one of four fundamental forces of nature. Some assert that gravitational force must be mediated by elementary massless particles called “gravitons”, similar to electromagnetic force being mediated by photons. And, since electromagnetic force can behave as a wave, then gravitational force can also behave as a wave. Electromagnetic waves travel at the speed of light and so do gravitational waves. Detectable gravitational waves result from spectacular collisions of black holes, pulsars, and other extremely dense cosmic objects. Gravity is so strange that some describe it not as a thing unto itself but, rather, as a consequence of curvatures in space/time brought about by masses.

In any case, and whatever gravity is, we’d better be glad we have it. There wouldn’t be any “us” without it. Neither would there be any stars or planets. The universe would be very different.

Conversation 8

Me: “Are you crazy? Why do you have to keep telling me about the universe? I want to know about gravity here on Earth and in its immediate vicinity. How does it affect where things are and, therefore, our bases for describing locations? What about the map I want to make”?

Guru: “It’s true. I do seem to get distracted by esoteric matters. You ought to hear me go on and on about quantum mechanics”.

Me: “If you don’t shut up about this stuff and, rather, start telling me things I actually need to know, I’m going to stop listening to you”.

Guru: “I promise we will get around to ‘mean sea level’ soon”.

Earth’s Gravity

If we think of gravity as a force, we must remember that a force can be described as a vector, with magnitude and direction, acting at a single point. On the other hand, if we think of gravity as acting throughout a volume of space, containing an infinite number of points, we might call the phenomenon a “gravity field”. Earth’s got one, as do all the planets, satellites (natural or artificial), the Sun, asteroids and comets, you and I, every rock, every bird, and all other objects having mass. These gravity fields interact with one another to have effects on their objects. Many of the effects are readily observable. Earth’s gravity field and mine work together to keep my feet on the ground. Earth’s gravity field and that of a GNSS satellite work together to keep the satellite in a carefully-defined orbit. Earth’s gravity field and those of the Sun and Moon work together to produce periodic ocean tides that we monitor to help us find mean sea level. Those same three gravity fields also produce periodic “Earth tides” that cause Earth’s elastic crust to move in and out almost as if the planet is breathing. These tidal movements of Earth’s crust are different than the motions described in Chapter 2. Many of these things are familiar to us and do not seem strange.

Here’s the rub. Earth’s gravity is not merely an attractive force due only to mass. Some refer to the attractive force alone as “gravitation”. Then, “gravity” is a combination (a vector addition) of the attractive force and a different “centrifugal” force, that is repulsive and arises from Earth’s rotation. The attractive force draws objects towards Earth’s center of mass and the repulsive force drives them away from Earth’s spin axis. See Figures 1 and 2, where relative vector magnitudes are represented by lengths of arrows.

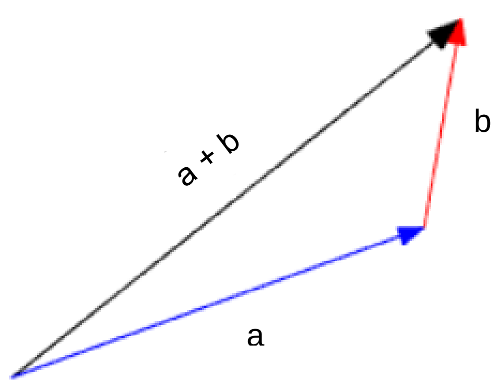


Figure 1.
Vector Addition: Vector b is Added to Vector a by Placing Its Tail at the Head of a and then Constructing Vector a + b from the Tail of a to the Head of b

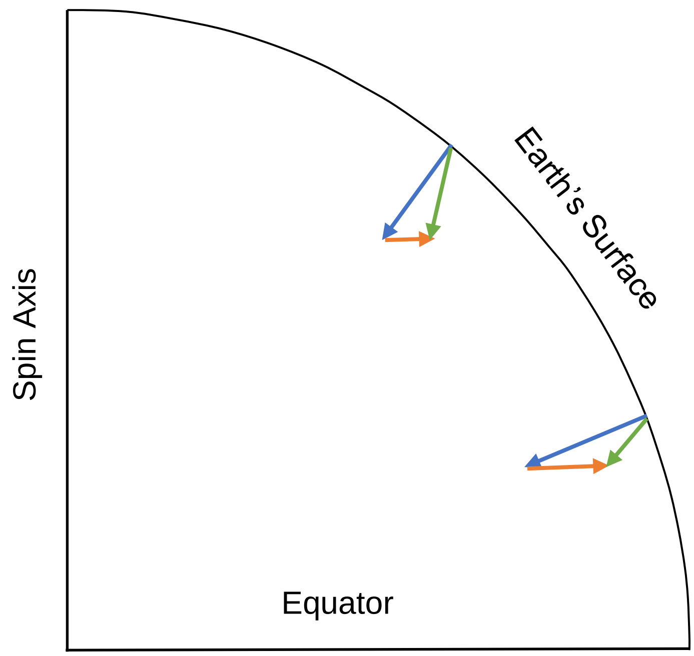


Figure 2.
Force Vectors: Gravitational (Blue), Centrifugal (Orange), and Gravity (Green)

Centrifugal forces cause Earth to resemble an ellipsoid, bulging at the equator and flattening at the poles. Gravitational forces decrease with the square of their distances from Earth's center of mass and centrifugal forces increase with their distances from Earth's spin axis. Fortunately, repulsion never overcomes attraction, and nothing goes flying off into space unless we make it happen by applying another force such as the thrust of a rocket engine.

Conversation 9

Me: "I'd say the gravity of gravity is quite grave. It's not a grave, it's just grave. Isn't the English language wonderful"?

Guru: (No response).

Equipotential Surfaces

The magnitude of the gravity vector at any point can be expressed as the potential energy of unit mass at that point. All points having the same gravity potential (there are an infinite number of them) form a single, closed "equipotential" surface in three dimensions. Since gravity potential depends on distance from Earth's center of mass, and there are an infinite number of possible distances, then there are an infinite number of equipotential surfaces. None of these surfaces intersects any other because no single point can have more than one gravity potential. On the other hand, equipotential surfaces are not parallel. One cause is gradual reduction of the centrifugal component from the equator to the poles. In addition, even though equipotential surfaces are smooth, each has its own undulations due to nonuniformity of Earth's density, with concentrations of mass deflecting gravitational forces. Even the varying topography of Earth's surface causes nonuniform spatial behavior among equipotential surfaces. Some equipotential surfaces intersect Earth's surface. Some are entirely above it. Some are entirely below it. See Figure 3.

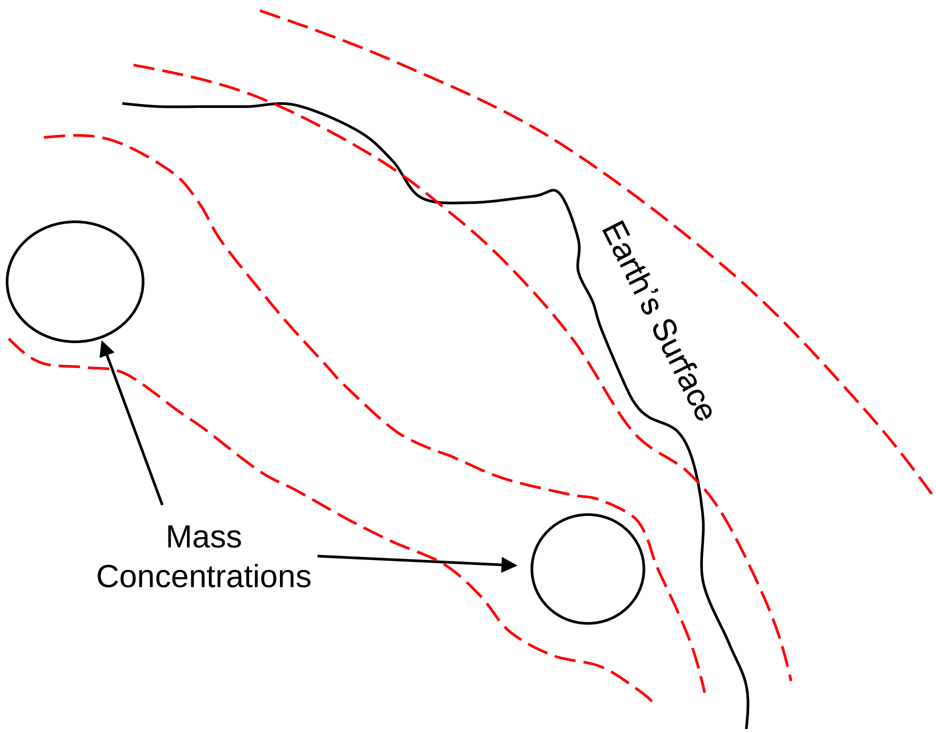


Figure 3.
Some Equipotential Surfaces (Red Dashed Lines)

Equipotential surfaces are real but intangible. We can describe and locate them but we cannot see or touch them. They are physical in nature but they have no substance. Equipotential surfaces aren't like reference ellipsoids. Those ellipsoids are not only intangible, they are neither physical nor even real. They are mathematical constructs, made up for our own convenience.

Conversation 10

Me: "How do you get off telling me what's real and what isn't? Do you mean to say that gravity is different everywhere I go"?

Guru: "Yes, unless you are travelling on an equipotential surface and that's not very easy to do. Even then, the direction of gravity will change in a three-dimensional reference frame. In fact, if you are standing up, gravity is different at the top of your head than at the bottoms of your feet".

Me: "What has this to do with the map I'm trying to make"?

Guru: "When you are standing up, the top of your head is at a different elevation than the bottoms of your feet, right"?

Me: "Yes".

Guru: "Your map is going to have elevations on it, isn't it"?

Me: "Yes".

Guru: "Read on".

Plumb Lines

The gravity vector at any point on an equipotential surface is perpendicular to that surface and defines the direction of "vertical," or "up" and "down," at that point. The "plumb" line containing the point is a continuous, vertical line of force. That is, a plumb line is perpendicular to all equipotential surfaces it intersects. Because equipotential surfaces are not parallel, plumb lines are curved. See Figure 4. Ultimately, all plumb lines meet at Earth's center of mass.

One Very Special Equipotential Surface

There is one equipotential surface, called the "geoid," that best fits global mean sea level. Some references make no distinction between the geoid and mean sea level but there actually is a difference. If we eliminate waves and tides, the surface of the open ocean would still not perfectly conform to an equipotential surface because atmospheric pressure variation moves water, the density of water varies with temperature and salinity, and prevailing winds push around high-volumes of water. Sources of inputs (e.g., rain, rivers, melting ice) do not coincide with areas where water is lost by evaporation. The sea surface is not in equilibrium although it is always seeking it (Bomford, 1985). So, we use the geoid instead of mean sea level. What do we use it for? Well, as a vertical datum…the basis for elevations.

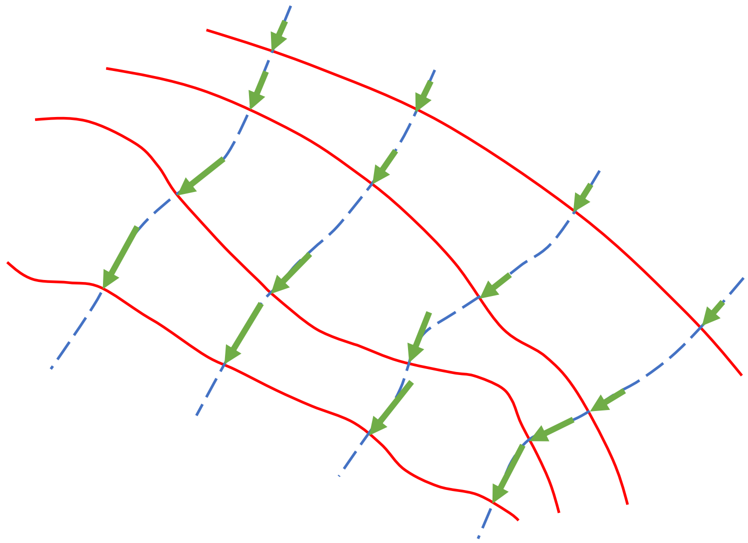


Figure 4.
Equipotential Surfaces (Red), Plumb Lines (Blue), and Gravity Vectors (Green)

Conversation 11

Me: “Alright! Now we’re getting somewhere! How do we locate the geoid and how do we reference elevations to it, especially if we want to use GNSS”?

Guru: “Take a breather for a few minutes, hours, or days. Then read Chapter 4”.

References

Bomford (1985), Geodesy 4th Edition, Clarendon Press, Oxford, England.

CHAPTER FOUR

NORTH AMERICAN-PACIFIC GEOPOTENTIAL DATUM OF 2022 (NAPGD2022)

Note on Terminology

"Elevation" is a general term meaning "the distance of a point above a specified equipotential surface; the distance is measured along the plumb line between the point and the surface".

"Orthometric height" is "elevation" with the specified equipotential surface being the geoid.

The Glossary of the Mapping Sciences (ASCE, ACSM, ASPRS (1994)), when defining "elevation", adds "The (equipotential) surface is understood to be the geoid unless some other surface is specified, and the elevation is then the orthometric (height)". The glossary's expanded definition of "elevation" is used in this document. Herein, "elevation", without explicit mention of equipotential surface, implies the geoid as the equipotential surface and, therefore, is synonymous with "orthometric height".

Conversation 12

Me: "So, you are supposed to tell me how we locate the geoid and how to reference elevations to it".

Guru: "Let's do it".

Finding the Geoid and Defining "Elevation"

There are a variety of devices and methods for locating the geoid and determining its size, shape, and gravity potential. Perhaps the most fundamental method, which gets us part way there, is measuring the rise and fall of ocean tides over a period of years at specific points. The devices used are called "tide gages" and typically include a graduated staff fixed in position in the water. When the location of local mean sea level on the staff is determined, a level instrument and a second, mobile graduated staff are used to determine elevations above local mean sea level of nearby “tidal bench marks”. These tidal bench marks are well-monumented and are usually connected to those at other tide gages by “differential leveling lines” which are strings of relatively closely-spaced points whose consecutive elevation differences are measured using level (horizontal) lines of sight and vertically-held graduated staffs. Differential leveling networks crisscross North America, establishing and connecting bench marks that for many decades were the vertical reference points (vertical control) used for mapping.

The elevation of a bench mark tells us the distance from the mark to the geoid as measured along the plumb line. However, the geoid is often mapped with respect to the reference ellipsoid of a horizontal or 3-dimensional datum. The geoid and ellipsoid do not coincide, with the separation, which varies from place to place, being referred to as “geoid height”. To determine the geoid height at the horizontal location of a bench mark, we measure the mark’s ellipsoid height (see Figure 1), typically with GNSS. This was done at numerous bench marks across the USA during NGS’s recent Height Modernization Program.

The elevation of point p, also called its “orthometric height”, is the length of the curved plumb line from the geoid to point p (positive in the opposite direction of gravity). As we know, the ellipsoid height of point p is measured along the ellipsoid normal, which is a straight line (again see Figure 1). So, the simple equation in Figure 1 is really an approximation, but it is a very good one. Notice that Figure 1 shows the geoid above the ellipsoid. This is opposite the case for Wisconsin where all geoid heights are negative. Figure 1 was developed to most-easily illustrate the included equation.

In addition to, or instead of, measuring elevation differences and ellipsoid heights, we can measure gravity or just its potential. “Gravimeters” measure potential directly, either at a single point (absolute) or as a difference between two points (relative). Gravimeters can be ground-based, shipborne, airborne (e.g., NGS’s GRAV-D), or satellite-based (e.g., European Space Agency’s GOCE). In all cases, a gravimeter’s position must be known to produce maps of gravity potential, so gravimeters are often coupled with GNSS receivers. GNSS receivers can also be coupled with inertial measuring units (IMUs) to measure gravity. An IMU typically has accelerometers mounted along three mutually perpendicular axes and gyroscopes to stabilize its orientation in space as it moves. NGS’s “Gravity for the Redefinition of the American Vertical Datum” (GRAV-D) is an underway, 15-year effort to gather a dense collection of gravity measurements (with gravimeters, GNSS, and IMUs) across all of the terrestrial and littoral United States and its territories (see Figure 2).

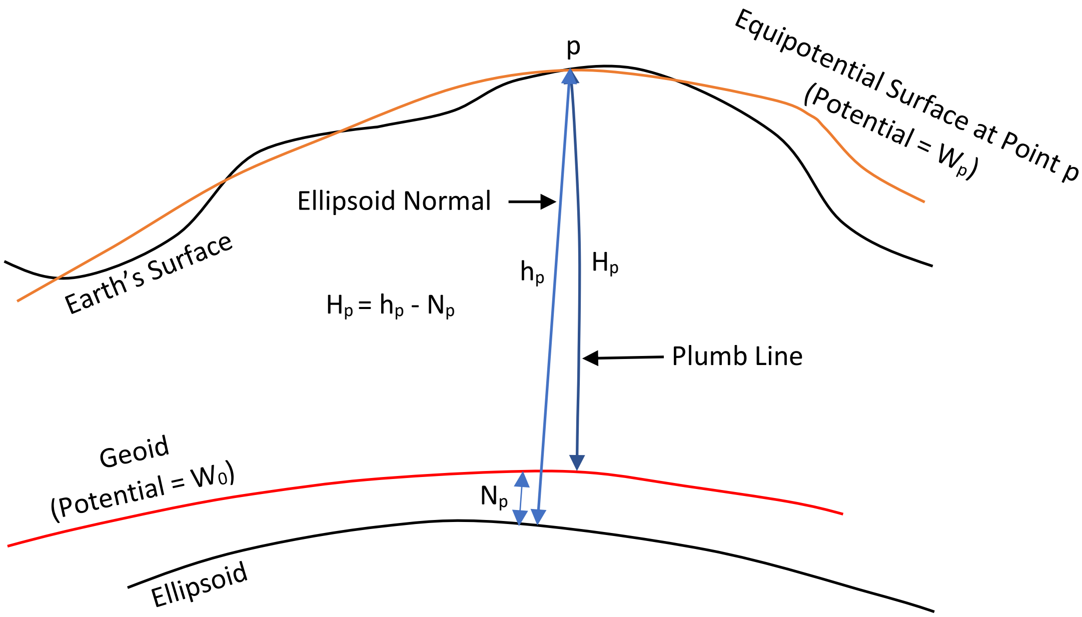


Figure 1.
Elevation or Orthometric Height (H_p), Ellipsoid Height (h_p), and Geoid Height (N_p) at Point p

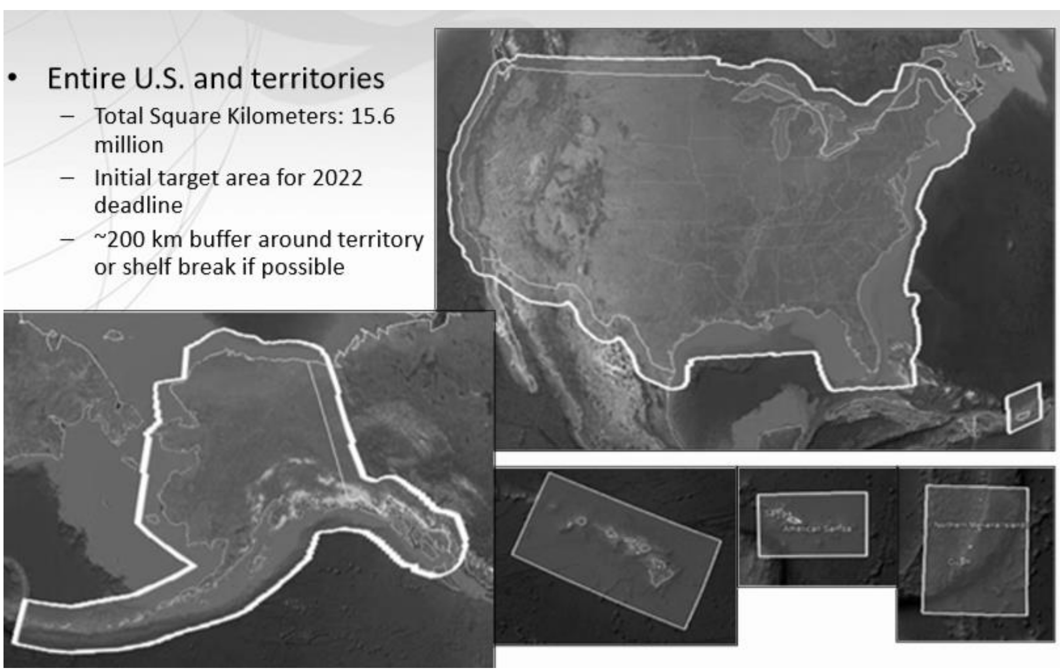


Figure 2.
Scope of GRAV-D Data Collection
(From https://c.ymcdn.com/sites/www.njspls.org/resource/resmgr/2018_HANDOUTS/GRAV-D-new_datum-NJSPLS-2018.pdf)

Gravitational forces can also be computed from well-tracked satellite (e.g., GNSS) orbits. Another technique, satellite altimetry, uses precise orbital data and advanced radar to map ocean topography which is by definition very close to the geoid. NASA's recent GRACE mission consisted of two satellites continuously measuring the distance between them using microwave ranging. Changes in the distance were used to compute gravity. The satellites' orbits were also measured by GNSS, so gravity maps could be produced.

Conversation 13

Guru: "There are all kinds of ways to measure gravity. With very recent advances in technology, we can even measure gravity with clocks".

Me: "Either you are bonkers or you're about to go Einstein on me again, aren't you"?

Guru: "Yeah…and for good reason. If your mind hasn't already been blown, wait until you hear what I'm about to say. Over 100 years ago, Einstein reasoned that there must exist everywhere a phenomenon he called "gravitational dilation of time" and he was right. The greater the gravitational force, the more slowly time passes. Remember how we talked about expanding and shrinking the durations of hours, minutes, and seconds? Well, that's what actually happens when we move around! Intuition, experience, and language begin to fail us. What do we actually mean by "time" and "duration"? Doesn't it seem they should be independent of reference frame? Well, they aren't.

Two identical clocks, under different gravitational forces, keep different times. We can use their time differences to compute differences in gravitational forces acting upon them! Last year, a new kind of atomic clock was constructed. It is so accurate that determination of its performance is actually constrained by our imperfect knowledge of Earth's gravity field! The new clock isn't as portable as most gravimeters…but that's being worked upon".

Me: "I'm sorry. I just never knew that relativity was relevant to surveying and mapping. Besides, from what you are telling me, I can infer that my head is slightly older than my feet because it's been above them most of the 'time'".

Guru: "That's right. I told you there were some strange things going on. Here's something a little more down to Earth…because of the physics at any given point p, there is a mathematical relationship between its gravity potential (W_p) and its elevation (H_p). Either can be computed from the other if the potential of the geoid (W_0) and the curvature of the plumb line are known (see Figure 1 again). However, it's quite a task to come up with an accurate enough estimate for the curvature of the plumb line. The curvature is smooth but it is not single-valued. It depends upon local mass distributions which are not uniform along it".

Finding the Geoid and Defining "Elevation" (Continued)

Gravity potential measurements are point data. To make a map from them and position the geoid over an entire area of interest, we use the measurements to model the geoid and reference it to a horizontal datum, thus producing derived geoid heights (see Figure 3). Such modeling is also used to produce uniform grids (e.g., GEOID12) for interpolation of geoid height at any point. Since we measure elevations from the geoid, Figure 3 can be thought of as a negative elevation map of the GRS 80 ellipsoid in the IGS08 (2005.00) Reference Frame.

Determination of elevations above a vertical datum in the field often begins with a search for a starting bench mark, followed by differential leveling to points of interest, perhaps to establish more-densified vertical control for mapping. Reliable NGS bench marks are becoming harder to find as more and more of them suffer from subsidence, uplift, deterioration, or destruction. If we could use GNSS to establish elevations of new starting bench marks locally, we wouldn't have to search forever and then run long differential leveling lines to get to our project sites. To do this, we need a very good model of the geoid so we can compute an elevation at any point from its measured ellipsoid height. That leads us to NAPGD2022.

NAPGD2022

NAPGD2022 will replace the North American Vertical Datum of 1988 (NAVD 88), all its subsequent adjustments, and all other nationally-recognized vertical datums in the United States. NAVD 88 is known not to best fit global mean sea level. The only elevation held fixed during NAVD 88's initial adjustment was that of a tidal bench mark referenced to local mean sea level at a tide gage in Quebec. From east to west across the United States, accumulating error caused the NAVD 88 datum surface to slope downward, away from the geoid. Figure 4 indicates approximate changes in elevation from NAVD 88 to NAPGD2022. Across Wisconsin, approximate elevation changes between the two datums range from -0.8 meter to -0.6 meter.

An additional consideration is that use of NAVD 88 requires either recovery of bench marks referenced to it or application of a geoid height model such as GEOID12 which is referenced to NAD 83. As mentioned, NAVD 88 bench marks continue to diminish in both number and reliability while future geoid height models will be referenced to NATRF2022 or IGS reference frames.

NAGPD2022 will consist of a number of datasets, the most significant of which, for Wisconsin geospatial professionals, will be a grid of geoid heights spaced at 1-minute intervals of latitude and longitude. This grid of geoid heights will be referred to as "GEOID2022" and will be accessed, primarily, through GNSS, CORS, and OPUS. NGS will also furnish software for accessing GEOID2022 and interpolating from it.

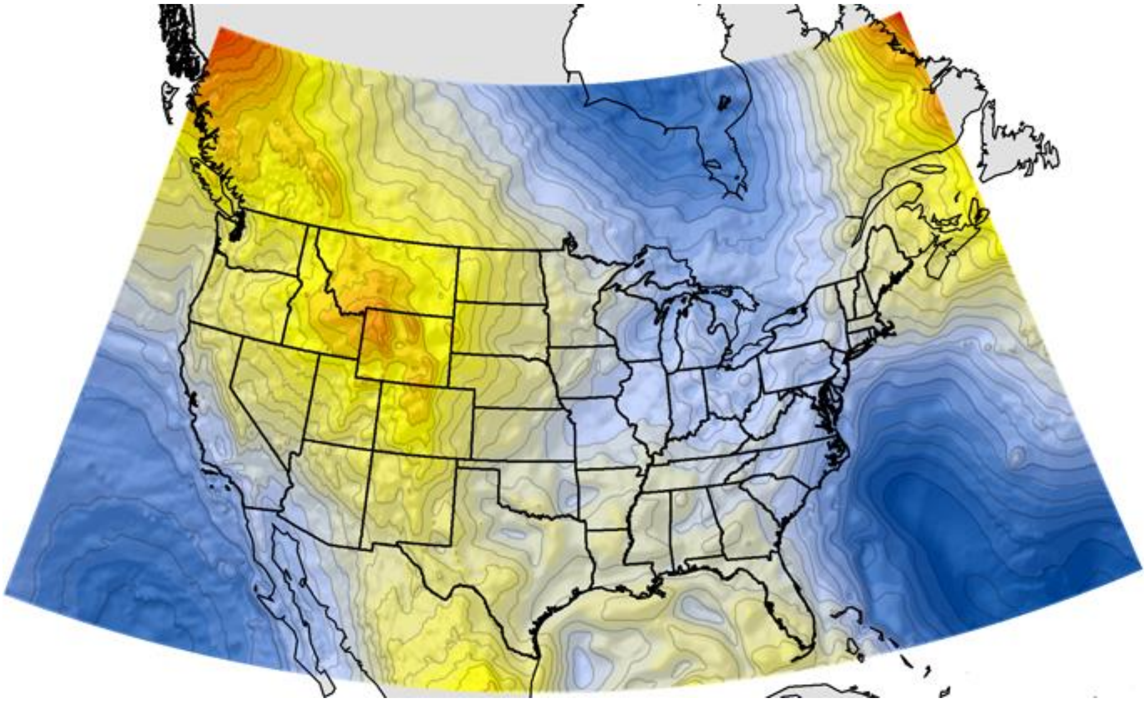


Figure 3.
Isolines of Geoid Height Referenced to GRS 80 Ellipsoid in IGS08 (2005.00) Reference Frame
(Blue Indicates Geoid Below Ellipsoid; Red Indicates Geoid Above Ellipsoid)
(From https://geodesy.noaa.gov/GEOID/GEOID12B/)

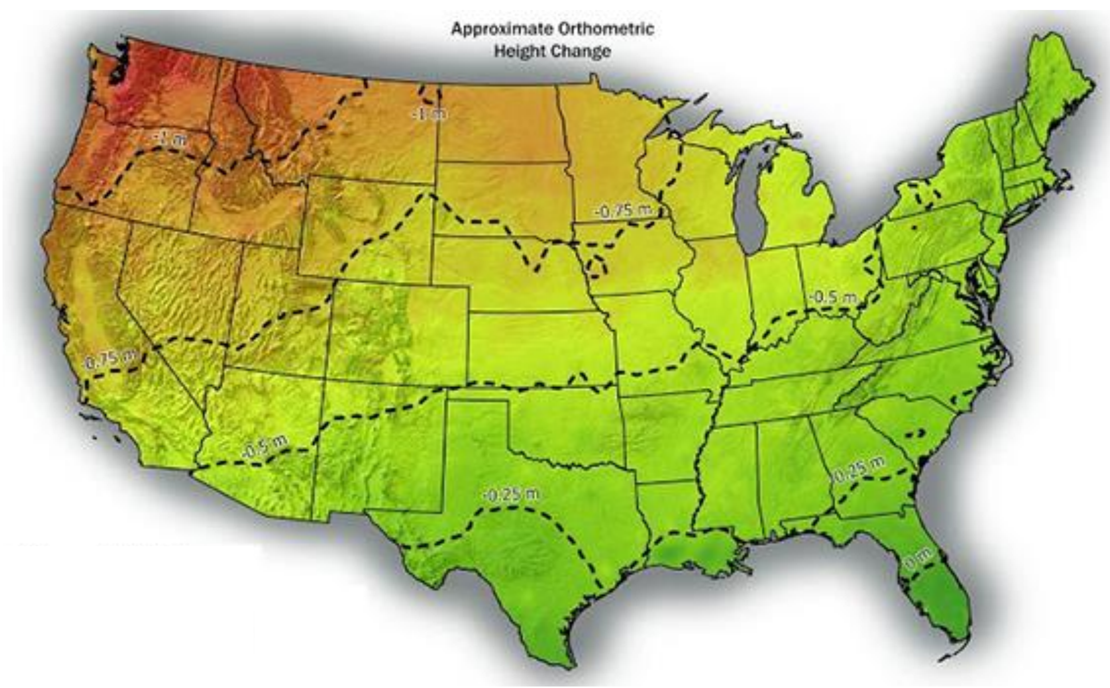


Figure 4.
Approximate Change in Elevation from NAVD 88 to NAPGD2022
(From https://www.ngs.noaa.gov/datums/newdatums/images/Orthometricheights.jpg)

Software for transforming elevations from NAVD 88 to NAPGD2022 will also be furnished. GEOID2022 will be derived from a gravity model which, itself, will be developed from many millions of measurements made with some of the techniques mentioned above. GEOID2022 will best fit global mean sea level at the epoch of NAPGD2022 and will be referenced to the GRS 80 ellipsoid on the then-current IGS reference frame.

The Dynamics of Everything

Earth's gravity field and, along with it, the geoid are constantly changing. For one thing, global mean sea level is rising. There are also redistributions of mass taking place from causes such as earthquakes and other episodic seismic events, glacial melting, regionalized periods of drought or monsoon, and secular glacial isostatic adjustment (which includes not only crustal uplift and subsidence but also more massive changes in Earth's viscous but fluid mantle). The location, size, shape, and potential of the geoid are dynamic. So, what do we do about this if we want the geoid to be our vertical datum?

NGS is addressing this problem with a multifaceted approach. All component data sets of the datum will have matching version numbers that change when modeling theory is improved significantly or significant changes in data become available (e.g., multiple or very large blunders are detected or large numbers of additional measurements are made). For example, official roll-out names include "NAPGD2022v01" and "GEOID2022v01".

Also, GEOID2022 will have both static and dynamic components. The static component will be coupled with global mean sea level at the initial epoch of version 1 but decoupled thereafter. That is, the static component of GEOID2022 will not rise with global mean sea level until a threshold value of 20 centimeters is reached some 60-70 years from now. At that time, an entirely new datum (i.e., NAPGDyyyyv01) will be developed and GEOIDyyyyv01 will be made to best fit the then-current geoid.

The dynamic component of GEOID2022 will be developed by monitoring on-going episodic and secular changes to the geoid, other than the rise of global mean sea level. The monitoring program will utilize a number of measurement techniques described above and will be based upon international agreements and funding.

These aspects of NAPGD2022 will cause all spatial references to it to be time dependent. Furthermore, we know that coordinates tied to NATRF2022 will also be time dependent, especially ellipsoid heights.

Conversation 14

Me: "Now wait just a minute here. How in the ##@&! am I supposed to get elevations from GNSS with all these things changing all the time"?

Guru: "It seems to be a bit of a challenge, eh? Most of the detail will be disguised by software, but we both know that we had better know what software is doing if we are to use it correctly and effectively…and then be able to interpret the results. I will try to put it all together in Figure 5, which is grossly exaggerated to show subtle changes over time (much like many of my previous figures have been grossly exaggerated to show detail)".

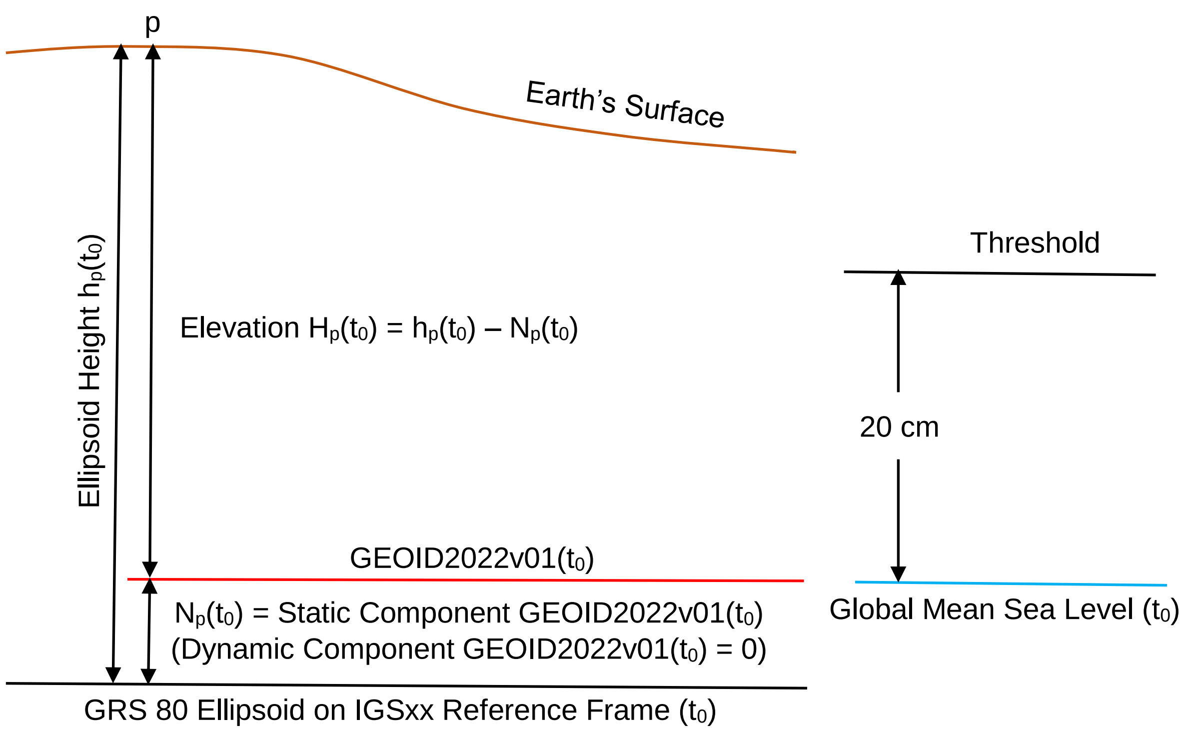


Figure 5a.
Time = t_0: Roll-Out Epoch for GEOID2022
Static Component GEOID2022v01(t_0) Best Fits Global Mean Sea Level (t_0)
Dynamic Component GEOID2022v01(t_0) = 0
Elevation $H_p(t_0) = h_p(t_0) - N_p(t_0)$

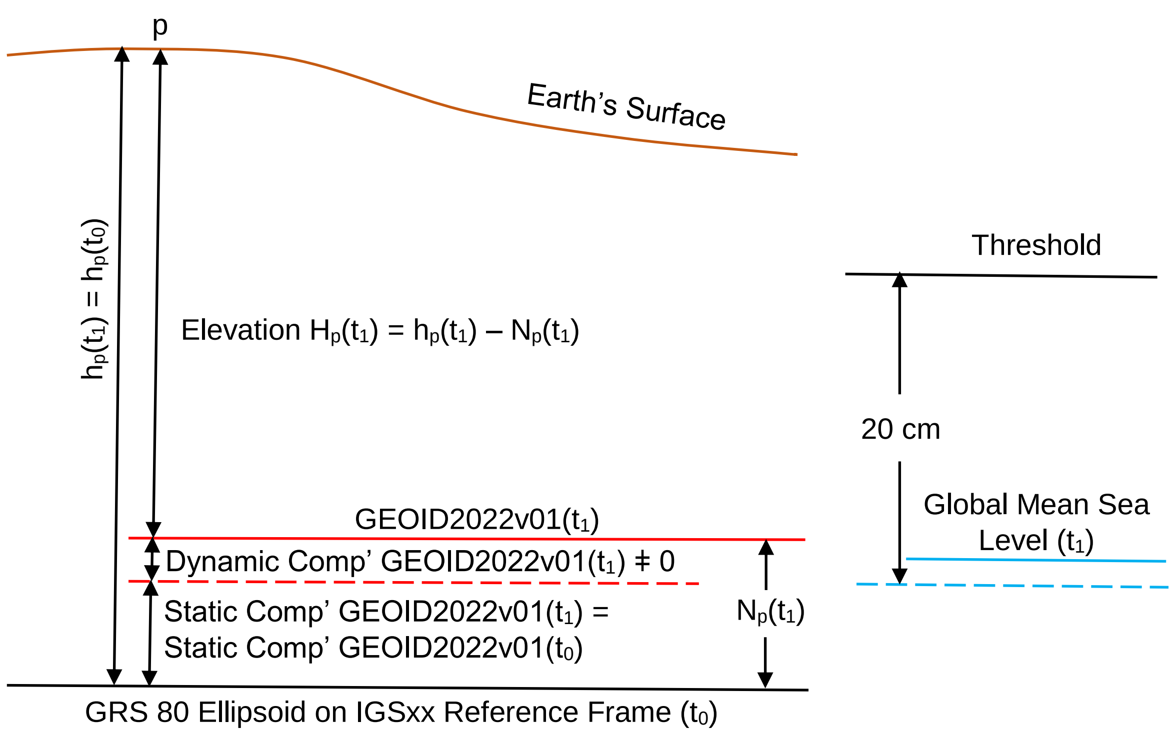


Figure 5b.
Time = t_1: Dynamic Component GEOID2022v01(t_1) ‡ 0
Static Component GEOID2022v01(t_1) = Static Component GEOID2022v01(t_0)
Elevation $H_p(t_1) = h_p(t_1) – N_p(t_1)$ ‡ $H_p(t_0)$
Global Mean Sea Level Rises
(Dashed Lines Indicate Previous Positions)

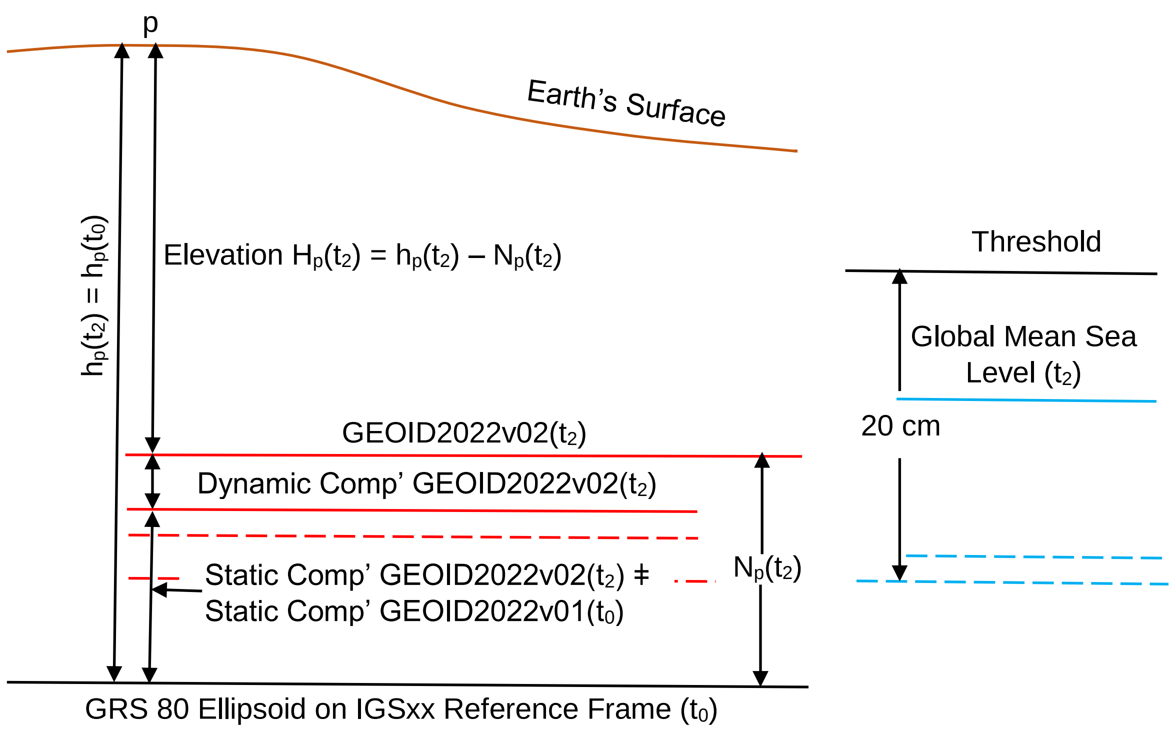


Figure 5c.
Time = t_2: Static Component Changes, Bringing about a New Version Number (GEOID2022v02)
Dynamic Component GEOID2022v02(t_2) ≠ Dynamic Component GEOID2022v01(t_1)
Elevation $H_p(t_2) = h_p(t_2) - N_p(t_2) \neq H_p(t_1) \neq H_p(t_0)$
Global Mean Sea Level Continues to Rise
(Dashed Lines Indicate Previous Positions)

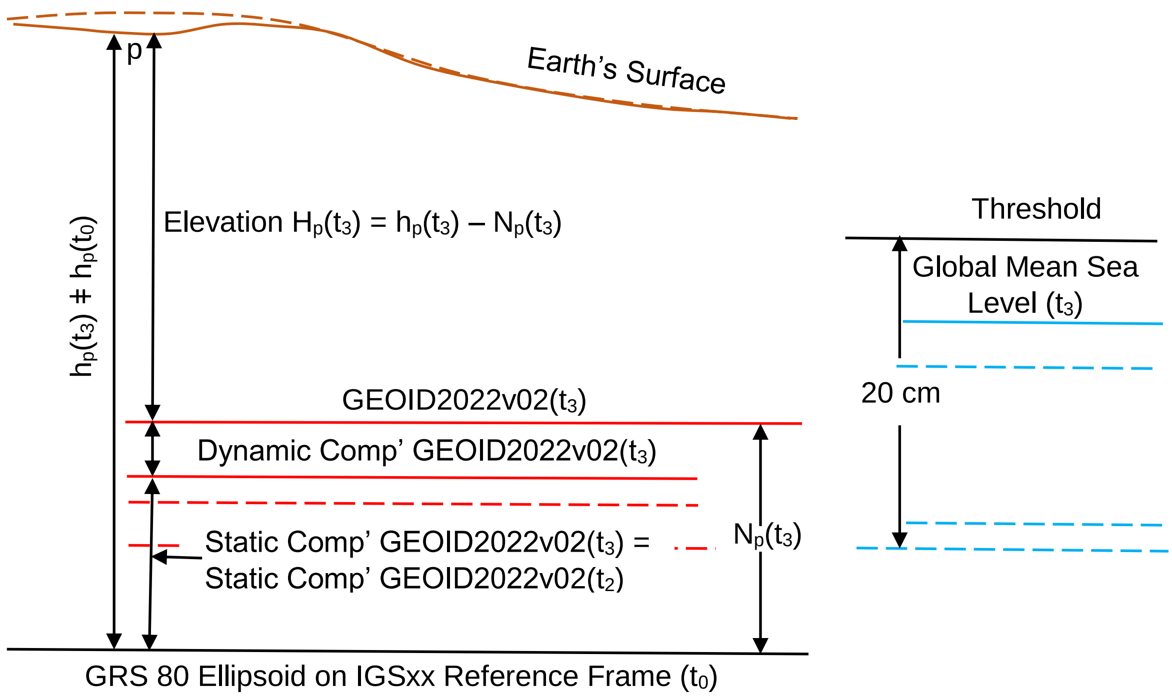


Figure 5d.

Time = t_3: Ellipsoid Height of Point p Changes Due to Local Subsidence

Static Component GEOID2022v02(t_3) = Static Component GEOID2022v02(t_2)

Dynamic Component GEOID2022v02(t_3) ‡ Dynamic Component GEOID2022v02(t_2)

Elevation $H_p(t_3) = h_p(t_3) - N_p(t_3) \ddagger H_p(t_2) \ddagger H_p(t_1) \ddagger H_p(t_0)$

Global Mean Sea Level Continues to Rise

(Dashed Lines Indicate Previous Positions)

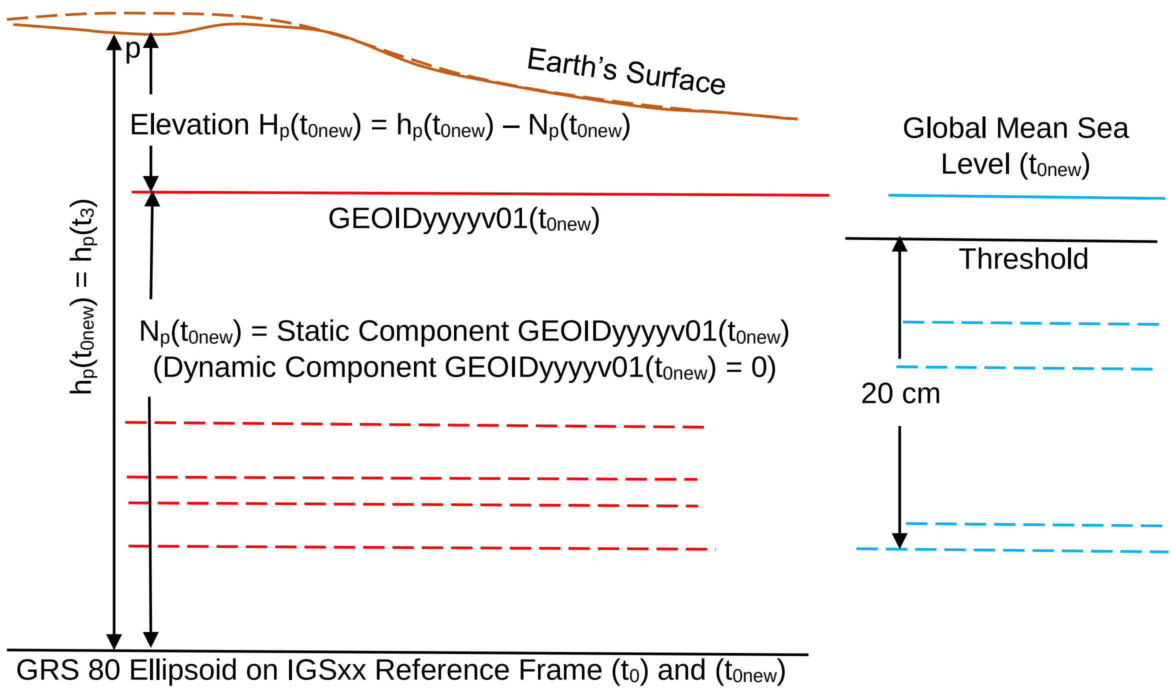


Figure 5e.

Time = t_{0new}: Global Mean Sea Level Rises above 20 cm Threshold, Bringing about a New Datum (NAPGDyyyyv01)

Static Component GEOIDyyyyv01(t_{0new}) Best Fits Global Mean Sea Level (t_{0new})

Dynamic Component GEOIDyyyyv01(t_{0new}) = 0

Ellipsoid Height of Point p Is Unchanged from t_3 unless a New Geometric Datum Has Been Created

Elevation $H_p(t_{0new}) = h_p(t_{0new}) - N_p(t_{0new}) \ddagger H_p(t_3) \ddagger H_p(t_2) \ddagger H_p(t_1) \ddagger H_p(t_0)$

(Dashed Lines Indicate Previous Positions)

Conversation 15

Guru: "Sorry that Figure 5 grew into Figure 5 (a,b,c,d,e). So, does Figure 5 help"?

Me: "I guess so. Sometimes I wish I was back in the 1950s using transit, tape, Dumpy level, and level rod. I even knew how to make a map with a plane table back then. The only computers were room-sized and resided at the Pentagon and NASA. No spatial databases, GIS, CAD or digital navigation systems for cars and fishing boats. Calculators were mechanical devices. We looked up trigonometric function values in great big, thick books. We drafted with pen and ink, straight edges, compasses, dividers, LeRoy lettering sets, and mylar. The only satellites up there were "Sputnik" and a few of its successors. Life seemed easier then, but everything we did took longer (and it wasn't because gravity was stronger). NGS was called the 'United States Coast and Geodetic Survey'. USGS was around and so was BLM".

Guru: "I too like reminiscing, but it's usually about Isaac Newton, Carl Gauss, Friedrich Helmert, Max Planck, Albert Einstein, Werner Heisenberg…those kinds of people and what they accomplished…and also, the wonderful scientists and thinkers who have come after them, on up to the present…along with all that has been learned about Earth, the solar system, the universe, and our place and role in them. Oh well, let's get back to business. We are almost finished. How about we conclude by discussing…"

Guru is interrupted by Me: "How about you go back and count all the acronyms appearing in this brief book? I mean, the ideas are hard enough to follow. Throw in the acronyms, and I need both a glossary and a thesaurus to make sense of it all. Plus some text books too".

Guru: "Sorry about that. I guess it's just another sign of the times. Was any of this helpful for your job and for making your map in a more knowledgeable way"?

Me: "Yes, for sure. I am grateful. Let me conclude by telling you something I am certain of: I'm glad I'm not a guru".

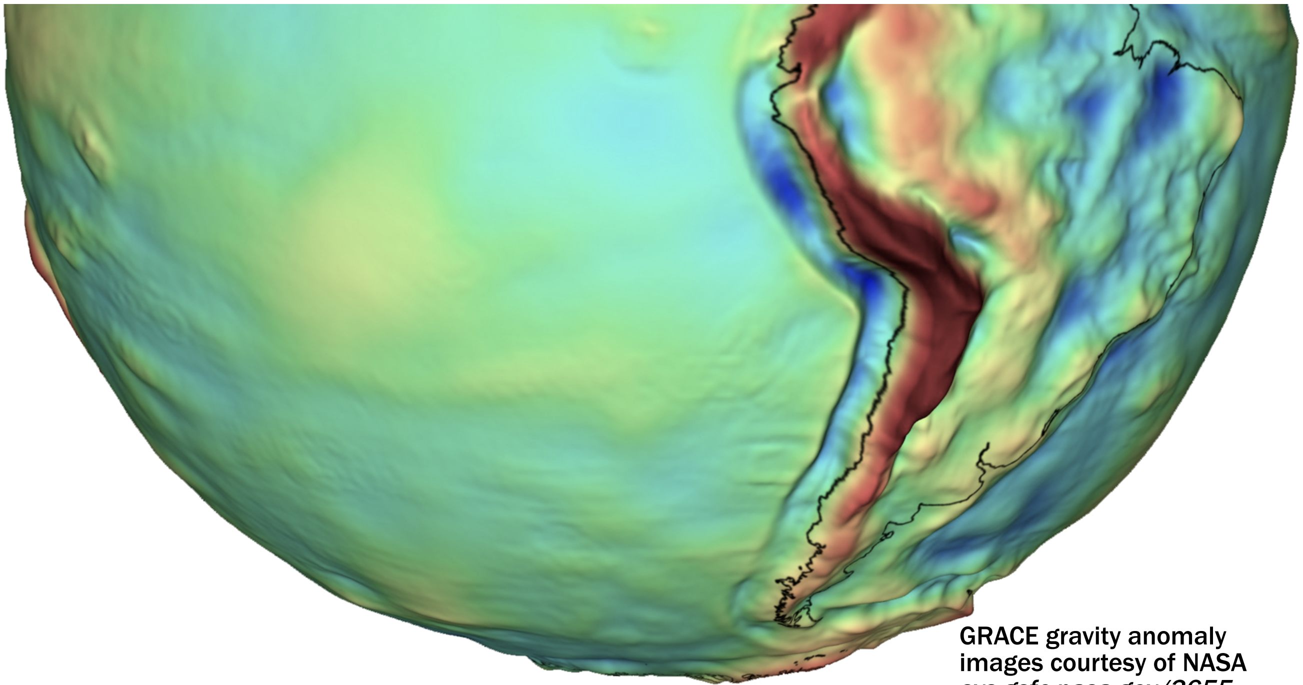
Guru: "Let you conclude? I'm the guru! Yesterday (10 April, 2019), it was announced that the first image of a black hole had been captured. So, national network news broadcasters were saying things like 'At the event horizon of a black hole, gravitational forces are so strong that time actually stops'. And you told me I was crazy for saying we could measure gravity with clocks".

Me: "Uh…"

References

ASCE, ACSM, ASPRS, (1994), Glossary of the Mapping Sciences, Bethesda, MD and New York, NY.

NOAA, (2017), "Blueprint for 2022, Part 2: Geopotential Coordinates," Technical Report NOS NGS 64.



GRACE gravity anomaly
images courtesy of NASA
svs.gsfc.nasa.gov/3655

WSRS2022

Wisconsin Spatial Reference System 2022 Task Force

www.sco.wisc.edu/community/wsrs2022