THE CENTER OF THE UNITED STATES AND OTHER APPLICATIONS OF CALCULUS TO GEOGRAPHY

DAVID RICHESON

The centroid, or center of mass, of a bounded planar region is the location at which that region would balance if it were made of a flat material of uniform density. The concept of the centroid is typically presented to second semester calculus students as an application of the integral. Most textbooks provide few, if any, real-world applications of the centroid, although some students may see the idea of the centroid later applied to probability.

As a refreshing and different application of the centroid I assigned the following project to students in my calculus classes: find the geographic center of the conterminous United States (ignore Alaska and Hawaii). Inspired by the high quality solutions and the enthusiastic responses that I received, I suggest expanding the project so that it includes several other topics from calculus. The objectives are as follows.

(1) Estimate the area of the U.S.
(2) Locate the geographic center of the U.S.
(3) Locate the geographic median of the U.S. (the point that divides the nation into equal area north-south regions and equal area east-west regions).
(4) Find the population center of the U.S.
(5) Find the population median of the U.S.
(6) Estimate the perimeter of the U.S.

The United States as a mathematical region

Our first task is to represent the United States mathematically. Again, for simplicity we consider only the conterminous United States. First, transfer the map of the United States to graph paper. Pick enough points on the map to describe sufficiently the boundary of the country, then transfer these coordinates to the CAS. Fit functions to the northern and southern borders of the country. Finally, use these two functions and the CAS to compute the desired quantities. (For the calculations that accompany this paper see [12].)

Despite the simplicity of this checklist, there are a few challenges that we face. First, the United States is not bounded by the graphs of two functions. We can safely ignore small geographic features such as Long Island and Cape Cod, but we cannot ignore Florida's eastward protrusion into the Atlantic Ocean on a coast that slopes to the south-west. As a result of Florida's orientation some vertical lines intersect the boundary more than twice. There is simple and elegant solution to this problem: rotate the map. With the map oriented so that Florida points straight down, the country can be reasonably represented as a region bounded by two functions.

The second hurdle to overcome is the problem of fitting a curve to the complicated boundary of the country. While it is possible to find a polynomial of high degree that fits the data points exactly, its graph would have wild oscillations that do not resemble the underlying boundary. Thus, to have an accurate fit we break the boundary into small pieces and create a piecewise function that fits the boundary.
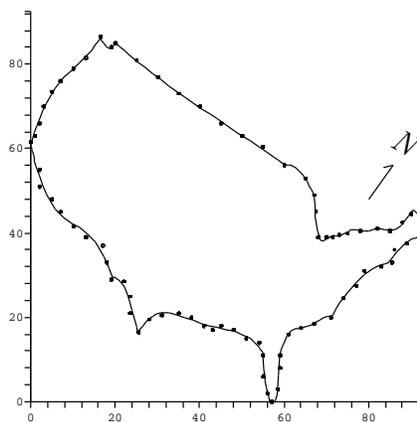

FIGURE 1. The U.S. as a region bounded by the graphs of two functions.

In Figure 1 we see the country bounded by the graphs of two functions: the northern function, $N(x)$, and the southern function, $S(x)$ (although not literally accurate, we refer to up, down, right and left as north, south, east and west). The piecewise functions $N(x)$ and $S(x)$ consist of six and seven polynomial functions respectively. We denote the western and eastern extremities of the country as $x = W$ and $x = E$. We assume that the U.S.-Canadian border runs through the middle of Lakes Superior, Huron, Erie and Ontario.

Before proceeding to our calculations we should pause and discuss one of our assumptions. What are the the consequences of using a flat

map of the United States? Because of this projection, objects in the north are distorted and appear proportionally larger than they would in the south. We compute the distance between Miami and Los Angeles and find that each unit on our graph paper corresponds to 28.8 miles. However, doing the same for Boston and Seattle we find that each unit is 31.8 miles. This distortion places the centroid of the country farther north than it should be. Similarly, the computed lengths of the U.S.-Canadian and U.S.-Mexican borders will also suffer. For the sake of simplicity we ignore this distortion, but it is a detail about which we should be aware.

The area of the United States

As an easy exercise, we can now compute the area of the United States. Assuming that the side length of each square on the graph paper corresponds to $\alpha = 30.3$ miles, the area of country is

$$\alpha^2 \int_W^E \left(N(x) - S(x) \right) dx.$$

Using the CAS to compute the integral we find the area to be approximately 3,084,000 square miles, only 7% larger than the accepted figure of 2,870,083 square miles [3].

The geographic center of the United States

The center of the United States may seem to be nothing more than an item of trivia, but in some circumstances its location may be very important. The idea of centrality and the definition of "center" was important to the founding fathers at the end of the eighteenth century [1]. At this time they were trying to choose a central location for the nation's new capital. The northern states wanted it to be near the population center, arguing that it was important that it be near the people. They asserted that the population center was in Pennsylvania, northeast of the Susquehanna River (the census of 1790 showed that it was actually in Maryland). Wanting a location farther south, the southern states pushed for the capital to be established near the geographic center of the thirteen states. They argued that it was important for the center of the government to be as close as possible to every part of the nation. (Some Northerners argued that the capital should be near the economic center.) Citing these facts and others, they chose Washington, D.C. to be the nation's capital.

This historical anecdote illustrates the importance of the problem of finding the center of the United States. It also illustrates that there

may be many different definitions of the center. We begin with the *geographic center*, the centroid.

The centroid of the region bounded by the curves $y = f(x)$ on the top, $y = g(x)$ on the bottom, and $a \leq x \leq b$ is the point $(\bar{x}, \bar{y})$ where

$$\bar{x} = \frac{1}{A} \int_a^b x(f(x) - g(x))\, dx, \quad \bar{y} = \frac{1}{A} \int_a^b \frac{1}{2}(f(x)^2 - g(x)^2)\, dx, \text{ and}$$

$$A = \int_a^b (f(x) - g(x))\, dx.$$

With these integrals and the functions $N(x)$ and $S(x)$ we use the CAS to compute the geographic center of the United States. We find that it is located just south of the Kansas-Nebraska border in Phillipsburg, Kansas (see Figure 2). Phillipsburg is 40 miles west of the accepted location of the geographic center, Lebanon, Kansas [4].
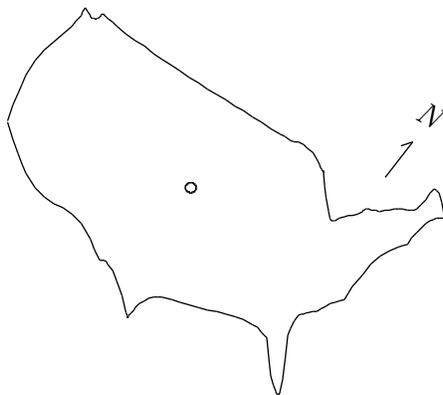

FIGURE 2. The geographic center of the United States.

While the government acknowledges that Lebanon is the geographic center of the conterminous United States, there is no officially sanctioned marker at this location. However, in 1940, with the hope of creating a tourist attraction in their small town, the Hub Club of Lebanon hired engineers to find the exact location of the geographic center and placed a monument there. A hotel was opened near the monument, but the tourist draw was smaller than expected and the hotel closed within a few years [8].

In general it is not true that any line through the centroid divides the region into two regions of equal area. For instance, it is not necessarily the case that the north-south and east-west lines that divide the country into regions of equal area pass though the geographic center. We call the intersection of these two lines the *geographic median*. The

geographic median depends on the choice of axes. We cannot divide our (rotated) map vertically and horizontally to find the geographic median. We must use lines parallel to the lines of longitude and latitude.

We now explain how to find the east-west line that runs through the geographic median; the north-south line is obtained similarly. On the graph paper, we measure the slope, m, of an east-west line. We must find the intercept, k, so that the line $l_k(x) = mx + k$ divides the country into two regions of equal area. Notice that on the west coast the east-west line intersects the graph of $N(x)$ and on the east coast it intersects the graph of $S(x)$. Given k, the CAS solves the equations $l_k(x) = N(x)$ and $l_k(x) = S(x)$ for x. Denote these x-values by $x = L_k$ and $x = R_k$ respectively. Then, by evaluating the following expression, we find the fraction of the area that lies north of line $y = l_k(x)$

$$\frac{1}{A}\Big(\int_{L_k}^{R_k} (N(x) - l_k(x))\, dx + \int_{R_k}^{E} (N(x) - S(x))\, dx \Big)$$

(as before, A is the total area of the country). We need two integrals because part of the region is bounded on the bottom by $l_k(x)$ and part by $S(x)$. We vary k until the computed quantity is $1/2$.

Since the United States is roughly rectangular we expect the geographic center and the geographic median to be close to one another. Still, we were shocked to discover that, by our calculations, the geographic median is less than 13 miles to the northeast of the geographic center (see Figure 3).
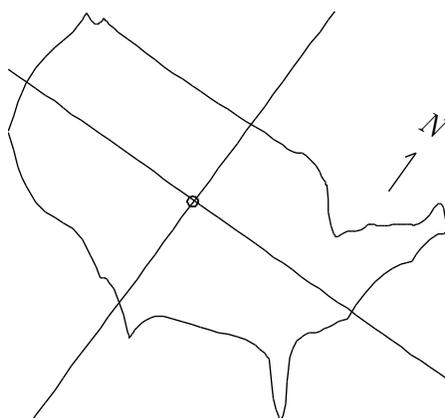

FIGURE 3. The geographic center (the circle) and the latitude and longitude lines through the geographic median of the U.S.

There are other noteworthy geographic centers in the United States. After Alaska and Hawaii became states, Lebanon, Kansas, became the "historic center of the United States." The geographic center of the fifty states is Castle Rock, South Dakota [4]. Rugby, North Dakota, is the geographic center of the North American continent [3]. A U.S. Coast and Geodetic Survey benchmark in Meades Ranch, outside of Osborne, Kansas, and 40 miles south of Lebanon, marks the *geodetic center* of the United States, sometimes called the North American Datum. This arbitrarily chosen central location serves as the reference point for all North American surveys, and all property lines and all local, state, national, and international boundaries are defined in reference to it. The geodetic center was established in Osborne in 1891; it is a site on the National Register of Historic Places.

The population center of the United States

The *population center* of a region is the centroid of the distribution obtained by replacing each person by a point with unit mass. To approximate this distribution for the United States we use the 75 most populated metropolitan areas (as reported by the 2000 census, [11]) as points and use each city's population as its mass. These metropolitan areas account for approximately 180,140,000 people, roughly 65% of the total population of the conterminous United States

The federal Office of Management and Budget is charged with defining metropolitan areas at the time of each census. Doing so enables all federal agencies to collect, share and compare statistical information about such areas. A metropolitan area is described as "an area containing a recognized population nucleus and adjacent communities that have a high degree of integration with that nucleus" [10]. A metropolitan area may consist of a city and its suburbs or it may include a group of cities. For instance the Baltimore-Washington D.C. region is one metropolitan area.

Recall that the center of mass, $(\bar{x}, \bar{y})$, for a set of n points in the plane, the ith of which has mass m_i at (x_i, y_i), satisfies

$$\bar{x} = \frac{1}{M} \sum_{i=1}^{n} m_i x_i \text{ and } \bar{y} = \frac{1}{M} \sum_{i=1}^{n} m_i y_i,$$

where $M = \sum_{i=1}^{n} m_i$ is the total mass. We place these 75 metropolitan areas on our graph paper and transfer the coordinates and the populations to the CAS. Using the centroid formulas we find the population center to be in Montgomery City, Missouri, approximately 70 miles west-northwest of St. Louis (see Figure 4).

FIGURE 4. The 75 most populated metropolitan areas
and the population center (the box-shaped point).

The reader may wonder what effect the rest of the population has
on the location of the population center. We computed the population
center for the to 50 most populated metropolitan areas before we com-
puted it for the top 75. The inclusion of these 25 metropolitan areas
added approximately 18 million people, or 6% of the total population,
yet the population center moved less than 11 miles. Thus one hopes
that including the remaining 35% would not cause the population cen-
ter to move too far.

Following the 2000 census, the U.S. Census Bureau announced that
the population center was in Edgar Springs, Missouri, approximately
90 miles south-southwest of our location, Montgomery City [2], [9].
However, they include the residents of Hawaii and Alaska. Also, their
computation technique is different from ours. To compute the cen-
ter, they subdivide the United States into 8,000,000 regions of equal
area and count the population in each. They treat the center of each
region as a point-mass and use a discrete centroid formula for points
distributed on a sphere that is analogous to the planar version that we
used [4].

The *population median* is the intersection of the north-south and
east-west lines that divide the country into regions of equal population.
Again, we use the same 75 metropolitan areas. With the help of the
CAS we compute the fraction of the population living north of a given
east-west line (or west of a north-south line), and we vary the intercept
until that proportion is 1/2. Doing so, we find the lines shown in Figure
5. They meet in Brookville, Indiana, which is 30 miles northwest of
Cincinnati and 65 miles southeast of Indianapolis. Notice that each of

the quarters do not necessarily contain the same population, but both
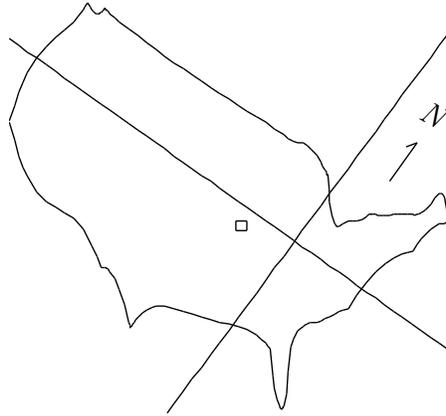pairs of opposite quarters do.

FIGURE 5. The population center (the box-shaped
marker) and the latitude and longitude lines through the
population median of the U.S.

The U.S. Census Bureau reports that the population median of the
United States (including Alaska and Hawaii) is in Van Buren Township,
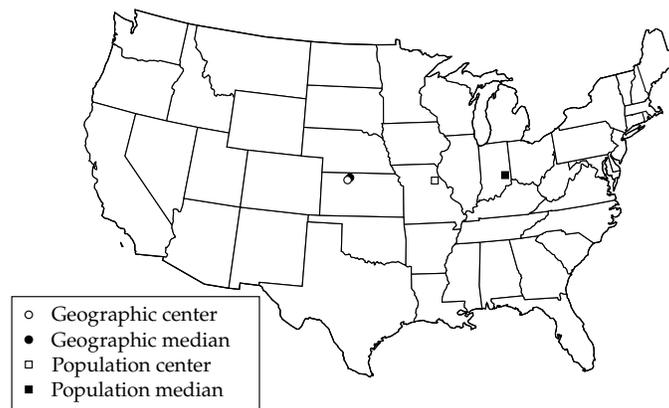Indiana, 85 miles southwest of Brookville [4].

FIGURE 6. Four centers of the United States.

It is an interesting class project to track the population center and
the population median through history. There are population data
available for each decennial census from 1790 to 2000 (see [7] for the 100
largest cities in each census year). A map from the U.S. Census Bureau,
[5], shows the population center moving from Chestertown, Maryland
in 1790 to Edgar Springs, Missouri in 2000 following a westward, then

southwestward path. Another map, [6], shows the population median migrating southwest from Troy, Ohio, in 1880 to the 2000 location of Van Buren Township, Indiana.

The perimeter of the United States

Finally, we compute the perimeter of the United States. We know that the length of a curve $y = f(x)$ for $a \leq x \leq b$ is given by

$$L = \int_a^b \sqrt{1 + (f'(x))^2}\, dx.$$

Using the functions $N(x)$ and $S(x)$ (and using α to convert to miles) we find the perimeter of the United States to be 8,878 miles: the U.S.-Canadian border is 3,283 miles, the U.S.-Mexican border is 1,386 miles, and the lengths of the Atlantic, Gulf, and Pacific coasts are 1,719, 1,315, and 1,175 miles respectively.

According to the 1963 Boundary Agreement, the length of the U.S.-Mexican border is 1,933 miles. By the U.S. Geological Survey the length of the U.S.-Canadian border (excluding the Alaskan border) is 3,987 miles [3].

The lengths of the jagged coasts are more difficult to compute than the relatively smooth national borders. The National Oceanic and Atmospheric Administration (NOAA) publishes two sets of data regarding the length of the coasts. One measurement they title "coastline" and the other they title "shoreline."

The *coastline* is the length of the coast measured using a "ruler" of length 30 minutes of latitude. It is the length of the general outline of the coast, ignoring small bays, peninsulas, and islands. Using this method NOAA found the Atlantic, Gulf, and Pacific coastlines to be 2,069, 1,631, and 1,293 miles respectively [3].

The *shoreline* is the length of the coast obtained by measuring every feature, including offshore islands, bays, sounds, and mouths of rivers (up to the head of the tidewater or to the location where the tidal waters narrow to 100 feet). In 1939-40, using the most accurate charts then available, NOAA found the shorelines to be 28,673, 17,141, and 3,863 miles for the Atlantic, Gulf, and Pacific coasts respectively [3].

Regardless of which of these techniques we use to measure the coasts, our computed perimeter is a significant underestimate. Using the coastline figures the perimeter of the United States is 10,913 miles and using the shoreline figures it is 54,059 miles. The error of our estimate is 19% using the first figure and 84% using the latter!

A crude approximation of the boundary does not significantly affect the enclosed area or the location of the centroid, but it can have a huge effect on the perimeter. The length of the boundary depends to a very high degree on small, local fluctuations. We are reminded of the analogy likening a rocky coast to a fractal; certain fractal curves, such as the Koch snowflake, can enclose a finite area, yet have infinite length. Thus, we must question whether our problem is well defined. What do we mean by the perimeter of the United States? In response to the question, "What is the perimeter of the United States?" one should reply with the question, "What size ruler should I use?"

Conclusions

This project succeeds as a teaching tool on several levels. It provides calculus students with real-world applications that are interesting and easy to understand. It gives the kind of results with which they can wow their friends at the lunch table. Meanwhile, the project forces them to use important calculus concepts: the area formula, the various centroid formulas, and the arc length formula. It gives students an opportunity to work with real data, not contrived data that are created for the simplicity of calculations. They see the usefulness of technology for solving complicated problems. Finally, they are forced to examine their assumptions (using a flat map and not a globe, sampling a percentage of the total population, using polynomials to model the jagged coastline, etc.) and see the consequences of these assumptions.

References

[1] K. Bowling. *The creation of Washington, D.C.: the idea and location of the American capital*. George Mason University Press, Fairfax, Virginia, 1991.

[2] F. Broome. *2000 U.S. Population centered in Phelps County, MO*. U.S. Census Bureau, Washington, D.C., April 2, 2001.

[3] R. Famighetti, editor. *The world almanac and book of facts 1999*. World Almanac Books, Mahwah, NJ, 1998.

[4] Geography Division. *Centers of population computation for 1950, 1960, 1970, 1980, 1990 and 2000*. U.S. Census Bureau, Washington, D.C., April 2001.

[5] Geography Division. *Mean center of population for the United States: 1790 to 2000*. U.S. Census Bureau, Washington, D.C., March 30, 2001.

[6] Geography Division. *Median center of population for the United States: 1880 to 2000*. U.S. Census Bureau, Washington, D.C., 2001.

[7] C. Gibson. *Population of the 100 largest cities and other urban places in the United States: 1790 to 1990*. U.S. Census Bureau, Washington, D.C., June 1998.

[8] The Lay of the Land: The Center for Land Use Interpretation Newsletter. *The geographical center of the lower 48 United States, at Lebanon, Kansas*. The Center for Land Use Interpretation, Culver City, CA, Spring 1999.

[9] NOAA News. *NOAA's National ocean service marks the USA's population center*. National Oceanic and Atmospheric Administration, Washington, D.C., May 2, 2001.

[10] Office of Management and Budget. Standards for defining metropolitan and micropolitan statistical areas; notice. In *Federal Register*, volume 65, No. 249, pages 82227–82238. Washington, D.C., December 27, 2000.

[11] Population Division. *Ranking tables for metropolitan areas: population in 2000 and population change from 1990 to 2000*. U.S. Census Bureau, Washington, D.C., April 2, 2001.

[12] D. Richeson. Complete calculations for this paper using Maple. http://www.dickinson.edu/∼richesod/centerus/.

DICKINSON COLLEGE, CARLISLE, PA 17013
E-mail address: `richesod@dickinson.edu`