(Re)introducing the equilateral age-period-cohort demographic surface

Tim Riffe

Department of Demography
University of California, Berkeley

August 2, 2014

Abstract

The use of demographic surfaces based on time proportionality between age, period, and cohort axes is introduced as an alternative to standard *Lexis* proportions in order to eliminate distortion of the cohort perspective and ensure comparable time scales. Resulting images are legible and interpretable in a similar way to standard demographic surfaces, but entail no visual bias in the visual comparison of slopes between temporal dimensions.

This paper is about demographic data visualization and the link between the *Lexis*[1] diagram and demographic surfaces. Any demographic quantity cross-classified by age, period and cohort (APC), or any two thereof, is candidate to be plotted as a demographic surface. Usually, such a surface is projected onto an age-period cartesian grid, with age on the y axis and calendar year on the x axis. The plane is then filled with color gradients or contour lines to specify value intervals for the quantity in question, the third dimension. It may be a misnomer to refer to such figures as surfaces, since these do not create an illusion of three-dimensionality (as per a stereogram or perspective plot) but the analogy invites the use of one's spatial intuition in understanding the data. This family of demographic plot is a subset of heatmaps or contour plots.[2] One of the main points of plotting a demographic surface is to view data over all three major time axes. The demographer often searches for cohort patterns in the data, diagonal ridges or trend shifts, that would be less apparent if viewed as a series of lines along the age or period year dimensions.

Demographic surfaces are almost always rendered following the cartesian time-axis conventions of the *Lexis* diagram, which forces a unity aspect ratio between calendar years and age years along the age and time axes. Cohorts are represented by diagonal lines at 45°, where a single cohort year stretches across the hypotenuse of the standard Lexis triangle, making it around 40% longer on paper than an age or calendar year. While unproblematic in didactic line diagrams, this cohort distortion can cause visual bias when comparing the period and cohort dimensions of data. I reintroduce a coordinate transform that ensures a 1:1:1 relationship between all three temporal dimensions in the APC surface, thereby aiding visual interpretation of surfaces. This is the same graphical relationship originally rendered by Lexis as a diagram, and thereafter in the famous stereogram of Luigi Perozzo (1879), but never applied to the graphical representation of demographic data in contemporary graphics.

Cohort distortion in the *Lexis* diagram

The *Lexis* diagram conforms to a set of principles on orientation and proportionality, including a unity aspect ratio between age increasing upward on the y-axis and calendar years increasing rightward along the x-axis. A unity aspect ratio means that one year in the age direction is equal to one year in the period direction when drawn. Birth cohorts start with years on the x axis at age 0 and increase upward and to the right at a 45° angle (segment AC). Imaginary life lines of individuals within the population follow along the cohort direction. A side effect of this 45° angle is that the cohort dimension is stretched on paper by about 40% (scaled by $\sqrt{2}$). Demographers are well aware of this distortion, but it does not make the *Lexis* diagram any less useful.[3]

These conventions conform to demographers' 'data-intuition', as demographic data are typically collected annually, and the most-used indicators are period in-

---

[1]in italics throughout because Wilhelm Lexis was only one of many potential namesakes (Vandeschrick 2001, Keiding 2011)

[2]Edmund Halley is often cited as the maker of the first contour map (an isogonic map), and he also holds a place in the history of demography for the development of lifetables.

[3]For example, the HMD Methods Protocol (Wilmoth et. al., 2007) makes explicit mention of the $\sqrt{2}$ distortion of the cohort dimension in diagrams.

dicators. A clear right angle between period and age lends itself to illustrating period indicators. Resulting cohort dimension distortion is largely innocuous when drawn as a diagram in outline form, since *Lexis* diagrams do not represent data values, but rather data structure and relations. Distortion only arises when plotting data values represented by color gradients, contour lines, or similar graphical devices. To illustrate the nature of this bias, consider a line graph, for instance a log mortality curve over age. Figure 1 shows US year 2010 log male mortality (Wilmoth et. al., 2007), with the principle axis on the bottom. The stretched axis on the top represents the exact degree of distortion that the cohort dimension is subject to on Lexis surfaces. From this it is clear that the nature of this distortion impedes the visual judgement of slopes, but not levels.

Figure 1: Translation of the visual bias in the standard *Lexis* surface.
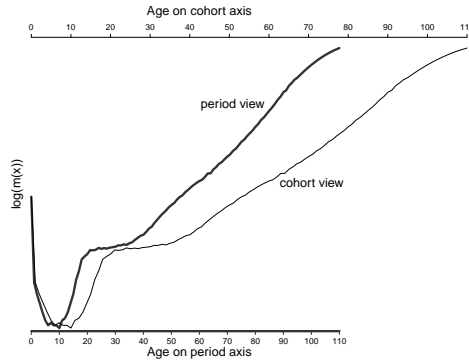
Modern computing has facilitated the visualization of large tables of demographic rates cross-classified by age, period and cohort using surface plots, an idea popularized by Arthur and Vaupel (1984). Software packages have been designed to do this with demographic data, such as `LEXIS` by Vaupel et al. (1987) or the `Lexis` program by Andreev (1999), which recognize and graphically represent the inherent triple-classification of the underlying data. A variety of general surface functions for gridded data (e.g., age-period only) in widely used statistical programming languages[4] that can be coerced to accurately represent standard demographic surfaces. Surface plots produced in this lineage have followed the design principles of the *Lexis* diagram, entailing distortion of the cohort dimension in favor of a right angle between the age and period dimensions.[5]

Surface plots are about summarizing data; the human eye can find patterns and relationships in the data when represented as a surface. Demographic surfaces are a diagnostic tool, but more importantly they offer tremendous potential to illustrate and inspire stories from the data that aid understanding, and that can even sometimes make intuitive difficult-to-grasp concepts, such as tempo and quantum. The utility of demographic surface plots is therefore different

---

[4]such as `levelplot()` in the `lattice` package (Sarkar 2008) for gridded data in the R language (Ihaka and Gentleman 1996), or the `contourf()` function in `MATLAB` (MATLAB 2010), among many others.

[5]Rearranging the data, a less common rendering of age in the y-axis and cohort in the x-axis with a unity aspect ratio, in which case the period dimension is stretched by the same amount.

from but complimentary to that of *Lexis* diagrams. Cohort distortion in surfaces is at times a problem because the viewer is implicitly invited to judge the three temporal perspectives.

An equilateral coordinate system

One can ensure visual comparability between the age, period and cohort perspectives by mapping data to equilateral triangles rather than to *Lexis* right triangles. In this way, one year older in the age direction is equal to one year lived in the cohort direction and one calendar year in the period direction. Under this system, age-period squares become age-period diamonds[6], and so forth. The idea of using equilateral triangles to ensure proportionality in APC renderings is not new. Lexis (1875) first entertained the use of 60° angles[7], later developed by Lewin (1876) and implemented in the famous stereogram of Perozzo (1880)[8]. Perozzo's stereogram is the only instance I have found of a data representation under equilateral coordinates. This feature of the rendering is not necessarily obvious to the viewer, due to the use of three dimensional perspective and the low angle of view. The present proposal for rendering such a surface is in keeping with the present-day convention of age increasing upward, rather than downward.

The arithmetic required to transform standard *Lexis* surface coordinates into equilateral coordinates on a cartesian plane while maintaining a 1:1 scale is straightforward. Using p to denote year, a for age and c for cohort, we can produce x and y coordinates for the standard and equilateral triangles seen in Figure 2. Figure 2 compares a diagram of age-period regions (any given p and a) divided into two APC triangles using the standard *Lexis* proportions (left) and an equilateral proportions (right). Points A, B, C and D map to each other in both depictions.

The segment AD marks the beginning of year p, and BC marks year $p + 1$ in both diagrams. In the equilateral depiction, however, years slant upward and to the left at 120°. The age dimension is read horizontally in both depictions, segments AB (a) and DC ($a + 1$), and its length remains 1. Segment AC marks cohort c in both depictions, dividing the upper triangle ACD (when $p - c - a = 1$) from the lower triangle ABC (when $p - a = c$). Segment AC changes its angle from 45° in the standard depiction to 60° in the equilateral depiction. Importantly, its length changes from $\sqrt{2}$ in the standard drawing to 1 in the equilateral drawing, which is the entire justification for this transform.

In short, to translate from standard to equilateral *Lexis* coordinates for any given (p, a) pair, subtract $\frac{a}{2}$ from p and scale a by $\frac{\sqrt{3}}{2}$ (or $\sin 60°$).

---

[6]60-120 degree rhombuses

[7]See Lexis (1875) page 13, and Figure 2

[8]At the time of this writing I have been unable to locate Lewin (1876), but it is included in the discussion by Keiding (2011).
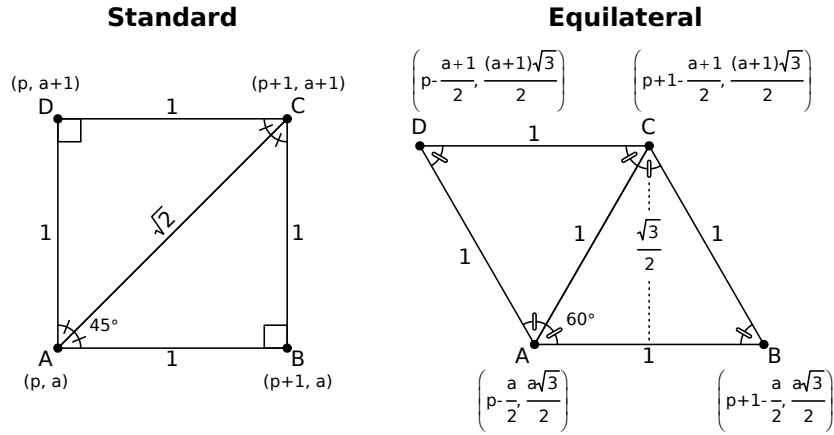
Standard

Equilateral

Figure 2: Standard *Lexis* versus equilateral coordinates.

Standard versus equilateral APC demographic surfaces

A demographic surface composed of tessellating APC equilateral triangles (or AP, PC or AC diamonds) is not radically different or exotic from one made of the more commonly used APC *Lexis* triangles (or AP squares, or PC or AC parallelograms); it appears slanted 30° to the left, rather than as a rectangular surface. Compare two renderings of Czech APC fertility rates extracted from the Human Fertility Database (HFD, 2011) in Figure 3.

The equilateral demographic surface in Figure 3 resembles the better-known standard demographic surface closely enough that one generally familiar with interpreting surfaces requires little effort to understand it equally well. For this particular Czech fertility surface, one appreciates the visual transform of the post-velvet revolution fertility changes (say, from 1992 to 2005) in the equilateral rendition. The 1975 cohort, for instance, stands in relatively sharp relief against the 1974 cohort in the age range 17 to 23 in both the standard and equilateral forms, but its relative importance (say, relative to the year 2000 ASFR) can be assessed with fidelity only in the later. The same advantage holds for each inter-cohort postponement decrease in the age-range 17-23 in the 1974-1977 cohorts, and for inter-cohort fertility increases above age 30 in the 1969-1977 cohorts (and potentially beyond). Period shifts, such as 1964-1965 or 1979-1980 are just as eye-catching in the equilateral surface as they are when viewed at 90° in the standard surface. On the whole, the equilateral surface from Figure 3 enhances one's understanding of the underlying data.
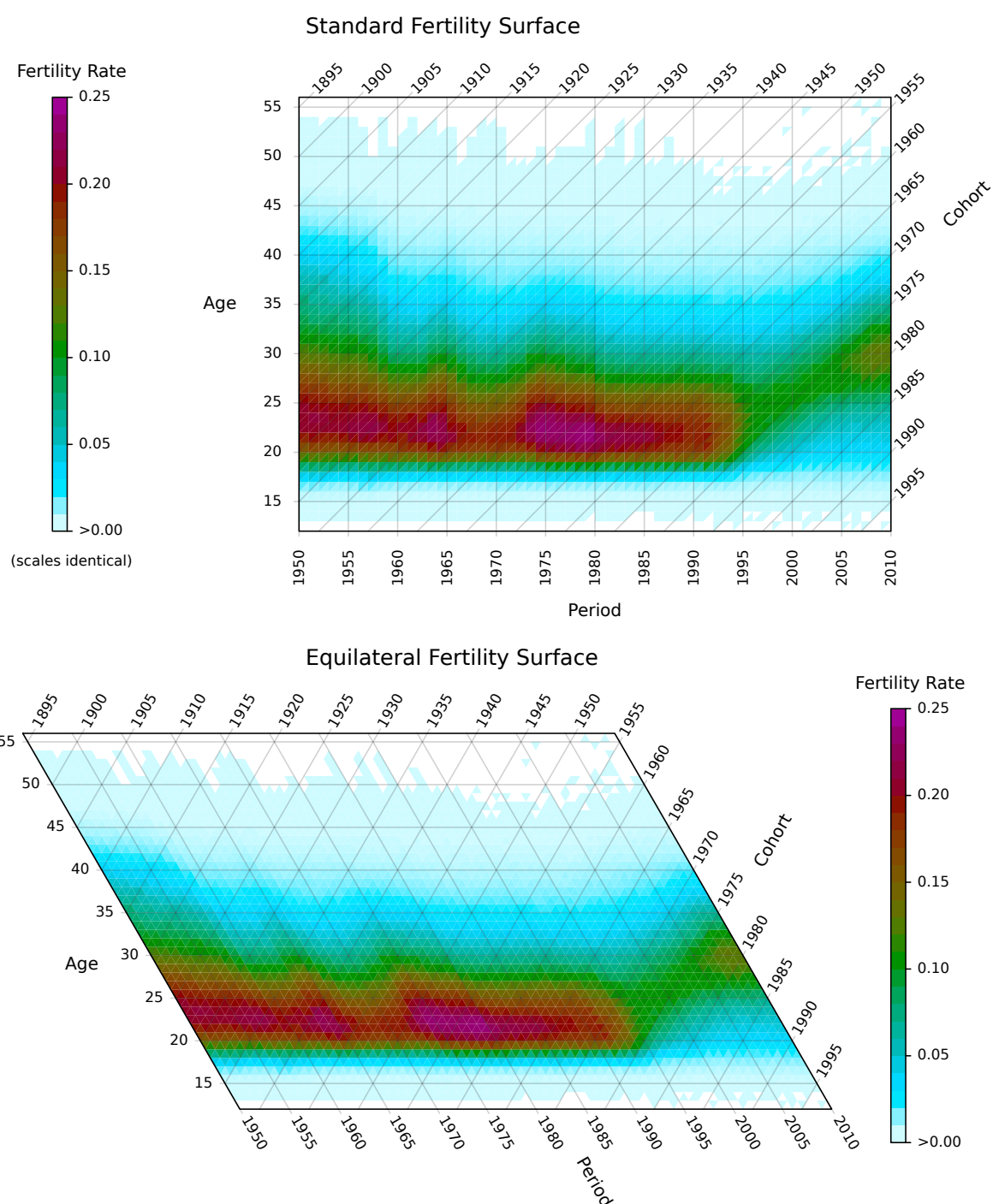
Figure 3: Czech Age-Period-Cohort specific fertility rates, 1950-2009 (HFD). Standard *Lexis* proportions and coordinates (above) versus equilateral (below). Surfaces are plotted on 1:1 scales and share the same color legends and value intervals.

The color scale used in Figure 3 also adds value by interpolating over the HCL color space (Ihaka 2003, Ihaka et al. 2011) rather than the more frequently used RGB or HSV color spaces, either of which would entail irregularly perceived color intensity, potentially leading to false edges or other optical illusions (Zeileis et al. 2009). This HCL ramp equalizes perceived distance between colors by forcing a uniform colorfulness (chroma) and equidistant hues. An approximately linear luminance gradient ensures that the surface remains intelligible, albeit less sharp, both for color blind viewers and on grey scale devices. Further work is needed to establish best practices for the use of color in demographic surfaces in general and the most frequently used subtypes of APC surfaces, i.e. mortality and fertility, in particular.

Conclusions

Equilateral Lexis surfaces overcome the problem of cohort distortion found in standard APC surfaces, and they are worth consideration for use in both diagnostic and didactic settings, minimally for the sake of comparison. The small tweak to the cohort dimension that sparked this endeavor has not been overlooked in recent applications of *Lexis* diagrams, but the coordinate transform presented here appears never to have been implemented for APC demographic surfaces, save for the original work of *perozzo1880della*. The R code used to create the surfaces in Figure 3 is freely available in the online appendices to this paper, and an R package for making both standard and equilateral demographic surfaces is currently in preparation.

References

K. Andreev. Overview of the program lexis 1.0. *S. Heinzel and T. Plesser*, 1999.

W.B. Arthur and J.W. Vaupel. Some general relationships in population dynamics. *Population index*, 50:214–226, 1984.

Human Fertility Database. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria), December 2011. URL www.humanfertility.org.

R. Ihaka. Colour for presentation graphics. In K. Hornik and F. Leisch, editors, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing Vienna Austria*, 2003.

Ross Ihaka and Robert Gentleman. R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, pages 299–314, 1996.

Ross Ihaka, Paul Murrell, Kurt Hornik, and Achim Zeileis. *colorspace: Color Space Manipulation*, 2011. URL http://CRAN.R-project.org/package=colorspace. R package version 1.1-0.

N. Keiding. Age–period–cohort analysis in the 1870s: Diagrams, stereograms, and the basic differential equation. *Canadian Journal of Statistics*, 39(3): 405–420, 2011.

J. Lewin. Rapport sur la détermination et le recueil des données relatives aux tables de mortalité. *Programme de la neuvième session du Congrès International de statistique à Budapest I*, pages 295–361, 1876.

W.H.R.A. Lexis. *Einleitung in die Theorie der Bevölkerungsstatistik*. KJ Trübner, 1875.

MATLAB. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts, 2010.

L. Perozzo. Della rappresentazione graphica di una collettivita di individui nella successione del tempo. *Annali di Statistica*, 12:1–16, 1880.

Deepayan Sarkar. *Lattice: Multivariate Data Visualization with R*. Springer, New York, 2008. URL `http://lmdvr.r-forge.r-project.org`. ISBN 978-0-387-75968-5.

C. Vandeschrick. The lexis diagram, a misnomer. *Demographic Research*, 4(3): 97–124, 2001.

J.W. Vaupel, B.A. Gambill, and A.I. Yashin. *Thousands of Data at a Glance: Shaded Contour Maps of Demographic Surfaces*. International Institute for Applied Systems Analysis, 1987.

J.R. Wilmoth, K. Andreev, D. Jdanov, DA Glei, C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, and P. Vachon. Methods protocol for the human mortality database. *University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock*, 2007.

A. Zeileis, K. Hornik, and P. Murrell. Escaping RGBland: Selecting colors for statistical graphics. *Computational Statistics & Data Analysis*, 53(9):3259–3270, 2009.