

# **Lexis Fields**

**(authors redacted)**

## **Abstract**

### **BACKGROUND**

Lexis surfaces are an established visualization technique to show how a given value changes over age and time. Vector fields are a two-dimensional representation of two variables: direction and speed (or force). Phase planes represent limits and thresholds of differential equations, such as the relationship between first and second derivatives.

### **OBJECTIVE**

We aim to increase the dimensionality of patterns shown on the Lexis surface by placing a vector field on the Lexis surface. We also relate the phase planes to the Lexis surface by way of the Lexis field.

### **RESULTS**

We show Lexis field. using different combinations of visual encodings, such as color, contour layering, and angle, length, and curvature of field elements. These instruments enable information layering that is not otherwise possible on standard Lexis surfaces.

### **CONCLUSIONS**

Lexis fields extend the analytic power of the Lexis surface, and these can be designed to display information at higher densities than commonly made Lexis surfaces.

## 1. Introduction

Lexis surfaces are a graphical tool used to display data on the Lexis coordinate plane, a Cartesian plane that is also a simplex relationship between age, period, and cohort. Surfaces are often displayed as heat maps, contour maps, perspective plots, or variants of these things (?). Various kinds of quantities, such as raw magnitudes, differences (?), ratios (?), intensities, proportions, derivatives (?), and even compositions (?) can be displayed on Lexis surfaces in order to put age, period, cohort, or other patterns in relief.

Vector fields are a graphical form generally used to display variation in speed, direction, or force over a plane (?). Point estimates of these quantities on the plane are often represented with segments or arrows, where length may be proportional to a function of magnitude (force, speed), and angle indicates direction, potentially disambiguated with an arrowhead or articulated as a curve.

Maps in general combine layers of categorical, continuous, and symbolic information on a common spatial projection. Lexis surfaces in contrast almost exclusively display one visual layer at a time. We propose to enrich Lexis surfaces by adding a visual layer of quantitative information coded symbolically as a vector field, and we liken this to cartographic information layering. We propose a fusion of Lexis surfaces and vector fields, *Lexis fields*, as a tool to display variation in relationships between variables over age and time. A Lexis field may either be shown atop on a Lexis surface, representing two map layers — a true Lexis map — or as a single-layer stand-alone visualization.

We give an overview of constructing a Lexis field with an application. Our example explores the relationship between remaining life expectancy and the standard deviation of remaining lifespan over age and time based on all available populations in the ? from 1950 onward. Variations in aesthetics of the field over age and time reveal macro-patterns between demographic variables. This example and others discussed motivate how Lexis fields could be used in other areas demographic research.

## 2. Lexis field construction

It makes sense to plot a Lexis field if data contain a relationship that can be summarized with a line, a simple curve, or similar, that varies by age and/or time. Constructing a Lexis field involves several design choices, which can however be codified into four basic steps, which we summarize in Fig. 1. The steps to do so are outlined in the following description, referenced to regions of Fig. 1.

**A** Determine a grid over age and time on which to condition data selection. For example, a five-year grid implies  $5 \times 5$  Lexis cells. Data may be selected from multiple populations on the same reference grid or may consist in a

selection of attributes from a single underlying Lexis surface. Presumably two variables are required.

**B** Abstract a model from the data, such as a linear, parabolic, or similarly simple relationship.

**C** Translate the model to the characteristics of a line segment, which we refer to as a field *pointer*. For example, the pointer angle can relate to the model relationship, and length and color may be determined by other aspects of the model, such as  $r^2$ .

**D** Draw the segment in the corresponding Lexis grid cell, repeating all steps for each cell in the diagram.

### 3. Application

We select all HMD data available for females after 1950. For each life table we calculate two additional columns: the standard deviation of remaining lifespan  $sd(x)$ , and the coefficient of variation of remaining lifespan  $CV(x)$ . Each Lexis field element is based on the relationship between  $sd(x)$  and remaining life expectancy  $e(x)$  in  $1 \times 5$  Lexis cells<sup>1</sup>, as summarized by bivariate linear regression over the data points in each cell. The regression results used for each Lexis cell in resulting Lexis fields are identical, but the translation of regressions to field pointers varies between designs. We offer four examples of Lexis field designs, each drawn on  $5 \times 5$  Lexis subplots.

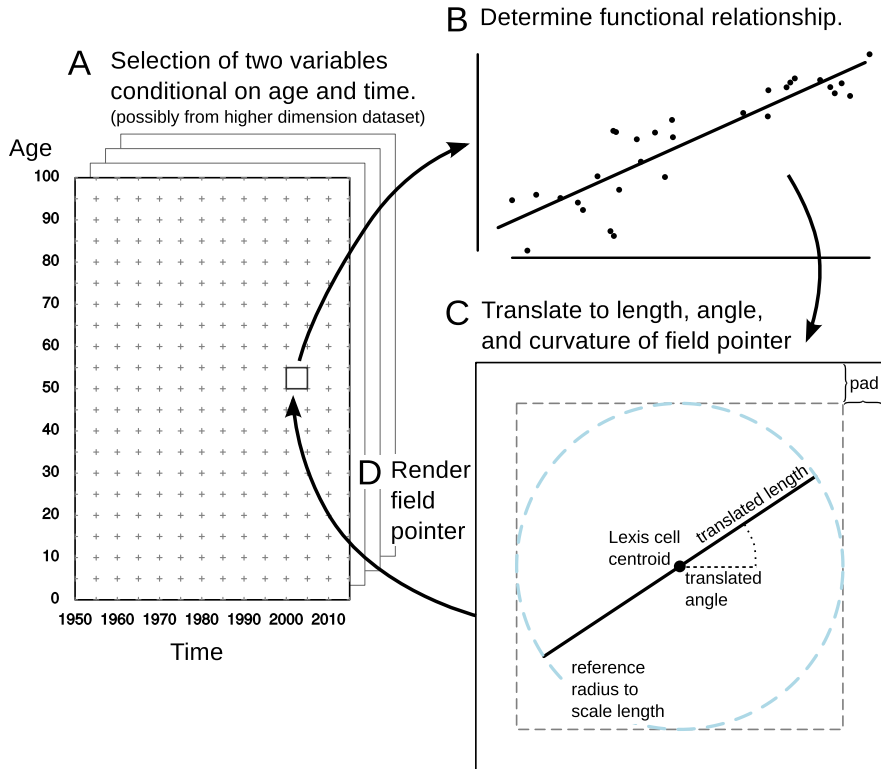
The first of these, Fig. 2, is a bare-bones Lexis field that serves to illustrate the underlying concept. This display draws each regression slope as a line segment of equal length (4 “years” long) and centered on each Lexis cell. The slope of each pointer is identical to regression slopes, which may be justified in this case because  $e(x)$  and  $sd(x)$  are both in year units. This is the truest and most literal depiction of how these regression slopes vary over age and time among HMD females, and nothing more. From this figure we can see, for example, that there is some age where the relationship turns from negative to positive, which increased slightly over time. Slopes dampened in younger ages around the 1980s, but have since increased (except infants).

The second version, Fig. 3, draws slopes multiplied by two, with pointer lengths scaled proportional to the between-population interquartile range of the  $e(x)$  and  $sd(x)$  values used in regressions<sup>2</sup>, and line weight and grayscale proportional to the  $r^2$  of the

<sup>1</sup>Data points included in regressions are single ages evenly divisible by five for each of the five years included in a Lexis reference cell.

<sup>2</sup>Specifically, pointer length is proportional to the central spread of the relationship between  $e(x)$  and  $sd(x)$ , approximated as  $\sqrt{IQR(e(x))^2 + IQR(sd(x))^2}$

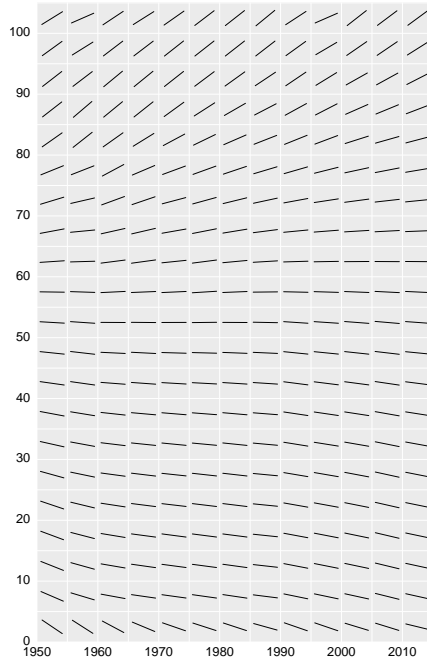
**Figure 1:** A diagram depicting the translation of functional relationships in data conditioned on age and time to visual encoding on a Lexis field. **A:** Condition data selection on age and time. **B:** model the functional relationship in data subset. **C:** Translate the model to field elements, ‘pointers’, using angle, and possibly also length, curvature, thickness, etc to encode model qualities. **D:** Populate the Lexis plane with field pointers to create a Lexis field.



regression fit. Segment lengths are therefore indicative of the spread in the data, while higher Pearson's  $r$  results in more contrast in the field.

The final example, Fig. 4 is a true Lexis map. The base of the map is a filled contour plot of the mean (over HMD populations) coefficient of variation of the conditional remaining distribution  $CV(x)$  in each single age and year. This map is redundantly coded with a sequential color palette and labelled contours, which liberates the surface from an

**Figure 2:** Lexis field:  $sd(x)$  by  $e(x)$  linear fits, with slopes drawn equal to regression line slope coefficients. Pointer lengths are all equal.

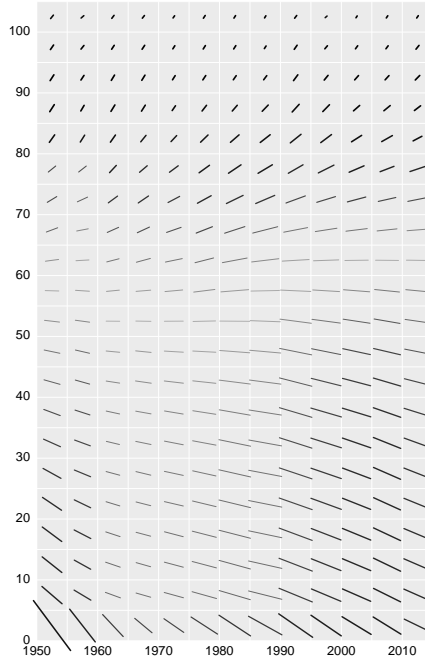


explicit color legend. The same field from Fig. 4 is layered atop the CV surface, achieving a true layered map. In principle, one could represent variation in the slope of some other regression over age and time as a contour plot, with the present field atop, thereby layering comparable information. However, the present example serves to illustrate layering with the field.

## 4. Discussion

We suggest the use of vector fields on the Lexis surface, introducing the notion of a Lexis field, which is a standard vector field on a regular Lexis grid over age and time. This instrument allows researchers to overlay Lexis surfaces in complex ways. Doing so may imply a lower degree of age and time resolution for each of the variables overlaid — for

**Figure 3:** Lexis field:  $sd(x)$  by  $e(x)$  linear fits, with slopes exaggerated by 2. Pointer length is proportional to the diagonal of the IQR box around  $sd(x)$  and  $e(x)$ , while grayscale and segment width are proportional to Pearson's  $r$ .

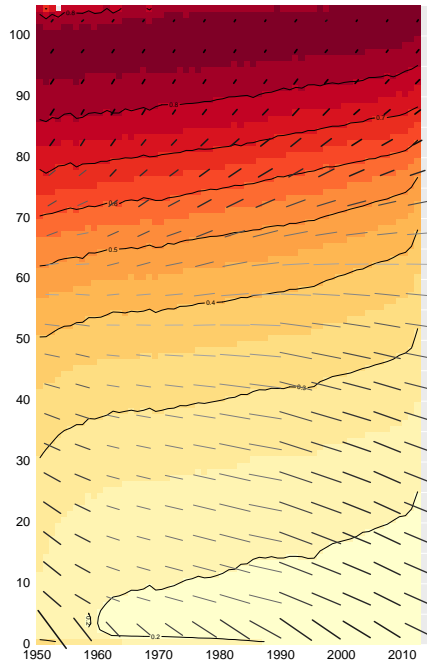


example we used a 5x5 grid in our application. We do not evaluate the tradeoff between the potential clarity of disjoint graphs versus between-surface comparisons using Lexis fields, but researchers may wish to experiment using the reproducible code we provide.

We demonstrate some of the designer freedom in translating data into the elements of a Lexis field, as well as an instance of Lexis map layering. These examples serve to illustrate the construction of Lexis fields, but do not pretend to be *best practice* Lexis surfaces in terms of visual design or legibility. It is our sense that the patterns revealed in Figures~2-4 are accessible to the viewer and lend themselves to substantive interpretation.

Even the composite surfaces of ? display layered information as a single visual layer. Small multiples of Lexis surfaces (for example panel plots of Lexis surfaces) on the other hand constitute a de-layering, as these are spatially disjoint. Comparisons between plots

**Figure 4:** A Lexis map: Lexis surface of mean CV as a filled contour plot plus a Lexis field of  $sd(x)$  by  $e(x)$  linear fits, slopes exaggerated by 2. Pointer length is proportional to the diagonal of the IQR box around  $sd(x)$  and  $e(x)$ , while grayscale and segment width are proportional to Pearson's  $r$ .



then require an extra step of attention from the the viewer as one cross-references patterns or values at specific coordinates of age and time. More dimensions will imply more such attentive processing work from the viewer.

This visual instrument indeed arose in practice in an attempt to investigate the apparently mechanical relationship between lifespan variation and average length of life with a macro view. Evidence suggests a negative relationship between lifespan variation and life expectancy when measured at young ages such as 0 or 15 (cite Vaupel et al 2011, Colchero et al 2016, Smits and Monden, Wilmoth Horiuchi 1999, Nemeth 2017). This means that at the aggregate level, as life expectancy increases (in low mortality countries),

length of life becomes more equal, and this relationship appears to hold up generally between human populations (Colchero). Studies point out that the age used for truncation is important (Nusselder & Mackenback 1996 and Robine 2001, Engelman 2010). Our visual approach reveals that the relationship starts as strongly negative and somewhere near age 55 reverses to become positive. A crossover was already documented by Myers and Manton (1984), but it has not previously shown in published research in such a systematic and comprehensive way, possibly because standard ways of looking at these data would have required dozens or hundreds of graphs. We aim to fill this gap by proposing a way to visualize these complex patterns in a single plot.

Fig.4 simply serves to illustrate that Lexis fields can be layered with traditional Lexis surfaces that are color coded, increasing the information and pattern density on the Lexis plane with little drawback in terms of legibility. Although patterns in data may be much more complex than can be expressed with linear models, these simple model fits can be thought of as regular samples from the complex space implied by data, such that the pattern revealed on the Lexis field is still revealing. On the other hand a field may be derived from a single underlying pattern rather than a series of regressions on different populations or variables. For example, ? recommends the use of phase diagrams to represent the rate of change of the hypothetical life course implied by period fertility curves. This construct could be translated to a Lexis field in a straightforward way, with pointers mapping to the notions of acceleration and velocity. Certainly variants of vector fields could be used to intuitively display other demographic phenomena and components of demographic change, and these do not need to adhere to the Lexis grid. Indeed, we suggest wider use of fields based on demographic relationships in spatial settings.

## 5. Conclusions

We describe the construction and use of vector fields on the Lexis plane. We argue that this technique can increase the information density and scope displayed on the Lexis surface. We also argue that fields are easily made compatible with Lexis surfaces drawn as filled contour plots, which lends them to map layering. We suggest alternative field designs and uses. In sum, displaying a larger variety of demographic quantities on the Lexis plane and increasing the information density on the Lexis plane using layering techniques such as fields should broaden the scope of demographic exploration and sharpen the instruments of demographic pattern detection.



## **6. Reproducibility**

Code used to produce the experimental visualizations here is available in a public repository:

`https://github.com/timriffe/MacroShape/DR`

## **7. Acknowledgements**

We wish to thank Alyson van Raalte for helpful comments received that improved this manuscript.