



Quadratic Trends: A Morphometric Tool Both Old and New

Fred L. Bookstein^{1,2}

Received: 3 July 2023 / Accepted: 28 November 2023 / Published online: 1 February 2024
© The Author(s) 2024

Abstract

The original exposition of the method of “Cartesian transformations” in D’Arcy Thompson’s *On Growth and Form* (1917) is still its most cited. But generations of theoretical biologists have struggled ever since to invent a biometric method aligning that approach with the comparative anatomist’s ultimate goal of inferring biologically meaningful hypotheses from empirical geometric patterns. Thirty years ago our community converged on a common data resource, samples of landmark configurations, and a currently popular biometric toolkit for this purpose, the “morphometric synthesis,” that combines Procrustes shape coordinates with thin-plate spline renderings of their various multivariate statistical comparisons. But because both tools algebraically disarticulate the landmarks in the course of a linear multivariate analysis, they have no access to the actual anatomical information conveyed by the arrangements and adjacencies of the landmark locations and the distinct anatomical components they span. This paper explores a new geometric approach circumventing these fundamental difficulties: an explicit statistical methodology for the simplest nonlinear patterning of these comparisons at their largest scale, their fits by what Sneath (1967) called quadratic trend surfaces. After an initial quadratic regression of target configurations on a template, the proposed method ignores individual shape coordinates completely. Those have been replaced by a close reading of the regression coefficients, accompanied by several new diagrams, of which the most striking is a novel biometric ellipse, the circuit of the trend’s second-order directional derivatives around the data plane. These new trend coordinates, directly visualizable in their own coordinate plane, do not conduce to any of the usual Procrustes or thin-plate summaries. The geometry and algebra of the second-derivative ellipses seem a serviceable first approximation for applications in evo-devo studies and elsewhere. Two examples are offered, one the classic growth data set of Vilmann neurocranial octagons and the other the Marcus group’s data set of midsagittal cranial landmarks over most of the orders of the mammals. Each analysis yields intriguing new findings inaccessible to the current GMM toolkit. A closing discussion suggests a variety of ways by which innovations in this spirit might burst the current straitjacket of Procrustes coordinates and thin-plate splines that together so severely constrain the conversion of landmark locations into biological understanding. This restoration of a quantitative diagrammatic style for reporting effects across regions and gradient directions has the potential to enrich landmark-driven comparisons over either developmental or phylogenetic time. Extension of the paper’s quadratic methods to the next polynomial degree, cubics, probably won’t prove generally useful; but close attention to local deviations from globally fitted quadratic trends, however, might. Ultimately there will have to emerge a methodology of landmark configurations, not merely landmark locations.

Keywords Quadratic morphometric trends · Deformation grids · Geometric morphometrics · The geometry of ellipses · The orders of mammals · Vilmann’s neurocranial octagons · Two-point shape coordinates · Second derivatives

Introduction

The main methodological thrust of this paper can be understood as a companion to both Bookstein (2022) and Bookstein (2023a) that brings the original suggestion of Sneath (1967) back into geometric morphometrics (GMM). Its principal practical suggestion can be distilled down to the pair of diagrams in Fig. 1. They present a

✉ Fred L. Bookstein
fred.bookstein@univie.ac.at

¹ University of Vienna, Vienna, Austria

² University of Washington, Seattle, WA, USA

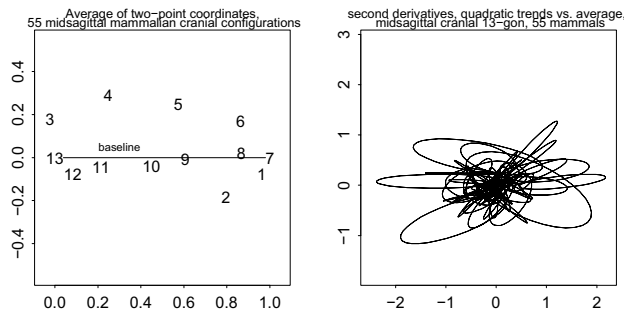


Fig. 1 Principal methodological theme of this paper: a novel ordination of the *quadratic trend descriptor* apposite to the landmark data sets of geometric morphometrics, here as fitted in “[Revisiting a Mammal Cranial Data Set](#)” section to a 13-point midsagittal cranial landmark configuration for each of 55 mammal specimens from 23 orders. Each ellipse traces the second derivative of the quadratic trend fit around the circle of directions from the sample average with respect to a convenient posteroanterior baseline (posterior foramen magnum to premaxilla). (left) The average of the 55 configurations of 13 midsagittal cranial landmarks in this example. The thirteen midsagittal cranial landmarks are as follows: 1, anterior symphysis of mandible; 2, posterior symphysis of mandible; 3, inion sagittal; 4, frontal-parietal sagittal; 5, frontal-nasal sagittal; 6, tip of nasal sagittal; 7, tip of premax sagittal; 8, premax-maxillary sagittal; 9, maxillary-palatine sagittal; 10, posterior palate; 11, basisphenoid-basioccipital; 12, anterior foramen magnum; 13, posterior foramen magnum. (right) All 55 ellipses of directional second derivatives for the 55 quadratic trend fits from the configuration at left. Thus this is a scatter of ellipses, each one a summary of one polynomial regression. Compare Figs. 23 or 34

novel ordination of the *quadratic trend descriptor* intended to help rebuild our current toolkit for GMM’s landmark data sets. The paper’s main empirical contribution, “[A Simple Example: The Vilmann Neurocranial Octagons](#)” and “[Revisiting a Mammal Cranial Data Set](#)” sections, is a new pattern analysis filling a major gap (the representation of spatial gradients) in the current GMM toolkit. This first figure conveys the basic idea of the new ordination method: conversion of two-dimensional landmark configuration data into an explicit representation of just their quadratic trends.

The data set on which the figure is based comprises the 13-point midsagittal subset of a 35-landmark cranial configuration originally exploited in Marcus et al. (2000). By restricting attention to just this unpaired subset, the pedagogy can be managed using purely two-dimensional displays, making dissemination much easier. While the Marcus presentation relied on Procrustes shape coordinates, the registration here, in keeping with the recommendation of Bookstein (2023a), instead uses a two-point coordinate system (Bookstein, 1986, 1991) with baseline from posterior foramen magnum to tip of premaxilla in this sagittal plane. (The two-point approach permits referring any report of a trend model to an explicitly anatomical registration rule.) Each comparison is from the configuration in the left-hand

panel here, the average of these 55 representatives of 23 mammalian orders, to one of the individual configurations.

The conventions of this right-hand panel of Fig. 1, one step in the exemplary analysis of “[Revisiting a Mammal Cranial Data Set](#)” section, are unusual. The reader is probably used to reports of sample variation of landmark configurations or their shapes in the form of scatters of Boas coordinates or shape coordinates like Fig. 19 later on. Where ellipses appear in the textbooks, they are representations of the Gaussian model of a bivariate bell-curve distribution—the locus of some constant Gaussian probability density around a sample mean in two dimensions. The iconics of the right-hand panel of Fig. 1 are different. The figure is genuinely a scatter of ellipses—a total of 55 of them—each of which represents a single specimen by the large-scale gradients of a quadratic trend fit, a second-order polynomial regression on the sample average of their Cartesian coordinates after that two-point registration. And the role of any individual ellipse in this context is unexpected: it does not parametrize a distribution over a sample, but instead the suite of six regression coefficients (r, s, t, u, v, w in the later notation of this paper) summarizing one single specimen at a time. The geometry of these ellipses will be introduced in due course, and then a useful typology that will allow some of them to be directly interpreted as coding particularly simple deformations, namely, bilinear maps. We will see that when these curves are annotated by the corresponding directions of the transects with which they align, transects across the original organismal image, they lead to useful pattern inferences inaccessible from earlier GMM approaches.

Back in Fig. 1, each ellipse there, when appropriately annotated as in later, more detailed diagrams, will convey one of those six-parameter representations of the next geometric term after the uniform: a least-squares fit to a quadratic polynomial explicitly encoding the fitted trend’s second derivatives—the trend for increments to accelerate or decelerate—along every linear transect of the form. Linear multivariate analysis of these fits can proceed either by analysis of their six coefficients or instead by an equivalent eight-coordinate data set, the “cardinal directions” (easterly, northeasterly, northerly, northwesterly) of their ellipses, likewise to be introduced in “[Geometric Fundamentals](#)” section; but nonlinear analyses offer considerable power as well. There result new pattern analyses of these trends per se, the curving of coordinate lines whose graphical power has so intrigued all of us since D’Arcy Thompson.

The ellipses are just part of the report of those quadratic trend analyses. When the data of each 13-gon are fitted as a quadratic trend over that two-point average, the simplest polynomial extension of the conventional approach to a uniform (linear, affine) model (as notated and diagrammed in “[Geometric Fundamentals](#)” section), there result the 55 tableaux sampled here in Fig. 20 through Fig. 22 and Figs. 31 through

33; the entire collection is to be found in the Supplement to Bookstein (2023b). Each specimen's extended diagram has four panels: a Cartesian grid of the fitted trend, a polar grid of the same, a tracing of a half-unit circle deformed quadratically, and, as collected in the right-hand panel of Fig. 1, the quadratic part of each fit visualized as explained in “Geometric Fundamentals” section by the ellipse of its second derivatives in every direction with respect to that two-point baseline.

The theme of visible curving driving these composite graphics was replaced (unfortunately, in my opinion) over the development of today's GMM by an inappropriate surrogate arising instead from linear multivariate analyses (particularly principal-component analyses) of the otherwise disarticulated shape coordinates of the Procrustes approach. An earlier graphical innovation, my thin-plate spline deformation grid, has proved insufficient to restore the missing articulation with the anatomical sciences. (The relation of the new analysis to the thin-plate spline is one major topic of this paper's concluding Discussion.) In their place I foreground these ellipses as a first step in the replacement of the current GMM toolkit by a successor capable of generating hypotheses more closely aligned with the language of organismal anatomy. A praxis for feature analysis of the gradients and other nonlinear aspects of form-comparisons, such as this article sketches, could be a crucial component of the return of GMM to any future state-of-the-art biometric toolkit for either evolutionary or developmental studies.

Beginning with D'Arcy Thompson

The new biometric methodology this paper introduces realizes a very old programme: the production of hypotheses about the causes or consequences of organic form from geometric observations of those forms as represented in line drawings. Such a thrust is over a century old, deriving from the first edition (1917) of the celebrated essay *On Growth and Form* by the polymath D'Arcy W. Thompson. For three-quarters of a century after that initial provocation—right through 1993—the literature of this purpose remained disorganized, a range of approvals and disapprovals every decade or so without much of a consensus. D'Arcy Thompson's much-quoted original example still stands as the clarion announcement of his purpose, and as he was a master of English prose style, it is best to quote him in his own words. From the most readily available edition (1961, abridged by John Tyler Bonner), pp. 275–276 and 300–301:

The deformation of a complicated figure may be a phenomenon easy of comprehension, though the figure itself have to be left unanalyzed and undefined. This process of comparison, recognizing in one form a definite permutation or deformation of another, apart

altogether from a precise and adequate understanding of the original 'type' or standard of comparison, lies within the immediate province of mathematics.... When the morphologist compares one animal with another, point by point or character by character, these are too often the mere outcome of artificial dissection and analysis. Rather is the living body one integral and indivisible whole, in which we cannot find, when we come to look for it, any strict dividing line even between the head and the body, the muscle and the tendon, the sinew and the bone. Characters which we have differentiated insist on integrating themselves again, and aspects of the organism are seen to be conjoined which only our mental analysis had put asunder. The co-ordinate diagram throws into relief the integral solidarity of the organism, and enables us to see how simple a certain kind of correlation is which had been apt to seem a subtle and a complex thing. But if, on the other hand, diverse and dissimilar fishes can be referred as a whole to identical functions of very different coordinate systems, this fact will of itself constitute a proof that variation has proceeded on definite and orderly lines, that a comprehensive 'law of growth' has pervaded the whole structure in its integrity, and that some more or less simple and recognizable system of forces has been in control. It will not only show how real and deep-seated is the phenomenon of 'correlation' in regard to form, but it will also demonstrate the fact that a correlation which had seemed too complex for analysis or comprehension is, in many cases, capable of very simple graphical expression.

And then his most celebrated example, still on t-shirts to this day:

[One DWT figure] is a common, typical *Diodon* or porcupine-fish, and [in another DWT figure] I have deformed its vertical co-ordinates into a system of concentric circles, and its horizontal co-ordinates into a system of curves which, approximately and provisionally, are made to resemble a system of hyperbolas. The old outline, transferred in its integrity to the new network, appears as a manifest representation of the closely allied, but very different looking, sunfish, *Orthogoriscus mola*. This is a particularly instructive case of deformation or transformation.

No, it is not “particularly instructive” in any contemporary sense—for one thing, neither circles nor hyperbolas are among the curves that characterize either the thin-plate splines of the current consensus or the quadratic trends explored in this paper. But Thompson's insistence on a “simple graphical expression” has fired the imagination of many

of us over the century since his argument first appeared, while a like number of counterarguments have appeared to caution the enthusiasm of the same readers. An early counterargument was Karl Przibram's (1923, p. 14), insisting that no comparisons of this sort could be regarded as biologically credible unless and until they could be reproduced repeatedly in an experimental setting (such as the laboratories of his Vienna 'Vivarium,' Müller, 2017). Huxley (1932) promulgated the differential model $\log(y) = a \log(x)$ for distances x and y only to realize that if such a model applied to, e.g., the sides of a rectangle, with different a 's, it didn't apply to the diagonal. Medawar (1945) circumvented the difficulty by diagramming only the sides of rectangles, not their diagonals, in his contribution to a Thompson Festschrift using human body growth as an example (but see also Richards and Kavanagh (1945), the contribution following on his in the same volume). Sokal and Sneath (1963) present several earnest attempts in the spirit of Thompson but end up concluding (pp. 82–83), "No general and simple methods seem yet to have been developed for extracting the factors responsible for such transformations.... It is not easy to see how many separate factors are needed to express more complicated examples" where "not only are the grid lines deformed in several ways, but the deformation is different in different parts of the skull. What would be useful would be a way of extracting the minimum number of factors that would account for the difference in form.... It would probably be at first necessary to mark operationally homologous points [today's 'landmarks'] on the diagrams before feeding them into the computer."

Four years later, this same Peter Sneath published the first attempt to properly quantify the issues here. He adopted a technique then under active exploitation in geology, the method of trend surfaces (now a component of the subdiscipline known as geostatistics) to apply to two variables over the same map, which could be taken as the coordinates of the corresponding points of the image of another organism entirely. Sneath (1967) offers a "partial solution" to the problems set down in his earlier collaboration with Sokal. He claims success in a first goal, "to estimate numerically the overall similarity between two figures, i.e., the gross difference in shape," and proceeds to demonstrate using "sagittal sections of four hominoid skulls." The "overall similarity" he suggests is close to Procrustes distance as the current consensus has it, and the "trend surfaces" he computes for his quadratic and cubic examples are the same as those of this paper. Unfortunately, he pays no attention to the coefficients of those formulas, noting only the displacements at each landmark in turn. Each of the six comparisons among his four specimens is diagrammed and verbalized separately. There is no summary ordination, merely a hope that the sum of squared differences in the coefficients might serve as some sort of inverse index of "phenetic affinity."

This was roughly the state of the art 10 years later when I published my doctoral dissertation (Bookstein, 1978) that, picking up on an alternate theme of the 1940s, attempted to customize a coordinate system for the description of the change per se, not the forms being compared. But the biorthogonal grid method had the same flaw as Sneath's: it applied to only a single transformation—a single pair of forms—at a time.

The "Morphometric Synthesis" of 1993

Shortly afterwards, however, a series of innovations resulted in the "morphometric synthesis" that this paper is now trying to replace. The first of these contributions was my announcement of two-point shape coordinates (Bookstein, 1986), a statistical space that supports applications of conventional tactics like mean difference and regression to the shape of configurations of points by algebraic manipulations of their Cartesian coordinates in a rigorous, theorem-governed way. Just after that came a 1988 conference (Rohlf & Bookstein, 1990), the journal announcement (Bookstein, 1989), and finally the book form (Bookstein, 1991) of the thin-plate spline model for deformation, which explicitly converts landmark-driven shape comparisons to grids over the picture of the organism. Meanwhile Kanti Mardia and John Kent (University of Leeds) had tightly tied a crucial additional multivariate tool, principal components analysis, to coordinate data via the Procrustes shape space Kendall (1984) had announced via the mathematical literature. The fusion of these latter two tools in Jim Rohlf's NTSYS software package was the core of the NATO Advanced Studies Institute of 1993 organized by Leslie Marcus at Il Ciocco, Italy (Marcus et al., 1996). It was there that Rohlf and I announced the "morphometric synthesis," which by then included several other extensions (to semilandmarks, to symmetry). At about the same time, the corresponding rigorous probability theory that had been developed in parallel by Mardia, Kent, and their students Ian Dryden and Colin Goodall was beginning to appear in the formal statistical literature (cf. Kent & Mardia, 1994), a theory soon codified in the graduate textbook *Statistical Shape Analysis* (Dryden & Mardia, 1998, second edition, 2016). The synthesis emphasizes Procrustes shape coordinates over the two-point version because of their much greater suitability for principal components analysis with respect to Procrustes distance.

This serendipitous confluence became the core computational engine for data graphics and ordination of samples across a wide range of biological investigations, particularly in anthropology and paleontology, fields not susceptible to Przibram's challenge of laboratory confirmation. Over time it has become the theme of pedagogy directed at biologists over a wide range of levels of sophistication. Whether course

or workshop, most of these curricula favor the same shared workflow—gather your landmark coordinates, convert to Procrustes shape coordinates by John Gower’s (1975) Generalized Procrustes Analysis, carry out any of the popular linear multivariate analyses there (principal components, canonical variates, multivariate analysis of variance or covariance, partial least squares) on samples (and, more recently, on their phylogenetic contrasts), diagram the analyses by thin-plate spline, and publish.

But over the course of this development Thompson’s original goal, the pursuit of simple explanations hinting at meaningful hypotheses, was subordinated to a reversed logic in which extant hypotheses were “tested” using morphometric arithmetic. Philosophers of science often refer to this pejoratively as the context of confirmation, not discovery. Today, 30 years on, the synthesis is overdue for major revision. The multivariate aspects of this context have been revolutionized in most other sciences by the advent of techniques of machine learning and artificial intelligence, while GMM’s tools have remained pretty much where they were created thirty or more years ago. To put matters bluntly, GMM techniques no longer produce surprising findings any more, findings that lead to unexpected hypotheses about the causes or consequences of organismal form. Thin-plate splines do not often meet Thompson’s criterion of being simpler than the data that drive them. (This point is another theme of my closing Discussion.)

Thus it is time to revisit Thompson’s original goal, the production of interpretable diagrams simpler than the anatomies they compare. Bookstein (2023a) has already noted how the Procrustes method, by prohibiting the investigator from rotating a Thompsonian coordinate grid, interferes with the generation of optimally simple accounts of findings. This paper is the second step in this programme: the construction of an explicit statistical method for the transformation grids that Sneath could already compute, by imitating what the geologists were doing, but did not know how to compare or extend to ordinations of samples. Once the new method is adjoined to an accessible statistical package, our community might discover its strengths and weaknesses quite a bit more rapidly than was the case for Procrustes tools and the thin-plate spline.

Purpose and Contents of This Paper

Like the original announcement of the thin-plate spline deformation (Bookstein, 1989), this paper has two distinct purposes: to teach the mathematics driving the new praxis for ordinating quadratic growth-gradients and other quadratic trends, and also to present a pair of potentially classic examples, one involving growth and the other involving adaptive radiation, in order to hint at the kinds of biometric questions that might now enjoy the possibility of explicitly

geometrical answers along with the diagrams that help convert those answers into biological hypotheses. The outline of the rest of this paper is as follows. “[Geometric Fundamentals](#)” section, which is mostly elementary college geometry, retrieves a simple fact about parabolas that can be made to apply to the quadratic case of the regressions Sneath was already demonstrating half a century ago. It may surprise the reader that the same elliptical shape we’ve used since the 1880s to characterize linear regression has an equally promising role to play in these simplest *nonlinear* regressions, the quadratic trend fits. (I also demonstrate that Thompson made a completely avoidable mistake back in 1917 when he failed to consider polar coordinate grids as well as Cartesian grids for the diagramming of his comparisons.)

“[A Simple Example: The Vilmann Neurocranial Octagon](#)” section shows how the method affords a reanalysis of one aspect of a classic data set (the growth of Vilmann’s neurocranial octagons) in such a way as to lead automatically to a report of its features, a report that turns out to match one of the special cases surveyed in “[Geometric Fundamentals](#)” section. “[Revisiting a Mammal Cranial Data Set](#)” section is a more challenging reanalysis, one well beyond the usual bounds we set on diversity of data sets that will yield morphometric sense: the sample of over 50 mammal skulls originally assembled for just this sort of challenge by Marcus and his colleagues in 2000. The analysis by Marcus et al., in my view, was not fully a success—the finding was only that Procrustes analysis had something to say about mammal phylogeny, but not what that message actually was. Once the quadratic trend is fitted to landmark configurations, the shape coordinates per se are completely ignored in favor of the features of those six-dimensional derived formalisms instead. A closing Discussion, “[Discussion](#)” section, is, I hope, the foreshadowing of a much more nuanced, more demanding use of GMM to generate actual biological understanding. It touches on the deeper question of exactly what quantities are appropriate for minimizing by some empirical biomathematical algorithm, but also on the importance of a pre-existing language of graphical reportage that must be present from the beginning whenever a quantitative method of describing biometric morphology is under development.

Geometric Fundamentals

A Simple Fact About Parabolas

Much of the geometry needed for the new approach to quadratic trend analysis is elementary. A first underlying principle is simple indeed: the midpoints of chords of a fixed span over a parabola all lie at the same distance from the parabola underneath. Writing the parabola as $y = x^2$, this is the identity $((x + 1)^2 + (x - 1)^2) / 2 - x^2 = 1$, which is the same as the

second differences of a parabola are constant

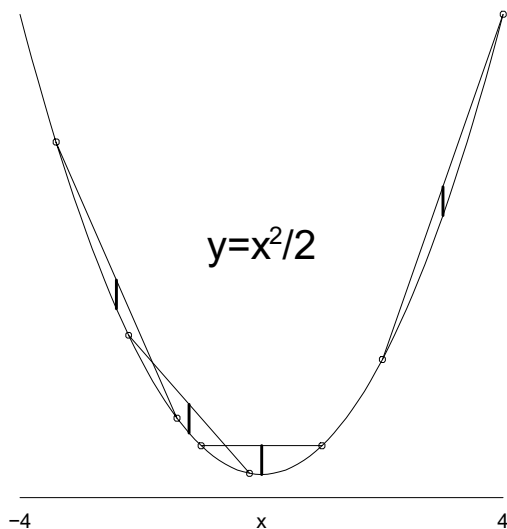


Fig. 2 An ancient identity: the second difference of a parabola is a constant, equalling the second derivative of the curve's formula. The proposition is easily confirmed once both sides are divided by two: invariance of the heavy segments, each the height over the parabola of the midpoint of any chord whose projection on the x -axis is a fixed interval (here, two units)

coefficient of x^2 in the parabola's formula. Figure 2 confirms the identity for several chords all of projected length 2 over the parabola $y = x^2/2$. As you see from the equality of all the heavy vertical segments, the midpoint of each chord between x 's separated by 2 is at height $\frac{1}{2}$ over the curve, the same as the coefficient of x^2 in the formula—half of the second derivative in question, and constant everywhere along the parabola. The identity is more familiar to the mathematician after being multiplied by two: it is the equality of the second *derivative* of the parabola with its *second difference*, the formula $((x+1)^2 - x^2) - (x^2 - (x-1)^2) = 2 = \frac{d^2}{dx^2}(x^2)$.

A fact to be exploited in the sequel is the dependence of these second differences on the scaling of the figure as a whole. If both horizontal and vertical in Fig. 2 are divided by a factor a , the curve that was $y = bx^2$ is now $(ya) = b(xa)^2$, or $y = abx^2$, so that both the second differences and their interpretation as second derivatives are *multiplied* by the factor a . Then to restore the original quantification we need to divide computed contrasts by that same factor a . For two-point shape coordinates, the application to follow, scaling is division by baseline length, and so the correction in connection with Fig. 13 will involve division of computed second differences among the resulting two-point shape coordinates by baseline length as it varies over alternative two-point analyses. The reason for preferring polar to rectangular coordinates in most of this text is that

second differences of a quadratic trend lie on an ellipse

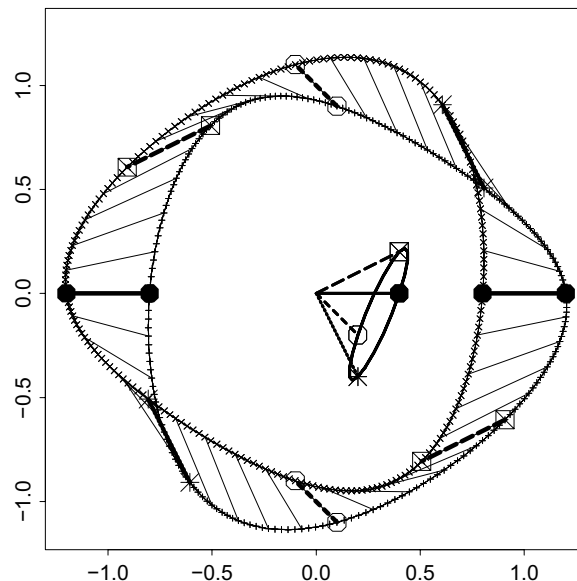


Fig. 3 The same for a two-dimensional (quadratic) trend fit. Symbols for highlighted points along the curves will be introduced in the next figure. Curve of + signs: the effect of the quadratic trend function $Z(x, y)$ on a unit circle around the origin of coordinates. Curve of \times signs: the opposite curve, effect of the function $-Z(-x, y)$. When x and y are replaced by $-x$ and $-y$, only the linear part is altered; the quadratic terms do not change. The segments between corresponding points of Z and its reflection in the origin trace twice around a new ellipse, drawn in the center, that represents the second derivative of the transformation accounting for either of the outline curves. The points of that ellipse can thereby be identified by the directions of those second derivatives (the four symbols here, corresponding to northerly, northeasterly, easterly, and southeasterly transect directions). The center of this ellipse is at the point $(0.3, -0.1)$ corresponding to the values $(r+t, u+w)$ from its formula, the solid black dot is at the point $(2r, 2u) = (0.4, 0.0)$, etc. See text

when we compare locations of deformed points that are originally ends of diameters of one single circle, all the baselines are the same length, so that no corrections for their ratios are necessary.

In Two Dimensions: Ellipses and Their Cardinal Directions

The same proposition, appropriately reinterpreted, turns out to apply to quadratic trends (which, direction by direction, *are* parabolas) in two or any other higher count of dimensions. In every direction, the second *difference* is equal to the second *derivative* in that direction, which, for any quadratic trend (such as the quadratic regression of a target configuration on some template), is constant over the whole picture being mapped. But in higher dimensions another useful property emerges: the value of this second directional derivative lies on an exact ellipse (in

two dimensions; in 3D it will be an ellipsoid). Figure 3 confirms this by a clever choice of coordinate systems borrowed from the morphometric methodology to be introduced presently. Begin with either lopsided oval curve, the one drawn in tiny plus signs or the one drawn in tiny times signs. Either serves as an arbitrary pedagogical choice of an example quadratic trend as rendered (and this step is crucial) by its effect on a unit circle as in the various examples of “A Simple Example: The Vilmann Neurocranial Octagons” and “Revisiting a Mammal Cranial Data Set” sections. Here that curve represents the example

$$(x, y) \rightarrow Z(x, y) \\ = (x + 0.2x^2 - 0.1xy + 0.1y^2, y + 0.0x^2 - 0.3xy - 0.1y^2) \quad (1)$$

of no particular symmetries or other idiosyncrasies and with the identity as its linear term. In this notation, Z is a mapping function applying everywhere in the original digitizing plane. One curve in the figure is the effect of that deformation on what was originally the unit circle before deformation; the other curve is the reflection of that first curve in the origin $(0, 0)$, the replacement of the value Z by $-Z$. Notice that beyond the fixed linear term (x, y) here, the formula includes six decimal coefficients, one each for x^2 , xy , and y^2 for each of the two Cartesian coordinates of a grid point in two dimensions.

The quadratic trend function leaves the origin $(0, 0)$ unchanged, so, copying the formula for the parabola, the second difference across any diameter of the unit circle is just the sum of the values $Z(x, y)$ and $Z(-x, -y)$ in the direction $(x, y) = (\cos \theta, \sin \theta)$ over the full circle of angles θ corresponding to points on the circle. But because all the quadratic terms in the formula are invariant when both x and y are multiplied by -1 , the second difference $Z(x, y) + Z(-x, -y)$ is the same as the simple difference $Z(x, y) - Z'(-x, -y)$ where Z' is the function $-Z$, reflection of the function Z in the origin. In this way Fig. 3 has converted an *average* of deformed points, like the ends of the chords on the parabola in Fig. 2, into the *difference* of a pair of deformed points. These are the short chords drawn as straight lines in the figure.

Figure 3 draws the values of the function Z with tiny + signs and those of Z' with tiny × signs, and their difference is drawn in light segments every 9° and heavy segments at the four cardinal directions (lines at 0° , 45° , 90° , and 135° to the horizontal), identified by the same four symbols that will be assigned to them in later figures of this paper. The final step in constructing this figure transported all of these chords from their positions around the deformed circles to vectors out of the center instead. (As you go around the circle, you see that each of those little chords is encountered twice.) After this translation to the origin $(0, 0)$ of both coordinate systems, these heavy segments are four out of the continuum that traces the second-difference ellipse

that losslessly, indeed redundantly, encodes the equation of the quadratic trend formula (1) driving it. While the curves of + signs and × signs are obviously not ellipses, the curve of their differences in this mixed registration must be, for the same reason that the second differences of an ordinary parabola (Fig. 2) must be constant. Note that each point of the ellipse arises from two diametrically opposite segments of the original circuit (the second differences in opposite directions on the same transect are the same) and that that each point of that inside ellipse is associated with a specific bipolar direction on the starting curve. The four symbols of the legend in Fig. 3 mark four of these directions, the *cardinal directions* at intervals of 45° from the horizontal of the original coordinate system. This is the main reason for using two-point coordinates in this new toolkit: because the directions of the analysis can be read right back onto the original organismal images, numerical aspects of the fitted trends can be directly interpreted as gradients over the organism.

Why does a procedure like this give ellipses for its directional second derivatives? Consider the second derivative of the deformation in Eq. (1) along the direction of the unit vector $(\cos \theta, \sin \theta)$. This will be the second derivative with respect to h of the deformation of points $(h \cos \theta, h \sin \theta)$ along this direction. But because h^2 is a factor of every term in the second-order polynomial of that equation, that second derivative reduces to double the coefficients of h^2 , which make up the vector $2(r \cos^2 \theta + s \cos \theta \sin \theta + t \sin^2 \theta, u \cos^2 \theta + v \cos \theta \sin \theta + w \sin^2 \theta)$ with $r = 0.2, s = -0.1, t = 0.1, u = 0.0, v = -0.3, w = -0.1$. Recall three elementary identities: $\cos^2 \theta + \sin^2 \theta = 1$, $\cos^2 \theta - \sin^2 \theta = \cos 2\theta$, and $\cos \theta \sin \theta = \frac{1}{2} \sin 2\theta$. Using them, we see that as θ rotates that expression for the second derivative is just the deformed point $((r + t) + (r - t) \cos 2\theta + s \sin 2\theta, (u + w) + (u - w) \cos 2\theta + v \sin 2\theta)$, which is a linear deformation of a circle onto an ellipse centered at $(r + t, u + w)$, as the angle 2θ goes around a unit circle twice. And the cardinal directions themselves are linear in these quadratic coefficients r through w : the horizontal cardinal $(\frac{\partial^2 x'}{\partial x^2}, \frac{\partial^2 y'}{\partial x^2})$ is $(2r, 2u)$, the one along the diagonal $(1, 1)/\sqrt{2}$ is $(r + s + t, u + v + w)$, etc.¹

Notice that the points of the ellipse representing the second derivative of the quadratic trend are not aligned out of that ellipse's center with the directions of the second derivative they account for. Because the circuit of transects of the original organismal image must go around the ellipse twice, for cardinal directions that lie at 90° in the data space, like x and y of the coordinate data, the corresponding second derivatives lie at 180° on the ellipse—they are at opposite ends of one of its diameters. That is why the the symbols

¹ The symbol ∂ , read “del,” stands for the partial derivative function, the derivative of a function of two or more variables when only one at a time is allowed to vary.

plotted around the ellipse are needed: they label a few of the points with the diameter (i.e., the transect direction) of the original circle that generated them.

The easiest way to apply the strategy of Fig. 3 to a configuration of landmarks is to follow the advice of Bookstein (2022) by switching from a Cartesian coordinate system to a polar system. This does not affect the computation of polynomial trends, only the graphics by which their changes are represented as gradients. (The question of whether there is *any* reason to prefer Cartesian coordinates in morphometrics is an interesting one, but it falls outside the scope of this paper; for an introduction from the last century, see Bookstein 1981.) Fig. 4 shows how the analysis in Fig. 3 can be attached to a comparison of landmark configurations in order to represent the quadratic trend of interest as an ellipse that can be submitted to explicit multivariate statistical analysis. For this example (and all the others of this section) the template is taken simply as a 3×3 grid of squares one unit on a side, a template that will be deformed by a quadratic formula written out in advance instead of being computed by a regression. The heading presents the coefficients of that quadratic trend in the same order as the example in Fig. 3 or formula (1). Here that formula is

$$(x, y) \rightarrow Z(x, y) \\ = (x + 0.06x^2 + 0.28xy - 0.07y^2, y + 0.24x^2 - 0.03xy + 0.06y^2)$$

as printed over the upper left panel. I cobbled this together to be mainly the xy term of the x -coordinate regression together with the x^2 term of the y -coordinate regression, with a little effect of the other four terms. At above right is the better rendition of this same deformation, now as a polar coordinate grid (Bookstein, 2022). The subtractions that produce the chords of ellipses like the one in Fig. 3 are not chords between landmarks of the data set, such as are commonly encountered in other GMM methods, but instead chords between ends of deformed diameters of any selected circle of the polar system.

Below in the figure is the conversion of the polar plot to the ellipse tracing all of its directional second derivatives. Here they are drawn as sums of pairs of evaluations of Z instead of differences between Z and $-Z$ as in Fig. 3, but the algebra is exactly the same. Around the enlargement of the deformed unit circle from the upper-right panel are vignettes of four of these second differences plotted with the same four different symbols as in Fig. 3. For these little schematics, any radius of a polar coordinate system will do; the delta-shaped outline here corresponds to 1.1, the outermost radius in the upper right panel, and the second derivative along each radial direction is the difference between the sum of the endpoints of that deformed diameter and twice the deformation of the center of the polar system (which here remains at $(0, 0)$).

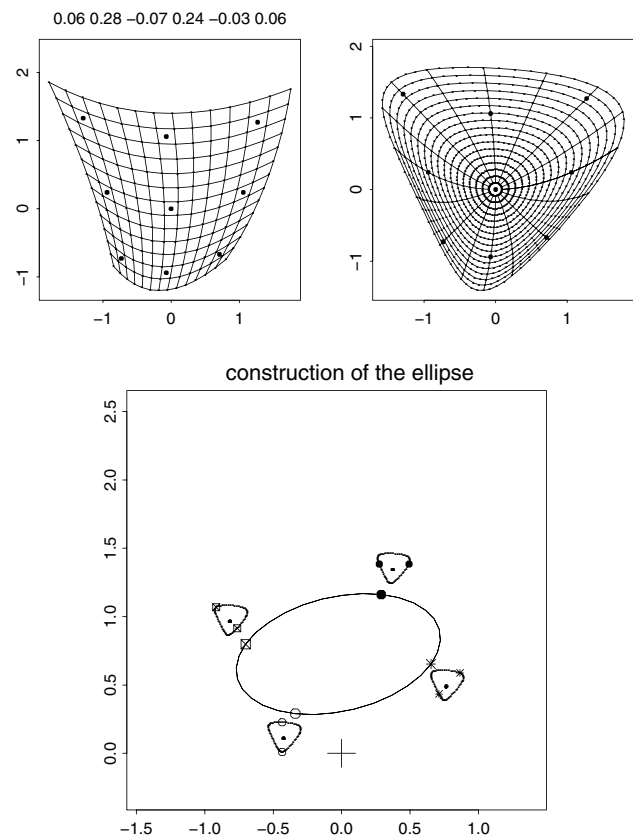


Fig. 4 Schematic of the specimen-by-specimen analysis, here exemplified by a pure quadratic leaving the point $(0, 0)$ fixed at the center of a 3×3 template. (above left) The target is an exact quadratic deformation of the template according to the trend with the coefficients printed above the diagram, represented in Cartesian coordinates over a 3×3 grid of cell size 1. (upper right) The same transformation now diagrammed in polar coordinates up through radius 1.55. (below) The ellipse of second derivatives of that quadratic trend, constructed as the simple sums of points at opposite ends of a deformed diameter of the image of the unit circle as they deviate from the point at the origin, or, equivalently, twice the average of that pair of points with respect to the origin. The + marks the origin of coordinates; note that it is *not* the center of the ellipse. Symbols on the ellipse mark the same cardinal directions as they did in Fig. 3: a solid disk for the horizontal transect, an open circle for the vertical, an asterisk for the transect from northwest to southeast, and a cross in a square for the transect at 90° to that one, southwest to northeast. The figure is to a different scale from Fig. 3 because it uses a polar radius of $1.1\sqrt{2}$ instead of 0.5 so as to embrace the diagonals of the 3×3 landmark template

In other words, each point of the ellipse is the vector difference between the center of the polar system, which here is stationary, and the sum of the two dots terminating the locations after quadratic deformation of the diameter in the direction of interest for a circle of polar coordinates of some convenient radius. For instance, in the little scene at upper center of this lower panel, representing the horizontal cardinal direction, the sum of the two originally horizontal

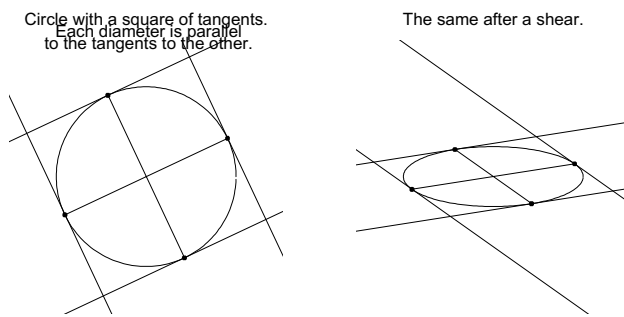


Fig. 5 Conjugate diameters of an ellipse are pairs of diameters that were perpendicular in the circle from which the ellipse emerged as a shear. On the circle, each diameter is parallel to the tangents to the circle at the touching points of the *other* diameter, and this parallelism is unaffected by shear operations. (left) A circle with its circumscribed square. (right) The corresponding configuration after the circle is sheared into an arbitrary ellipse

endpoints (larger dots) after this quadratic deformation falls mainly above the center of the polar system (small dot), leading to the displacement of the big black dot from the plus sign in the larger-scale scene. Again these four highlighted differences are the four *cardinal directions* whose statistics will concern us in the examples of “A Simple Example: The Vilmann Neurocranial Octagons” and “Revisiting a Mammal Cranial Data Set” sections. The figure clarifies the symbols that will be used to differentiate them: a bullet for the horizontal diameter, an open circle for the vertical diameter, an asterisk for the diameter joining southwest and northeast on the template, and a times sign in a box for the diameter at 90° to that one, joining northwest and southeast. Note how the second derivatives in the x and y directions lie at opposite ends of a diameter, and likewise those for the $x + y$ and $x - y$ directions, and that these diameters are what the geometer calls *conjugate*, meaning that each one is parallel to the ellipse’s tangent at the points of the other. (This property is the affine generalization of the situation for perpendicular radii of a circle.) See Fig. 5 and its caption.

Put more abstractly: the six-dimensional feature space of quadratic trend grids of arbitrary landmark configurations over a common template involves the same distribution of derived data as the nominally eight-dimensional feature space of cardinal directions of an ellipse inscribed on the same picture plane. The Euclidean formulas for distance are moderately different in those two spaces, but, obviously, neither one is “incorrect”—both are reasonable.

Trend Formulas with Just One or Two Terms

Because there are six free coefficients in formulas like (1) for the quadratic trends to be examined, it is worth drawing their effects singly and, more importantly, in pairs. Figure 6 uses Cartesian coordinates to show each single term

twice, once with a positive coefficient and once with a negative coefficient. The more realistic Fig. 7 switches to polar coordinates to show all of the possibilities involving *two* of these terms—several of the actual examples to follow in “A Simple Example: The Vilmann Neurocranial Octagons” and “Revisiting a Mammal Cranial Data Set” sections will resemble one of these. The four panels here that look like stacks of circles with shifting centers correspond to projected images of a circular paraboloid, one of the standard quadric surfaces (Hilbert and Cohn-Vossen 1952/1952).

For the analysis of the Vilmann growth process in “A Simple Example: The Vilmann Neurocranial Octagons” section and for some of the mammal examples in “Revisiting a Mammal Cranial Data Set” section, we will need the rendering of the transformations of Fig. 6 by the grid protocol of Fig. 4. The resulting twelve “ellipses,” Fig. 8, are actually line segments through the origin. I mean this literally: if you carry out the construction of Fig. 4 on each frame of Fig. 6, frame by frame you arrive at the corresponding configuration in Fig. 8, where every ellipse has collapsed into a straight line, upon which some pair of cardinal directions turn out to be overprinted at the same point. Each of these lines, then, represents the effect of treating one of the single-coefficient models in Fig. 6 by the algorithm set down graphically in Fig. 4.

At this point we have already arrived at a diagram whose salient features can be used for reports of empirical findings wherever any *two* of the six coefficients of the quadratic trend map dominate the other four in absolute value. Figure 9 supplies such an ellipse for each two-coefficient panel in Fig. 7. (Again the computation is precisely the same as set out explicitly in Fig. 4.) Those that are not points or circles appear as lines oriented at either 0° or 45° to the horizontal and vertical.

“Ellipses” are Sometimes Points or Lines

Let us look a little more closely at Figs. 8 and 9. In Fig. 9, two of the frames display single points (overlaid with all four of the cardinal symbols). These are the frames labelled “ $+x^2.x$ and $+y^2.x$ ” and “ $+x^2.y$ and $+y^2.y$,” meaning, configurations with the coefficients of x^2 and y^2 in one direction equal to each other and all other coefficients zero. Why is this the case? The answer can be phrased either algebraically or geometrically. Geometrically, note in Fig. 8 that the “ellipse” for the regression with single term rx^2 , upper left corner panel, is a line segment with the point for the x -direction at one end, the point for the y -direction at zero, and both points for the xy and $-xy$ directions halfway between. For the analogous regression with just the term ty^2 , lower left corner, the ordination of directions is the same except the solid disk and the open disk are reversed. (The letters r and t here, as

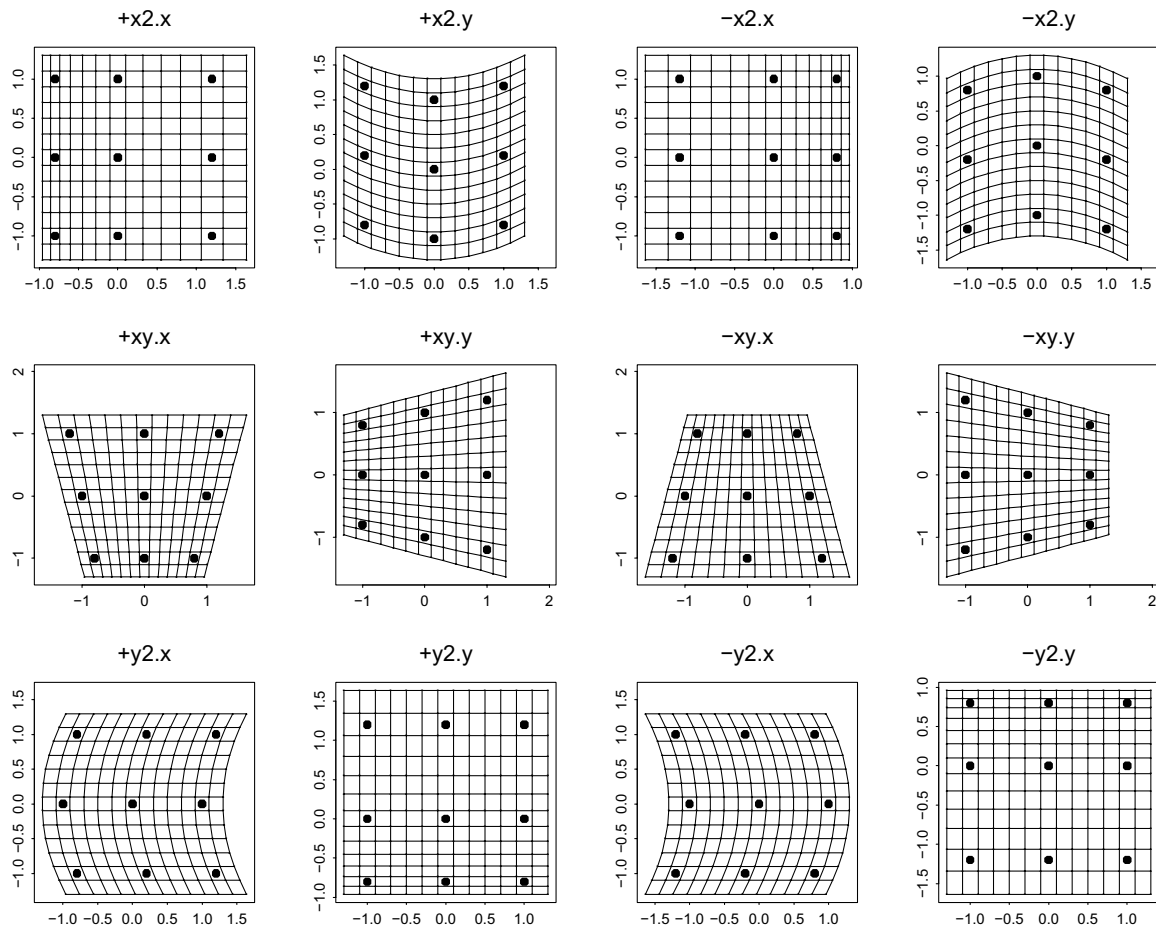


Fig. 6 The six degrees of freedom of a quadratic trend fit, each plotted with both signs over a 3×3 template, in Cartesian coordinates. In the panel labels, “ $x2.x$ ” stands for a regression term rx^2 with $r = 0.2$ for the x -coordinate of a deformation, and likewise $y2.x$ is a term ty^2

in the caption to Fig. 6, correspond to terms in the regression formulas of Eq. (3) to come.) The sum of these two configurations replaces both of the disks by their average and leaves the other two symbols right where they are, but that is the same place as the average of the two disk symbols. Hence after averaging, all four symbols land in the same place, a point on the x -axis but not on the y -axis, as in row 1, column 2 of Fig. 9. Algebraically, the transformation $rx^2 + ty^2$ with $r = t$ sends all of the points $(x, y) = \pm(\cos \theta, \sin \theta)$ to $r(\cos^2 \theta + \sin^2 \theta) \equiv r$, independent of the angle θ .

Likewise one can confirm that combinations like the one with $r = s$ and all other regression coefficients zero, upper left panel of Fig. 9, yield “ellipses” that are lines with both of the second-derivative coefficients for the mixed derivatives at zero and the other two atop one another on the x -axis. There are four such combinations, versus two for the point-type of the previous paragraph. Geometrically, this is the sum of the labelled line segments in the first two rows

for the deformed x ; similarly $x2.y$ and $y2.y$ for ux^2 and wy^2 ; and finally $xy.x$ and $xy.y$ for terms sxy or vxy in one of the two coordinate regressions. Letters here correspond to terms in the regression formulas of Eq. (3) to come

of column one of Fig. 8. Algebraically, these points are the averages of terms $r(\cos^2 \theta \pm \sin \theta \cos \theta)$, where the second term cancels over the \pm operation and the first one simply tracks the squared cosine function over its angular range.

But there is another way to get a straight line “ellipse” in this approach: any of the single-term regressions set down in Fig. 8. There are two different types of this ordination, depending on whether the single term is the mixed expression sxy or vxy (row 2 of the figure) or instead one of the pure squares (row 1 or row 3). In the latter case, one of the disks is at the origin, the other disk is at some remove, and the symbols for the other two cardinal directions overlap halfway between. In the former case, both of the disks are at the origin, and the second derivatives for the $\pm 45^\circ$ directions lie on equal and opposite vectors.

The two pure types of ellipses, points and lines, combine, so that any of the lines in Fig. 8 or Fig. 9 can be shifted to accommodate any point representing equal coefficients r and

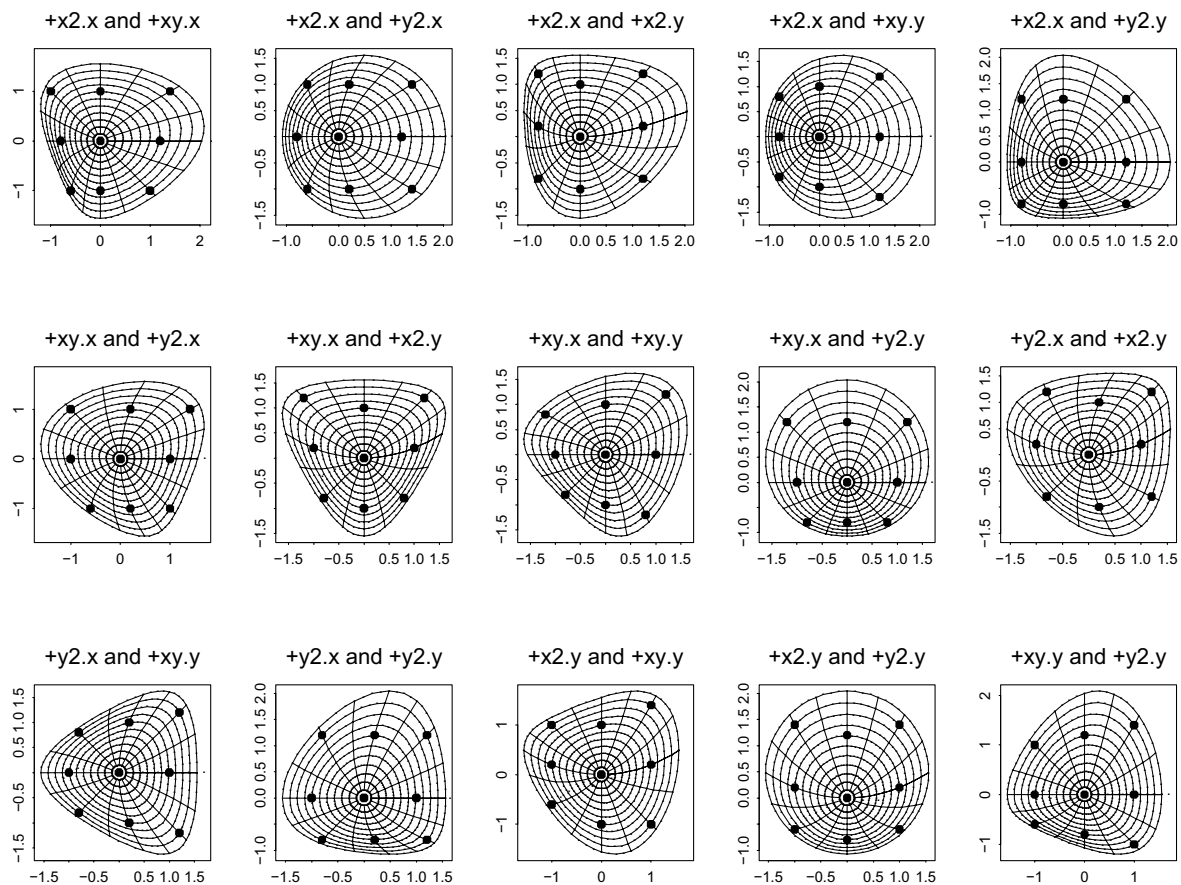


Fig. 7 All pairwise combinations of the separate panels labelled with plus signs in Fig. 6, now plotted more appropriately in the polar coordinate system of Bookstein (2022). Similar figures could be drawn for combinations of +-, -+, or --

t or *u* and *w*. We will see examples of all these intriguing configurations in the next two sections.

For data in 3D, as I mentioned above, the sphere of directions would lead to a surface of second derivatives taking the form of an ellipsoid, not an ellipse, and analysis would proceed using all thirteen cardinal directions (three edge directions, six face diagonals, four body diagonals), not just the four of this presentation.

A Simple Example: The Vilmann Neurocranial Octagons

The taxonomy of examples in “Geometric Fundamentals” section can serve as a typology of ideal types for the understanding of individual examples. This section does so for a familiar textbook data set, the “Vilmann neurocranial octagons” tracing around the midsagittal neurocrania of close-bred laboratory rats radiographed in the 1960s

by the Danish anatomist Henning Vilmann at eight ages between 7 and 150 days and digitized some years later by the New York craniofacial biologist Melvin Moss. This version of the data is the one explored in my textbook of 2018: the subset of 18 animals with complete data (all eight landmarks) at all eight ages. The concern in this section, an extension of the corresponding analysis in Bookstein (2023a), begins with the contrast of the Procrustes-averaged shapes for the age-7 and age-150 animals (only the averages, no consideration of covariances).

Starting from a Regression Instead of from a Model

The quadratic maps in “Geometric Fundamentals” section were all synthetic, in the sense that they illustrated an exactly quadratic correspondence

$$(x, y) \rightarrow (x, y) + (rx^2 + sxy + ty^2, ux^2 + vxy + wy^2) \quad (2)$$

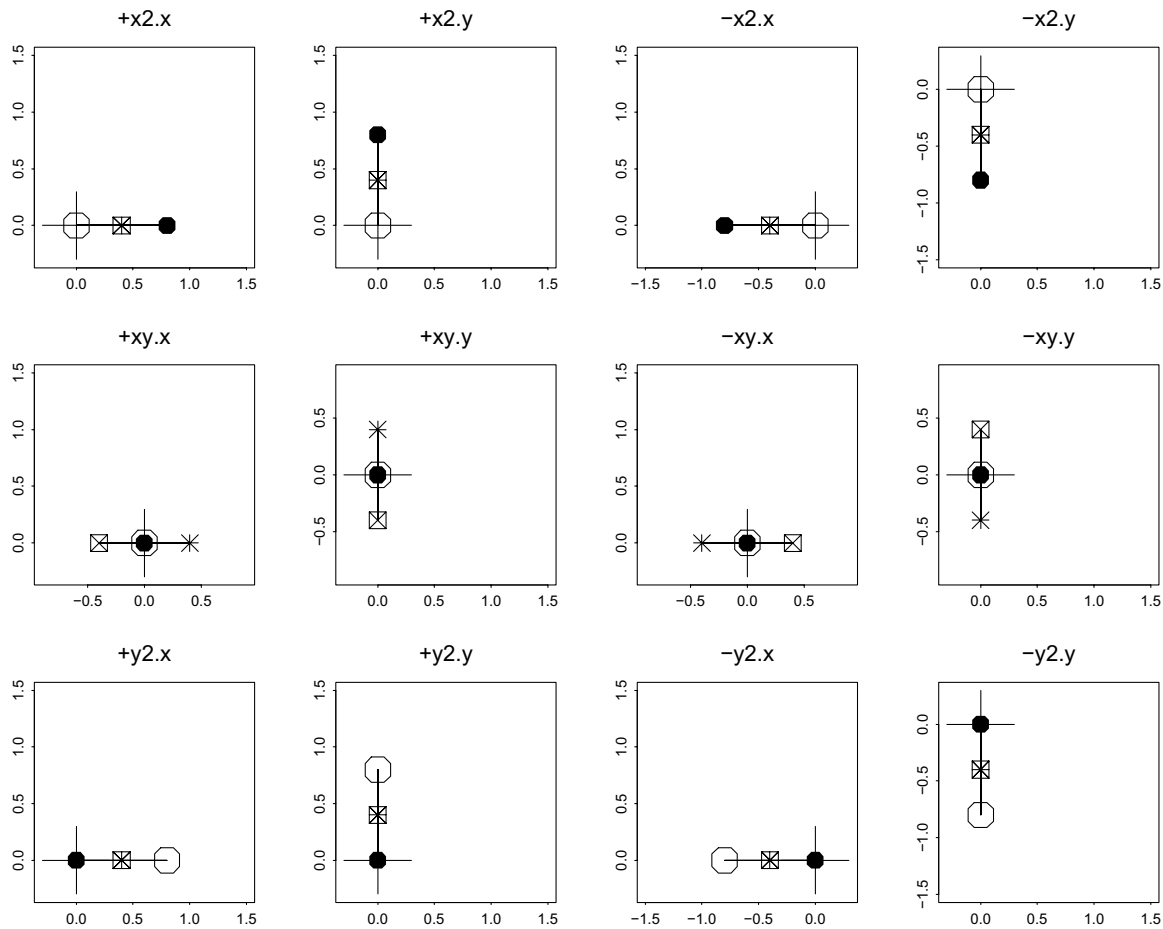


Fig. 8 Single-term prototypes for the quadratic trend ellipses: the second-order derivative analysis for each of the frames in Fig. 6. Plus sign: origin of coordinates. Other symbols show the second deriva-

tives in the four cardinal directions as in Fig. 4, except where some overprinting has been necessary

between two configurations of nine landmarks, one of which was an exactly Cartesian grid. This was the case even when the ultimate visualization, as in Figs. 7 or 9, was not itself Cartesian. We identified these coefficients r through w with half the second derivatives of the resulting mapping, but for any empirical study those coefficients need to be produced by some arithmetical manipulations based in the actual data. As Sneath suggested so long ago, that arithmetic is the standard least-squares analysis that applied statisticians in a great range of different disciplines exploit when it is adjudged sensible to “fit a model by least squares”: they are the coefficients r through w of the more highly parameterized regression model approximating the target configuration (x', y') as an exact polynomial function of the template configuration (x, y) , the formula $(x', y') = (a + bx + cy + rx^2 + sxy + ty^2, d + ex + fy + ux^2 + vxy + wy^2)$.

Our task is to minimize the sum of squares of discrepancies of this predictor with the target locations (x', y') over the template configuration: minimizing

$$\sum_{(x,y)} |(x', y') - (a + bx + cy + rx^2 + sxy + ty^2, d + ex + fy + ux^2 + vxy + wy^2)|^2 \equiv |(x', y') - Q(x, y)| \tag{3}$$

(this will be our definition of Q) in which all twelve coefficients are calculated to minimize the sum of squared lengths of the error term in the complex plane. Each coordinate is then itself the result of an ordinary multiple regression $x' \sim a + bx + cy + rx^2 + sxy + ty^2$, etc. In the examples of “A Simple Example: The Vilmann Neurocranial Octagons” and “Revisiting a Mammal Cranial Data Set” sections we ignore the values of the constants a through c (and likewise the d, e, f that characterize the analogous regression for y'), examining only the coefficients of the quadratic part, which, according to the geometry of “Geometric Fundamentals” section, can be interpreted unambiguously as half the second derivatives of the fitted quadratic trend: at every point of the picture plane, we have $2r = \frac{\partial^2 x'}{\partial x^2}$, $2s = \frac{\partial^2 x'}{\partial x \partial y}$, $2t = \frac{\partial^2 x'}{\partial y^2}$, and similarly for the second partial derivatives u, v, w of y' . The

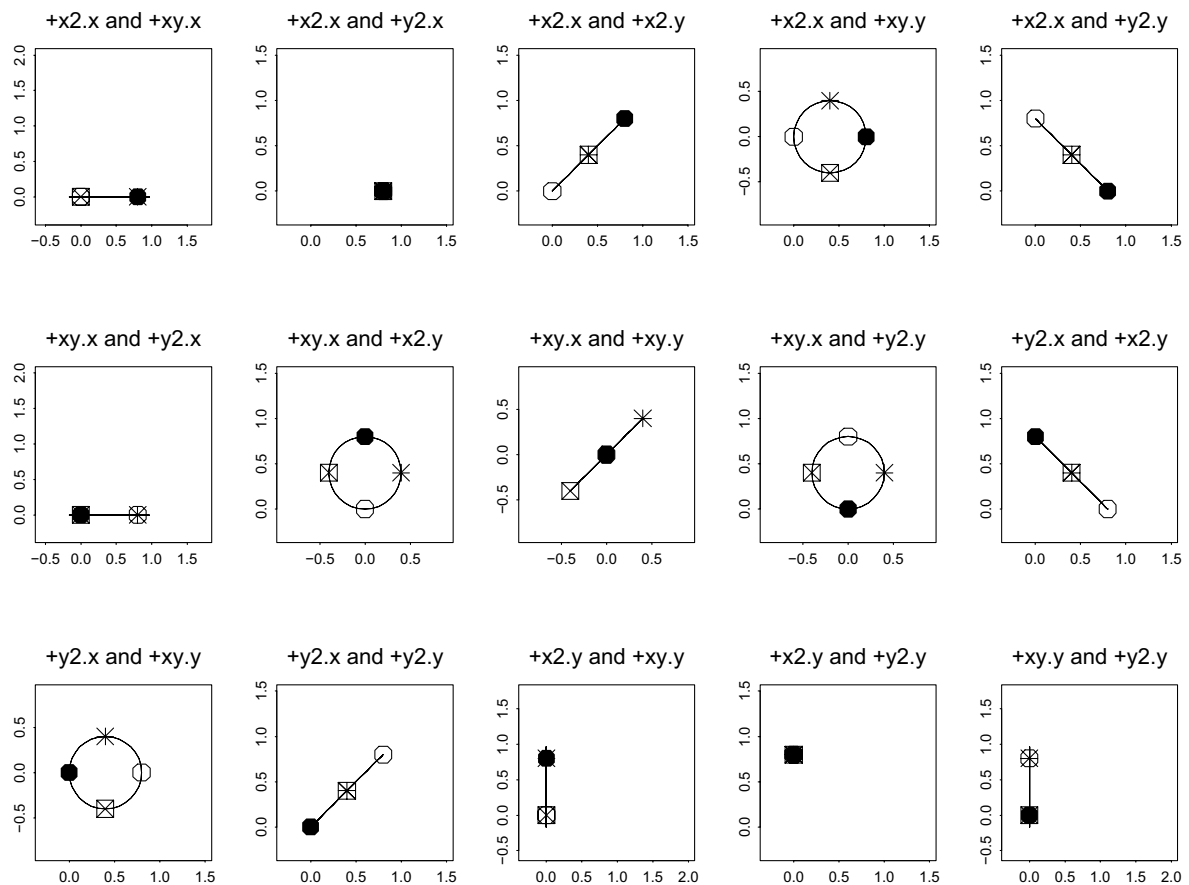


Fig. 9 Ellipses for the transformation grids in Fig. 7 are all either circles, line segments, or (surprisingly) single points. As in Fig. 8, the second directional derivatives in the x -direction and y -direction and their bisectors are plotted with the usual four symbols

calculus of the complex plane allows us to combine these two ordinary multiple regressions into the one quadratic trend analysis in two dimensions minimizing the sum of *both* families of squared errors, the one for x' and the other for y' , because of the Pythagorean mystery that what we perceive as distance on the picture is actually the square root of the sum of these two squared arithmetical differences. (This observation is certainly not original; it is already explicit in Sneath’s paper of 1967, and it lies at the core of my earlier publication (Bookstein, 2023a) on this same theme of polynomial trend analysis.)

Near the end of the Discussion, “Discussion” section, I will return to this convenient equivalence. Until then, it is simply assumed that it makes biological sense to consider the parameters r through w to be sensible quantifications of what the biologist’s eye would already see as one meaningful aspect of a composite characterization of the difference in form of two organisms, each as represented for the purposes of that comparison by a configuration

of finitely many landmarks. When landmarks are closer together, which is not the case for this example, one may think of the quadratic regression as a specialized version of a *smoothing*—a projection of the x - and y -coordinates of every landmark configuration on the five predictors x , y , x^2 , xy , y^2 derivable from the template. It is not the ordinary sort of smoothing of an image, convolution with a Gaussian, but a representation within one shared specifically smoothed subspace of second derivatives all constant.

Rotating Coordinates Helps Interpretation

Under the assumption that landmark configurations can yield meaningful sets of coefficients r through w , Fig. 10 begins at upper left with one version of the conventional analysis of this quadratic trend, the straightforward graphical comparison of the Vilmann octagon averages at age 7 to the quadratic trend prediction at 150 days in a biologically meaningless coordinate system (the principal axes of the

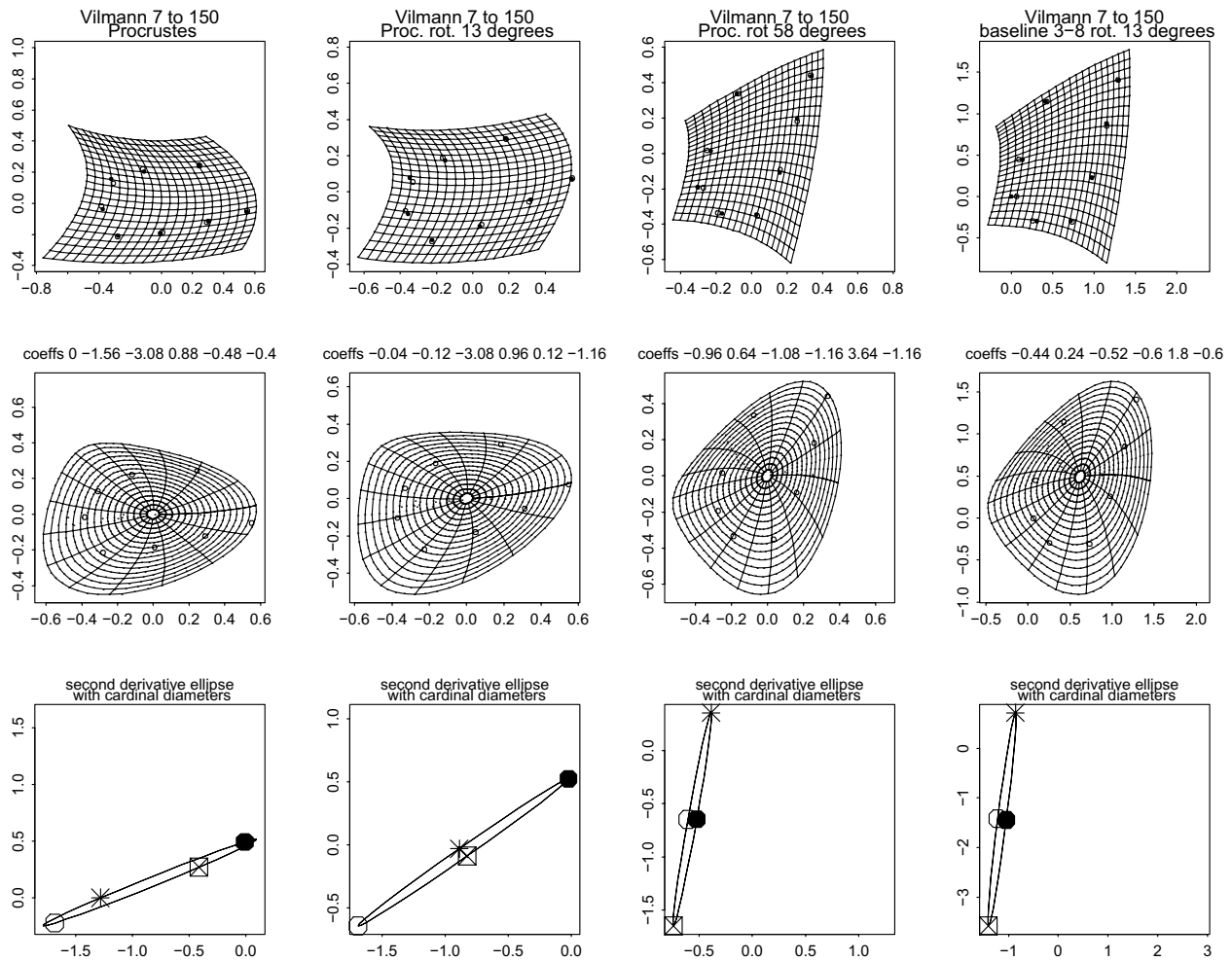


Fig. 10 Analysis of the growth of the Vilmann neurocranial octagons averaged over the usual sample of 18 laboratory rats aged 7 days to 150 days. Columns, left to right: conventional Procrustes pose, rotation by 13° to superimpose the second derivatives along the diagonal directions (1, ±1), rotation by 58° that superimposes the second derivatives $\partial^2/\partial x^2$ and $\partial^2/\partial y^2$ along the coordinate axes instead, and rotation by 13° of a two-point registration (landmark 3 to landmark 8, Interparietal Point to Sphenoöccipital Synchronosis) yielding exactly the same interpretation but without any use of the word “Pro-

crustes” or any of the corresponding formulas or algorithms. Rows, top to bottom: Cartesian representation of the quadratic trend fit, polar-coordinate rendering of the same, and the second-directional-derivative ellipse (always the same shape) with the four cardinal directions highlighted. In the upper two rows, the locations to which the points of the age-7 template are deformed by the grid are plotted in open circles; in the top row, their actual age-150 averages are shown as well by the solid circles

Procrustes average of the two age-specific means). From the quadratic fit (open disks) of the trend’s deformation of the age-7 average configuration to the age-150 average (filled disks) it is clear that the trend method has captured nearly all of the relevant geometric signal here; the problem is rather to state in simple words and coefficients what the meaning of that signal actually *is*. (Analogous grids could have been produced using principal component 1 or the regression of the octagon’s shape coordinates on Centroid Size—see, in general, the range of contexts of this example in Bookstein (2018)—but the thrust in this section is the interpretation of the single comparison of one pair of averaged configurations.) Rotations of these grids have already been examined

in the companion paper to this one (Bookstein, 2023a), but parameterization via their second-derivative ellipses is new here.

Indeed the ellipse here, bottom left in the figure, appears to be close to a special case. It has essentially only one dimension of variation—the minor axis is of length close to zero. We are free to rotate the coordinate system so that that minor axis falls on a meaningful alignment of the coordinate system in which the trend analysis was couched: explicit variation of the orientation of the Cartesian system used to convey the trend. Column 2 of Fig. 10 offers one alternative, a rotation of 13°, for which that null diameter connects the second differences at $\pm 45^\circ$ to the baseline. From the formula

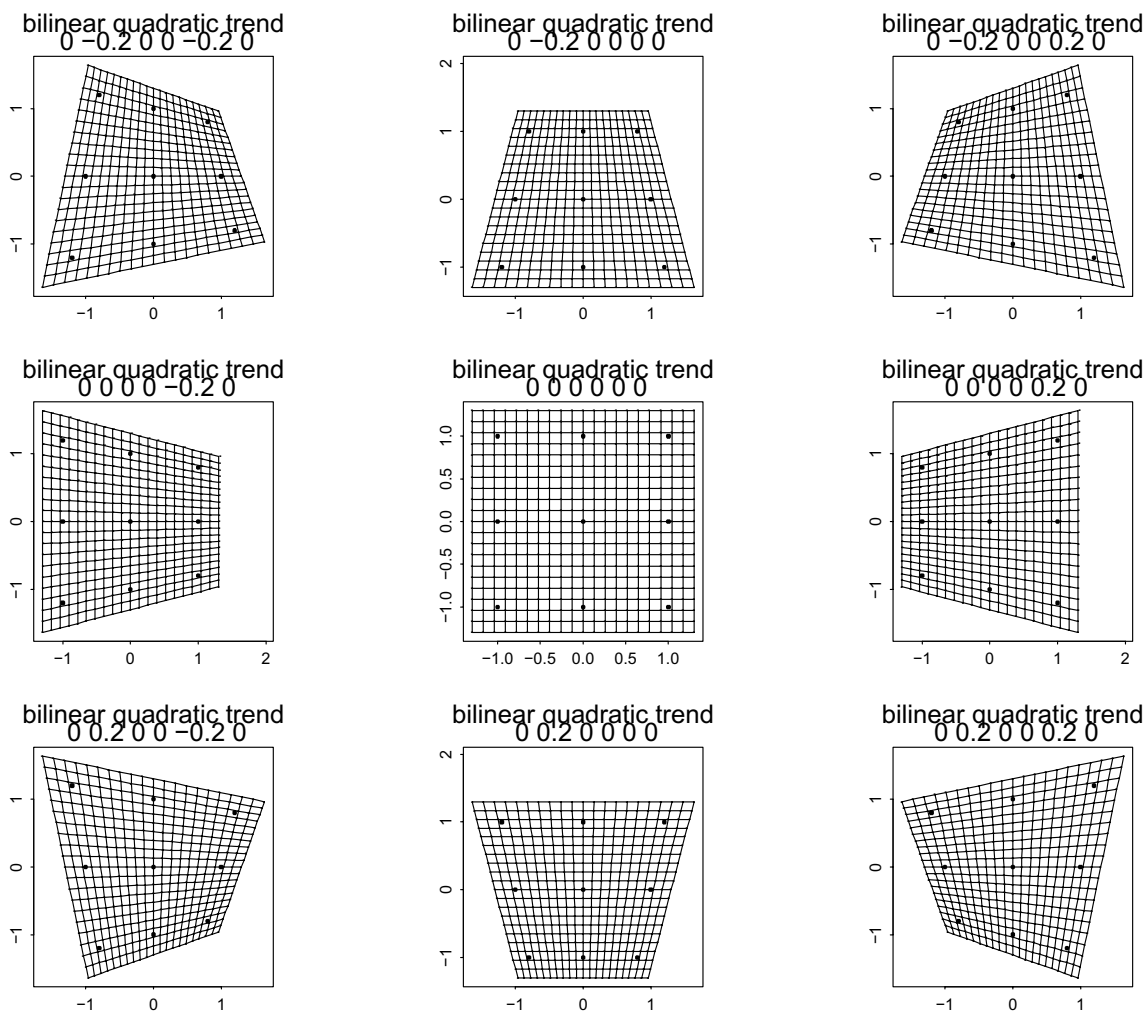


Fig. 11 The *pure bilinear trend* of Bookstein (2023a) is the quadratic trend of this paper with parameter string $(0, \pm 2, 0, 0, \pm 2, 0)$ as in any of the corner instances here

(25.3.26) of Abramowitz and Stegun (1964) it follows that the mixed second-order partial derivative of this quadratic trend is close to 0 for both the x - and the y -coordinates of the target configuration: the map is just a superposition of two processes each looking like any of the diagonal-dominant frames in Fig. 9. (We can detect the dependence on just x^2 and y^2 in the top row, second column, of Fig. 10, where neither system of grid lines is distinguishable from parallel translations of the same parabolic curve.)

That rotation zeroed the mixed partial derivatives. A different rotation, at 45° to that one, will shift the vanishing diagonal of the ellipse from the diagonal canonical direction to the cardinal direction with $\partial^2/\partial x^2$ equal to $\partial^2/\partial y^2$ for both dimensions of the target configuration. In this representation, furthermore, the ellipse has rotated close to orientation with a different, equally salient ideal type: it is nearly aligned with one of the coordinate axes of the plot. And in yet *another* potential special case, the uppermost point of this ellipse,

for the second derivative in the $(1, 1)$ direction, is close to the $(0, 0)$ of this diagram. After rotating 45° to sum-and-difference coordinates, then, we find ourselves close to the situation in the lower right panel of Fig. 9, an “ellipse” that is just a line, horizontal or vertical, anchored near the origin of its coordinate plane.

An appropriate summary of this finding’s dominant feature would thus concentrate on that single mixed derivative. The situation (Fig. 11) is the one I described decades ago (Bookstein, 1985) as the *bilinear map* leaving two families of straight lines straight. (The linear term of this trend fit cannot alter the straightness of those lines, although it may well modify their angle from the 90° characterizing their relation to the highly symmetrical template of a square.) In Bookstein (2023a) the interpretation as a bilinear map was an inference from the grid diagram itself. Here, by contrast, it has been derived *analytically* as an observation about the

near-degeneracy of the ellipse that explicitly represents the coefficients of that same quadratic regression.

The bilinear map has an unexpectedly simple verbal report: opposite boundary segments are transformed linearly and the mapping deforms intersections of proportional transects across the template quadrilateral to intersections of the two new sets of segments connecting proportional aliquots of the target quadrilateral. Hence the map transforms two sets of straight lines on the template (in Fig. 11, a square) into two other sets of lines that are likewise straight (but no longer parallel) on the target image. Most other straight lines map into parabolas. The grid at upper right in Fig. 10 is close to the prototype in the second row, third column of Fig. 11.

Note that the analysis in Fig. 10 involved no thin-plate spline, nor did it rely on any details of the Procrustes analysis driving the configurations of the leftmost three columns in the top row. The finding is unchanged except for a translation and rescaling of the ellipse when, imitating the analysis in Bookstein (2023a), we abandon the Procrustes framework for a two-point (Bookstein coordinate) representation as in the fourth column. The Procrustes procedure per se added nothing to the biological interpretation here, and in fact it seriously interfered with the interpretation of the finding, inasmuch as freedom to rotate the coordinate system of the reporting grid is crucial to understanding the deformation. But how is that rotation to be described? Calling a pose a “58-degree rotation from Procrustes” is not helpful when that Procrustes pose itself bears no biological reference: such a reference position, conventionally aligned with the first principal axis of the landmarks of the template, is not accessible to the biologist’s intuition. In contrast, the figure’s description of the identical pose as 13 degrees from a specific interlandmark segment is a clear instruction. In terms of the prototypes in “Geometric Fundamentals” section, the analysis here is closely aligned with a combination of just two frames: for the vertically extended ellipse, the combination in row 2, column 2 of Fig. 8; for the left shift of all those second derivatives in the x -direction, column 2 of row 1 of Fig. 9.

Effect of Baseline Choice

The analysis in the rightmost column of Fig. 10 rests on a seemingly arbitrary choice of baseline for the two-point construction: the segment from the Interparietal Point to the Sphenoöccipital Synchrondrosis. (This was the ultimate recommendation of the earlier analysis in Bookstein 2023a.) From Fig. 10 we see that the ellipses of interest are invariant in size and shape, but only rotate with the coordinate system. That is reassuring, but it is more important to see the extent to which the analysis is stable against changes in the selection of the pair of points against which the baseline is

ellipses, 10 baselines

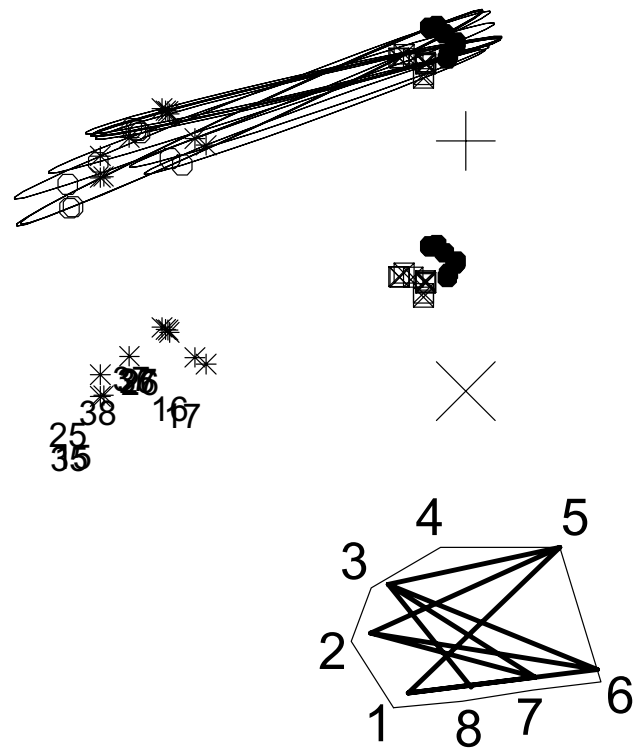


Fig. 12 When rotated back into the original digitizing coordinate system, second-derivative ellipses of the longest baselines align with one another extremely well. (top) The quadratic trend ellipses to the ten selected baselines, with the standard four directional second-derivative symbols from Fig. 4. Big plus sign, (0, 0), the origin of coordinates (all second derivatives zero). (middle) The same with ellipses suppressed and the y -direction second derivative point replaced by its two-digit baseline code. To avoid confusion the big plus sign is replaced by the big \times sign. (bottom) The ten baselines: every join of one of the landmarks 1, 2, 3 to one of the landmarks 5, 6, 7 in the geometry of the average age-7 configuration, along with the 3–8 baseline from Fig. 10

constructed. Figure 12 continues the reassurance by superimposing those second-derivative ellipses for nine more different baselines, as shown in the inset diagram: not only Interparietal to Sphenoöccipital but also every segment linking one of Basion, Opisthion, or Interparietal Point to one of Bregma, Sphenoëthmoid Synchrondrosis, or Intersphenoidal Synchrondrosis.

At the top are drawn all ten of the resulting ellipses after they are rotated into the orientation at right in Fig. 10 and rescaled to accommodate variation of baseline lengths simply by dividing by chord hemilength as explained near the top of “Geometric Fundamentals” section. The plus sign is the origin of this plot, which is even more favorably

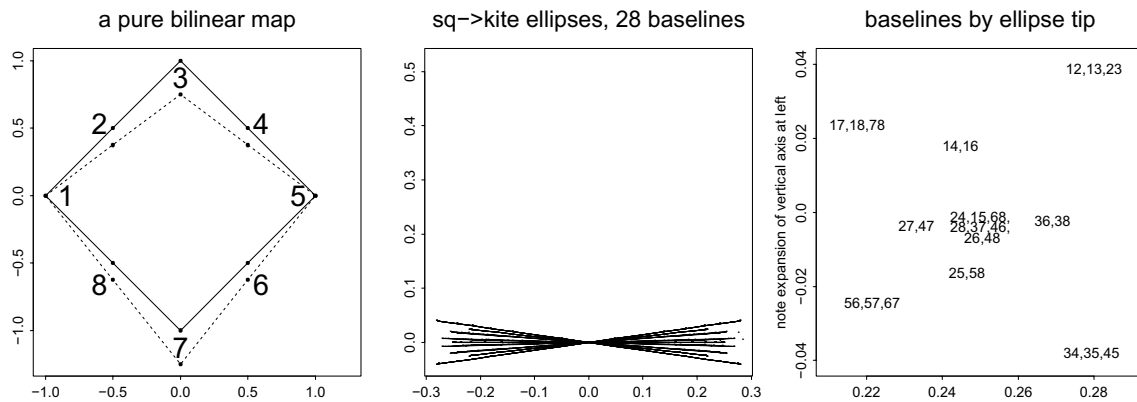


Fig. 13 Geometry of baseline choice for a bilinear deformation. (left) A pure bilinear map, square to kite, with eight landmarks as numbered. (center) The corresponding 28 “ellipses” (in this special case, straight lines), which fall into only nine versions. (right) The nine

placed than the example in Fig. 10, i.e. closer to zeroing the second derivative in the x -direction. In the middle, the same cardinal directions are displayed without the elliptical arcs connecting them, using the usual symbols for all except $\partial^2/\partial y^2$, which is labelled instead by the pair of landmarks serving as the baseline. Clearly the points for $\partial^2/\partial x^2$ are tightly clustered, and also those for the $(x + y)$ direction. Those for the second derivative in the y direction or the $(x - y)$ direction show more scatter on this plot but the scatter would not affect the interpretation of the growth pattern as bilinear. The baseline used at right in Fig. 10 is the one numbered “38” here, which appears near the middle of the distribution of the $\partial^2/\partial y^2$ points of these ten alternatives.

Figure 12 focused on the quadratic Vilmann growth analyses for a selection of longer baselines. But length, which (on an isotropic model) is inverse to digitizing error of the two-point registrations per se, is not quite the correct criterion for this choice even among the options that optimize the grid line representations of that quadratic fit: there needs to be a concern for spatial position as well. Figure 12 itself hints at this when we note that among those longer baseline choices, those along the cranial base, positions “16” and “17” in place of the icon for $(\frac{\partial^2 x'}{\partial y^2}, \frac{\partial^2 y'}{\partial y^2})$, seem to trend differently from “25” and “35” near or along the upper calvarial margin.

To investigate this tendency more clearly, let us turn to a prototype that is precisely quadratic without error. The left panel of Fig. 13 shows this configuration: a template that is exactly square, deforming symmetrically into the shape of a kite as in Figure 13 of Bookstein (2023a). In addition to the four corners I have highlighted four more pairs of landmarks at precisely the midpoints of the edges of either form. (Recall that the bilinear transform

variant analyses, by baseline, plotted as the tip of the ellipse of positive x -coordinate: four triples in an outer trapezoid, four pairs in an inner rhombus, and a core of eight at the correct location, for baselines passing through the centroid. See text

is by definition linear along these edges.) There result $8 \cdot 7/2 = 28$ possible baselines. In the center panel of the figure are the corresponding 28 “ellipses,” each one now a straight line as in Fig. 8. For the purposes of this comparison, as was done previously in connection with Fig. 12, they have all been rotated back to the original coordinate system of the left panel, then rescaled to correct for the division by baseline length per se. After standardization this way, it appears that there are only nine options among these 28 ellipses. Although they align quite well as regards their central tendency, it is their variability, not their trend, that concerns us here.

In the panel at right, which plots the right-hand endpoints of these ellipses, a clear spatial pattern emerges *among the 28 baseline choices themselves*. For each edge of the original template, the three baseline choices supply very nearly the same ellipse tip. These four triples lie in four different locations that together make up a mildly nonrectangular trapezoid with corners corresponding to the 1–7 edge, the 1–3 edge, the 3–5 edge, and the 5–7 edge (clockwise at the corners of this panel). Inside this quadrilateral, and aligned with its *diagonals*, are four pairs of ellipse tips closer to the center that correspond to what chess players would call “knight’s moves” over the template: baselines connecting any corner of the square to the midpoint of one of the edges opposite (e.g., 27 and 47). There remain eight baselines out of the original 28, all of the tips of which cluster very closely right at the center of this scatter, the “correct answer.” Of these eight baselines, four are the actual central diameters of the template (15, 26, 37, 48) while the other four, taking advantage of the symmetries of this particular configuration, lie parallel to one of the template’s diagonals (24 and 68 parallel to 15, 28 and 46 parallel to 37).

Thus, just as longer baselines are less aleatory in the presence of a real quadratic trend, clearly preference should go

to baselines that pass near the centroid of the template per se. In the analysis of the Vilmann growth data, the selected baseline, 3–8, is analogous in its position to 2–8 of Fig. 13; the preferred coordinate system in Fig. 10 rotated this baseline by 13 degrees as explained there. In the mammal skull analysis to come in the next section, the selected baseline is nearly the longest possible choice, and furthermore passes nearly directly over the centroid of the full sample scatter of shape coordinates.

Evidently, when analytic results like these are rotated back into an appropriate digitizing coordinate system, second-derivative ellipses of the longest baselines all representing the same quadratic comparison align quite well. Put another way, the Procrustes rotation per se has no scientific meaning—any requirement that this (or any other) orientation be standardized by some such least-squares procedure as part of an informative GMM dataflow makes no biometric sense. In a better toolkit, it is not rotation of individual specimens that would be standardized, but instead those analyses will be highlighted which, like the quadratic trend ellipses here, do not depend on rotation—for which the rotation does not much affect the arithmetic of findings but so substantially affects the cogency of their reports. It is not that the Procrustes method disagrees with these versions – its ellipse, in column 1 of Fig. 10, agrees with these. But nor did that Procrustes orientation gain us anything over the registration Moss originally applied to Vilmann’s radiographs 40 years ago. We want analyses for which Procrustes orientation, or any other orientation prior to analysis, is *irrelevant* to the reportage. That frees us to explore the grammar of the template coordinate grid per se, which can be a crucial component of a biological interpretation.

For Genuinely Longitudinal Data

Growth changes computed from comparisons of average forms in samples of contrasting age have standard errors of estimate based on standard formulas of multivariate theory. We could have computed such an estimate, for instance, for the comparison in Fig. 10. But for data arising from true longitudinal designs, such as Vilmann’s, we can visualize the variability of these ellipses as actually observed one case at a time. The following example also serves to demonstrate how to interpret the major axis of one of these ellipses when it is not aligned with a cardinal direction, the way it was in Figs. 8, 9 and 10.

Figure 14 displays the quadratic trend ellipses for a selection of eleven out of the possible 28 age-to-age comparisons of the 18 Vilmann neurocranial octagons. In a context of growth analysis there is no purpose to division by any size measure, whether a specific interlandmark distance or the summary Centroid Size. Thus the analysis here, following the recommendation of Bookstein (2022),

uses the raw Cartesian coordinates from the original data archive as published in Bookstein (1991, Appendix 1), centered on Bregma and registered on the direction toward Lambda but not altered in scale from Vilmann’s original neuroontogenograms. So these second-derivative analyses will emerge in a shared physical scale. The originally archived coordinate data were apparently in units of 10μ , so in keeping with the formula for baseline length correction in the main text all my second-derivative computations have been rescaled by a factor of 1000 in order to be expressed in the more intuitive unit of cm^{-2} . The top two rows of the figure show each of the comparisons between successive observations of the same animal, 7 days to 14, 14 days to 21, ..., 90 days to 150; all seven panels are to the same axes. With these conventions, second derivatives for the age-to-age comparisons range from 0 through 3 in absolute value.

Only for the last of these age-to-age plots, age 90 days to 150, does the variation appear to be symmetric and homogeneous about a central tendency (in this example, the zero vector). Every other frame shows obvious deviations from this expectation—ellipses that are outlying in orientation or length or even that completely fail to overlap others of the sample, deviations most striking in the comparison of the age-7 octagons to their age-14 homologues. This particular display, in the upper left panel, suggests that a growth analysis of the sample should eschew any reference to the age-7-to-14 segment, but instead should begin at the later age. Analogously, if the transition from age 90 days to age 150 has mean quadratic component zero (middle row, rightmost panel), these last 60 days of development might well be adding only random error to a longitudinal analysis.

Better perhaps, then, to consider the developmental topic of “growth” to involve an informed choice of starting and ending ages, not simply the full range afforded by the original experimental design. In the lower row of Fig. 14 are four of these alternative end-to-end analyses, starting with either the age-7 configuration or the age-14 and ending at either age 90 or age 150 days. (These bottom four panels are to a doubled range, corresponding to the larger temporal scope of second derivatives, which now range up to just over ± 6 .) Clearly these four alternative “quadratic trends of overall growth” differ in their intrasample variability. The two alternatives beginning at age 7 days, far left and center left panels, include some wildly deviant individual analyses that owe to the obvious inhomogeneity of the corresponding age-7-to-14 analyses of the panel at upper left. So it is a reasonable decision to launch the longitudinal analysis at age 14 days rather than age 7. But also, comparing the center-right and far-right panels of this same lower row, there is clearly more noise in the lengthier of these two longitudinal ranges. If one intended to describe some homogenous morphogenetic process, observations past 90 days seem uninformative.

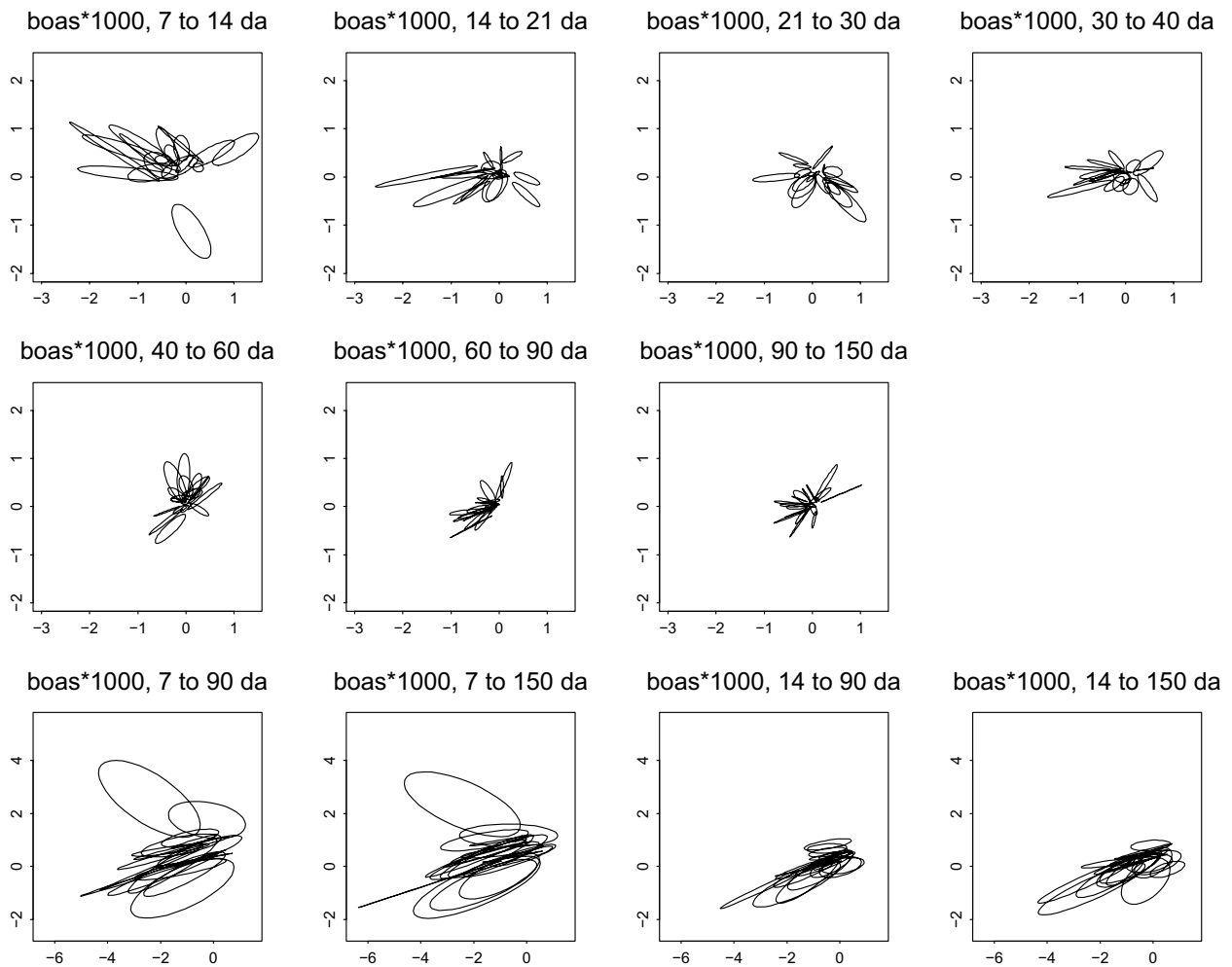


Fig. 14 Quadratic trend ellipses for diverse age-to-age comparisons of the individual specimens of the Vilmann neurosagittal data set, using the original registration by Melvin Moss archived in Bookstein (1991). Top row and middle row, the seven comparisons across

successive observations. Bottom row, four alternative “end-to-end” representations. The panel second from left in the bottom row is the specimen-by-specimen deconstruction of the 18-animal analysis in Fig. 10

Hence out of this series of eight ages of observation, the most informative longitudinal analysis will plausibly be to compare the data from the second observation, age 14 days, to the data from the seventh, age 90.

Figure 15 examines this choice more closely. The upper left panel thus displays this specific pair of age means, still scaled in units of 10μ , but rotated now by 19.3° clockwise so that the corresponding quadratic trend ellipse for the pair of age-specific averages, upper right panel, has its principal axis horizontal. (The original data were registered on Bregma and oriented with Lambda to the left; the rotation approximately corresponds to a baseline from landmark 2 to landmark 5, IPS–Brg, but this is not an analysis of any such two-point coordinates,)

In this upper right panel, the plus sign indicates the (0, 0) of these second-derivative coordinates. It is near one end

of the ellipse here, which lies horizontally, indicating an analysis close to one of the ideal types in Fig. 8.

The 19.3° rotation of the Cartesian system has induced a corresponding rotation in each of the animal-specific quadratic ellipses from Fig. 14 as in Fig. 15’s lower left panel. (Because these individual ellipses are based on regressions on nearly the same template, the trend analysis of the mean is very nearly the same as the mean of the individual trend analyses.) It is clear that these generally align with the mean analysis, upper right panel, but differ in both their vertical position (the second derivative in the vertical direction at upper left) and also the left-hand endpoint of their long axis (to be discussed in connection with Figs. 17 and 18 below). The final panel of this figure, lower right, makes this variability explicit by plotting the projection of all 18 individual ellipses on the axes of the mean analysis. The range

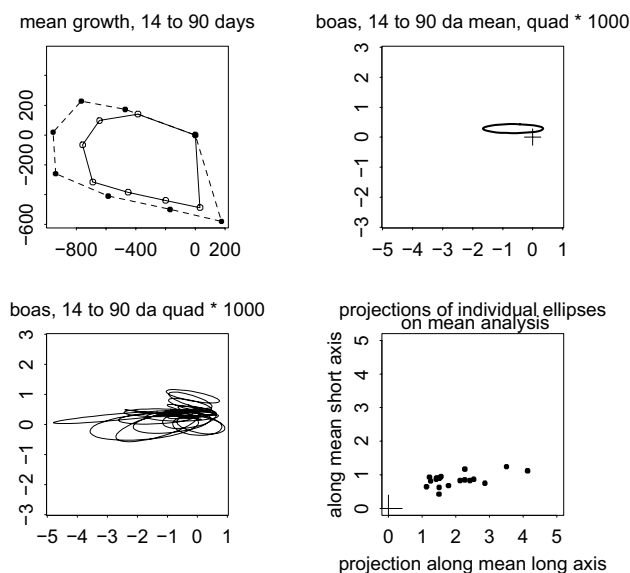


Fig. 15 More detailed analysis of the third alternative in the bottom row of Fig. 14. Upper right, rotation to a horizontal position of the quadratic trend ellipse for the age 14 to age 90 mean configuration. The plus sign locates the (0, 0) of the coordinate system here; points of the ellipse near it correspond to directional transects of the quadratic trend that have very low second derivative. Upper left, equivalent Boas superposition after the corresponding 19.3° clockwise rotation of the raw data. Lower left, the eighteen ellipses for individual animal comparisons after this same rotation. Lower right, projections of each of the individual ellipses on the axes of the mean ellipse above. The plus sign is still at (0, 0) even though the horizontal axis is reversed from the panel above

of variation of this projection along the long axis direction of the quadratic comparison of means is just about fourfold (with the extremes arising from animals 2, 6, and 9). It not only always dominates the variation along the minor axis at 90° to it but also is uncorrelated with that orthogonal component. (This pair of descriptors appears not to be Gaussianly distributed.)

If this neurosagittal octagon persists in its information content across more sophisticated analyses, particularly in three dimensions, it might well be a physiologically valid indicator at the individual level—a potential quantitative growth descriptor suited for studies of experimental challenge. So the quadratic trend analysis has resulted not only in a sample average analysis that greatly simplifies the geometry, but also perhaps a new specimen-by-specimen one-dimensional parameter, the individual animal’s extent of participation in this pooled growth gradient direction. Figure 16 shows the corresponding quadratic fits for all eighteen of these experimental animals individually. The grid for the average, at lower right here, is clearly very close to a bilinear transformation as explained in “A Simple Example: The Vilmann Neurocranial Octagons” section. Most of the individual grids share this bilinearity with the average.

Those that do not are for the specimens of greatest nonlinearity—animals 2, 6, 9, the same that had the extremes of major axis length in Fig. 15.

But we are not yet finished with this more detailed growth analysis—we have not fully reconstructed the annotations implicit in Figs. 4 or 10, the association of each point of the ellipse with its own specific direction upon the actual plane of the organismal data. This built-in correspondence between quadratic ellipses and their associated grid diagrams suggests a further graphical protocol for interpreting the ellipses as legible, reportable features of the fitted quadratic grids. Figure 17 demonstrates this protocol by enhancements of the pair of displays for animal 2, the specimen of greatest left extension of its ellipse in Fig. 15 and greatest apparent rightward bulging of the grid in Fig. 16. As the left panel of Fig. 17 indicates, the left extremum of the ellipse for animal 2 is for the quadratic regression coefficient oriented at 65° to this approximate Opi–Brg baseline. Orthogonal to this, the right extremum, for transects oriented at –25°, involves second derivatives much closer to zero (i.e., first derivatives—forms and orientations of the little grid squares—that change very little along such transects). The left extreme, A, is close to the y-cardinal direction (open circle), meaning that the main feature of this quadratic gradient is close to the value of the second derivative $\frac{\partial^2 x'}{\partial y^2}$, the gradient of the deformed x-coordinate along the rotated y-axis. Over the full sample in Fig. 16, it is the curvature of these originally vertical grid lines that varies most from panel to panel. Grid lines of the other set, those that were horizontal before deformation, remain almost perfectly straight and evenly spaced along the left and right boundary segments of the landmark octagon: the signal of a bilinear transformation as encoded in the near-linearity of this ellipse (and also their sample average, Fig. 15 upper right).

The coefficients r through w (Eq. 3) of the fitted quadratic trend, printed above this ellipse plot, are strongly patterned. The rightmost three, for the second derivatives in the y-direction, can be treated as effectively zero, as can the leftmost one, the -0.67 for $\frac{\partial^2 x'}{\partial x^2}$, for now. The remaining two are nearly equal, so the situation is very close to the prototype already introduced in the second row, first column of Figs. 7 or 9, albeit with minus signs. The structure of the situation is thus summarizable with just three parameters: the coefficients s and t from Eq. (3), plus the angle of rotation from the originally digitized coordinate data.

At the right in the figure is an explicit visualization of these gradients along the two specified transects. (The pair here were drawn through the centroid of the target landmarks, but these quadratic trends are invariant along any parallel transect and along any interval upon such a transect.) Reading from the ellipse’s computed coordinates, direction A has second derivative (–4.79, 0.05)—this is clearly visible

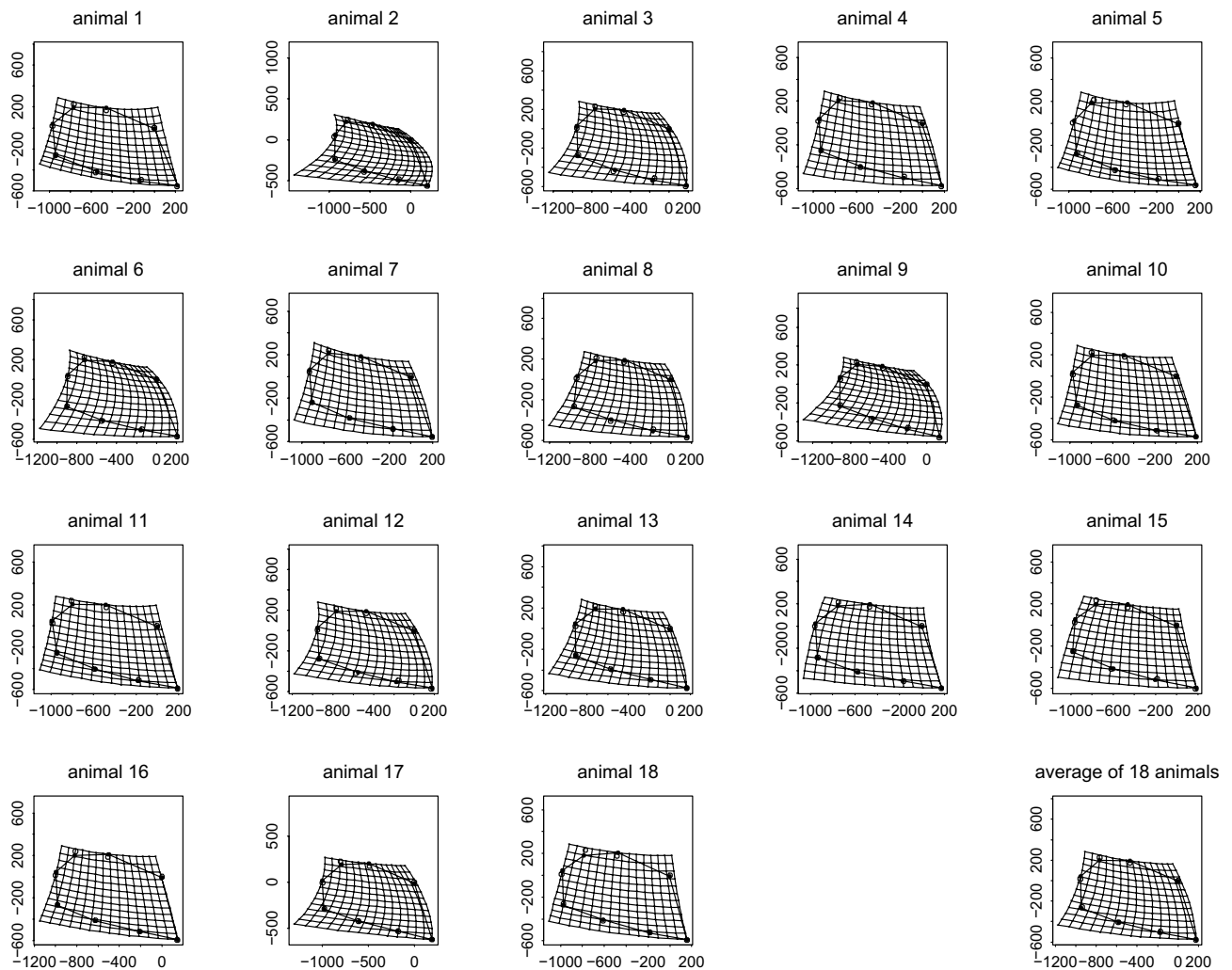


Fig. 16 Quadratic trend grids (linear terms included) for each of the eighteen experimental rodents individually, and, in the final panel, for their Boas average (Fig. 15, upper left)

in the severe reversal of shears of the little grid cells from the lower margin to the upper margin of this grid rectangle. (There is no similar feature in the first partial warp of the corresponding thin-plate spline.) Explicitly, the value of -4.79 is the acceleration (per cm^2) of the x -shift from the starting squared grid along any transect in that 65° direction. Direction B, by contrast, seems uninteresting (as a feature) but may perhaps be interesting as a developmental invariant (as it seems to be common to all eighteen of the experimental animals here—see the next figure).

The interpretation of the grid’s visual features is straightforward in terms of those coefficients printed above the ellipse. We understand the last three immediately as all approximately zero, implying that the originally horizontal grid lines are deformed into what are still nearly straight and parallel lines, with the original horizontal

spacing now slightly attenuated toward the right (that coefficient -0.67 for $\frac{\partial^2 x'}{\partial x^2}$). The largest of the six coefficients, -1.98 , is for $\frac{\partial^2 x'}{\partial y^2}$; it quantifies the bending of the original coordinate lines of constant x into left-opening parabolas. The other substantial coefficient, -1.68 , is for $\frac{\partial^2 x'}{\partial x \partial y}$. It specifies how the slope $\frac{\partial x'}{\partial x}$ along the deformed originally horizontal coordinate lines shrinks (i.e. has a negative $\frac{\partial}{\partial y}$) as their original vertical coordinate rises. This can be easily confirmed, for instance by counting the original grid cells crossed horizontally by the segments IPS–Brg and Bas–ISS here, which are parallel and the same length in this age-90 specimen. They originally traversed 12 and 9 coordinate cells, respectively, and thus grew in a ratio of 3:4 over these 76 days. Thus the whole pattern is

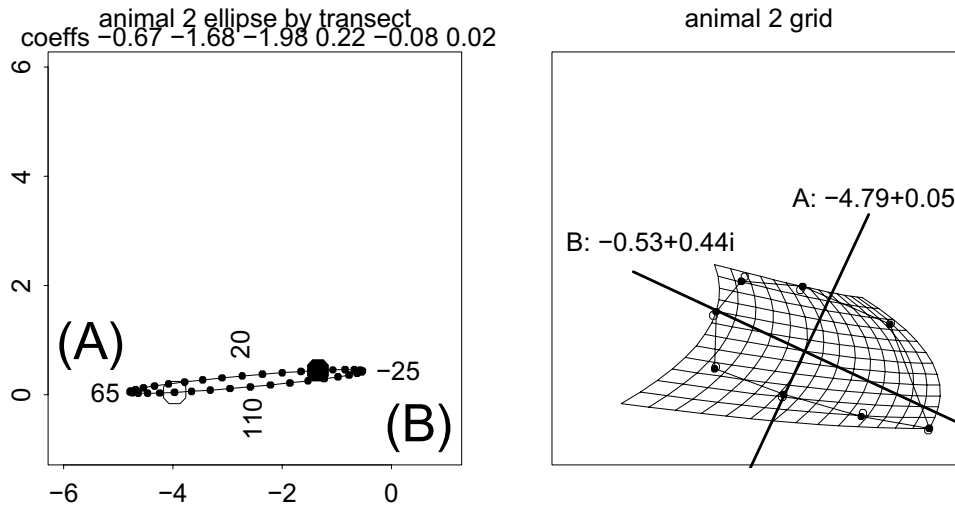


Fig. 17 How to read a quadratic trend ellipse: example of a frequently encountered type. (left) The ellipse for animal 2 (whose grid bulges the most to the right in Fig. 15), with points marked at transect orientations every 5° . The extreme points A, B of the major axis are labelled by the corresponding transect orientations, in degrees counterclockwise (positive) or clockwise (negative) of the horizontal after the rotation in Fig. 15. The second derivative is bipolar, so directions 65° and -115° are the same, likewise -25° and 155° . The large open dot, as usual, is the fitted second derivative vector along the direction of the y-axis after this rotation; the large filled disk is the same along the rotated x-axis. (right) Corresponding extremal gradient directions,

drawn over the quadratic trend grid itself. Transect A has a slightly greater negative second derivative in the direction of the grid's x-axis than the precisely vertical transect indicated by the large open circle near it in the left panel. (In passing, this panel illustrates a point set down in slant type near Fig. 3 of “Geometric Fundamentals” section: the directions of the transects borne by these ellipses are not the same as the directions of the ellipse's points out of its center. Here the direction of greatest second derivative is roughly north by northeast, but the point representing that transect on the ellipse lies at the end of its major axis, which is almost exactly horizontal in relation to the ellipse's center.)

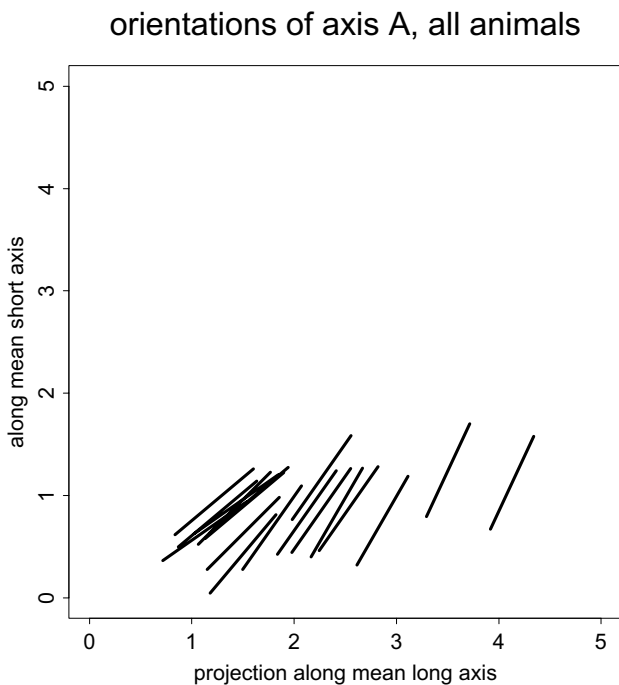


Fig. 18 The analogues of Axis A in Fig. 17 for all eighteen of the Vilmann animals, scattered as edgels on the projection plane from the lower right panel of Fig. 15. The homogeneity of this orientation of maximum second derivative is striking, particularly as the actual value of that maximum is so variable (Fig. 15). In a different study design, it might well be referred to as a character

reportable as the superposition of a quadratic bulge toward the right up the horizontal midline of the octagon (the texture of those vertical parabolas) together with a simple bilinear transformation extending the cranial base by some 33% with respect to the parallel segment at the top of the cranial cavity.

This analysis, while suggestive, pertains to just one of the eighteen Vilmann animals. Comparison of the grid in Fig. 17 to that at lower right in Fig. 16 suggests that the bilinear component is shared by the average even as the extent of that rightward bulging varies greatly over Vilmann's little sample. We would do well to check whether the alignment of these extremes around the circuit of the ellipse's reference directions (recall Fig. 4) persists over the full sample. Figure 18 confirms this hope quite elegantly, by plotting the direction of transect A over each point of the lower right panel in Fig. 15. It is clear that these directions are highly aligned over the entire rightmost two-thirds of the scatter, from ellipse major axis length 2.0 upward; only the six shorter ellipses rotate from this shared orientation, and that by only 10° or 20° . Thus this direction of greatest second derivative appears to be indeed an invariant of the normal growth of these animals, while the direction at 90° seems canalized in this respect. I know of no other morphometric analysis capable of unearthing so subtle a regularity of this classic data set. Analogous finds could well arise from the

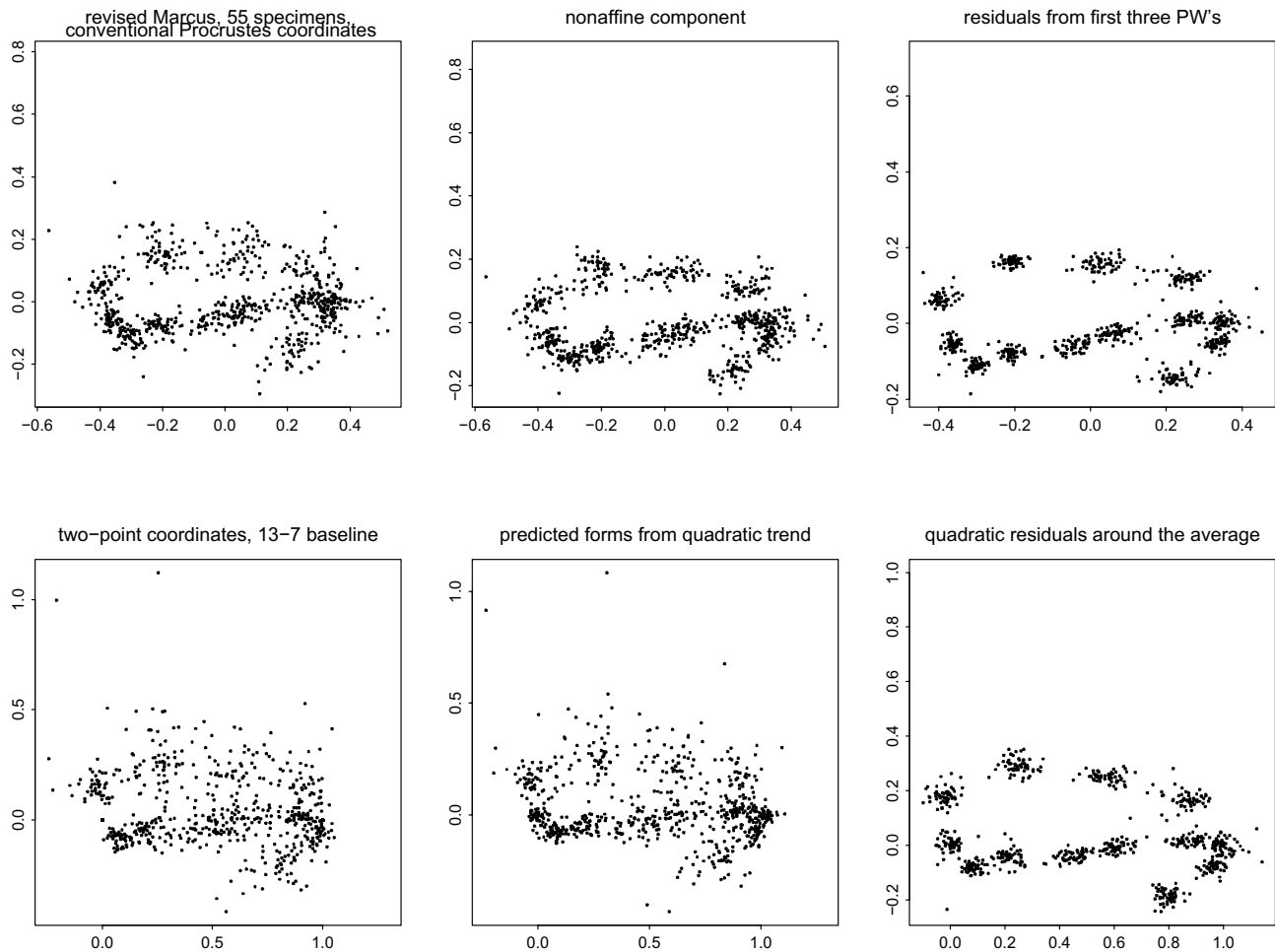


Fig. 19 A selection of scatterplots relevant to the information in Fig. 1 and those following in this section. All panels pertain to the distribution of the 13-landmark configurations for the 55 representatives of mammalian orders as regressed individually upon the template shown in Fig. 1. (upper left) The conventional Procrustes coordinates for these 55 13-landmark configurations. (upper center) The nonaffine component of that Procrustes scatter, adjusted for the linear (affine) aspect of their variation around the average from Fig. 1.

knockout experimental data associated with animal models for human birth defects.

Revisiting a Mammal Cranial Data Set

The data set on which Fig. 1 is based is the 13-point mid-sagittal subset of a 35-landmark cranial configuration that arose as a revision of a data set originally exploited in Marcus et al. (2000). Its 55 specimens, all but the dolphin and the hyena from a representation kindly sent to me by Erika Hingst-Zaher in 2013, omit the deer-pig Babirusa from the 2000 sample but insert specimens of Kangaroo, Man, Sheep, and Boar. (This particular revised data set was previously analyzed in Bookstein (2018, 2019) by a

(upper right) Residuals from the first three partial warps of the conventional Procrustes-spline toolkit. (lower left) The two-point coordinates of the same raw data set to the longitudinal baseline used in Fig. 1 and later figures. (lower center) Predicted forms from the quadratic trend analysis, showing roughly as much variation as the original data at their left. (lower right) Residuals from the quadratic trend analysis of this paper

method different from the approach put forward here. The 2000 article also notes that the selected Elephant skull, along with those of Walrus and Manatee, had to be a young one in order to fit into their digitizing apparatus.) Fig. 19 scatters the landmark data via various analytic displays pertinent to their two-point shape coordinates as in Fig. 1. In the upper right panel, the count of three partial warps the residualizations from which are scattered is set to correspond to the six degrees of freedom for the analysis by quadratic trend in the panel below it. Each partial warp score involves two degrees of freedom because in the GMM formulary it is a complex number. For an introduction to this way of notating the Cartesian plane see pp. 370–371 of Bookstein (2018), or, for a much deeper pedagogical guide, Chapter 2 of Mumford et al. (2006).

A note on notation. When, as in this example, a large sample of target forms is to be regressed on the same template, the regressions can be written out all at once in a compact matrix expression of what the packages call a “multivariable regression.” But that notation, while useful for the programmer, does not simplify the exposition case by case—unlike the situation for principal components analysis, writing out all the shape coordinates of all the cases in one data matrix is not an insightful step toward understanding how their quadratic trends produce ellipses from the circle of directions—and does not substantially aid the task of visualizing the individual polar-coordinate grids and parameterizing the second-order ellipses at the core of the new methodology. Nor does the matrix notation help in the extension mentioned in “Discussion” section to phylogenetic contrasts, where the template would be different at every branch point of the phylogeny. So that alternative notation is not written out here; experienced R coders can construct it straightforwardly from Eq. (3) for themselves. For a similar reason, this paper is not accompanied by the Splus code that generated all the figures, a total of nearly 12,000 inexperienced, unoptimized, uncommented lines. Early adopters of the quadratic-trend method should instead choose to rely on professional programmers for navigating the analytic geometry of all these variants.

The Four-Panel Dashboard

The Vilmann exegeses of “A Simple Example: The Vilmann Neurocranial Octagons” section dealt mainly with one single transformation at a time, as in Fig. 10, or with its within-species variation, Figs. 14 through 18. To extend this approach to multispecies samples it is helpful to have a formal dataflow, as laid out in the dashboards to follow. For the main text I have selected one exemplar (Bear) near the average of these configurations as assessed by the net 22-dimensional Euclidean distance (two coordinates for each of the eleven moveable points) from the sample average in the two-point coordinate system diagrammed at left in Fig. 1, another (Ondatra, muskrat) near one of the extremes of this Euclidean distance, and a third one, Man, of parochial interest to most readers. For the full set of all 55 of these, please consult the Supplement to Bookstein (2023b). To understand the layout of any of the 55, review the three examples to follow here. (I am not claiming that this distance measure, or indeed any distance measure whether or not a sum of squares, makes any sense as a focal numerical quantity in a GMM analysis—only that this particular sorting was a useful way of generating the helpful list of three specimens in the dashboards of Figs. 20 through 22.)

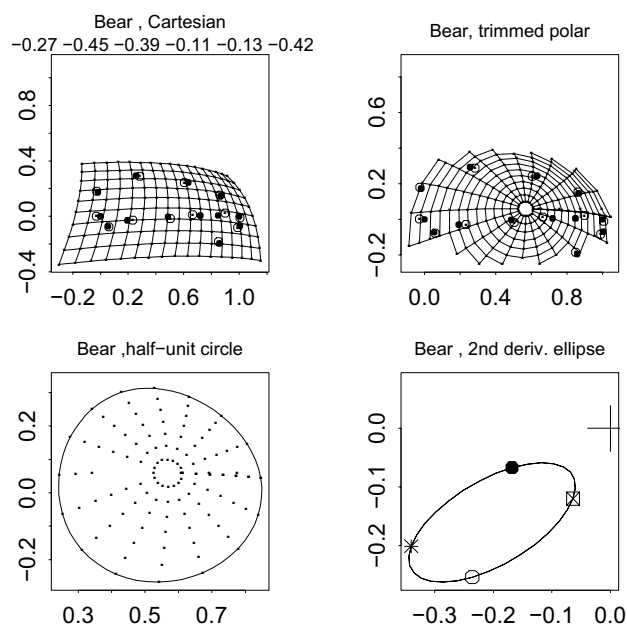


Fig. 20 Dashboard for the quadratic trend fit of the average two-point shape coordinate configuration (Fig. 1) to the Bear configuration, at a summed squared distance of 0.028 (fourth-closest) to the grand average of Fig. 1’s two-point coordinates in this data set. Comparing the axis scales of the lower-right panel here to those of the same panel in the next two figures, we infer that the ellipse here is not far from the prototype of one single point discussed in connection with Fig. 9: not much quadratic warping at all

The key to Figs. 20 through 22 is as follows. At upper left is a conventional Cartesian plot of the quadratic fit to the specimen from the average in Fig. 1, with the coefficients r through w of the formula Q , Eq. (3), printed above. Solid dots, observed shape coordinates; open circles, predictions of the quadratic fit (including linear terms that do not contribute to the second derivative). At upper right the same deformation is rendered in polar coordinates around the centroid of the template. (This grid has been trimmed to avoid large expanses devoid of landmark data.) At lower left is a close-up of the polar deformation normalized to linear term (x, y) (as in Fig. 4) within which is traced the warping from a circle of radius 0.5 (so that the corresponding chord length is 1.0—the second differences are the second derivatives) in the template; the sums of differences of ends of diameters from the center here are the loci that, taken all together, comprise the ellipse at lower right. Cardinal directions of these diametral comparisons are keyed as in Fig. 4. As none of the ellipses here are centered at $(0, 0)$ except by accident, each diagram marks the origin of its coordinates (the purely linear transformation) by a large + sign.

Reporting Ellipses by Their Cardinal Directions

As rendered in Fig. 1 the ellipses of “Geometric Fundamentals” section capture five degrees of freedom (e.g., center coordinate pair, major axis orientation, both axis lengths) of each quadratic fit’s six degrees of freedom (three coefficients for the horizontal shape coordinate, three more for the vertical). What is missing from the five is the angulation of the original Cartesian axes upon the ellipse. Here in Fig. 23, which enhances a subset of the information in the right-hand panel of Fig. 1, nineteen of the 55 ellipses that lay at or near the margin of the superposition there are named with the sixth degree of freedom indicated by symbols for the cardinal directions corresponding to the four directional second derivatives specified in the figure legend, icons already introduced in “Geometric Fundamentals” and “A Simple Example: The Vilmann Neurocranial Octagons” sections. (Given any one of these along with the unlabelled ellipse as in Fig. 1, you have all four by the combination of their properties of central symmetry and conjugacy: directional derivatives along the baseline and at 90° to the baseline lie at opposite ends of a diagonal of their ellipse; the derivatives along the ±45° directions lie on the conjugate diagonal.) For each specimen, as explained in “Geometric Fundamentals” section, these directional derivatives apply across the entirety of the fitted quadratic grid.

Figure 24 conveys the same information less redundantly, by eliminating the elliptical curves entirely, leaving only the 19 pairs of cardinal directions introduced in “Geometric Fundamentals” section. Like the vector of six coordinates introduced above in Eq. (1), the quartets of points in two dimensions bear a total of six degrees of freedom (since the midpoints of the two pairs characterizing each specimen must occupy the same location—a total of two linear constraints). The second derivatives along the baseline (filled circles in Fig. 24) are tightly clustered around (0, 0), not because the average shape of this configuration is highly elongated in this direction but because the observed range of the explicit second derivative vector ($\frac{\partial^2 x'}{\partial x^2}, \frac{\partial^2 y'}{\partial x^2}$) is relatively limited. This is not a function of the choice of baseline direction in Fig. 1—it is a fact about the mammals in the Marcus sample, not an artifact of the two-point method. Second derivatives in the perpendicular direction (open circles), derivatives vertical in this two-point system, are much more widely scattered. We shall see that the ordination of ellipses (trend gradients) here is moderately invariant to the choice of that baseline among reasonable alternatives. Nevertheless the diversity of these ellipses—the diversity of directional second derivatives of a quadratic trend fit unchanging in its template and in its least-squares formulation—is both novel and remarkable, and may be telling us something important about the evolvability of cranial form across Mammalia.

Principal Components of Coefficients; of Cardinal Directions

To explore the diversity of these quadratic trend fits it is useful to begin with the complete display of their six-dimensional space. Figure 25 presents a view in the form of three two-dimensional projections each pairing one of the three coefficients—for x^2 , y^2 , or xy —for both of the Cartesian coordinates of this midsagittal plane. Plainly the lists of the extremal forms on these three panels are distinctive, in fact, nearly nonoverlapping: for x^2 , Baboon, Giant Anteater, Echidna, and Ornithorhynchus (platypus); for y^2 , the two cetaceans and again the Giant Anteater; and for xy the nearly collinear series Hyrax–Sheep–Man facing Beaver at the opposite corner of the scatter. We will shortly see echoes of all these positions in the multivariate analyses to follow.

Figure 26 deals with diverse multivariate statistics of this new ordination by quadratic trend fits. The upper pair of ordinations, trend coefficients versus cardinal directions, are based in exactly the same six-dimensional vector space to two different bases, one ($rstuvw$) from the regression coefficients themselves, Fig. 25, and the other from the redundant eight-coordinate cardinal directions (the full sample of 55, including the subsample of 19 in Fig. 24). The seven-species convex hulls of their scatters are identical as lists and nearly identical as shapes up to an affine transformation. But neither is acceptably in alignment with either of the Procrustes versions below: canonical correlations with the first two dimensions of the cardinal-diameters version are 0.8417 and 0.5779 vis-à-vis the full Procrustes shape coordinate space and 0.9018, 0.3576 with the nonaffine subspace there, which is ostensibly aiming at the same goal. Indeed the principal components (PC’s) of the canonical diameters are badly misaligned with their analogues in the Procrustes nonaffine setting—the correlation matrix of the two versions of the first three PC’s is

$$\begin{pmatrix} 0.091 & 0.613 & -0.609 \\ -0.473 & -0.307 & -0.358 \\ -0.798 & -0.005 & -0.086 \end{pmatrix}.$$

In the score plots, note, for instance, how much greater the separation of Hyrax from its neighbors is in the Procrustes plots compared to its relatively tame location in either quadratic-trend PCA.

Back in Fig. 25, note the cluster of seven taxa at the right in the panel for $\frac{\partial^2}{\partial y^2}$: Pangolin, Aardvark, Lesser and Giant Anteaters, Armadillo, Elephant Shrew, and Man. (We saw the same seven points as the cluster of open circles at the right in Fig. 24.) As Fig. 23 confirmed, six of these seven lie near tips of ellipses that are elongated only in this direction, whereas the seventh ellipse, for Man, differs greatly in its size, in its centering, and in its axis ratio—in humans, gradients parallel to and perpendicular to the baseline do not dominate the quadratic trend description, whereas they are indeed the principal

descriptors of the other six taxa in this cluster. Here in Fig. 26, note this same cluster of seven taxa in the principal-component plot of cardinal directions at upper right—these seven points extreme on PC1 span nearly the entire range of PC2. In the conventional plot of Procrustes shape coordinates, lower left, the cluster substitutes two different taxa for the shrew and Man, and in the scatter for the nonlinear Procrustes subspace (lower right panel), which purports to represent much of the same information as these ellipses do, there is no analogous cluster at all. Clearly the quadratic approach does better than the Procrustes approach of Marcus et al. at adumbrating the visual similarities among this particular six-taxon subset (the mammals with sharp, elongated snouts).

In conventional Procrustes-spline GMM, the results of a multivariate analysis are often diagrammed as a scatterplot amplified by drawings at each end of each axis that interpret the scatterplot's axes as deformations. For scatters more heavy-tailed than bell-shaped, like the ones here, the rendering by endpoints of axes is not as helpful as the more complete rendering by the full circuit of directions over the scatterplot (especially as the species anchoring this circuit around the outside of the hull are so suggestive of the overall variability of the class). In Figs. 27 and 28 the plotted locations of these orienting specimens are identified by the first four or five letters of their long names, as in Table 1. These eight would outline the convex hull of the scatter were it not for the position for Elephant (“Elef”), but as explained in Marcus et al. (2000) this specimen had to be a juvenile in order to fit into their digitizing apparatus. Grids are generated by reconstructing the coefficients $\frac{\partial^2}{\partial x^2}$, etc., of the quadratic trend from the entries for the four cardinal second derivatives in the scatterplots of Fig. 20 ff.² Grids opposite one another are not inverse maps but instead renderings of opposite coefficient vectors in their quadratic trend formula. Figure 27 renders these first two principal components as their effects on the 55-landmark average every 30° out of that average at an arbitrary multiple.

Figure 28 represents the same twelve quadratic trend grids in the less familiar polar coordinate system we have

² To convert a predicted set of cardinal direction endpoints, such as one of these principal component loading vectors, into a quadratic trend, one must produce the values of all six parameters of that trend. The coefficients of x^2 and y^2 for both of the target coordinates are explicit in the locations of the endpoints of the predicted cardinal direction at angles 0° and 90° to the baseline. The remaining two coefficients, pertaining to the mixed partial $\frac{\partial^2}{\partial x \partial y}$ in the two dimensions of the target configuration, can be derived as half the difference of the two directional derivatives in the $(x + y)$ and $(x - y)$ directions (Abramowitz and Stegun 1964, formula 25.3.26). The multiplier here of $\frac{1}{2}$ adjusts the coefficient $\frac{1}{4}$ of their formula 25.3.26 to account for the facts that the radii of the circle of diameters here is only $\frac{1}{\sqrt{2}}$ of the spacing of the grid points in that formula and that the dependence of the estimated derivative on these differences goes as the inverse square of their separation.

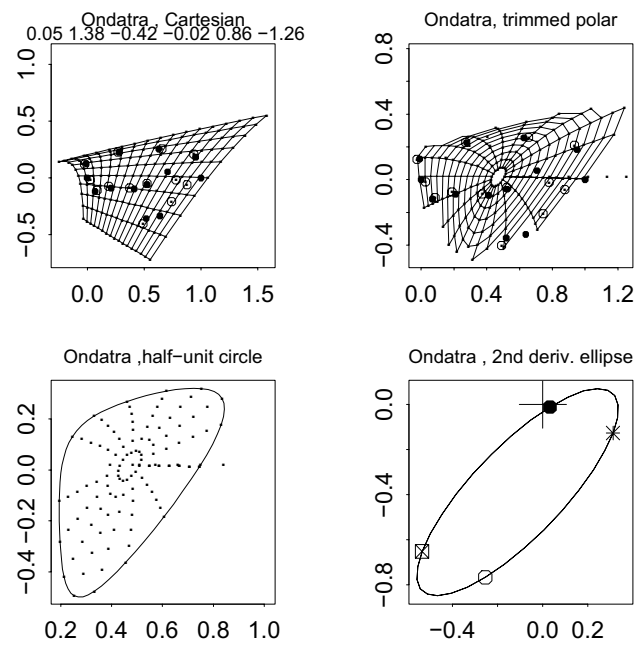


Fig. 21 For the *Ondatra* (muskrat), distance 0.34 (fifth-largest) from the full sample average in the system of Fig. 1 (left). Obviously the closed curve in the panel at lower left (and likewise in the same position in every other diagram of the same design) is not an ellipse, but still the curve in the lower right panel is elliptical, as is the prototype near the center of Fig. 3

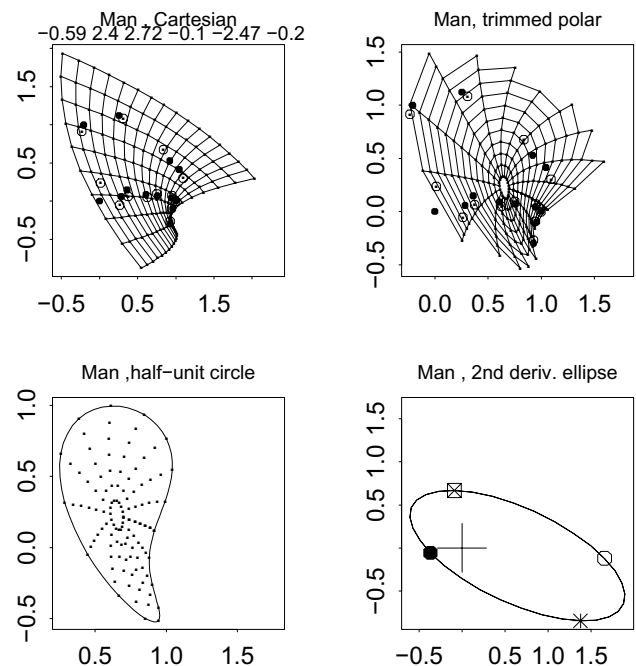


Fig. 22 For the human specimen, farthest from the average of Fig. 1 (left) at sum of squares 1.90, owing mostly to the extreme positions of the inion and the frontal-parietal suture here. The behavior of this second derivative in the vicinity of the mandible is interesting; it will concern us in more detail in Fig. 29 below and in the “teardrop” analysis of “Discussion” section

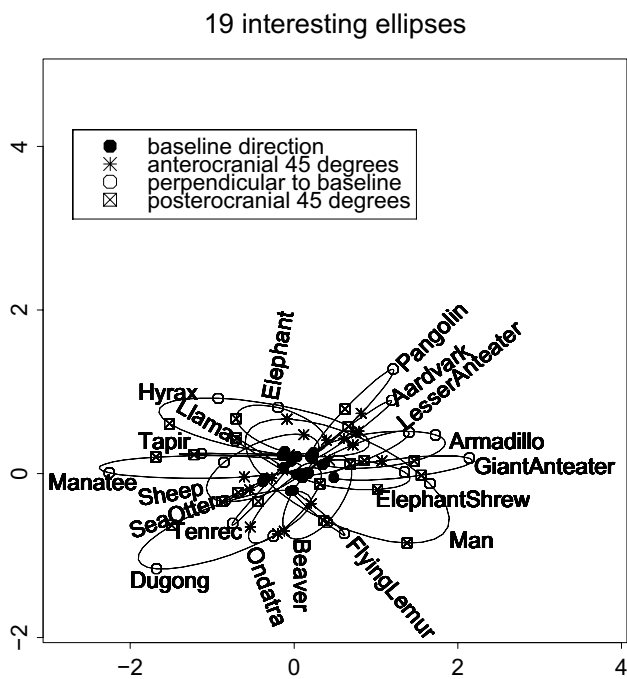


Fig. 23 Nineteen ellipses from Fig. 1 (right) that lie at or near the margin of the distribution there. See text

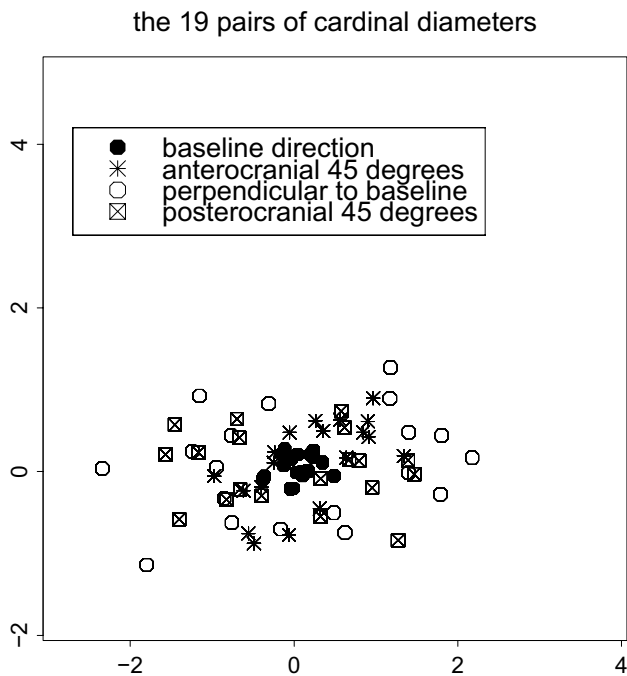


Fig. 24 An even less cluttered rendering: just the cardinal directions from Fig. 23

already encountered in previous figures. What is intuited in Fig. 27 as relative enlargement of the upper left (neurocranial) quadrant of the grid in the direction on which Man

lies, for instance, is seen more cogently here in Fig. 28 as a rotation apart of the relevant *radii*—compare the angulation of “ribs” in this region between the directions of Man and Manatee, for instance. From either of these two figures it is clear that the vertical reorientation of the human face so clear in Fig. 22 is not captured by this first pair of quadratic trend principal components. It may be seen instead in Fig. 29, the analogous plot just for principal component 4 alone, on which Man is a striking outlier. The effect of this component is both to shorten the lower jaw and to straighten the facial angle with respect to the rest of this midsagittal configuration, two characters among the familiar synapomorphies of *Homo sapiens*. At the other end, the substantial negative scores for Beaver and its neighbors Hystrix (porcupine), Capybara, and Ondaatra are consistent with the large residuals at the landmarks of the Ondaatra lower jaw in Fig. 21.

Ellipse Axes, Ellipse Centers

In Figs. 1, 23, or 24 there is strikingly variability in several aspects of these ellipses. The orientation of their cardinal diameters with respect to the Cartesian directions of the two-point registration has concerned us in connection with Fig. 10 of the previous section, but there is additional information in the lengths of the ellipses’ own semiaxes. (A semiaxis is the distance from the ellipse’s center to one of the endpoints of its axes; it is half the axis length per se.) Fig. 30 is an ordinary scatterplot of these two lengths for each of the 55 specimens in this data set. (As an additional quantification one might consider the product of these two distances, which, when multiplied by π , is the area of the ellipse.)

Some observations are clear from the figure. *Homo sapiens* has by far the largest of the minor semiaxis lengths (and also the largest ellipse area). But Hyrax, Manatee, and Dugong all have nearly the same *semimajor* axis length—the range of variation in one specific direction across the midsagittal cranium. These three examples are thus quite directional in their range of second derivatives (cross-cranium gradients of derivative). In the other direction, Baboon appears to have the most isotropic of these distributions—its ellipse of second derivatives is closest to a circle.

At the other extreme, the animal called Tasmanian Wolf (TasmW) here, a thylacine, has the smallest of these ellipses—it is closest to a purely linear transformation of the average configuration in terms of these quadratic trend fits. But several other species along the lower border of Fig. 30—notably Tapir, the primates Sifaka and Gorilla, and also Cheetah, Sea Otter, and Capybara—have ellipses that almost reduce to the flattened lines of Fig. 8. (We saw this pattern already in “A Simple Example: The Vilmann

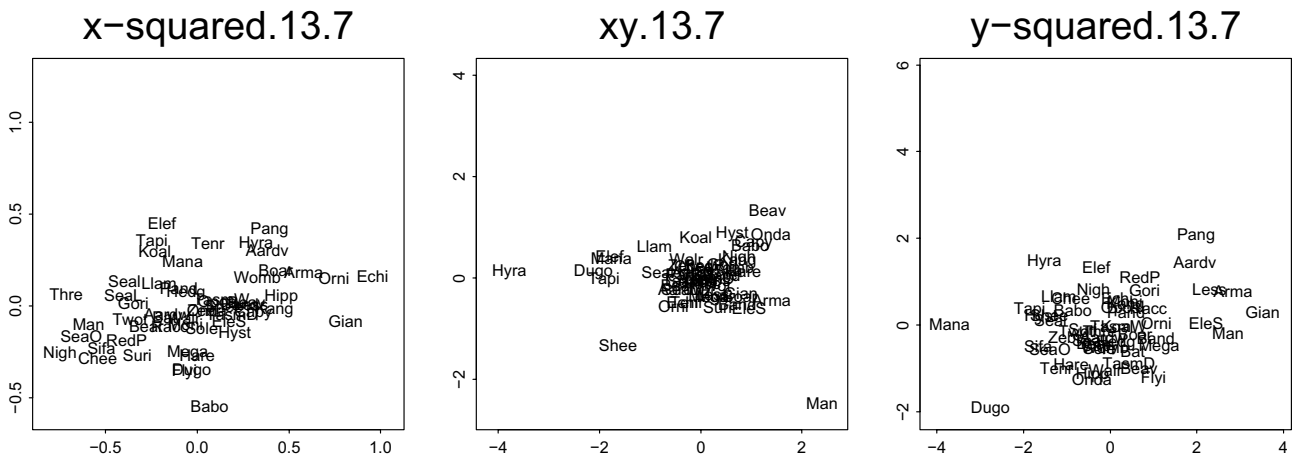


Fig. 25 Scatters of the six dimensions of quadratic trend coefficients in three pairs, one for each of the patterns in the rows of Fig. 6. Specimen short names are as in Table 1: the full text strings except for the two beginning with Aard (five letters each) and the following four:

TasmW, Tasmanian Wolf; TasmD, Tasmanian Devil; Elef: Elephant. EleS: Elephant Shrew. Note the difference in scale between the x-squared panel and the other two, confirming the impression from the subsample in Fig. 23 or Fig. 24

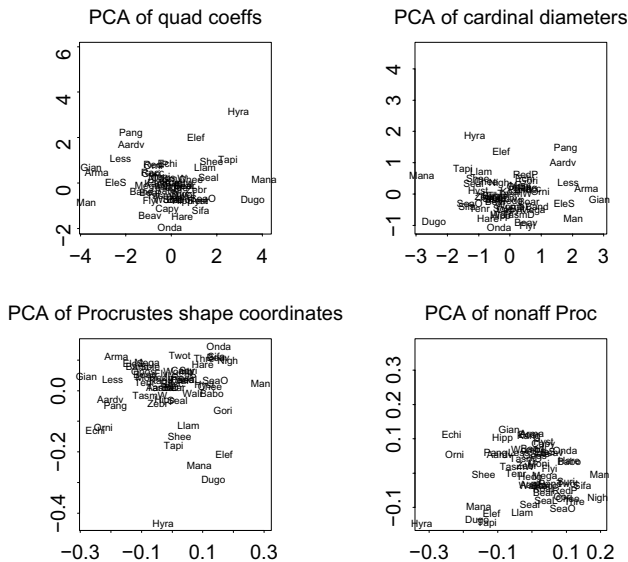


Fig. 26 Four principal component analyses of the 55 × 13 Marcus et al. (2000) data set, 2013 version, plotted by the scatter of their first two scores. The upper left panel shows the first two principal components of the six dimensions laid out in Fig. 25. The principal component analysis at upper right considers all 55 of the quartets of which 19 were displayed in Fig. 24

Neurocranial Octagons” section for the growth analysis of Vilmann’s laboratory rats, a rodent in the same family as the capybara.) The unidimensionality of the ellipses along this border suggests an actual *constraint* of their evolution—e.g., for the tapir there may be considerable developmental reorganization involved in liberating the nose for its preferred herbivore diet. In contrast, the position of Man in

this scatter is consistent with our formal classification as a neotenus species, one whose heterochronies have not yet had “developmental time” to emerge. On the fourth PC axis, Fig. 29, Man is most different from Beaver; in the PC plots of higher explained second-derivative variation, Fig. 26, we are instead most contrasted with Tapir, Hyrax, Manatee, and Dugong. Thus the principal component analysis of Fig. 26 ff. seems quite sensitive in its numerical pattern to the pattern of semiaxis lengths in Fig. 30. Geometrically this is no surprise, as the axis lengths are confounded with the separation of their endpoints, which, whenever they arise as cardinal directions, are within the scope of linear combinations of these reference features.

Figure 31 looks more closely at the specific situation of Tapir here, the rightmost point along the lower margin of Fig. 30. This particularly skinny ellipse is oriented along the horizontal in Fig. 1, the long axis of the template configuration as a whole. The y term in its circuit of second derivatives nearly reduces to a single value (coefficients $u = w$ and $v = 0$) and its x term has two equal coefficients, s and t , which is one of the special cases discussed in connection with Fig. 9 (specifically, the negative of the configuration in row 2, column 1 of Fig. 9). The constancy of the regression $ux^2 + vxy + wy^2$ for the warped y-coordinate implies a constant positive gradient in every direction, hence, the increasing second derivative in x as one passes one’s eye up the page—these grid parabolas become steadily sharper with height. The dominance of the open disk and its neighbor in the lower-right panel corresponds to a principal direction of increasing separation aligned halfway between their directions—roughly the direction at 20° counterclockwise of horizontal. This effect is *negative*, meaning that the spacing of verticals is closer toward the right side of the grid (positive

Table 1 Table of specimens

Long name	Plotting icon
Aardvark	Aardv
Aardwolf	Aardw
Armadillo	Arma
Baboon	Babo
Bandicoot	Band
Bat	Bat
Bear	Bear
Beaver	Beav
Boar	Boar
Capybara	Capy
Cheetah	Chee
Dugong	Dugo
Echidna	Echi
Elephant Shrew	EleS
Elephant	Elef
Flying Lemur	Flyi
Giant Anteater	Gian
Gorilla	Gori
Hare	Hare
Hedgehog	Hedg
Hippopotamus	Hipp
Hyrax	Hyra
Hystrix	Hyst
Kangaroo	Kang
Koala	Koal
Lesser Anteater	Less
Llama	Llam
Man	Man
Manatee	Mana
Megachiroptera	Mega
Monito	Moni
Nightmonkey	Nigh
Ondatra	Onda
Opossum	Opos
Ornithorynchus	Orni
Panda	Pand
Pangolin	Pang
Raccoon	Racc
Red Panda	RedP
Sea Lion	SeaL
Sea Otter	SeaO
Seal	Seal
Sheep	Shee
Sifaka	Sifa
Solenodon	Sole
Suricata	Suri
Tapir	Tapi
Tasmanian Devil	TasmD
Tasmanian Wolf	TasmW
Tenrec	Tenr

Table 1 (continued)

Long name	Plotting icon
Threetoed Sloth	Thre
Twotoed Sloth	Twot
Walrus	Walr
Wombat	Womb
Zebra	Zebr

Specimen names are as tabulated in Marcus et al. (2000) with exceptions as noted in the text

PC1 versus PC2 of the cardinal diameters

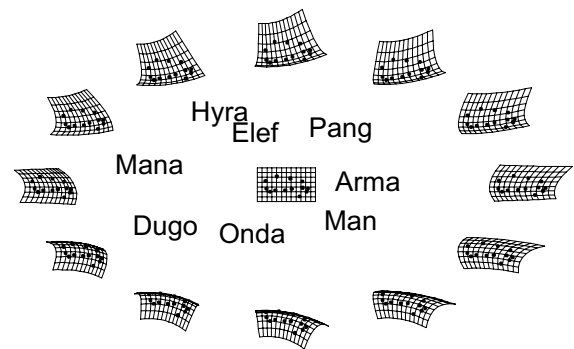


Fig. 27 Interpretation of the axes of the upper right scatter in Fig. 26 by extrapolations of the pertinent quadratic trend grids (in Cartesian format) every 30°. The square grid at center bears the landmarks of the 55-specimen average from Fig. 1. Affine terms are omitted

PC1 versus PC2 of the cardinal diameters

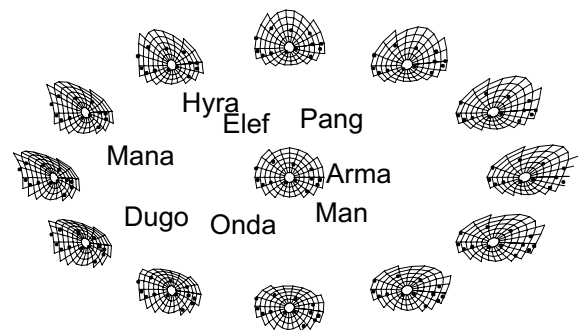


Fig. 28 The same with the axes interpreted as trimmed polar coordinates instead, after the fashion of Fig. 20 ff. Again affine terms have been omitted

x's) than the left, principally in the direction aligned with the longer diagonal of the deformed rectangles projecting from the form as in the panel at lower left. In the polar plot, upper right panel, this gradient of radial spacing is particularly clear. That the roof of the calva is shortened is obvious, but that could have been accomplished in any of several

geometrically different ways insofar as the midpoint of that reduced calva stayed fixed over the cranial base, shifted anteriorly, or shifted posteriorly. Here it apparently shifts mainly posteriorly, as one clearly sees in the polar grid plot at upper right.

The simplicity of this analysis strongly suggests that one search for a simple “cause”—a simple functional gradient accounting for the transformation of the template (a plausible ancestral state) with the constraints $s = t, u = w, r = v = 0$ of the regressions here. It affords an interesting contrast with the corresponding analysis for an entirely different animal, the manatee, Fig. 32. There are strong similarities here—juxtaposition of the 45° and 90° pairs, horizontality of the ellipse in the coordinate system of Fig. 1. Notwithstanding that the rank-order of the six regression coefficients r through w is different, the overall impression of this quadratic trend is remarkably similar, though more intense in the sea creature. That “teardrop” shape of the half-unit circle plot, lower left panel in both of these figures, may well be a novel character parameterizing the feasibility of extreme forms across a wide subset of the mammalian radiation. (Think of the teardrop as the combination of suppressed angular spacing along with enhanced radial spacing in the same polar sector—concentration of an intensified second derivative of a trend coefficient in one relatively narrow direction together with a suppression in the perpendicular direction.) The temptation to refer to this situation as a canalization should perhaps not be resisted.

Either Tapir or Manatee, similarly extreme according to the quadratic-trend feature space, can be profitably compared to a contrasting situation, a rotation at 45° of those cardinal directions. In the upper left panel of Fig. 26 the PC1–PC2 score for the elephant (“Elef”) is relatively isolated along the vertical (PC2) axis, whereas the point for Manatee is where one expects it (far out along the PC1 axis). This divergence of positions corresponds to the appearance of the polar-coordinate panels at upper right in Figs. 31 versus 33 below. For the manatee, there is a sharp convergence of polar radii along a northwesterly direction; for the elephant, along a northerly direction instead. The elephant’s half-unit circle plot, lower left in Fig. 33, shows a version of the teardrop pointed vertically rather than to the northwest, corresponding to the rotation of its cardinal directions around the ellipse at lower right. (There has also been an elevation of this ellipse above the horizontal axis of its panel, corresponding to the positive second derivative of the spacing of the horizontal curves in the upper left panel.) In keeping with Sneath’s vision of factors it would be reasonable to search embryologically for some putatively one-dimensional form-factor that accounts fairly simply for this shape change, in spite of its extremely large magnitude, as a plausible two-parameter change of the quadratic trend coefficients. Teardrops may indeed be one good candidate for the rhetoric of form-factors that Sneath hoped for in his 1967 article. I return to this possibility in the Discussion.

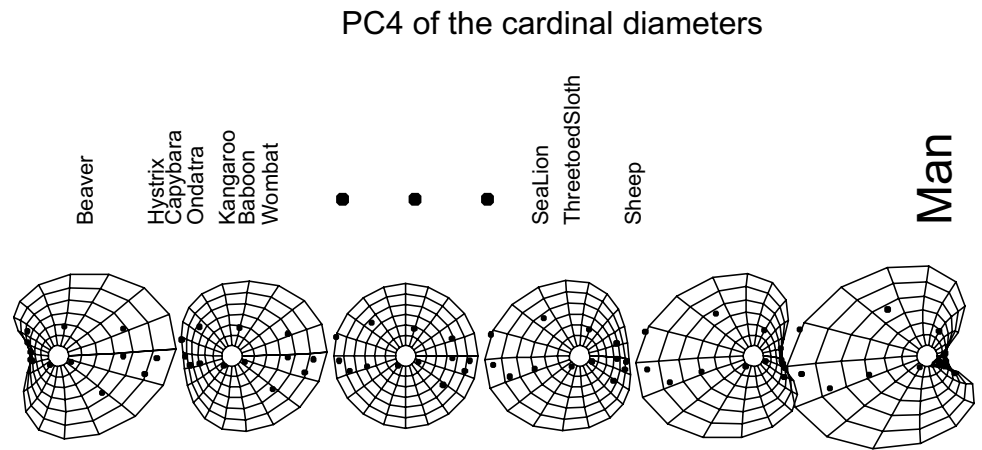
It is fair to inquire about the uncertainty of this lengthy series of analyses against that initial analytic decision, the choice of a longitudinal baseline. Figure 34 complements the right panel of Fig. 1 by an alternative for this highly diverse sample. It is a source of considerable comfort that the obvious features of Fig. 1 are replicated here to a great extent in spite of the quite different functional contexts of the baseline points chosen. (This alternate baseline makes an angle of 33.5° with the nearer of the axes of the usual baseline; the maximum possible such deviation is 45°.) In particular, the large ellipses are still large in general, and the general arrangement of these large shapes is similar on the page up to a linear transformation. The first two canonical correlations of the axis vectors between this pair of ordinations are 0.990 and 0.986, and the first four canonical correlations of the regression coefficient vectors (r, s, t, u, v, w) are 0.994, 0.989, 0.986, and 0.966. In short, the information content of the two data sets is nearly the same in spite of the quite different appearance of the two panels of the figure.

Complementary to the display of the intraspecimen range of second derivatives in Fig. 30, the axis lengths of their ellipses, is the information content of that ellipse’s center. Analytically this is the same as a morphometric quantification introduced decades ago, the concept of *roughness* explained in Bookstein (1978). The roughness of a grid transformation at a grid point is defined as the second-order approximation of the discrepancy (corrected for cell size) between the deformation of the centroid of a grid cell and the centroid of the deformation of its vertices. According to the formula on page 109 of the 1978 reference, this discrepancy is the vector whose x -coordinate is the sum of the two unmixed second partial derivatives of the x -coordinate of the deformation, while the y -discrepancy is the sum of the same two second derivatives of the y -coordinate. The mathematician would notate these components as the Laplacians ($\Delta x', \Delta y'$) where each Δ is the sum of the two unmixed second partial derivatives of the deformed coordinates separately: the expansion ($\frac{\partial^2 x'}{\partial x^2} + \frac{\partial^2 x'}{\partial y^2}, \frac{\partial^2 y'}{\partial x^2} + \frac{\partial^2 y'}{\partial y^2}$).³ So this Laplacian vector equals the sum of the vectors that are plotted with the filled disks and the open disks in Fig. 26, and that sum, in turn, is twice the average of those two disk locations, which is to say, the center of the ellipse. These are plotted in Fig. 35 (left) for each of our 55 species. (As a corollary, we uncover another characterization of the quadratic trend representation: grids having a roughness vector that is the same everywhere.)

This ordination is intriguing. Near (0, 0) we find the names of the species for which the deformation fitted by the

³ The operator Δ is sometimes written ∇^2 , where ∇ is the operator the lexicons name “nabla” that stands for either “gradient” (if it is applied to a scalar function) or “divergence” (if a vector). The Laplacian of a scalar function f is the divergence of its gradient, that is, $\Delta f \equiv \nabla \cdot (\nabla f)$, which, by an abuse of notation, gets to be written as ∇^2 .

Fig. 29 Principal component 4, ranging from Beaver to Man, with selected intermediate scores. The line of dots stands for the other 44 specimen names, which overlap as a solid block of ink when printed at any readable scale



2nd-derivative ellipse semiaxis lengths

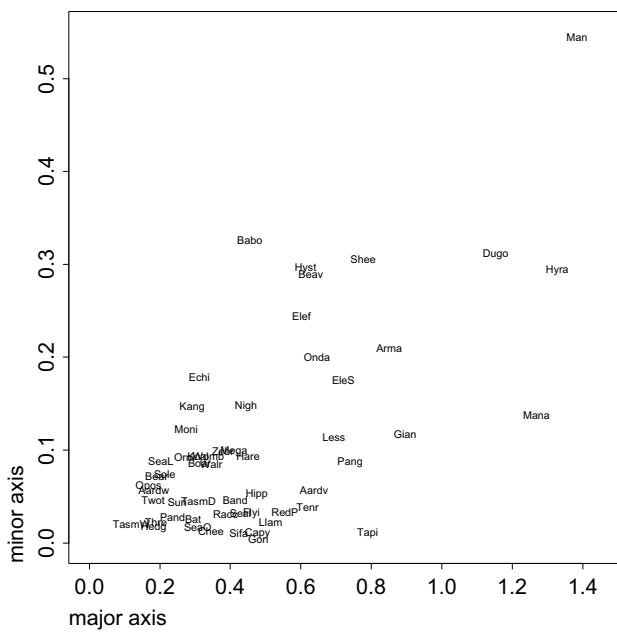


Fig. 30 A suggestive bivariate quantification: semiaxis lengths of the second-order derivative ellipses. See text

quadratic comparison to the average is most nearly linear: Koala, Hedgehog, Tasmanian Wolf, Wombat, Aardwolf, Panda, Capybara, Monito. Around the outline of the distribution are the species for which the center of the ellipse is farthest from (0, 0): Armadillo, Lesser Anteater, Elephant Shrew. and (most extreme in either baseline direction) Giant Anteater versus Manatee. All these extremes represent shifts along the x -axis, the baseline from Fig. 1, which is so much longer than the perpendicular skull height dimension. The species that are most shifted in the orthogonal direction are fewer—Hyrax, Flying Lemur, and Elephant – while Dugong, Pangolin, and Aardvark show substantial roughness in both directions. Any of these extremes might

PC4 of the cardinal diameters

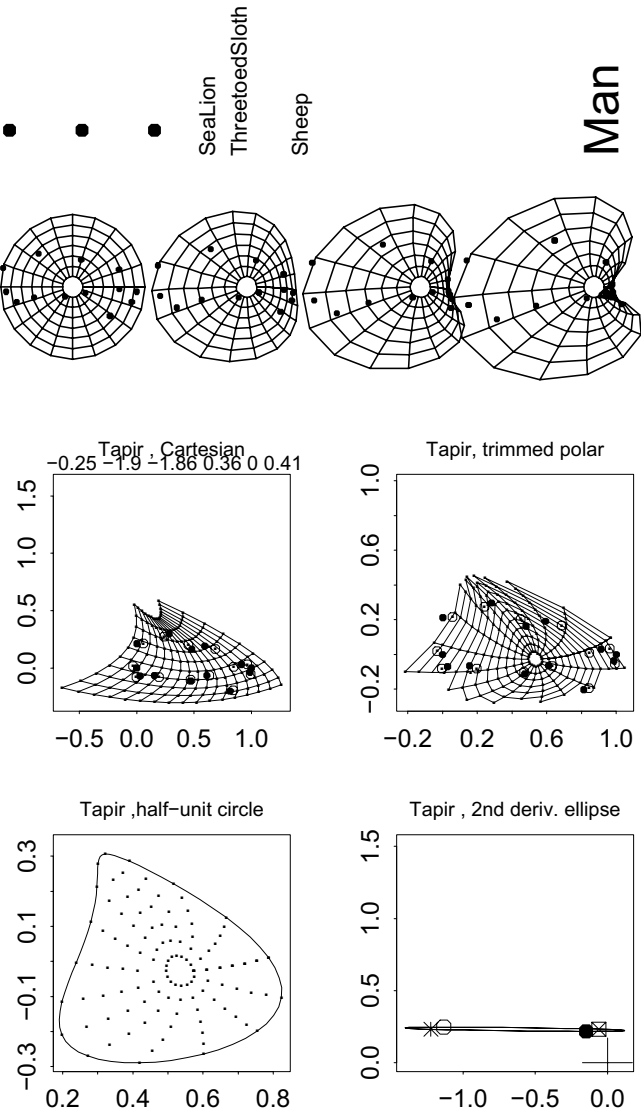


Fig. 31 Four-panel plot akin to earlier Figs. 20 through 22 for the tapi, the species furthest right along the lower margin of Fig. 30. See text

qualify as a morphogenetic innovation or synapomorphy in analyses of how these skull forms relate to their phylogeny. This interpretation is stable against the change of baseline in Fig. 34—the alternative scatter of ellipse centers, Fig. 35 (right), shows the same list of outliers. In neither scatter is the position of the center of the ellipse for Man particularly unusual, not even for a primate (see also Bookstein, 2018, Fig. 3.20c). It is not the average of these second derivatives over the circle of directions but their directionality per se that is unusual for our species among the mammals (Fig. 30).

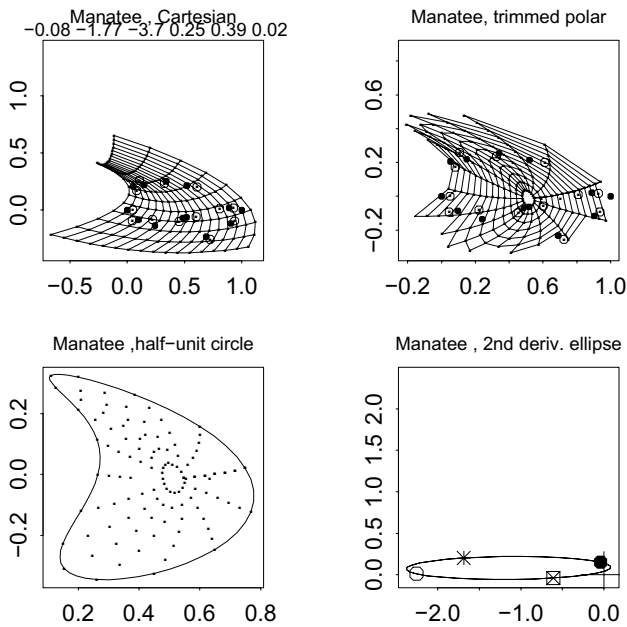


Fig. 32 The same for the manatee. Note the similarity to Fig. 31 in spite of the very different ecology of the creatures

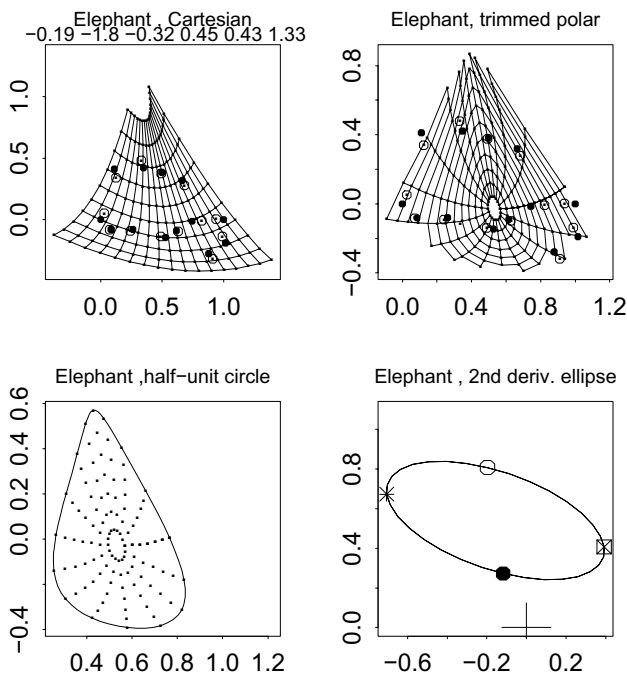


Fig. 33 The same for the elephant

Discussion

The analysis here of the 55-mammal data set is quite different from that published by Marcus et al. The original publication (Marcus et al., 2000) emphasized mainly the matrix

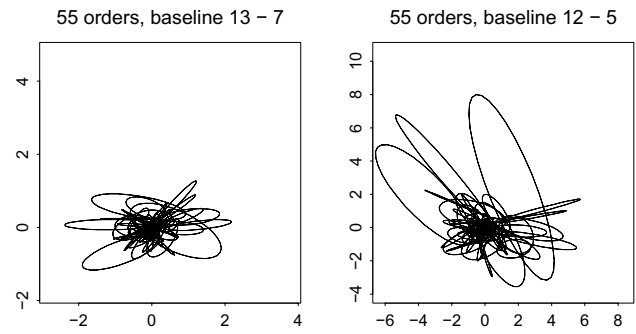


Fig. 34 Critique of Fig. 1 (right) by change of baseline to 12–5 in the numbering of Fig. 1 (left), anterior foramen magnum to fronto-nasal sagittal, which makes an angle of 33.5° with the baseline of the other figures in this section

of Procrustes distances among specimens (for a more extensive list of landmarks) along with the implications of those distances for a cladogram or a phylogeny. I have severely criticized Procrustes distance in a range of papers over the last decade (see Bookstein, 2015, 2016, 2018, 2021) and need not repeat the critique here except for a summary: Procrustes distance is not a meaningful biological quantity, as it depends too much on the investigator’s subjective choice of landmarks, it has divided out a potentially meaningful factor (size) using a biologically meaningless formula (Centroid Size, see Bookstein 2021), and, by treating all landmarks with complete algebraic symmetry, it offers no access to prior anatomical or functional insights that might direct interpretation of multivariate analyses of shape coordinates. For example, the ordination in Fig. 30 of the directionality of a quadratic trend, a quantity that likely relates to evolvability in more general senses, is inaccessible from the standard Procrustes or thin-plate multivariate algebra.

In general, quadratic regressions may be expected to organize some potentially important aspects of landmark configuration data more effectively than either the Procrustes maneuver or the various thin-plate spline visualizations of the resulting shape coordinates whenever the biological phenomenon under exploration (in “A Simple Example: The Vilmann Neurocranial Octagons” section, the pattern of rodent neurocranial growth; in “Revisiting a Mammal Cranial Data Set” section, the huge diversity of geometric arrangements of the adaptively radiated mammalian cranium) involves large-scale effects on these configurations. The scatterplot at lower right in Fig. 19 assured us that most of the variation over this class in this 13-landmark configuration is captured by the trends here that, beyond any linear terms, support the straightforward parameterization displayed in Fig. 1. The ellipses that organize these circuits of directional derivatives correspond to a space of just six degrees of freedom better for morphological interpretation (i.e. more anatomically organized) than either the thin-plate

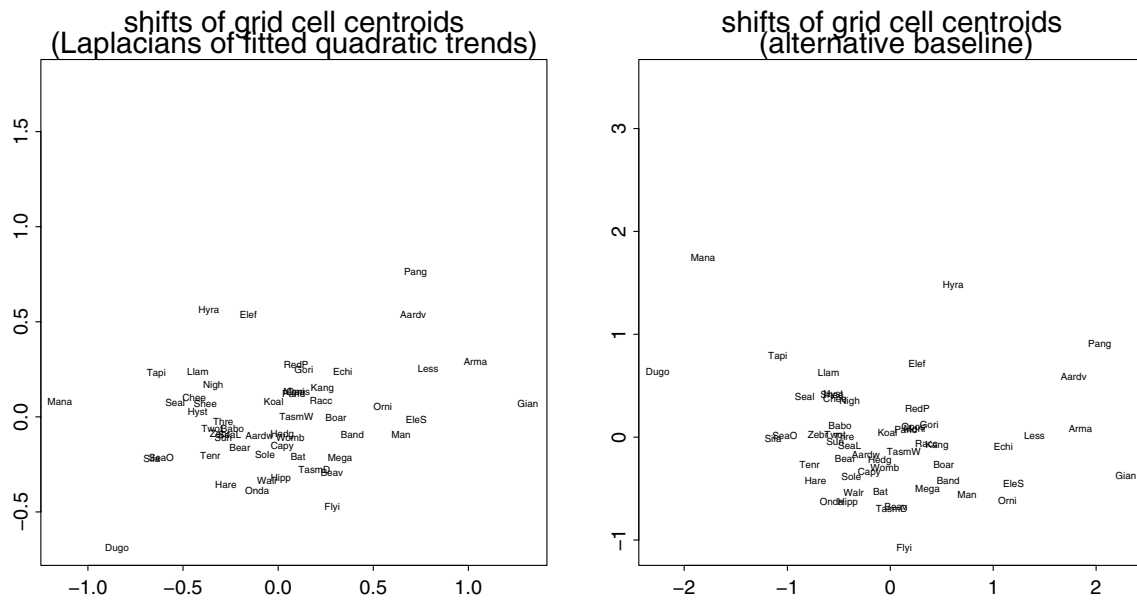


Fig. 35 Complement to Fig. 30: scatterplot of the centers of the ellipses in Figs. 1 or 34. The location of each point is the vector of separation between the quadratic prediction of the centroid of a grid square and the centroid of the prediction of the corners of that square. Labels

are the short names of the 55 species as in Figs. 25, 26, or 29. (left) For the baseline of Fig. 1. (right) For the alternate baseline at right in Fig. 34. The extreme taxa are the same

spline or the multivariate analysis of the whole set of 22 or 26 shape coordinates can manage. In the presence of diversity as great as what is laid out in Fig. 1, we must protect ourselves from the utter arbitrariness of the original data resource here, the finite scheme of disarticulated landmark or semilandmark locations. The quadratic regressions draw the original landmark data, whatever their count (as long as it is greater than 6), into a summary designed to suggest explanations of a morphodynamic or biomechanical sort, explanations plausibly associated with the embryology or ecology of these highly diverse genera.

In this way the quadratic trends rescue D’Arcy Thompson from the sheer obscurity of the method (if any) by which his line drawings were generated, while at the same time their distillation into those six-parameter ellipses of second derivatives rescues Peter Sneath from his unfortunate preoccupation with distances and sums of squares. Together these formalisms, one old and one new, may help to deprecate GMM’s current focus upon Procrustes shape coordinates and thin-plate splines, both of which seemed much more promising back in 1993 than they seem today. Manatee and Tapir are satisfactorily close in the principal component scatterplots of the regression coefficients and the cardinal directions, Fig. 26. But adjacency in principal component plots is much less informative than similarity in a-priori geometric subspaces dealing explicitly with a-priori patterns of landmark rearrangement, the spaces where patterns lead to morphodynamic or functional explanations. Then closeness

in principal component projections is much less informative than closeness in explicit character spaces defined a-priori—in the comparisons of Figs. 31 through 33 this is that “teardrop shape” of the half-circle plots at lower left.

We have seen how the parameters of this paper’s ellipses seem meaningful beyond their algebraic identity as linear combinations of shape coordinates. Three different presentations of the complete space of these shape representations have been introduced. One iconography, Fig. 25, is the full set of six regression coefficients in the appropriate Cartesian setting of three complex numbers. Another representation, Fig. 23, shows the ellipses in full relevant detail for a perspicuous subsample of 19 interesting forms, while a third view, combining Fig. 24 (the scatter of just the cardinal directions) with Figs. 30 and 35 (the geometric parameters of the ellipses per se), parameterizes these ellipses in geometric rather than algebraic terms. Each of these involves just six large-scale parameters regardless of the count $2k - 4$ of dimensions of the actual shape space, and each interpretation is intrinsically more cogent than anything offered in the conventional GMM toolkit. (Recall how “Geometric Fundamentals” section guided us in interpretations of the patterns of the largest two or three regression coefficients out of the sextet.) Restricting shape space to this six-dimensional subspace, in other words, achieves D’Arcy Thompson’s purpose—a simpler vision of shape comparison regardless of how complicated the template may be—while also serving as a potentially feasible character space such as Peter Sneath was hoping for: geometric patterns that could

be exogenously tested for meaning phylogenetically, morphodynamically or functionally.

But these adaptations of otherwise standard bivariate and linear multivariate strategies to the new quadratic regressions are clearly not the end of this programme. The examples in the preceding two sections suggest some further methodological speculations: whether it is worth going to the next order of polynomial trends (cubics rather than quadratics, the way Sneath did), what accounts for the considerable difference between splines and regressions (of which an example will be displayed in Fig. 39), and whether it is worth considering third *derivatives* of the landmark coordinates instead of their third *powers*. This composite thrust can in turn be approached in two different ways: by an alternative spline supplying replacement grids *in toto*, or instead, adapting a familiar tool from applied image analysis, by attending to the simpler task of turning those teardrops into vector-valued characters. Following brief excursions into each of these extensions I will proffer a tentative summary of their implications for morphometric method: we need to become much more clear about what is meant by “a configuration of landmarks” over multiple organisms and how to justify the design of any such configuration.

When Would a Cubic Fit be Worth Pursuing?

It is a necessary characteristic of quadratic trend fits that the second derivative of the deformation in any direction be constant across the grid. Whatever the rotation of the original Cartesian system upon the template, any convex-downward grid line must reside within a family of parallels all likewise convex-downward, and similarly for convex-upward, convex-leftward, or convex-rightward options. Of course real examples do not have the algebraic perfection of the explicit models in Eq. (3)—this paper makes no assumption that any shape comparison is *actually* a quadratic (or any other) polynomial—but there is, as noted in Bookstein (2023a), a considerable cognitive cost to the passage from quadratic to cubic fits: not just the additional $2 \cdot 4 = 8$ degrees of freedom of the least-squares fits, but also the visual complexity of the deformed Cartesian axes that result.

Especially in a context of biomechanical or morphodynamic interpretation, it is convenient to have a glossary of potentially meaningful refinements of that quadratic model. The cubic alternative originally broached by Sneath (1967) embraces two attractive options: the appearance of an S-shape in a deformed Cartesian axis curve (in geometric language, a point of inflection of the curve where the tangent changes sides and the curvature changes sign), or the appearance of a U-shaped gradient in spatial derivatives *along* some deformation of a Cartesian grid line. In fact these prototypes were already visible in Sneath’s Fig. 25, but went unremarked there.

These cubic extensions turn out to arise empirically in the data set just reviewed from the quadratic point of view, the diversity among 55 mammalian exemplars. The same grid-trimming maneuver we have already seen in Figs. 20 through 22 can be applied instead to the more familiar Cartesian grid style as restricted to the interior of the convex hull of the mean landmark template in Fig. 1. (As Bookstein (2023a) notes, trimming the resulting deformation grids is now essential if their curves are not to fold illegibly over one another outside the outlines of the form under study.) There results the 133-point template at left in Fig. 36. We shall exploit this template to explore the effect of regressing the target’s x and y shape coordinates on *nine* predictors instead of the five involved in the quadratic analyses: not only x , y , x^2 , xy , and y^2 of the template but also four more terms x^3 , x^2y , xy^2 , and y^3 —in other words, to fit *cubic* trend grids instead of the quadratic trends of Eq. (3). I have not been able to construct a geometry analogous to the ellipses here for the eight-dimensional cubic terms of these regressions. Instead I treat the 133 vertices of these trimmed representations as a 266-dimensional set of variables for multivariate analysis (x and y coordinates for each of the 133 vertices), and pass the *difference* between the two regressions, cubic minus quadratic, to an ordinary principal-component analysis of their covariation over the 55 specimens of the data set.

The resulting analysis of 266 dimensions has only eight degrees of freedom. Eigenvalues in units of squared shape coordinates are 1.174, 0.570, 0.216, 0.160, 0.139, 0.110, 0.076, and 0.052—the other 257 round to zero to seven decimal places—and of the eight only the first two are potentially meaningful ordinations of the data, the other six being indistinguishable from spherical. Figure 36 shows these first two eigenvectors as grid deformations of the trimmed template, and Fig. 37 scatters the 55 specimens over the corresponding scores. There is no canonical setting of the sign for either of these components—they can point in either direction.

These two principal components are easily verbalized. The positive direction of PC1 (Fig. 36, upper center) represents the U-shaped cubic trend already mentioned: first derivatives small at the ends of a northwest-to-southeast axis but larger in the center of that same axis. The first derivatives in question here are in the x -direction (baseline direction) of this coordinate system. The negative direction $-PC1$ merits, of course, the opposite description. Perpendicular to it in the shape space of these 133 vertices and bearing just about half the variance as PC1’s is the second principal component (Fig. 36, lower row), which very clearly manifests the S-curve deformations among one of the coordinate axes, along with a component of that same northwest-to-southeast U-shape already noted in PC1 but now aligned in the y -direction (which the top row of the figure showed not to be involved in PC1).

The way our visual system processes these grids is distinctly different from its processing of the quadratic grids in Figs. 20 through 22. Algebraically the contrast is a simple resetting of an integer quantity—now it is the third directional derivatives, not the second derivatives, that are constant across the scene—but without any further graphical annotation we directly perceive the extrema at both ends of the ESE–WNW diagonal in the top row of Fig. 36 and, even more automatically, the horizontal S-curving present in either panel of the bottom row. That S-curve actually proffers two features, not just one: the misalignment of left-end and right-end apparent shears and the vertical concentration of dilation or reduction in the NNE–SSW direction in the center by comparison with the periphery. In this way features of cubic trends exploit a visual syntax that is fundamentally local, corresponding to the classical systematist’s search for “characters” that was one motivation for Sneath (1967). Our eyes disassemble the scene into a report by regions as well as gradients. The upper left panel of Fig. 36 permits a labeling of the anatomy of these regions and trends in terms of the classical subdivisions and directions of the mammalian skull; once that information is adjoined, any or all of these grid features might prove to have some morphodynamic, genomic, or phylogenetic basis in particular sample designs.

The 55 pairs of scores on these two PC’s are scattered by species short name (Table 1) in Fig. 37. Species that are phylogenetically close do not noticeably cluster well in this ordination—note, for instance, the discrepancies among the positions of Baboon, Gorilla, and Man or those for any of the sets of neighboring ellipses in Fig. 23. Figure 38 presents a selection of six of these cubic-minus-quadratic contrasts from different regions of the scatter. That for Bear is closest to the (0, 0) of this ordination; the corresponding deformed grid shows hardly any features at all—the cubic trend is indistinguishable from the quadratic in this case. PC1 is exemplified by the contrast of Zebra with either Man or SeaLion (the propinquity of which might be considered an embarrassment to any great-chain-of-being theorist); PC2, by the orthogonal contrast of Elephant with Seal (small “l” here, versus capital “L” for SeaLion). The grid for Gorilla (not shown) is mainly an intensification of that for Elephant, thus even more distant from that for Man.

Should we quantify the net import of this additional octet of regression coefficients by its effect on the residual sums of squares of the resulting analyses? Simple calculations yield a net mean square for variation of the full sample around the mean template equal to 0.1773 in squared shape coordinate units. The simplest geometrical analysis, the 4-d.f. uniform term of the thin-plate toolkit, reduces this mean square by 0.1033, to a residual mean square of 0.0740; the next 6 degrees of freedom of the quadratic analysis lower this by 0.0476, leaving a mean squared residual of 0.0264; and the final eight d.f. of the cubic extension explain 0.0195

of that, leaving, at the end, 0.0068 for the final six degrees of freedom. Thus the incremental contribution of each degree of freedom of the cubic fit is only $(0.0195/8)/(0.0476/4) = 20\%$ of the quadratic terms. Effects so minuscule are probably uninterpretable at any clade-wide level—it is likelier that only the individual diagrams along the lines of those in Fig. 38 will serve for evo-devo speculations. The gradient of decline in these explained sums of squares is neatly analogous to the decline in explanatory power of successive Legendre polynomials under a random-walk model for time series, Bookstein (2012), or the succession of partial warps of steadily higher specific bending energy under the intrinsic integration model of Bookstein (2015).

Thus a more intuitive alternative to ordination of these cubics by the coefficients of global deformation models, an approach that proved quite helpful for quadratic fits, is this focus on regional features. In view of the 18-dimensional complexity of cubic grids, extensions of trend analysis to higher dimensionalities than the quadratic should probably treat landmark or semilandmark shape coordinates not as a single homogeneous vector space but instead, following the suggestion of Bookstein (2023a), as an anatomical composite of regions diverse in their embryology and function—the evo-devo point of view, not the GMM one. In our running mammalian example, while some parameterizations of the quadratic analysis might lead to meaningful ordinations in particular clades, aspects of these cubic deformations and their principal components might better be interpreted as individual traits, to be interpreted embryologically or functionally, rather than as algebraic components of any sample-wide template-spanning shape space.

Trends Versus Interpolating Splines: The Contrasting Roles of Landmarks

Turn now from discussion of the dimensions of this simplified shape space to a more detailed examination of the grid figures themselves, the rendering of the picture area in-between the landmarks of the data set. Equation (3) of “A Simple Example: The Vilmann Neurocranial Octagons” section declares that the transformation grid we want for a quadratic trend is the minimizer of a sum of squares over what is represented as a prediction error at each landmark in turn: the displacement of the predicted location of a landmark from its location as actually observed in some specimen. That equation writes the numerical task here as the computation of numbers $a, b, c, d, e, f, r, s, t, u, v, w$ that together minimize

$$\sum_{(x,y)} |(x', y') - Q(x, y)|^2 \quad (4)$$

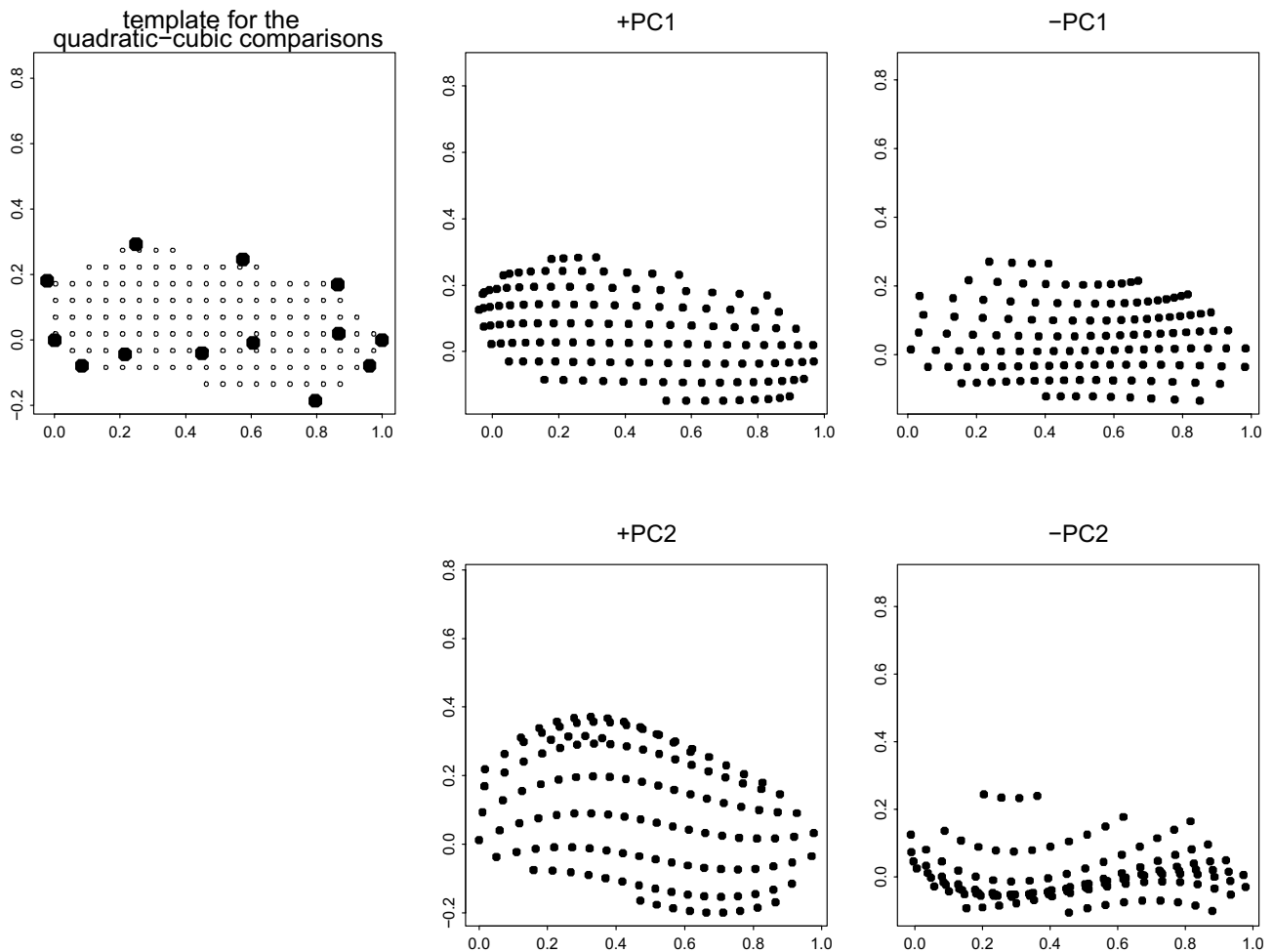


Fig. 36 The only two meaningful principal components of the difference between quadratic and cubic fits for the 55 exemplars of mammalian midsagittal cranial variability. In the guide to trimming, the panel at upper left, the smaller plotted points correspond to the 133

where Q is the quadratic trend function from Eq. (3), (x, y) and (x', y') are the Cartesian coordinates of template and target, respectively, the sum is over the list of the landmark locations themselves, and as before the vertical bars stand for ordinary distance on the picture plane. Six of the resulting values, the coefficients r, s, t, u, v, w , of Q 's quadratic terms, comprise one possible sextet of specimen descriptors (the paper touches on several others).

Remembering that the values r through w encode the actual second derivatives of the fitted mapping Q , it is instructive to juxtapose the assignment of minimizing expression (4) to the task that is solved by the conventional thin-plate spline as it has been exploited since the early 1990s throughout the GMM community. This is the task of choosing the function Q out of the restricted family of possible Q 's constrained a-priori to exactly match the values

grid vertices of the text while the larger dots locate the sample average two-point shape coordinates to the same baseline (Fig. 1) on which most of the other figures of this 13-gon analysis have relied. The eigenanalysis was of the locations of those 133 smaller dots

at the landmarks— $Q(x, y) = (x', y')$ at each pair of locations — that minimizes the quite different-looking expression

$$\int_{\mathbf{R}^2} \sum_{i,j,k=1,2} \sum \left(\frac{\partial^2 Q_k}{\partial x_i \partial x_j} \right)^2. \quad (5)$$

The symbol \mathbf{R}^2 here stands for the whole Cartesian plane, out to infinity, and the subscripts i, j, k range only from 1 to 2. The i and the j stand for the two Cartesian dimensions of the template, which can be the same (x and x , or y and y) or different (x and y or the opposite, which give the same $\frac{\partial^2 Q}{\partial x_i \partial x_j}$, as it happens), and k is 1 or 2 for the Cartesian coordinates of the target.

What is being minimized in Eq. (5) isn't the error of fit of the function Q at the landmarks—that error is identically

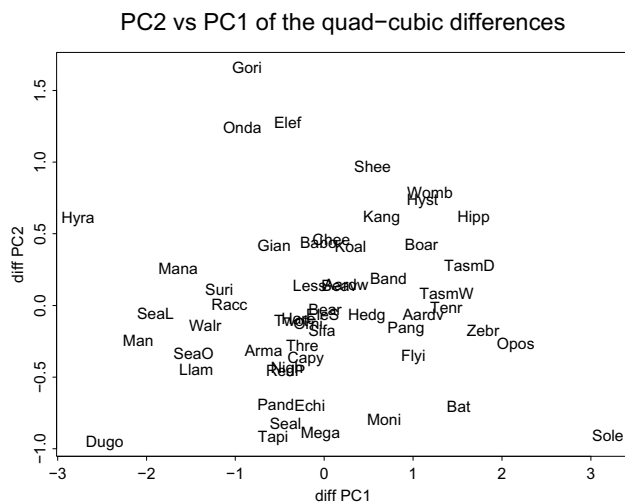


Fig. 37 Scatter of scores for 55 specimens on the principal components of Fig. 36. The ordination for Man appears to be mainly an intensification of that for SeaLion, although such an interpretation is unlikely to bear any evolutionary meaning

zero as the essential restriction on Q in the statement of this other problem—but is instead a property of those second derivatives in Eq. (5) that were taken to be constant coefficients in the formula for Q of expression (4): their own sum of squares, integrated (added up) over the whole picture, should be a minimum. This is no longer any sort of error at the landmarks—it is, rather, a quite different kind of error that would arise if we compared the function Q to one that had all of its second partial derivatives zero, which is to say, an exactly linear map. Formula (4) cares about the landmark locations; formula (5), about the grid in-between and all the way to the edges of the picture and beyond. The formulas that give us this second kind of optimizing Q are very clever indeed—mathematicians derived them only in the 1960s—but the values of their coefficients (the vectors $L^{-1}H$ of Bookstein (1989) and other presentations) are of no empirical interest at all, only the integral of summed squared second partials that engineers quickly recognized as the bending energy of an idealized thin plate. And in formula (5) the locations (x', y') of the target landmarks are not of any explicit interest, either: only that the function $Q(x, y)$ exactly reproduces each of them.

Then several formal relationships between Eqs. (4) and (5) are noteworthy. Either optimum is *rotatable*—when one of the Cartesian coordinate systems is rotated, both the minimizing spline (5) and the quadratic regression (4) rotate directly with (x', y') and inversely with (x, y) . And both approaches can be proven to supply unique global minimizers except under singular circumstances (e.g., all landmarks in a line). But the differences are more numerous, and more salient. Most importantly, there is no “error” at landmarks in (5), the quantity whose sum of squares is minimized in

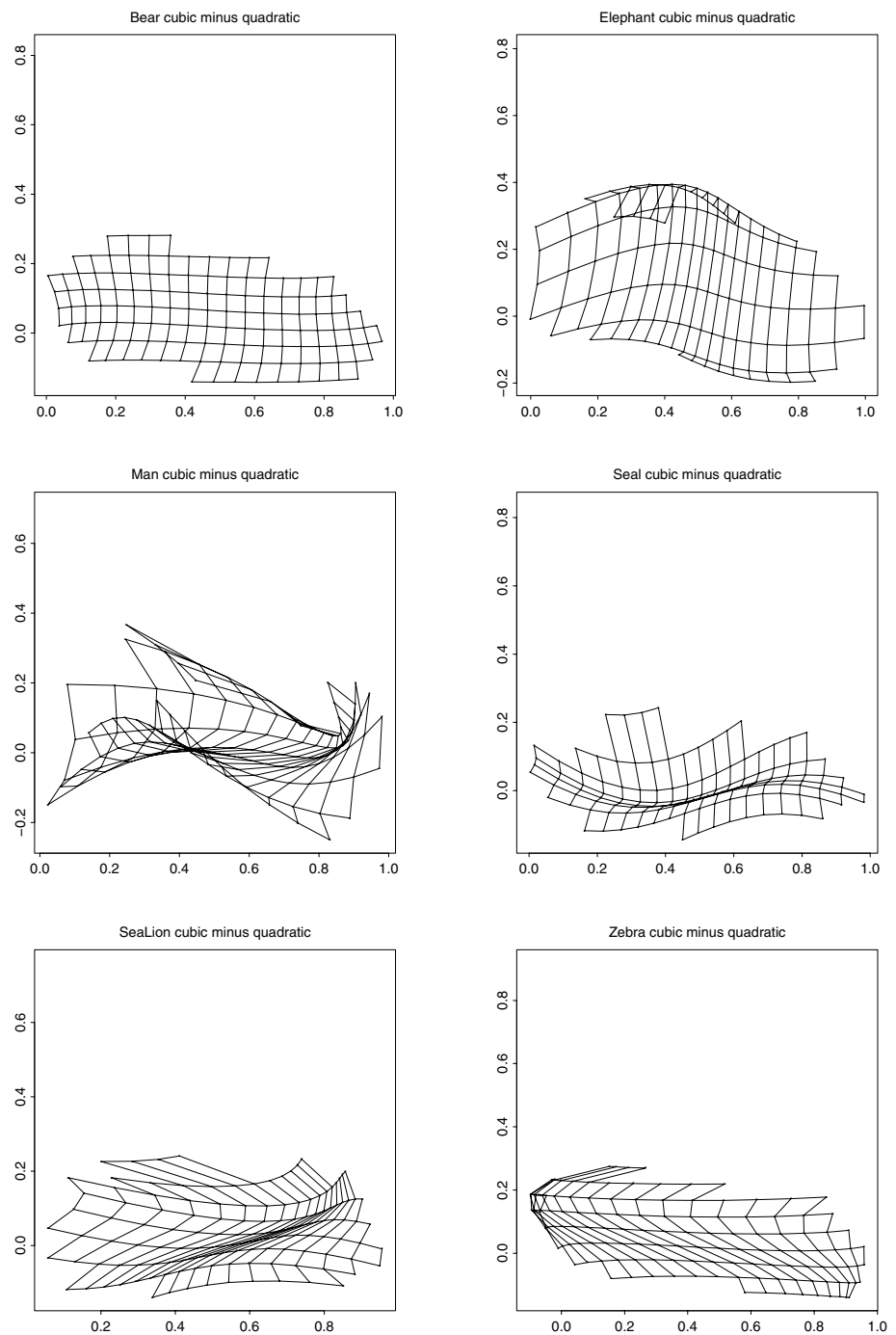
(4). Instead, what is minimized in (5) is the integral of the variation from point to point of what could be regarded as a model for variation of the regression coefficients b, c, e, f over position—the squared partial derivatives of the *linear* part of the fitting function Q , which constitute the second partial derivatives that the definition of the function Q in formula (3) sets to be constant everywhere. And while Eq. (4) involves a sum over the landmarks alone, expression (5) integrates over not just the interior of the landmark configuration but all the way out to infinity (which is why these maps *must* become linear toward infinity— r, s, t, u, v, w all have to drop to zero pretty fast for their integrals to be finite at all, let alone minimized).⁴

Sums of squares of position discrepancies are not equivalent to integrals of sums of squares of second partial derivatives either numerically or conceptually. The transformation grids produced by the two techniques can be quite dissimilar (for a relatively tame example, see the first two frames of Fig. 39—many other examples are on display in the figures of Bookstein 2023a). The coefficients of Q that serve as characters in the quadratic trend approach are mere nuisance variables in the spline approach, never examined or subjected to multivariate analysis until they are summed after each is multiplied by the very peculiar formula $r^2 \log r$ where r is the same Pythagorean distance as in formula (4), but now applied to an entirely different argument, driven not by the configuration of landmarks but instead by the continuum of gridded points themselves as they relate to the various data points of the template one by one. But in the quadratic-trend graphics, the error of fit landmark by landmark is explicit in the relation between the filled disks and the open disks in contexts like Fig. 20, while the six second derivatives of Q , hinted at in the variation of the shapes of the little grid squares across the diagram, become explicit when the polar coordinate construction is assessed by the second-difference method of Fig. 4.

In short, the two approaches to landmark configuration analysis, polynomial trend fits versus thin-plate splines, suit wholly different explanatory styles. The coefficients that the trend method uses to export biometric meaning are discarded immediately after computation by the spline. The second partial derivatives of Q , the relevant constants of the

⁴ But it is no help to restrict the region of integration in Eq. (5) to just the interior of the tissue outline. The resulting bending energy integral would no longer be a simple, elegant quadratic form (Bookstein 1989, 1991) in the registered landmark coordinates, and the bending spectrum of the spline itself would cease to be a useful rotation of the shape coordinate space per se, so that the interpretation of its features in terms of an ordering by spatial scaling would be lost. But GMM needs those tools in the course of subsequent multivariate analyses like the one in Fig. 19 or the extension to the BE-PwV plots, Bookstein (2015, 2018), that serve as one tool for local descriptions in the classic GMM toolkit.

Fig. 38 Quadratic-cubic contrasts, graphed as the 133-vertex grids trimmed to the template as in Fig. 36, for a selection of six exemplars from Fig. 37. That for Bear is closest to the (0, 0) of this scatter (recall Fig. 20 as well); the effect of PC1 is close to the contrast between the grids for SeaLion and for Zebra; that of PC2, for the contrast between grids for Seal and Elephant



quadratic trend fit, are the minimand of the spline method, which refers to the embedding space of the landmarks in a manner to which the regression equation has no access. And the error of fit of the trend method is forced to exactly zero in the spline method. This paper argues that, in view of the arbitrariness of the landmark lists driving GMM data sets from the outset, one of these assignments is much more conducive to subsequent biological insight than the other is.

The displays in Fig. 19 already permitted us to quantify the net modeling power of this paper's quadratic trend

analyses vis-à-vis an analogous attempt at large-scale modeling, the *partial warps* of the thin-plate approach (Bookstein, 1989, 2014, 2018). The upper right panel, for the spline-based approach, reduces the landmarkwise variances of the upper left panel by a factor of 0.65; the equivalent comparison along the lower row, for this paper's quadratic approach, reduces the summed variances by a factor of 0.85. The analyses in the upper row allow the coefficients of the prediction function to vary but require them all to attenuate sharply toward zero away from the origin; those in the

lower row leave those coefficients constant, instead optimizing the predictive accuracy of that single shared quadratic regression. Clearly the constant-coefficient approach inherited from Sneath (1967) dominates the varying-coefficient approach of the thin-plate consensus in this context of predictive accuracy per se. The discrepancy between the approaches is understandable—the regression approach is, after all, a least-squares optimization already, one that would become exact after 14 more terms are added, but these additional terms involve higher powers of x and/or y that are correlated with the quadratic terms already present and so are unlikely to add much accuracy overall; whereas the first three components of the spline’s principal warp decomposition are suboptimal for any decomposition except bending energy, but that is not a sum of squares of anything referring to the actual morphology of the organism, bounded as it is in extent.

A Different Thin-Plate Spline

The quadratic trends analyzed and diagrammed so far in this paper are regression fits that leave unexplained variation at each landmark. In the other approach to combinations of squares, the thin-plate spline formalism, the map is an interpolation rather than a regression, and the sum of squares that is being minimized is not a net prediction error but rather a sort of complexity, the version the literature calls “bending energy.” For the conventional thin-plate spline, bending energy is one quantification (albeit perhaps a peculiar one) of the departure of its Q from a linear map. But ever since the initial promulgation of these splines there has existed an alternative, the *quadratic thin-plate spline*, that likewise fits each landmark exactly, by analogy with our conventional thin-plate spline, but for which the drift term (the polynomial part that is not a function of the distance of a grid point from every landmark location in turn) is a quadratic map rather than a linear one. This is the thin-plate spline that minimizes a different kind of “bending” energy, namely,

$$\int_{\mathbf{R}^2} \sum_{i,j,k,l=1,2} \sum \left(\frac{\partial^3 y_l}{\partial x_i \partial x_j \partial x_k} \right)^2, \quad (6)$$

sum of the integrals of the squared *third* derivatives of the map (y_1, y_2) of x_1 and x_2 . And the additional contribution of each landmark, to be multiplied by some complex number, takes the form $r^4 \log r$ instead of $r^2 \log r$. (In other words, this so-called quadratic thin-plate spline minimizes the integral over the whole picture of the sums of squares of deformation’s *third* derivatives for all maps that exactly match the landmarks, whereas the cubic regression posits constancy of these derivatives over the whole picture plane.) For the theory of this spline, which is a close analogy to the usual

linear-trend spline of the conventional GMM toolkit, see Kent and Mardia (2022). I have exploited it once before, in Bookstein (2004), where, however, it was applied to a comparison of forms that differed hardly at all.

At upper right in Fig. 39 is the Vilmann comparison as in “A Simple Example: The Vilmann Neurocranial Octagons” section but using this exact fit with quadratic drift (not “trend”) curve analogous to the quadratic fit with regression errors that was displayed in Fig. 10. Clearly it is failing to cope with the situation along the cranial base (right-hand margin of the figure in this orientation), where it appears to be rolling up the paper on which the image is printed rather than telling us anything useful about the gradient of the interpolation there. The same impression of an unreal scrolling is apparent in applications to several of the Marcus mammalian crania. Three examples are displayed here, for three of the forms extreme in some of the panels in Fig. 26: Elephant, Beaver, and, of course, Man. In the present context of a highly radiated clade this alternative morphometric praxis does not appear promising—cubic regressions likely make better sense than minimization of third derivatives.

A Potential Character: The “Teardrop”

Toward the end of “Revisiting a Mammal Cranial Data Set” section I suggested that the emergence of a teardrop in the half-unit polar plot of second directional derivatives might actually be a potential character. That Sneath’s hope for a feature space was mostly ignored during the entire late-20th-century development of the current “morphometric synthesis” was an unfortunate oversight on the part of all its developers, including me. Figure 40 hints at a potential phylogenetic use of this quadratic trend approach in Sneath’s spirit. At left I have compiled all 55 of the half-unit outline curves from all the four-panel figures like the exemplars scattered throughout this essay. (Note that these curves have more information than just the quadratic trends—they combine both parts of the regressions in Eq. 3, not only the ellipses for the quadratic trend but also the linear terms with coefficients b, c, e, f .) There is a central skein of curves that differ little from circles, plus a variety of apparent deviations showing sharper curvature. It is convenient to parameterize these deviations by one classic measure of curvature, the change in direction between each of the line-elements of those deformed circles and their neighboring element, divided by the length of the shortest chord that spans the pair from its discretized directional oval. When we characterize our 55 species by the maximum of this improvised index, the top fifteen selections are an interesting list of species. The right panel of Fig. 40 draws these 15 deformed circles by themselves, now after a centering just to simplify the graphic, and labels each with its short species name at the vertex of maximum curvature using the index just explained.

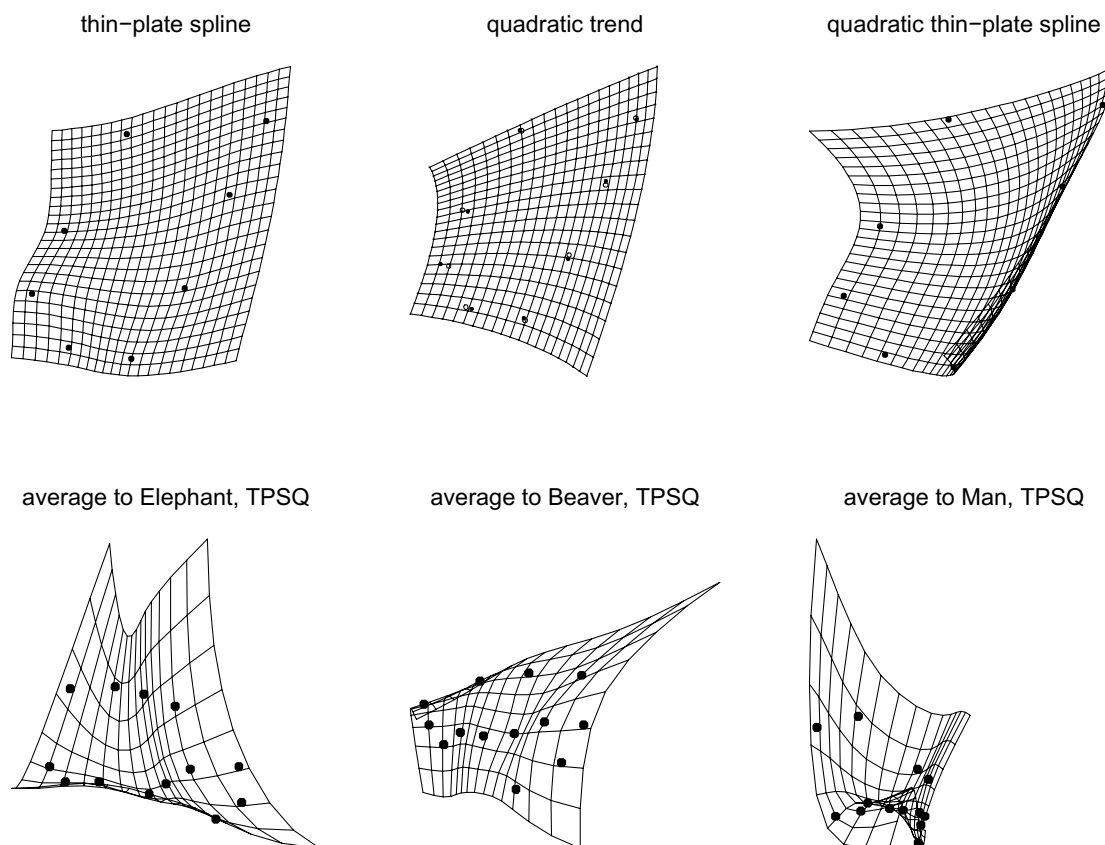


Fig. 39 Top row, three different versions of a transformation grid as applied to the fourth column of Fig. 10 (the 13° rotation of a carefully chosen two-point registration of the Vilmann octagons). (left) Conventional thin-plate spline, relaxing toward linearity outside the form. (center) The quadratic trend recommended for this application, simplifying the pattern of grid lines into a vector of six parameters

each a second derivative in some direction that is constant over the grid. (right) A different quadratic generated by an adjustment of the thin-plate spline formula itself. Bottom row, examples of this quadratic thin-plate spline (“TPSQ”) for three of the 55 cranial 13-gons in the revised Marcus mammalian data set. See text

The fifteen vertices can be reviewed in five groups that are quite suggestive of biological interpretations, as follows. Toward the top of the diagram are two forms, Sheep and Elephant, which peak in a direction perpendicular to the long axis of Fig. 1. Counterclockwise from them is a cluster of four forms of which two, Manatee and Tapir, aroused our special interest in “[Revisiting a Mammal Cranial Data Set](#)” section in connection with the zeroing of one of their ellipses’ semiaxes. A third, Dugong, is ecologically similar to Manatee, while Hyrax (associated with the most deformed polar circle out of all 55 of Marcus’s exemplars) is phylogenetically related to Elephant in the preceding cluster. Continuing counterclockwise, Ondatra (muskrat) and Beaver clearly overlap in their ecology. Our circuit next encounters a pair of singletons, Gorilla and Man; in view of their phylogenetic proximity, the separation might be attributed to *H. sapiens*’s neoteny. A final cluster of five species (Pangolin, Echidna, Aardvark, Lesser Anteater, Giant Anteater) collects

some of the forms with snout very highly compressed vertically (a cluster that would incorporate four more forms of the next five highest curvatures, including Elephant Shrew, Armadillo, Tenrec, and Bandicoot, not shown). Taken as a whole, the clustering here seems quite distant from a random pattern, but ought to be considered to represent some combination of niche and evolvability over the whole class of mammals. In my opinion, some less improvised version of this clustering may well justify deeper explorations.

Note how very nonlinear this collection of tools is, beginning with the nonlinearity of the predictors x^2, xy, y^2 from the template—the coefficients of the fitted quadratics are, of course, nonlinear in the template data already, based as they are in the inverse of their 5×5 covariance matrix. While the center of each fitted ellipse and also its cardinal directions are linear in those fitted coefficients, the organismally relevant graphic summaries—eccentricity of the ellipse, closeness of one of the endpoints of its major axis to $(0, 0)$ —are not. The polar graphic is linear in the regression coefficients

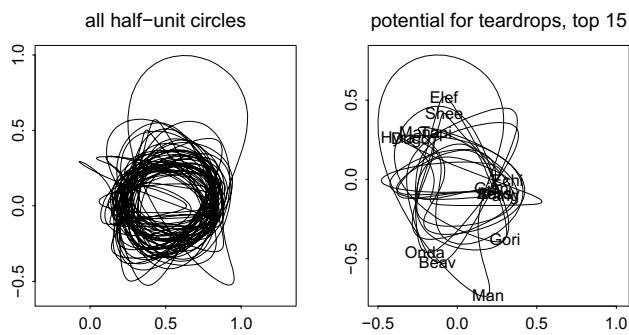


Fig. 40 Analysis of the polar plots of 55 specimen-specific quadratic trend regressions. (left) Overlay of all 55 instances. (right) The fifteen out of the 55 having the highest peak curvature scores as explained in the text. Each short species name is printed at the point of peak curvature. The illegible name list at upper left is Hyra(x), Tapi(r), Mana(tee), Dugo(ng); that at center right is Echi(dna), Gian(t Ant-eater), Less(er Anteater), Pang(olin), Aard(vark). Elef: Elephant. Shee: Sheep. Gori: Gorilla. Onda: Ondatra

but nonlinear in its argument $(x, y) = \frac{1}{2}(\cos \theta, \sin \theta)$, where the constraint locking the radius at $\frac{1}{2}$ is set to half the distance between the landmarks of that arbitrarily chosen baseline, so that the chord of these contrasts across the template's circle is constant in length at 1.0. The teardrop descriptor combined these with aspects of the *linear* part of those fitted quadratics: for each of the 55 specimens we located the maximum around the polar circuit of a one-dimensional summary of a two-dimensional regression prediction spanning the full configuration of landmarks but parameterized as a function of direction, not position, and inasmuch as what we are taking the maximum of is already a second derivative (for that is what our index of curvature is approximating), the result per se must be understood as having located the zero of a *third* derivative. In comparison with the hidden sophistication of this tactic, details of the Procrustes versus the two-point registration (a choice mandated in order that there be shape or form coordinates to be regressed) are trivial.

The teardrop regionalizes the reportage of a quadratic trend in a way partly analogous to the way a cubic trend analysis regionalizes, but with some differences. Instead of a separated pair of regions of the digitizing plane (PC1 in Fig. 36) or a pair of directions highlighting two different styles of features (PC2), the teardrop represents a combination of a single direction of extension with a diminished spacing in the perpendicular direction—in effect, a local feature of the global strain map borne by the ten-dimensional mixed linear-quadratic trend (coefficients $b, c, e, f,$ and r through w in Eq. 3) when rendered in polar coordinates this way. The rhetoric of this account of teardrops is analogous to an image processing tool quite different from the deformation-based approach of GMM: the *medial axis* or

symmetric axis of Blum (1973). Textbook reviews of this alternative include (Bookstein, 1991), Section 3.5, and, for the extension to three-dimensional data, Siddiqi and Pizer (2008). Blum argued that the way our visual system automatically constructs medial axes is built into the design of the brain's visual cortex. Similar innate algorithms might be responsible for both the perception of homogeneity of the quadratic trends of this paper and the unconformities of the cubic extension. We *see* these features with an immediacy that no visual processing of numerical tables can imitate. Algebraically, what we are examining is a vector of length 4: the two coordinates of that peak of curvature, the direction of the axis of the pinching there (medial axis of the feature, in Blum's sense), and the actual curvature of the deformed circle there.

Such visual processing algorithms must be nonlinear. (In particular, the four teardrop coordinates suggested in the previous paragraph are a strongly nonlinear dimension reduction of the ten-dimensional representation in equation (3).) The cardinal diameters are linear in the quadratic regression coefficients, and so their principal components, Fig. 25 through Fig. 28, are likewise accessible from the data set of these vectors, though the necessary matrix notations would be clumsy. But the extraction of the principal axes of the second-derivative ellipses, Fig. 30, was already beyond the capabilities of that standard toolkit, as the ellipses in question are not based on data-driven covariances, and the detection of the clusters in Fig. 40 is clearly beyond any standard GMM software system—it is, indeed, a task for the natural intelligence of the investigating biologist, one who is aware of the currently accepted phylogeny of these same specimens. Clearly linear multivariate analysis alone is not adequate preparation for the biometrics of landmark configurations. Of course there are many other branches of geometry applicable to biometrics as well—tensor calculus, differential geometry of surfaces, catastrophe theory, projective metrics, fractal geometry, hyperbolic geometry, to name just a few. That we have learned how to teach the linear vector geometry of covariance matrices is no reason to privilege it among this multitude of toolboxes.

Concluding Comment: The Relation of Landmarks to Grids, and to Biology

We are thereby brought face to face with a question that was implicit but mostly overlooked in the literature of thin-plate spline deformations from their first publication (Bookstein, 1989) on: what is the epistemology of these grid diagrams—what reality do they so persuasively appear to be claiming? To the extent that GMM is imagined a component of the science of biology rather than a subtheme of the artificial intelligence of shape recognition or classification

or the intellectual property of industrial biometric applications such as security or animal husbandry, I believe it is important to align with Thompson rather than the current GMM synthesis on this matter: to have the role of GMM be to *generate* hypotheses, not to test them, so that the job of the grid diagrams per se is specifically to serve as graphical metaphors of biological explanations that go on to be tested experimentally, as Przibram argued fully a century ago, or at least by explicit confrontation with data resources exogenous to morphology. The finding in Fig. 10, repeated in the upper central panel of Fig. 39, is thus a provocation to generate some explanation of why four of the six parameters of a quadratic trend fit are indeed close to zero in this growth data set. Likely, in view of the commonalities over the specimens of the closely related comparisons unearthed in Fig. 18, their vanishing, together with the resulting bilinearity of the growth deformation, is saying something important about the regulation of rodent neurocranial form; what might that bilinearity be announcing?

Similar to this concern about the biological meaning of the transformation grids is a concern for the meaning of the data that delimit them. Sneath (1967) proceeded past the six-coefficient stage of these quadratic transformations to consider the next level as well, the ten-coefficient stage that supplements the Q of Eqs. (3) or (4) with terms in x^3 , x^2y , xy^2 , and y^3 , and I suggested a tentative multivariate approach to this extension in Fig. 36 through Fig. 38. However, I think that in applications to most data sets of evolutionary or developmental design it would be ineffective to proceed with global summaries or derived parametrizations from this model or any other coefficient space of such high geometric complexity. Rather, the GMM community should turn to a serious, principled probe into what we mean by a “configuration of landmarks” in the first place. How are these lists constructed, and when do we decide to stop adding landmarks to them, or even to delete some? (As MacLeod 2017 wryly notes, “GM approaches to morphological analysis yield markedly different results depending on the morphological features sampled.”) The relation of landmark points to the theory of homology has hardly been explored since the pioneering work of Nicholas Jardine more than 50 years ago (Jardine, 1969). And what about semilandmarks, those representations of curving form that can number in the thousands or even higher now that powerful software can run almost autonomously on affordable multicore machines? Before extending the length of these coefficient vectors further I think the field ought to decide what it means to *be* a configuration, for instance, how and where boundaries are to be drawn between the global analysis of one composite form and a series of analyses of diverse component subforms followed by their synthesis using some non-morphometric method adapted from the lore of machine learning (MacLeod 2017 and references therein).

If *simplification of a grid report until it is comprehensible* is indeed the ultimate bioscientific role of a revised GMM, none of today’s standard GMM tools—not Procrustes shape coordinates, not their principal components, not the thin-plate splines that purport to visualize their multivariate patterns—are capable of supplying the appropriate rhetoric for the announcement of empirical pattern findings in a language suggestive of future organismal explanations. Their occasional usefulness in tasks of discrimination or classification notwithstanding, none of this technology accommodates information about the spatial arrangement of landmarks along with the spatial disposition of the curves that summarize the comparisons of the anatomical regions they delineate.

Neither the sum of squares that the Procrustes method minimizes nor the sum of squared loadings that the method of principal components maximizes comprises the appropriate currency of a biometric pattern analysis; nor does the bending energy of a thin-plate spline. What is required instead is an interpretation of the landmark configurations as expressions of causes or effects of their anatomical organization in life, prior to being digitized. It is the explicit calibration of the bilinear model in Fig. 10 or its validation over the sample in Fig. 23 that justifies the design of Vilmann’s octagon of landmarks.

This finding does not, of course, constitute an explanation of “the origin of form in force,” the way D’Arcy Thompson would have had it. It is only a striking new suggestion of how a properly morphogenetic investigation, one based in biomathematics and biophysics as well as geometry, might proceed to explore the hypothesis generated here on purely morphometric grounds. And ultimately the phylogenetic hints in the scatter of mammal skull trend ellipses, Fig. 23, and the teardrop analysis, Fig. 40, may serve the same role in justifying the midsagittal component (Fig. 1) of Marcus et al. cranial landmark configurations. Because the current GMM toolkit is not capable of serving biological science in this way, I put forward the quadratic technology of this article as a possible first step in liberating GMM from the current straitjacket of Procrustes shape coordinates and thin-plate splines into which the synthesis of the 1990s has inadvertently cast it.

Acknowledgements I am grateful to Joe Felsenstein (University of Washington), Norm MacLeod (Nanjing University), Jim Rohlf (Stony Brook University), editor Julien Claude, and another anonymous reviewer for extensive comments on earlier versions of the argument here. One of those earlier versions, Bookstein (2023b), was posted in March as a preprint to [bioRxiv](https://doi.org/10.1101/2023.03.15.532111) under the same title. Thanks as well to Erika Hingst-Zaher (Instituto Butantan, São Paulo) for the revised Marcus data set she shared with me 10 years ago in connection with the preparation of the work cited here as Bookstein (2018) as well as in “[Revisiting a Mammal Cranial Data Set](#)” section of this paper.

Author Contributions FLB did all the work reported here, and also all the writing. But the collection of mammalian cranial data came from

Erika Hingst-Zaher in 2013 as slightly modified from the data set of Marcus et al. (2001), of which she was one author.

Funding Open access funding provided by University of Vienna.

Data Availability The Vilmann data of “A Simple Example: The Vilmann Neurocranial Octagons” section were published in full in Bookstein (1991) and are available in the online resource of Bookstein (2018) as well as in Jim Rohlf’s morphmet data archive. Analogues to Fig. 20 for all 55 of the specimens in the data set of “Revisiting a Mammal Cranial Data Set” section may be found in the Supplement to Bookstein (2023b).

Declarations

Conflict of interest There has been no external support for the work reported here, and there are no conflicts of interest to be reported.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abramowitz, M., & Stegun, I. A. (Eds.). (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* Vol. 55, Applied Mathematics Series, National Bureau of Standards.
- Blum, H. (1973). Biological shape and visual science (Part I). *Journal of Theoretical Biology*, 308, 205–2887.
- Bookstein, F. L. (1978). *The measurement of biological shape and shape change*. Lecture Notes in Biomathematics (Vol. 24). Springer.
- Bookstein, F. L. (1981). Coordinate systems and morphogenesis. In T. G. Connelly, L. Brinkley, & B. Carlson (Eds.), *Morphogenesis and pattern formation* (pp. 265–282). Raven Press.
- Bookstein, F. L. (1985). Transformations of quadrilaterals, tensor fields, and morphogenesis. In P. L. Antonelli (Ed.), *Mathematical essays on growth and the emergence of form* (pp. 221–265). University of Alberta Press.
- Bookstein, F. L. (1986). Size and shape spaces for landmark data in two dimensions (With Discussion and Rejoinder.). *Statistical Science*, 1, 181–242.
- Bookstein, F. L. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 567–585.
- Bookstein, F. L. (1991). *Morphometric tools for landmark data: Geometry and biology*. Cambridge University Press.
- Bookstein, F. L. (2004). After landmarks. In D. E. Slice (Ed.), *Modern morphometrics in physical anthropology* (pp. 49–71). Kluwer Academic Publishers.
- Bookstein, F. L. (2012). Random walk as a null model for high-dimensional morphometrics of fossil series: Geometrical considerations. *Paleobiology*, 39, 52–74.
- Bookstein, F. L. (2014). *Measuring and reasoning: Numerical inference in the sciences*. Cambridge University Press.
- Bookstein, F. L. (2015). Integration, disintegration, and self-similarity: Characterizing the scales of shape variation in landmark data. *Evolutionary Biology*, 42, 395–426.
- Bookstein, F. L. (2016). The inappropriate symmetries of multivariate statistical analysis in geometric morphometrics. *Evolutionary Biology*, 43, 277–313.
- Bookstein, F. L. (2018). *A course of morphometrics for biologists*. Cambridge University Press.
- Bookstein, F. L. (2019). Reflections on the biometrics of organismal form. *Biological Theory*, 14, 177–211.
- Bookstein, F. L. (2021). Centric allometry: Studying growth using landmark data. *Evolutionary Biology*, 48, 129–159.
- Bookstein, F. L. (2022). Dimensions of morphological integration. *Evolutionary Biology*, 49, 342–372.
- Bookstein, F. L. (2023a). Reworking geometric morphometrics into a methodology of transformation grids. *Evolutionary Biology*. <https://doi.org/10.1007/s11692-023-09607-2>
- Bookstein, F. L. (2023b). Quadratic trends: A morphometric tool both old and new. *Posting to bioRxiv*. <https://doi.org/10.1101/2023.03.23.533997>
- Dryden, I. L., & Mardia, K. V. (1998). *Statistical shape analysis* (2nd ed.). Wiley.
- Gower, J. C. (1975). Generalized Procrustes analysis. *Psychometrika*, 40, 33–51.
- Hilbert, D., & Cohn-Vossen, S. (1952). *Anschauliche Geometrie*. Springer, 1932. Translated as *Geometry and the Imagination*. Chelsea (1952).
- Huxley, J. (1932). *Patterns of relative growth*. Dial Press.
- Jardine, N. (1969). The observational and theoretical components of homology: A study based on the morphology of the dermal skull-roofs of rhipidistian fishes. *Biological Journal of the Linnaean Society*, 1, 327–361.
- Kendall, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16, 81–121.
- Kent, J. T., & Mardia, K. V. (1994). The link between kriging and thin-plate splines. In F. P. Kelly (Ed.), *Probability, statistics, and optimization* (pp. 325–339). Wiley.
- Kent, J. T., & Mardia, K. V. (2022). *Spatial analysis*. Wiley.
- MacLeod, N. (2017). On the use of machine learning in morphometric analysis. In *Biological shape analysis: Proceedings of the 4th international symposium* (pp. 134–171).
- Marcus, L. F., Corti, M., Loy, A., Naylor, G. J. P., & Slice, D. E. (Eds.). (1996). *Advances in morphometrics* (Vol. 284). NATO Advanced Scientific Institutes Series.
- Marcus, L. F., Hingst-Zaher, E., & Zaher, H. (2000). Application of landmark morphometrics to skulls representing the orders of living mammals. *Hystrix*, 11, 27–47.
- Medawar, P. B. (1945). Size, shape, and age. In W. E. le Gros Clark & P. B. Medawar (Eds.), *Essays on growth and form presented to D’Arcy Wentworth Thompson* (pp. 157–187). Clarendon Press.
- Müller, G. B. (Ed.). (2017). *Vivarium: Experimental, quantitative, and theoretical biology at Vienna’s Biologische Versuchsanstalt*. Vienna series in theoretical biology. MIT Press.
- Mumford, D., Series, C., & Wright, D. (2006). *Indra’s pearls: The vision of Felix Klein*. Cambridge University Press.
- Przibram, H. (1923). *Aufbau mathematischer. Biologie Abhandlungen zur theoretischen Biologie* (Vol. 18). Verlag von Gebrüder Borntraeger.

- Richards, O. W., & Kavanagh, A. J. (1945). The analysis of growing form. In W. E. le Gros Clark & P. B. Medawar (Eds.), *Essays on growth and form presented to D'Arcy Wentworth Thompson* (pp. 188–230). Clarendon Press.
- Rohlf, F. J., & Bookstein, F. L. (Eds.). (1990). *Proceedings of the Michigan Morphometrics Workshop*. Special Publication No. 2, University of Michigan Museum of Zoology.
- Siddiqi, K., & Pizer, S. M. (2008). *Medial representations: Mathematics, algorithms and applications*. Springer.
- Sneath, P. H. A. (1967). Trend-surface analysis of transformation grids. *Journal of Zoology of London*, 151, 65–122.
- Sokal, R. R., & Sneath, P. H. A. (1963). *Principles of numerical taxonomy*. W. H. Freeman.
- Thompson, D'A. W. (1917). *On growth and form*. Cambridge University Press. Abridged and edited by J. T. Bonner, Cambridge University Press, 1961.