KENDALL, W. S. (1988). Symbolic computation and the diffusion of shapes of triads. *Adv. in Appl. Probab.* **20** 775–797.

LE, H.-L. (1987a). Explicit formulae for polygonally-generated shape-densities in the basic tile. *Math. Proc. Cambridge Philos. Soc.* **101** 313–321.

LE, H.-L. (1987b). Singularities of convex-polygonally generated shape-densities. *Math. Proc. Cambridge Philos. Soc.* **102** 587–596.

MARDIA, K. V., EDWARDS, R. and PURI, M. L. (1977). Analysis of central place theory. *Bull. Internat. Statist. Inst.* **47** (2) 93–110.

MILES, R. E. (1970). On the homogeneous planar Poisson point process. *Math. Biosci.* **6** 85–127.

MILES, R. E. (1974). A synopsis of 'Poisson flats in euclidean spaces'. In *Stochastic Geometry* (E. F. Harding and D. G. Kendall, eds.) 202–227. Wiley, London.

MOODY, J. E., TURNER, E. L. and GOTT, J. R. (1983). Filamentary galaxy clustering: A mapping algorithm. *Astrophys. J.* **273** 16–23.

O'NEILL, B. (1966). The fundamental equations of a submersion. *Michigan Math. J.* **13** 459–469.

O'NEILL, B. (1967). Submersion and geodesics. *Duke Math. J.* **34** 363–373.

SMALL, C. G. (1981). Distributions of shape and maximal invariant statistics. Ph. D. dissertation, Univ. Cambridge.

SMALL, C. G. (1982). Random uniform triangles and the alignment problem. *Math. Proc. Cambridge Philos. Soc.* **91** 315–322.

SMALL, C. G. (1988). Techniques of shape analysis on sets of points. *Internat. Statist. Rev.* **56** 243–257.

Further references will be found in the survey papers by Bookstein (1986), Kendall (1984, 1986, 1989) and Small (1988).

Comment

Fred L. Bookstein

The elegant metric geometry of David Kendall's shape spaces Σ_m^k is inherited from the Euclidean metric of the spaces $\mathbf{R}^m$ containing the original point data. In the applications he has sketched here, the points in $\mathbf{R}^m$ are independent and identically distributed (iid) and the metric in shape space, in turn, is symmetric in the points, a sort of spherical distance. Point data generated in other disciplines, however, are not always iid; different metrics may be appropriate to those applications. In this comment I justify a certain analysis of small regions of Kendall's shape space by using a metric quite different from the usual Euclidean-derived version, depict its relation to Kendall's metric and indicate the sort of inquiries it permits.

Morphometrics is the quantitative description of biological form. Its data can often be usefully modeled as sets of labeled points, or landmarks, that correspond for biological reasons from organism to organism of a sample (Bookstein, 1986). We say that these points are biologically homologous among a series of forms: they have identities—names—as well as locations in some Cartesian coordinate system. Any set of landmark locations has a "size" and a "shape" that may be construed according to Kendall's definitions. But the biological relations among different instances of such configurations partake of a feature space not effectively represented by the metric inherited from $\mathbf{R}^2$ or $\mathbf{R}^3$.

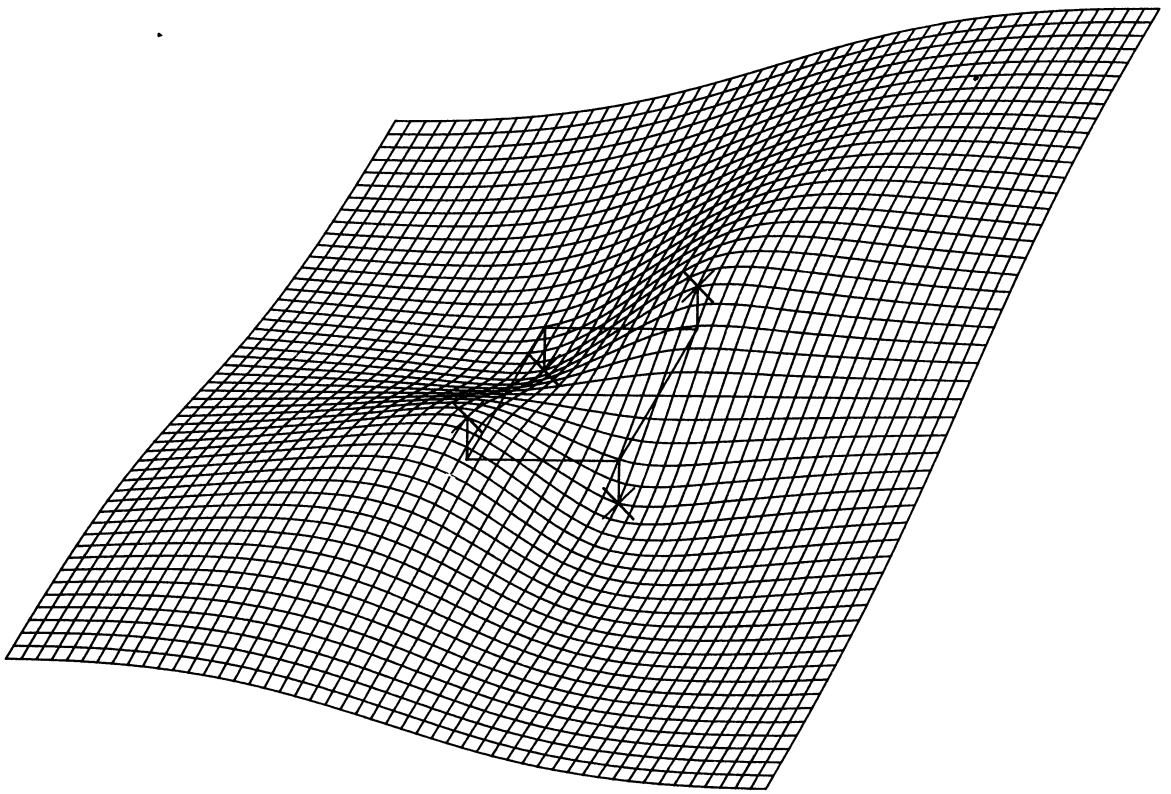
Fred L. Bookstein is Research Scientist, Center for Human Growth and Development, University of Michigan, Ann Arbor, Michigan 48109-0406.

In the biological context, my style of statistical analysis of shapes proceeds, as Kendall pointed out in 1986, within a tangent space of his Σ_2^k or Σ_3^k in the vicinity of a sample "mean form." (Small (1988) has an interesting comment on this construction.) The questions that in Kendall's applications are asked of an entire shape space—questions about concentration upon the "collinearity set," and the like—are replaced in morphometric applications by the more familiar concerns of multivariate statistical analysis: differences of mean shape, covariances involving shape or factors that may underlie shape variation.

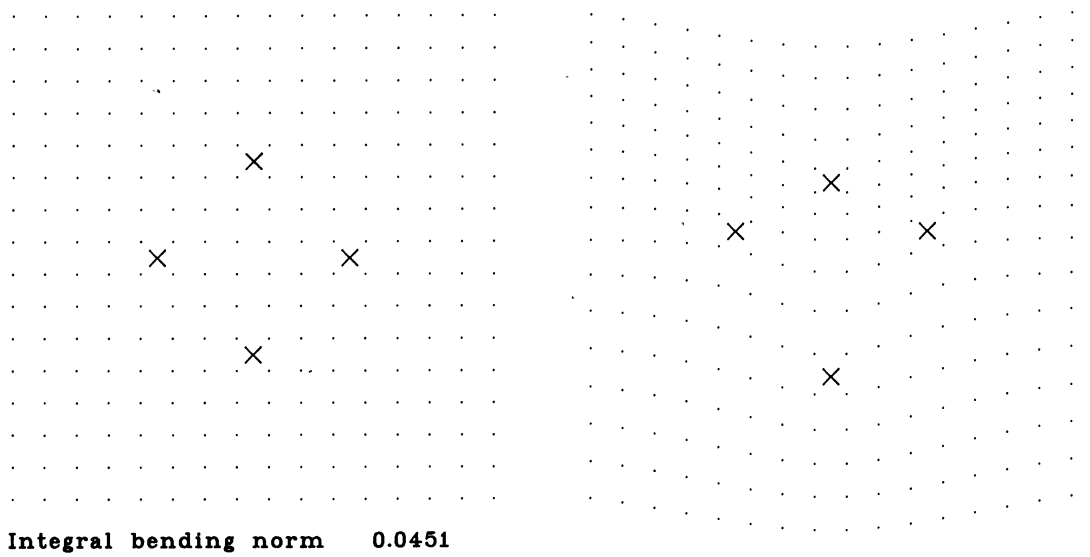
In the linearization of Kendall's shape space that applies to this tangent structure, the natural shape metric is an algebraic transformation of the "Procrustes metric," the ordinary summed squared Euclidean distance of two-point configurations after an appropriate optimizing rotation and scaling. But the Procrustes approach is not flexible enough fairly to represent biological structure within the context of multivariate statistical analysis. If two landmarks are typically close together, like the pupil of the eye and the outer corner, then we expect them to move together in their relation to more distant structures. The half-width and the orientation of the eye are more tightly controlled by diverse biological processes of regulation than is, say, the distance from the eye to the chin. These considerations lead one naturally to search out a shape metric that weights changes in small distances more heavily than changes in larger ones. In 1985, David Ragozin of the University of Washington suggested to me that the formalism of

thin-plate splines be considered in this connection. The suggestion has kept me quite busy ever since.

Consider (Figure 1A) an infinite thin metal sheet, originally flat but now displaced perpendicularly to itself at a discrete set of points. In the figure, the points are displaced by a rigid armature in the form of a square (here viewed in perspective). The end points of one diagonal of the square are displaced by



A



B

FIG. 1. *Thin-plate splines. A, the shape of an infinite thin uniform metal plate originally level, then displaced (by the contrivance in the diagram) upward at the ends of one diagonal of a square, downward at the ends of the other diagonal. The plate takes a shape describable as a superposition of fundamental solutions* $U(r) = r^2 \log r^2$ *of its partial differential equation, as given in the text. B, reinterpretation of the same scene as a deformation. The vertical coordinate of frame (A) has been combined with the in-plane Cartesian coordinate along one diagonal. The square mesh of dots on the left is transformed by this deformation into the meshwork on the right. The "bending energy distance" between the shapes in (B) is taken to be the (idealized) physical bending energy of the equivalent displacement pattern in (A).*

equal distances upward with respect to the rest of the plate; those of the other, downward by the same distance.

Producing these displacements of the plate required that physical work be done in the course of bending. The configuration that the plate "actually" occupies has the minimum of bending energy consistent with the four-point constraint applied by the armature. Let us introduce the notation $U(r) = r^2 \log r^2$. This function satisfies the equation $(\partial^2/\partial x^2 + \partial^2/\partial y^2)^2 U \propto \delta_{0,0}$, the equation governing bending of the plate when the displacements are sufficiently small and elastic effects within the plane of the plate can be neglected. Under these assumptions, the formula for the final form of the plate can be written in terms of the function $U(r)$: the plate takes the shape

$$z(x,y) \propto U(\sqrt{\{x^2+[y-1]^2\}}) - U(\sqrt{\{[x+1]^2+y^2\}}) + U(\sqrt{\{x^2+[y+1]^2\}}) - U(\sqrt{\{[x-1]^2+y^2\}}).$$

The functions $U(r)$ based at each of the four corners of the square are taken with coefficients +1 for the ends of one diagonal, −1 for the ends of the other, just as were the displacements of the armature.

The metric I am suggesting for shape change is equivalent to the bending energy of configurations like these. We realize it in the context of morphometrics by conflating the z axis (the perpendicular to the plate) with one of the in-plane directions, for instance, that of one diagonal of the armature. There results (Figure 1B) a deformation throughout the plane of that starting square. Reversing this procedure, the shape difference of the two sets of points shown, the square (before "bending") and the kite (after "bending"), becomes the bending perpendicular to itself of the square in Figure 1A, and is assigned a shape "distance" representing the energy of that bending, the work required to attach the armature to the plate. For small changes of shape, this measure is symmetric.

The computation of this bending energy for more general starting configurations of points than squares proceeds along the following lines (Bookstein, 1988a, b, 1989). Let $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, $\cdots$, $P_n = (x_n, y_n)$ be n points in the ordinary Euclidean plane according to any convenient Cartesian coordinate system. Write $r_{ij} = |P_i - P_j|$ for the distance between points i and j. Define matrices

$$K = \begin{bmatrix} 0 & U(r_{12}) & \cdots & U(r_{1n}) \\ U(r_{21}) & 0 & \cdots & U(r_{2n}) \\ \cdots & \cdots & \cdots & \cdots \\ U(r_{n1}) & U(r_{n2}) & \cdots & 0 \end{bmatrix}, \quad n \times n;$$

$$P = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \cdots & \cdots & \cdots \\ 1 & x_n & y_n \end{bmatrix}, \quad 3 \times n;$$

and

$$L = \left[\begin{array}{c|c} K & P \\ \hline P^T & 0 \end{array}\right], \quad (n+3) \times (n+3),$$

where ^{T} is the matrix transpose operator and 0 is a 3×3 matrix of 0's.

Let $V^T = (v_1, \cdots, v_n)$ be any n vector and write $Y^T = (V^T \mid 0\,0\,0)$. Define the vector $W^T = (w_1, \cdots, w_n)$ and the coefficients a_1, a_x, a_y by the equation

$$Y^T L^{-1} = (W^T \mid a_1\, a_x\, a_y).$$

Use the elements of $Y^T L^{-1}$ to define a function $f(x, y)$ everywhere in the plane

$$f(x,y) = a_1 + a_x x + a_y y + \sum_{i=1}^{n} w_i U(|P_i - (x,y)|).$$

Then the following three propositions hold:

1. $f(x_i, y_i) = v_i$, all i.
2. The function f minimizes the nonnegative quantity

$$I_f = \int\int_{\mathbf{R}^2} \left(\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 \right)$$

over the class of such interpolants. This is a constant multiple of the physical bending energy referred to above.

3. The value of I_f is proportional to

$$W^T K W = V^T (L_n^{-1} K L_n^{-1}) V,$$

where L_n^{-1} is the upper left $n \times n$ subblock of L^{-1}. This form is zero only when all the components of W are zero: in this case, the computed map is $f(x, y) = a_1 + a_x x + a_y y$, a linear or uniform transformation.

In the present application we take V to be the $2 \times n$ matrix

$$V = \begin{bmatrix} x_1' & y_1' \\ x_2' & y_2' \\ \cdots & \cdots \\ x_n' & y_n' \end{bmatrix}$$

where each (x_i', y_i') is a point "homologous to" (x_i, y_i) in another copy of $\mathbf{R}^2$. The resulting function f now maps each point (x_i, y_i) to its homologue (x_i', y_i') and is least bent (according to the measure I_f, integral quadratic variation over all $\mathbf{R}^2$, computed separately for real and imaginary parts of f and summed) of all such functions.

In effect, our metric is the bending energy of a four-dimensional thin plate; two dimensions of plate displaced in two "other" perpendicular directions. A similar computation can be mounted in three dimensions, using the basis function $U_3(r) = |r|$, to represent the bending energy of a "six-dimensional thin plate." In one dimension, $U_1(r) = |r|^3$ gives the ordinary cubic spline.

The matrix $L_n^{-1}KL_n^{-1}$, which will be called the bending energy metric in the discussion to follow, is not of full rank. It annihilates affine transformations (uniform shears) of the landmark configuration as a whole. Physically, these play the role of vertical shifts and uniform tilts of the real metal plate; under our assumption of no in-plane elasticity terms, these "deformations" proceed without any energy cost. In the context of the biological applications driving this discussion, those transformations are the uniform transformations lacking any local features (gradients, growth centers or the like). The biological meaning of bending energy is thus a sort of measure of local information required to specify the landmark reconfiguration.

In the context of shape change, the Procrustes metric becomes the summed squared point displacement from one form to another, which is to say, the summed squared point displacements normal to the thin plate in the two "additional" (nonphysical) directions. Because of the annihilation of all shears, the metric geometry of shape space induced by bending energy is quite different from that driven by the Procrustes metric. (For instance, all parallelograms are at energy distance zero from each other.) We may summarize the differences by computing an eigenanalysis of the bending energy metric with respect to the Procrustes. This is just the ordinary eigendecomposition of $L_n^{-1}KL_n^{-1}$. The relative spectrum of the bending energy matrix always has three zero eigenvalues whose eigenvectors span the affine transformations. The remainder of the spectrum is a series of what I have named principal warps, normal modes of deformation that may be ordered by bending energy per unit Procrustes displacement.

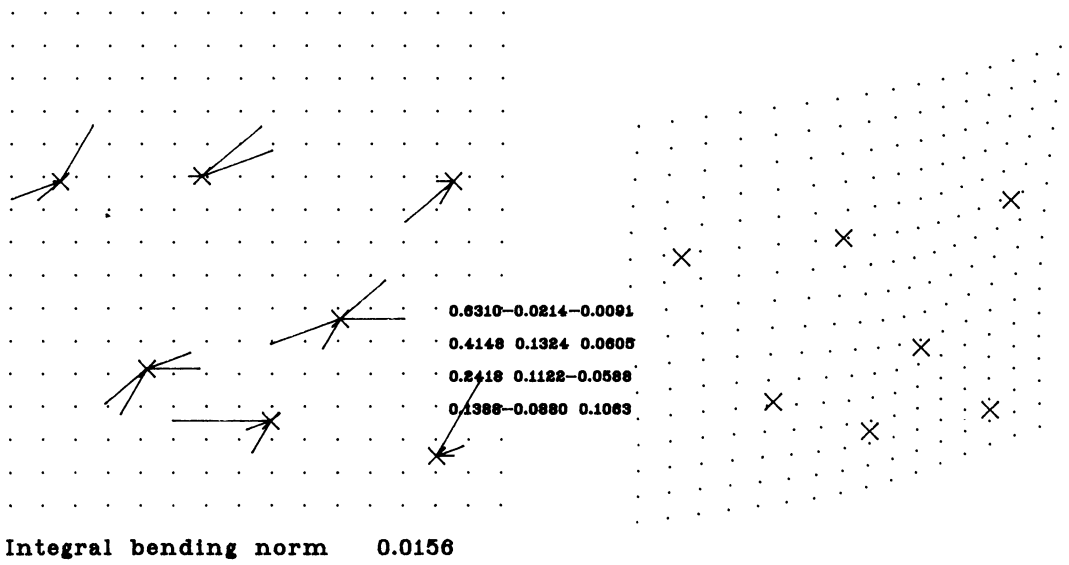
An example of this spectrum is shown in Figure 2 for a configuration of seven landmarks drawn from a study in craniofacial syndromology (see Bookstein, 1989, Section 7.5). Frame (A) shows the mean configuration of these seven points in a (biologically) normal sample and in a sample of 14 cases of Apert Syndrome. The eigenvectors of bending energy may be shown all at once as distinctive patterns of simultaneous in-plane displacement of all the landmarks in parallel (the segments at intervals of 20° in frame (A)). Alternatively, they may be considered individual patterns of displacement perpendicular to the picture, as drawn in frames (B)–(E). Notice that the eigenvectors of higher eigenvalue look more bent per net (Procrustes) vertical displacement, and that the larger principal bending modes somewhat resemble that in Figure 1, the square-to-kite bend, at diverse scales.

The annihilation of those three degrees of freedom for the affine changes alters the metric geometry of shape space severely. The set of forms at constant bending energy distance from a fixed form is not a hypersphere but a hyperellipsoidal cylinder heading out to "infinity" in the tangent space. Figure 2, B–E, may be reinterpreted as the semiaxes of sections of this cylinder. The cylinder is not general: to each thin plate pictured there correspond two equal semiaxes representing the same in-plane displacements rotated by 90°. Thus the full constant-distance locus is the extended Cartesian product of an infinite flat hyperplane by a series of circles of different radii. The changes in Figure 2, B–E, are all (Procrusteanly) orthogonal to three dimensions of generators, the space of all the affine transformations, which have bending energy zero regardless of Procrustes length. If the mean configuration of landmarks were different, so, too, would be this pattern of axes, the geometry of "normal sections" of the cylinder.

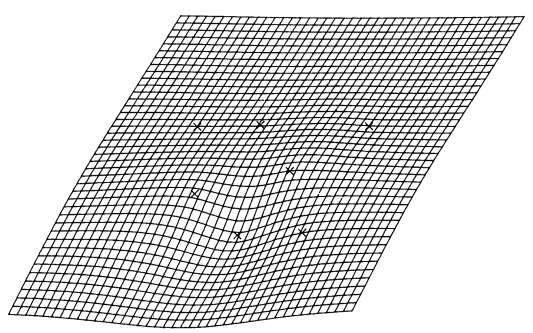
The affine features of change annihilated by this shape metric can be restored by separately measuring that aspect using the Poincaré hyperbolic metric (see Bookstein, 1986, 1989). There results a joint feature space with the same dimensionality as Kendall's Σ_2^k, but with a metric now Galilean in the small (cf. Yaglom, 1979), so that there is no possibility of "rotation" between affine and bending parts; they are incommensurate. This is analogous to the treatment of real (physical) spacetime in Newtonian mechanics, for which space is measured in centimeters and time in seconds, with no possibility of interconversion by purely geometric maneuvers. Any relation between space and time is coded instead in a Newtonian velocity, a vector of change in space per change in time. That is, the net vector spline $L^{-1}Y$ has no unitary "length" or "direction." Rather, its affine part has a geometry that may usefully be taken as hyperbolic in the large (cf. Small, 1988), and its nonlinear part W has a geometry that is cylindrical Euclidean. These two conceptually independent aspects of measurement relate only by a joint distribution, such as a covariance structure.

With the help of the bending energy metric, we can refer to localization of biological variability and to the apparent physical scale of that variability. Neither of these is possible using the Procrustes metric. A useful first step is the relative eigenanalysis of the covariance matrix of the shape coordinates $(Z_k - Z_1)/(Z_2 - Z_1)$ with respect to the bending energy metric. Some delicacies of the computation are discussed in Bookstein (1989, Section 7.6).

For example, Figure 3 presents an analysis of the positions of eight cranial landmarks (frame (A)) in the skulls of 21 7-day-old male rats. (The data were originally drawn by Henning Vilmann, the landmarks digitized by Melvin Moss.) Our concern is with large-scale shape regulation in this diadem of sutures around the rapidly developing brain. Frame (B) indicates the five principal warps of this configuration of landmarks

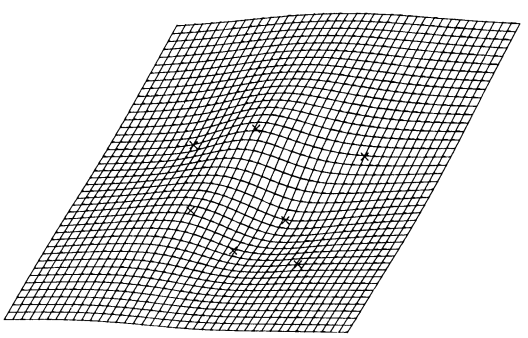


A



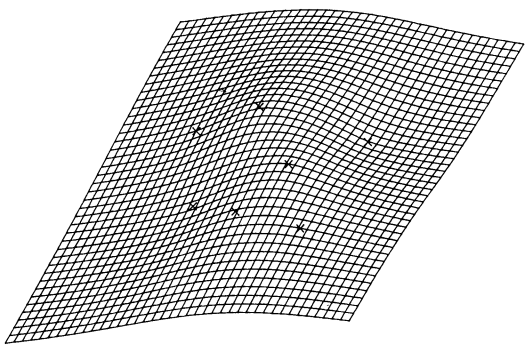
eigenvector 1

B



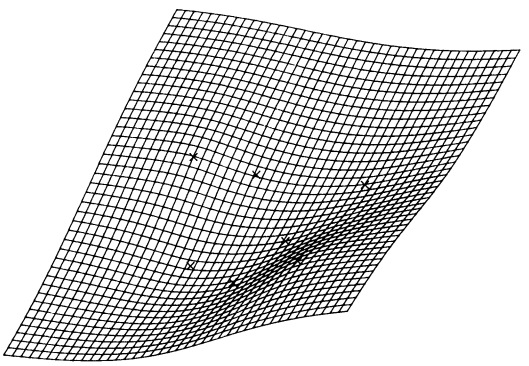
eigenvector 2

C



eigenvector 3

D



eigenvector 4

E

FIG. 2. *Spectrum of the bending energy with respect to the Procrustes metric. A, pattern of seven landmarks, and the four principal warps they generate, as described in the text. Segments along the horizontal: principal warp of largest bending energy,* 0.631 *(arbitrary units). At* 20° *counterclockwise of horizontal: eigenvector of second largest bending energy,* 0.415. *At* 40°: *third stiffest principal warp, eigenvector* 0.242. *At* 60° *counterclockwise of horizontal: eigenvector of smallest nonzero bending energy,* 0.139. *In the interpretation as deformation these multiply any two-vector to supply one displacement at each landmark. Also shown, by its effect upon the square grid at left, is the thin-plate spline mapping these seven landmarks to the homologous configuration on the right. The table in the center presents the loadings of the x- and y- components of this transformation upon each principal warp in turn: e.g., the y- component is mostly warp* 4, *the x- component the sum of warp* 2 *and warp* 3. *The data represent mean forms, as observed in lateral cephalograms, for a clinical sample of Apert Syndrome (right) and matched Ann Arbor normals (left). Landmarks, clockwise from upper left: sella, sphenoethmoid registration, nasion, anterior nasal spine, inferior zygoma, pterygomaxillary fissure (for definitions, see Riolo, Moyers, McNamara and Hunter,* 1974). *The "interior" landmark is orbitale. The transformation includes an affine component with strains of* 0.71 *along and* 0.93 *perpendicular to the direction at* 31° *counterclockwise of vertical on the left,* 37° *on the right. For the meaning of this example, see Bookstein* (1989, *Section* 7.5). *B–E, interpretation of each of the four eigenvectors as its own thin-plate spline: pairs of semiaxes of the bending energy cylinder.*

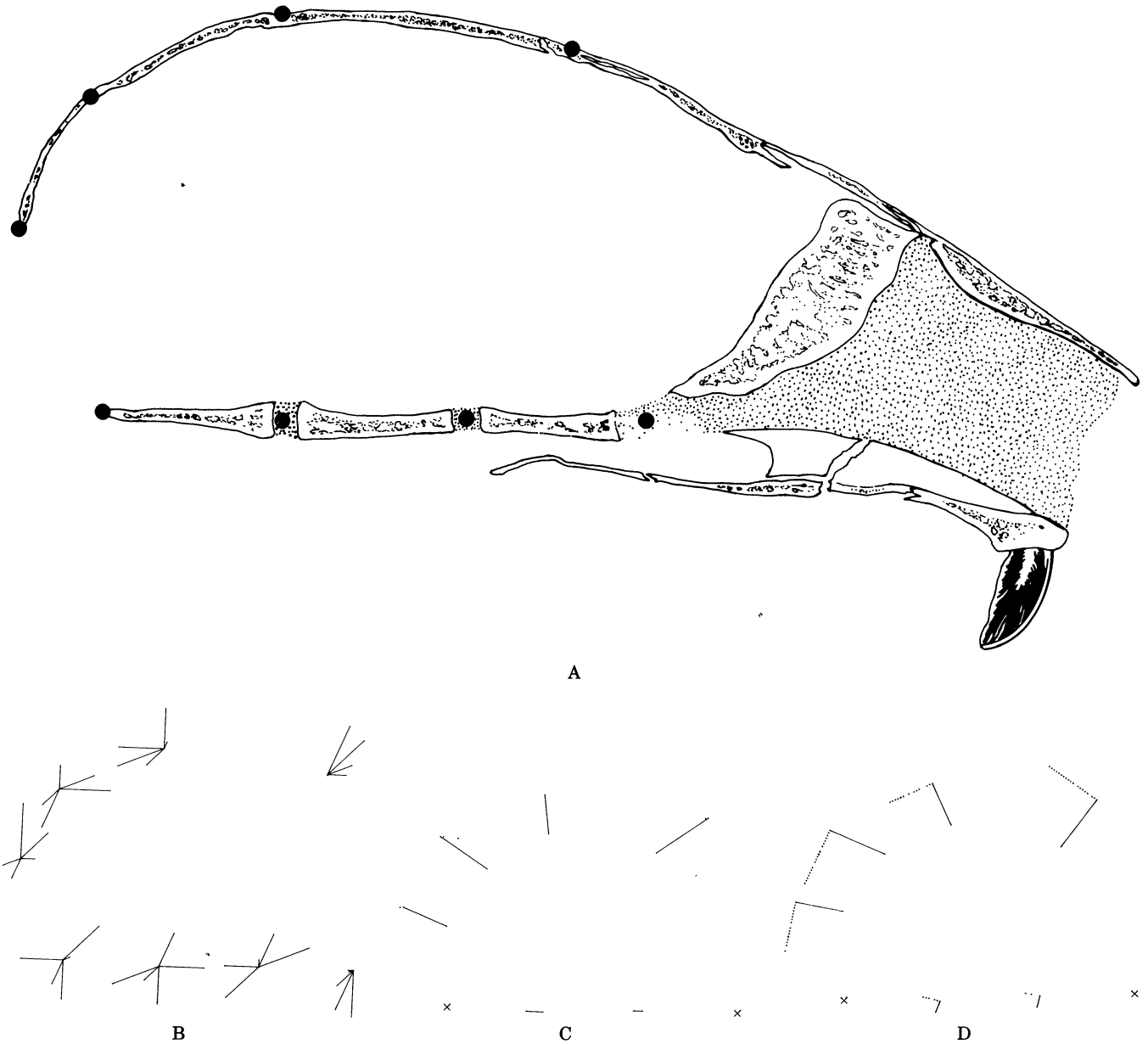


FIG. 3. *Relative bending energy component analysis of* 8 *cranial landmarks for* 21 *laboratory rats aged* 7 *days. A, midsagittal section: the landmarks* in situ. *B, the five principal warps, counterclockwise from the horizontal in decreasing order of bending energy (eigenvalues* 8.3, 6.2, 3.3, 1.8, *and* 1.1 *in arbitrary units). C, the last two principal warps as in-plane displacements in shape coordinate space (two landmarks fixed). Solid line, largest scale; dotted line, second largest scale. To each principal warp correspond two of these patterns, the one shown and its rotation by* 90° *at each landmark. D, the first and second relative eigenvectors of greatest sample variance relative to bending energy: the first and second components of scaled shape variance. Solid line, dominant relative eigenvector (eigenvalue* 0.0067); *dotted line, second relative eigenvector (eigenvalue* 0.0047).

after the fashion of Figure 2A, as parallel in-plane displacements, and frame (C) shows the last two of these (the bends of largest spatial scale) as in-plane displacements from the sample mean locations in shape coordinates (i.e., with two landmarks held fixed, as shown). The relative eigenanalysis of the sample shape coordinate covariance matrix with respect to bending energy has eigenvalues of 0.0067, 0.0047, 0.0003, $\cdots$, 0.0000. (There are ten in all, five principal warps times two Cartesian coordinates.) Figure 3D shows the first two of these relative eigenvectors (all others are trivial). The first of these connotes a generalized enlargement of upper cranial structures relative to the lower. The component is of highest relative eigenvalue in part because it has the lowest bending energy: compare the solid segments between frames (C) and (D). The second relative eigenvector, the dotted pattern in Figure 3D, represents rotation of

the upper margin of the braincase with respect to the lower margin. It is not at all equivalent to the second weakest principal warp (dotted lines, frames (C) and (D)). The bending energy eigenanalysis has extracted these large scale patterns of shape covariance by explicitly weighting empirical covariance patterns inversely to geometric localizability. Other equally plausible geometric patterns, such as bending of the upper or lower structures, are not observed to bear any sample variance.

The example suggests the descriptive possibilities inherent in accommodating the metric geometry of Kendall's shape space to a biological subject matter. One can imagine other modifications of the metric in response to other contexts than the biometric. For instance, one can imagine the statistical study of the positions of a robot arm. When the state of the linkage is coded by the coordinates of its joints, then because certain parts of the robot are rigid, an appropriate measure of "distance" would be somewhat altered from the Procrustes. In another sort of constraint, certain "landmarks" might represent the loci of curves in the data—boundary arcs not otherwise labeled—and would thus be "deficient" by one coordinate; again the Procrustes metric needs to be modified. In a study of schools of fish, or flocks of birds, an appropriate shape metric might be the Cartesian product of a biological shape space by a hydro- or aerodynamic one (for the V of migrating geese, for instance). Yet other modifications would arise when the points of Kendall's space are "colored" in classes whose separate patterns cannot be usefully studied without reference to their interpenetration, as in problems of multispecies ecology. These and other possibilities represent an enrichment of the metric geometry of shape space within the global purview pursued so sparely and elegantly by David Kendall.

ACKNOWLEDGMENTS

Preparation of this comment was partially supported by Grant GM-37251 from the National Institutes of Health to F. L. Bookstein. Christopher Small commented usefully on an earlier draft. I and many of the rest of us in shape statistics are grateful to Colin Goodall for arranging a workshop around Kendall's Wilks Lectures, Princeton, May 1987. Part of this comment was presented at that workshop in even more preliminary form.

ADDITIONAL REFERENCES

BOOKSTEIN, F. L. (1988a). Toward a notion of feature extraction for plane mappings. In *Proc. Tenth International Conference on Information Processing in Medical Imaging* (C. de Graaf and M. Viergever, eds.) 23–43. Plenum, New York.

BOOKSTEIN, F. L. (1988b). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Machine Intelligence*. To appear.

BOOKSTEIN, F. L. (1989). *Morphometric Tools for Landmark Data*. In preparation.

RIOLO, M. L., MOYERS, R. E., MCNAMARA, J. S. and HUNTER, W. S. (1974). *An Atlas of Craniofacial Growth*. Center for Human Growth and Development, Univ. Michigan.

YAGLOM, I. M. (1979). *A Simple Non-Euclidean Geometry and its Physical Basis*. Springer, New York.

Comment

Christopher G. Small

With a high standard of rigor David Kendall has given us an interesting survey of the theory of shape analysis that he has pioneered with the help of others over the last decade. This work is now of sufficient volume that the many topics discussed in this survey can be only briefly touched upon. I certainly hope that this paper is a stimulus to additional consideration of this topic by statisticians. It may well be that on future occasions the topologists will have to introduce their theory of shape with preparatory remarks to the effect that it is not to be confused with the growing statistical theory of shape.

At first glance, this paper might seem to have much in common with the differential geometric techniques in statistics that are associated with Amari (1985) and others. However, despite the abstraction of some of the theory, the methods of Kendall are essentially data analytic rather than model theoretic: the differential geometry is on the sample space not the parameter space. So how much differential geometry must the data analyst know in order to implement the techniques that are described in this paper? The answer is largely dependent on the amount of software

Christopher G. Small is Associate Professor, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.