ELSEVIER

Bulletin of Mathematical Biology, Vol. 58, No. 2, pp. 313–365, 1996
Elsevier Science Inc.

0092-8240/96 $15.00 + 0.00

0092-8240(95)00329-0

BIOMETRICS, BIOMATHEMATICS AND THE MORPHOMETRIC SYNTHESIS

■ FRED L. BOOKSTEIN
University of Michigan,
300 North Ingalls Building,
Ann Arbor, MI 48109-2007 U.S.A.

(*E-mail: fred@brainmap.med.umich.edu*)

At the core of contemporary *morphometrics*—the quantitative study of biological shape variation—is a synthesis of two originally divergent methodological styles. One contributory tradition is the multivariate analysis of covariance matrices originally developed as biometrics and now dominant across a broad expanse of applied statistics. This approach, couched solely in the linear geometry of covariance structures, ignores biomathematical aspects of the original measurements. The other tributary emphasizes the direct visualization of changes in biological form. However, making objective the biological meaning of the features seen in those diagrams was always problematical; also, the representation of variation, as distinct from pairwise difference, proved infeasible.

To combine these two variants of biomathematical modeling into a valid praxis for quantitative studies of biological shape was a goal earnestly sought though most of this century. That goal was finally achieved in the 1980s when techniques from mathematical statistics, multivariate biometrics, non-Euclidean geometry and computer graphics were combined in a coherent new system of tools for the complete regionalized quantitative analysis of *landmark points* together with the biomedical images in which they are seen.

In this morphometric synthesis, correspondence of landmarks (biologically labeled geometric points, like "bridge of the nose") across specimens is taken as a biomathematical primitive. The shapes of configurations of landmarks are defined as equivalence classes with respect to the Euclidean similarity group and then represented as single points in David Kendall's *shape space*, a Riemannian manifold with Procrustes distance as metric. All conventional multivariate strategies carry over to the study of shape variation and covariation when shapes are interpreted in the tangent space to the shape manifold at an average shape. For biomathematical interpretation of such analyses, one needs a basis for the tangent space compatible with the reality of local biotheoretical processes and explanations at many different geometric scales, and one needs graphics for visualizing average shape differences and other statistical contrasts there. Both of these needs are managed by the *thin-plate spline*, a deformation function that has an unusually helpful linear algebra. The spline also links the biometrics of landmarks to deformation analysis of the images from which the landmarks originally arose.

This article reviews the history and principal tools of this synthesis in their biomathematical and biometrical context and demonstrates their usefulness in a study of focal neuroanatomical anomalies in schizophrenia.

1. Introduction. Over most of this century, techniques for quantitative geometric study of organic form have fallen under one of two incommensurate headings. In one style of analysis, deriving from the biometrics of Karl Pearson and Sewall Wright, conventional multivariate techniques are applied to an undisciplined roster of quantifications of single forms. The only algebraic structures involved are those of multivariate statistics, limited mainly to partitions of sums-of-squares, diagonalizations of covariance matrices and solution of linear systems in which they are involved. No aspect of the geometric organization of the measures or their biological rationale is reflected in these statistical maneuvers. In particular, the geometry of the typical form plays no formal role in the analysis of variation.

In the other class of shape analyses, often associated with the name of D'Arcy Thompson, but actually dating from the discovery of perspective transformations half a millenium ago, changes of biological form are visualized directly as distortions of Cartesian coordinate systems that carry meaningful biological labels right along with the coordinate grid. Such analyses are inextricably graphical; generation after generation of devoted amateurs failed to produce any rigorous grammar for the quantitative aperçus to which they sometimes lead.

The incompatibility between these two main styles of quantification derives ultimately from a discrepancy between two fundamental metaphors by which biomathematical abstractions can sometimes acquire biotheoretical meaning. In the multivariate approach, prior biomathematical knowledge pertains to the quantities that characterize the single form. When forms are measured by ruler, for instance, it must simply be assumed that the interpoint separations are biotheoretically comparable; there is no way to test this axiom. In the second approach, comparability (called *homology* in this context) refers instead to the pairing of "corresponding" locations of bits of tissue; there is no practical way to test this axiom either. Whereas the first language makes no reference to biological processes, and thereby supplies no filter by which inappropriate comparisons might be suppressed, the second provides insufficient hints about the syntax of quantification, lacking, for instance, stategies for integrating descriptions across wide extents of a form or across the forms of a sample.

These two analytic traditions were pursued separately right through the 1970s, yet in the middle 1980s, without any premonitory ferment, the barrier between them was very rapidly breached by an unprecedented combination of algebraic and geometric tactics.

The breakthrough began—as breakthroughs in quantitative science often do—when it was realized that previous attempts at a morphometric synthesis had been struggling toward the wrong goal. What constituted the

appropriate subject matter for analysis of forms with labels was not, as I erroneously (if understandably) claimed in Bookstein (1978), the construction of a canonical coordinate system for D'Arcy Thompson's grids, which apply to only two forms at a time. Those grids were an artifact. Efforts at improving their quantification only distracted from the far more important task of constructing a canonical manifold for biomathematically salient aspects of shape across extended samples. In the tangent space of that manifold, any vector becomes a potential "shape process" that summarizes evidence from a sample. Conversely, any pairing of forms, as of individual specimens with their common average, becomes a descriptor that, after projection onto the tangent space, can be aggregated with other commensurate descriptors, examined for patterns and associated with explanations. Within this shape space, trends can be named and their statistical reifications can be assessed; the reliable diagrams that obtain of actual effects on real shapes are often very conducive to biomathematical insights.

To combine biomathematical with biometrical aspects of biological shape studies, it is first necessary to demarcate the boundary between them. Neither biomathematics nor biometrics taken separately is capable of this degree of insight. In suggesting an elegant graphic for the biologist's intuition of shape change, Thompson ignored the biometrics of shape description and shape variation. Conversely, in suggesting elegant techniques for the algebraic/statistical manipulation of shape descriptions, the multivariate school ignored the problem of making biotheoretical sense of the abstract combinations so produced. Once we could recognize prototypes for biomathematically sensible description of quantitative effects upon biometric shape, the analytic tools so long sought could be generated and annotated almost as quickly as the community could assemble exemplary data sets.

As it happened, the modern discipline of morphometrics was synthesized over a mere five years. The original insight occurred more or less simultaneously to three of us. (To the sociologist of science, that is a familiar signal that the field was ripe for metamorphosis.) At the same time that a paper of mine (Bookstein, 1984b) introduced two-point shape coordinates for triangles and showed how shape differences can be weighed by formal T^2 test, Goodall's (1983) dissertation derived the equivalent F-ratio while avoiding any size standardization. Meanwhile, Kendall (1984) announced the Riemannian structure of the global shape spaces, noting later, in one pregnant sentence, that my method, like Goodall's, pertained to the tangent spaces to his shape manifolds. Our joint publication in the first volume of *Statistical Science* (Bookstein, 1986, with commentary) signaled the convergence of all three of these notations upon one single foundation for the biometric analysis of landmark data in a biomathematically interpretable framework

of shape processes. This core of material has since been formalized further, in a different notation, in Goodall (1991). Meanwhile, one particular interpolation function—the thin-plate spline (Bookstein, 1989a)—turned out to support a feature space for these shapes in an uncanny way: A quadratic form, the computation of which involves *only* the mean landmark configuration, generates a biomathematically sensible basis for reporting and interpreting variations *around* that mean.

The most extensive exposition of this synthesis is my monograph (Bookstein, 1991). However, the Procrustes underpinnings of the synthesis are not emphasized there as much as I now would. Several proceedings volumes (Rohlf and Bookstein, 1990; Marcus *et al.*, 1993; Marcus *et al.*, 1995; Mardia and Gill, 1995) provide links to more classical languages of biometrics, biomathematics, mathematical statistics and systematics. The field remains in desperate need of a book-length primer; maybe some intrepid reader of this essay will be inspired to draft such a volume.

2. Biometric Analyses of Size and Shape Measures

2.1. *Historical background.* The sturdy algebraic structure that is linear multivariate statistics encourages applications from population biology through psychology and into the social sciences. Ironically, its techniques typically have arisen in response to specific challenges of size and shape analysis. For instance, the first quantitative study of human development was de Montbeillard's 1760 growth curve for the height of his son (see Boyd, 1980). Adolphe Quetelet relied on unsophisticated measures of height and weight to argue (quite fallaciously, of course) that normally distributed variates are generated by a single "true value" characterizing *un homme moyen*, a reified "average man." Additionally, Francis Galton's original example of regression fitted an ellipse to a cross-tabulation of height for 928 young men against the average of their parents' heights. Following Duncan (1984), I would suspect all this owes to the origin of these thrusts in the need for "social measurement" well before the idea of biometric statistics could be formulated. Generals and tailors needed to understand human size variability millenia before quantitative biology was more than an eccentric hobby.

Although all these were analyses of explicitly biometric data, the actual algebra of the techniques exploited (least-squares, correlations, fits to bell curves) does not articulate with any biomathematical models for the same data—the mathematics of cumulative height, for instance, as it is regulated by growth or inheritance. The power of biometric methods for broader applications is wholly by analogy. That multivariate analysis of covariance, for instance, works as well in political science as in agriculture owes to its

discarding the "bio" from the "metric" formalism at the outset. Outside of the specific morphometric context to be introduced presently, no multivariate biometric analysis offers any theory of where its measurements have originally come from.

Very early on in the development of multivariate statistics, it was realized that biometrics typically segregates multivariate algebra from the biomathematical context of the underlying quantifications. Unfortunately, both Galton and Karl Pearson subscribed to a mystical conflation of meaning between the fact of regression (that is, true, linear causation) and the convenient summary statistic of "co-relation." Out of this confusion emerged the British school of eugenics that has embarrassed so many of us since then (see Bookstein, 1995e). "When the British race is at risk" was no time for subtle questions about process. A century ago Pearson's colleague W. F. R. Weldon already was emphasizing the usefulness of a second formulation for pairs of biometric variables, such as alternate size measures of the same organism, in which the notion of causation was replaced by one of correlated response. The considerably different notion of "explanation" underlying this extension embodied the same least-squares logic for "combination of observations" that Gauss and Lagrange had worked out a century earlier (Stigler, 1986). That being so, the methods could be liberated from any biomathematical context, for instance, in G. Udny Yule's 1895 application of regressions for calibrating phenomena of social welfare.

In light of this double meaning, the understanding of regression and correlation in theoretical biology remained obscure until clarified by Sewall Wright in the 1920s. His method of *path analysis* (see Wright, 1968) applied a consistent language jointly to studies of inheritance of quantitative characters and to studies of multiple characteristics of the single organism. In this biomathematical strategy, observed correlations are the algebraic composite of patterns of mutual determination of data by observed or unobserved factors. Correlation coefficients are thus epiphenomena of the path coefficients—the only quantities in sight with a satisfactory epistemology. Wright's conception of the role of correlations in biometrics is still, in my view, the only coherent approach to their application in the biological sciences (see Bookstein *et al.*, 1985, or Bookstein, 1991).

While Wright was developing his rigorously causal models and enlarging their range to include the new abstractions emerging in population genetics by mid-century, the countervailing tradition of least-squares prediction was not dormant. Many intellectual developments sprang from the "general linear model" to play crucial roles in other areas of modern applied statistical practice, such as agricultural experimentation, econometrics, process control and psychometric test theory. None of these areas shares any theory with biomathematics nor are any of these extensions of Wright's

methods of much use in biometrics. Meanwhile, yet another development was arising in the context of morphometric data. The *locus classicus* for discriminatory analysis is a data set of four size measures of iris flowers. Fisher, in phrasing his problem of "discrimination" as the maximization of a certain ratio of statistical likelihoods, once again sequestered the algebraic core of the technique far away from any valid style of biomathematical reasoning: on what linear combinations of floral traits, indeed, is it sensible to suppose that selection can operate? Extensions of this method by Hotelling (1936) and others likewise continued to ignore biomathematical strictures while constructing the full range of canonical analyses of variance and covariance on which today's journeyman statistician relies.

By the 1960s, then, the biometrics of shape was entrapped in serious theoretical and practical difficulties. None of our techniques for shape analysis had any access to biomathematical theories. Our core statistical tactics—regression, factor analysis, canonical variates analysis—had arisen in the context of questions about shape, and yet in their current algebraic unfolding there was no possibility of any geometrical insights or interpretations at all. The question of whether the algebra of covariance matrices and design matrices did justice to the biological hypotheses investigated with their aid could simply not be posed.

For instance, the topic of Blackith and Reyment's (1971) book *Multivariate Morphometrics*, the first book to use the word "morphometrics" in its title, is actually the interpretation of matrix manipulations in vaguely functional biological terms. Variations in the metrology of morphometric data—lengths, angles, titres, proportions, in whatever combination—make no difference for these matrix mechanics: all quantities are thrown into the same crucible of canonical analyses and scatter plots. No discipline for the formulation of those variables could possibly arise in so pathological a context of tolerance. The authors of this canonical reference text saw no need to point out this lacuna or to include any diagrams dealing with the potential biomathematical meaning of their measures and their analyses.

In retrospect the difficulty is obvious: no matter how elegant one's matrix algebra, one cannot coherently interpret findings about size and shape without first saying what size and shape phenomena *mean*. From mid-century on, therefore, the occasional thoughtful biometrician attempted to modify the dominant matrix methods so as to encode some algebraic version of this biomathematical context. Easiest was the interpretation of size and shape that corresponded to one straightforward biomathematical model—*allometry*, the dependence of shape change upon size change. Rather more subtly than Wright had, Jolicoeur (1963), Hopkins (1966), Burnaby (1966) and others interpreted allometry as a single-factor biometric model, and suggested ways by which it might be calibrated using

empirical covariance matrices of log-transformed distance measures. Thus allometry can (sometimes) be not only detected but also described by variation of the coefficients of the first principal component of logarithms of size measures, analysis of "shape" can proceed (under fairly stringent conditions, and with limited power) using vectors of ratios of size measures, analysis of shape in a different sense, now no longer size-independent, can proceed by referring to residuals of the raw data from their allometric regressions and so on. This literature is summarized and assorted in Bookstein *et al.* (1985) and its semantics is dissected in Bookstein (1989b). Because single-factor models can be extended to distributions of indefinitely many measurements (as in classic psychometric test theory), it did not even matter that the dimensionality of such "size measures" was not finite.

Still, by about 1980 most of us were deeply troubled by the apparently ineluctable mismatch between the matrix operations of the dominant tradition, however modified for applications to the "and" of "size and shape," and the very reasonable sorts of questions about morphometric phenomena that had been asked of the raw data all along: where upon the organism, or its image, interesting patterns were to be unearthed, and whether sample covariance matrices offered any help in sharpening their detection and interpretation.

2.2. *A biometric shape space for landmark data.* Our field escaped from this impasse by learning how to restore the biomathematical context to the problem of shape description in a strikingly simple and explicit way. We had all construed geometric sizes as the primary stuff of biometrics. This was partly because size variables came, more or less, in the same units (Mosimann, 1970), so that their linear combinations and logarithms made sense, and partly because, in practice, primary shape data were an incoherent mass of form factors, angles and ratios subject to no apparent discipline. It had always been easier to treat this congeries as a family of flexibly derived descriptors emerging from subjective insight than as a "space" spanning some relevant finite-dimensional channel of biological description.

Yet our visual systems already construe shape as a primary observandum. We are hard-wired to recognize shapes as equivalence classes of forms under motions of the head: pedestrian changes of distance toward an object and rotations of the cranium on the atlas vertebra. This representation is finite dimensional whenever the primary datum is finite dimensional. To assure finitude, it is easiest to revert to the set of endpoints of the same rulers that were already measuring lengths, and close those sets. This leads directly to a formalism of landmark configurations as primary data (Book-

stein, 1978, 1991). More formally, a landmark configuration is a *discrete sample from a homology mapping across pairs of specimens*, a labeled series of points that are each homologous over an entire sample of organisms. The biometrical shape of a landmark configuration is best taken, unambiguously, as the equivalence class to which it is assigned under the group of Euclidean similarity transformations.

Multivariate statistics can be founded on interspecimen distances (cf. Gower, 1971) as easily as upon variables. This duality has long been recognized in particular applications, for instance, the interchangeability of "Q-mode" and "R-mode" factor analyses in psychometrics. Multivariate biometrical analysis of shape could postpone the generation of "variables," then, by beginning instead with a metric for distances among shapes taken, exactly as the eye takes them, as equivalence classes. Such a metric, a "distance" between labeled point sets that is invariant against similarity transformations, had been under development for some time in connection with other diverse sciences of space: the *Procrustes distance* between two landmark configurations. In the small, this distance is the sum of squared ordinary Euclidean distances between corresponding landmarks after each configuration is scaled to unit Centroid Size (sum-of-squares around its own centroid) and then one of the pair is rotated and translated upon the other so that that interspecimen sum-of-squares is a minimum (see Fig. 1). The preferred formula (Sibson, 1978; Kendall, 1984) is a transformation of that sum-of-squares that gives the resulting quantity the correct metric properties in the large. Specifically, for two-dimensional data it becomes identical to the Fubini-Study unitary-invariant metric for the ratios of $k-1$ complex numbers. Select representatives of the equivalence classes as vectors of complex numbers z_i with $\Sigma z_i = 0$ and $\Sigma z_i \bar{z}_i = 1$. Then the Procrustes distance between z and z' is $\arccos|\Sigma z_i \bar{z}'_i|$.

Without any further geometry we can already begin to carry out biometric analyses of shape. The *average shape* of a sample can be defined quite rigorously as the shape that has the least summed squared Procrustes distances to the individual shapes. Algorithms for computing this average are surprisingly simple. My favorite (Kent, 1994) is as the first principal component of the sample summed outer product Σzz^t, where zz^t is the $k \times k$ matrix whose ijth entry is $z_i \bar{z}_j$.

Kendall's (1984) great paper about Procrustes metrics and shape showed that the equivalence classes of two-dimensional landmark configurations under the similarity group form a smooth $(2k-4)$-dimensional manifold, the *shape manifold* Σ_2^k for k landmarks in two dimensions. Moreover, when shape distance is taken as exact Procrustes distance $\arccos|\Sigma z_i \bar{z}'_i|$, and *only* in that case, the shape manifold is the Riemannian manifold that expresses the original Euclidean geometry of the digitizing plane as partialed by the

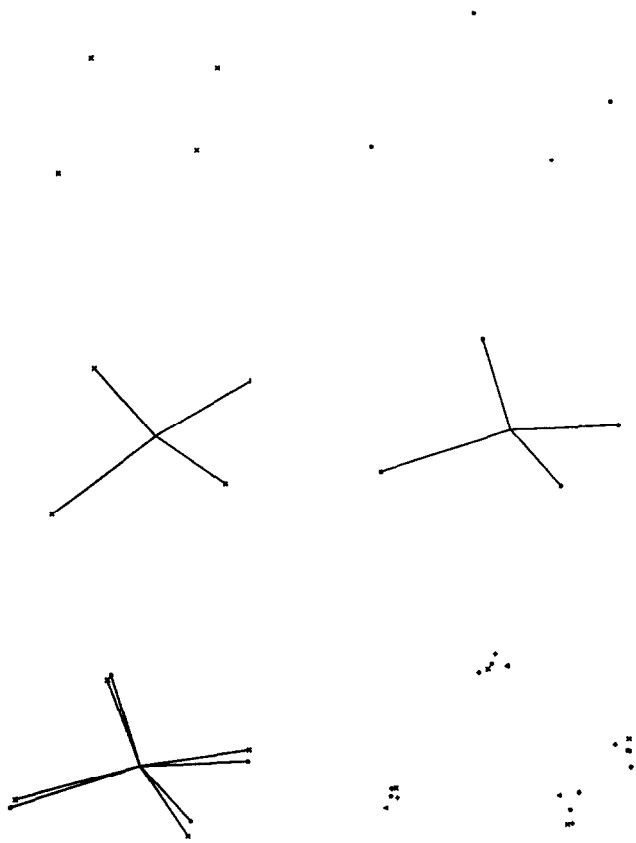

Figure 1. Geometry of Procrustes fits. Upper row: Two forms with four landmarks. Center row: Each form is scaled to unit Centroid Size (sum of squared distances from the centroid to the landmarks). Lower left: The centroids are superposed and then one form is rotated over the other to a position of least squared differences between the homologues. Lower right: If the fixed form is a sample average shape (see text), vectors of displacement in Procrustes space, shown here as four scatters around the average (dots), serve as one set of shape coordinates for the sample of forms. These coordinates total eight (four sets of two), which is four in excess of the actual dimension of the shape space of a quadrilateral.

action of the similarity group. That is, there is a *shared* geometry of shape space and the original Cartesian data such that shortest paths (geodesics) in the Cartesian space project down onto shape space as geodesics in the Procrustes metric there, circles around landmarks in the image plane project as circles around shapes in the shape manifold and so on.

Associated with any Procrustes average shape is the graphic (Fig. 1, lower right) that superposes every specimen of the sample over that average. We pursue the multivariate statistical analysis of sampling variation around the

sample average by treating the set of deviations of each landmark from the corresponding sample average as a set of *shape coordinates*, biomathematically promising measurements explicitly dealing with shape. In fact, the tie between these coordinates and the biometric notion of "variables" is closer than Kendall himself realized. In this representation, these coordinates have been removed from the shape manifold itself out to the *tangent space* to Σ_2^k at the averaged form just introduced. (This tangent space has dimension $2k - 4$: the original $2k$ Cartesian coordinates, less 4 for the free parameters of the similarity group.) Do not imagine this tangent space "classically" as a flat geometric object touching the shape manifold the way a plane touches a sphere. Better, following the modern point of view, construe it as a linear space of one-forms (differentials), the best linear approximation to the geometry of scalar functions of points of the actual manifold "in the vicinity of" a particular form. (The accuracy of this approximation as a function of diameter of a shape distribution is calibrated in Appendix 2 of Bookstein (1991).) In the biometric context, if the tangent space touches right at the Procrustes average form, these linear functions are precisely what we mean by biometric shape variables. One visualization thus refers both to shapes and to shape variables—the dual construction—in the selfsame diagram.

In this way, with Kendall's help, we unexpectedly solved the problem that had loomed ever since the time of Pearson: The geometry of biometric shape variables for landmark data is the geometry of the tangent space to Kendall's shape space at the average landmark shape. Shape coordinates are sets of directions in this tangent space—sets of shape variables—selected so as to represent each equivalence class of shapes once and otherwise for reasons of convenience or diagrammatic elegance. Because the tangent space itself inherits the Procrustes metric from the underlying manifold, in the vicinity of an average any shape variable specified by a formula, such as an angle or a ratio of two distances, has a Procrustes length per unit of shape value independent of its sample variance. Likewise, pairs of shape variables make a Procrustes angle independent of their sample covariance. (The Procrustes "length" referred to here—linearly extrapolated magnitude of the shape difference induced by a unit change in the linear combination under study—has nothing to do with actual sample range, which is normally 1 or 2 orders of magnitude smaller.) For triples of landmarks, this Procrustes geometry of variables was explored diagrammatically in Chapter 5 of Bookstein (1991).

Because the tangent space in which we are representing shape is linear, any other set of shape coordinates that preserves Procrustes distance must derive from these by an orthogonal transformation. That is, for the multi-

variate analysis of landmark shape in the vicinity of an average shape to preserve the Euclidean geometry of the original data under the action of the special similarity group—for our biometrics to correspond closely enough to the ordinary vernacular meaning of "shape" for Cartesian coordinate data—the basis we use for the tangent space to shape space must be generated as a Procrustes-orthogonal rotation of the Procrustes fits to the mean shape. For biometric tactics to make sense in terms of this intrinsic geometry of landmark shape, shape coordinates *must be* the superposition of specimen shape on the sample average shape or a Procrustes rotation of those superpositions. This is the unique statistical geometry that underlies the morphometric synthesis: the a priori metric geometry of shapes of landmark configurations in Euclidean space. Furthermore, because the Procrustes metric is a submersion of the original Euclidean one (Kendall, 1984), the distribution of shapes in this tangent space is (very nearly) spherical whenever the distribution of the original landmarks is characterized by independent circular Gaussians of the same variance (Mardia, 1995). In other words, this shape space has no "privileged directions" regardless of the average shape.

Figure 2 presents a typical example of the reduction of landmark configurations to a shape scatter in this way. The data are eight landmark points on the brain cases of 21 rats observed in lateral radiographs as they grew from 7 to 150 days of age. For more information on this sample, see Bookstein (1991). Most of the shape coordinates show a clear age trend; their joint biomathematical interpretation will concern us in a moment. In this setup, the scatters of all the shapes of a data set as fitted to their average together exhaust the space of shape variation embodied in those Cartesian coordinates. Write these *Procrustes-fit coordinates* for the shape of a typical sample form as the $2k$-vector $F = (r_1, s_1, r_2, s_2, \ldots, r_k, s_k)$, where the ith landmark is located at (r_i, s_i) after the Procrustes fit. The x-coordinates provide the subvector F_{odd}; the y-coordinates provide the subvector F_{even}.

Figure 3 introduces a second data set likely of greater potential importance: a set of 13 landmarks from nearly midsagittal magnetic resonance images of 14 patients with schizophrenia (of a variety of clinical types) from the Adult Psychiatric Unit, University of Michigan Hospitals, along with 14 images from age- and sex-matched patients found not to warrant a diagnosis of schizophrenia. The 13 landmarks generate points in a 22-dimensional shape space. Diagnosis was blind to these images, and the locating of landmarks (by Dr. John DeQuardo) was blind to diagnosis. For an interpretation of this finding in its neuropsychiatric context, see DeQuardo *et al.* (1995).

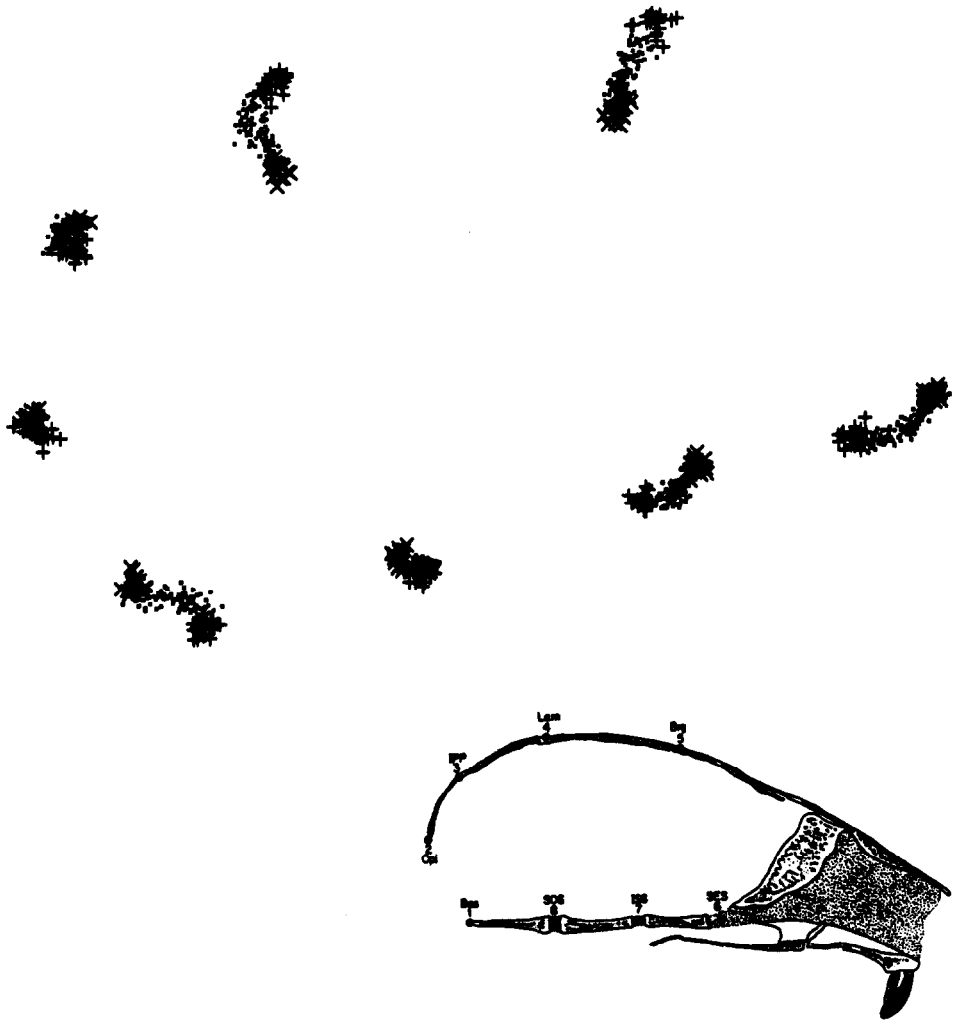

Figure 2. Shape space for eight landmarks from cephalograms of 21 rat skulls: Procrustes fit of all specimens to the grand mean shape, +, 7-day-old pups; ×, 150-day-old rats. Other ages were 14, 21, 30, 40, 60 and 90 days. Inset and data are from Bookstein (1991).

The Procrustes-fit coordinates of this sample (Figure 4) show much less local structure than the example for the growing rat skulls. The scatters, landmark by landmark, appear roughly circular and of the same radius. An enhancement of this plot, shown magnified for the central set of five landmarks, serves as a direct introduction to the problems of biometrical inference in high-dimensional data like these. The triangles represent shape coordinates for the schizophrenics; the dots, for the others. There

Figure 3. Nearly midsagittal brain scan of a normal subject, with 13 landmarks located. As numbered: 1, splenium of corpus callosum; 2, genu of corpus callosum; 3, uppermost point on upper boundary of arch of corpus callosum; 4, vertex; 5, tentorium at dura; 6, top of cerebellum; 7, tip of fourth ventricle; 8, bottom of cerebellum; 9, top of pons; 10, bottom of pons; 11, optic chiasm; 12, frontal pole (on the extension of a line from 1 through 2); 13, colliculus.

appear to be differences between the groups in the fitted coordinates of landmarks 1 and 13, differences significant separately by Hotelling T^2 at $p \sim 0.009$ and 0.012, respectively. However, we selected these out of all 13: the Bonferroni-corrected probability is thus not significant even at the 10% level. Continuing with the inspection of the right-hand plot, we notice that the separation of those mean vectors by group seems to be in opposite directions between the landmarks. Indeed, the vector that connects the two

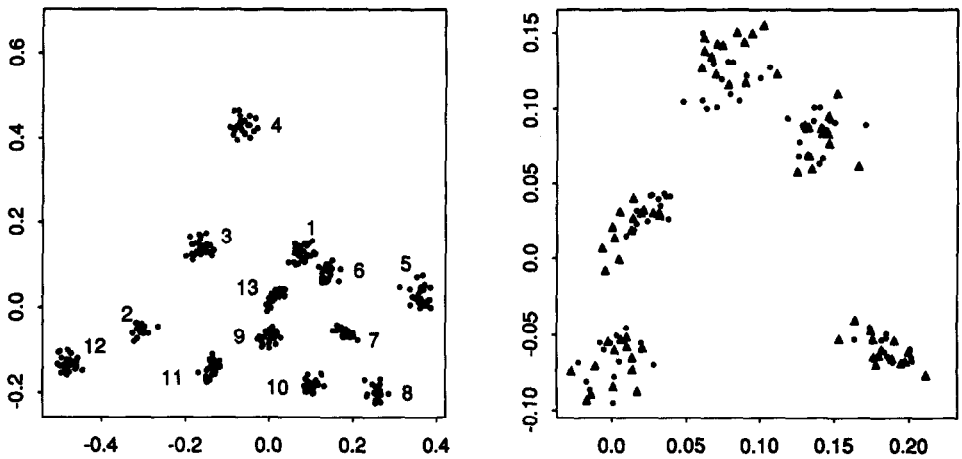

Figure 4. Procrustes registration for landmarks from 28 samples like Fig. 3 upon their average shape. Left: The complete data set of shape coordinates, rotated from radiological horizontal to align the longer principal axis of the average shape with the horizontal. Right: Enlargement of fits at landmarks 1, 6, 7, 9 and 13. Triangles, schizophrenic subsample. Apparent differences at landmarks 1 and 13 are statistically problematic (see text).

fitted landmarks shows a group difference significant by T^2 at a level of 0.004. However, we have selected this vector as one out of $13 \times 12/2 = 78$ pairs, leading, again, to insignificance after Bonferroni correction. While it is implausible that the most discrepant pair of group differences would be situated at adjacent landmarks, nothing in the Procrustes multivariate machinery yet allows us to adjust our probabilities accordingly. In other words, the biometrics of these scatters does not yet articulate with such reasonable biomathematical descriptors of shape phenomena as "the displacement of a landmark" or "increase along a segment between two landmarks." In fact, it is well known (Rohlf and Slice, 1990) that although the Procrustes-fit shape coordinates have the right sum-of-squares, the associated superposition, which is exactly what we are contemplating here, leads to misleading interpretations in most practical applications to shape processes.

We will restore biomathematical cogency to these manuevers by passing from the Procrustes-fit basis of $2k$ coordinates to a very carefully selected new basis that has the correct number $2k - 4$ of shape dimensions. If this change of basis is by a (Procrustes) rotation, the new basis will be *orthonormal*. Geometrically, this means a set of $2k - 4$ directions that all have Procrustes length 1 and are pairwise Procrustes-orthogonal. Statistically,

these are sets of shape variables whose differences sum in square to squared Procrustes distance and such that, for variation by circular Gaussian distributions of the same small variance around each landmark of the mean shape, all have the same variance with all covariances zero. Although the formula for the Procrustes metric itself does not explicitly incorporate the mean landmark configuration in any way, we are free to take the mean shape into account in choosing which rotation to apply. One particularly useful Procrustes-orthonormal basis—the only one yet constructed, in fact—is the set of partial warps that arise as eigenfunctions of the thin-plate spline we will use to visualize shape changes as deformation. The next section introduces this flexible interpolant and the spectrum of its energy form. First, however, I will review the tradition from which this rotated basis ultimately arose.

3. Biomathematical Studies of Shape Transformation

3.1. *Historical background.* What we borrow from the biomathematics of shape change is, of course, the visualization by transformation grid. Although this idea is usually associated with the famous treatise *On Growth and Form* by the British naturalist D'Arcy Thompson (1917), it is actually hundreds of years older than that. The first "transformation grids" reflect efforts of Renaissance artists to comprehend the variability of the human forms that they were just beginning to reproduce realistically. Figure 5, for instance, assembled from Albrecht Dürer's (1528) *Vier Bücher von Menschlicher Proportion*, explores diverse types of "transformation grid," both affine and localizable, in the effort to explore the limits of normal variation and the strategies of effective caricature. The semiotics is that of geometric perspective, but the information conveyed is wholly different: no longer the effect of a change of vantage-point, but a change of organism.

This formal theme of shape transformation as the explicit object of biometric discussion was most clearly set forth in the famous Chapter XVII of Thompson (1917), "On the Theory of Transformations, or the Comparison of Related Forms." Thompson's goal is distinctly old-fashioned and much too Platonic to articulate with biometrics without severe modification:

> [If] diverse and dissimilar [organisms] can be referred as a whole to identical functions of very different co-ordinate 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 recognisable system of forces has been in control. ...

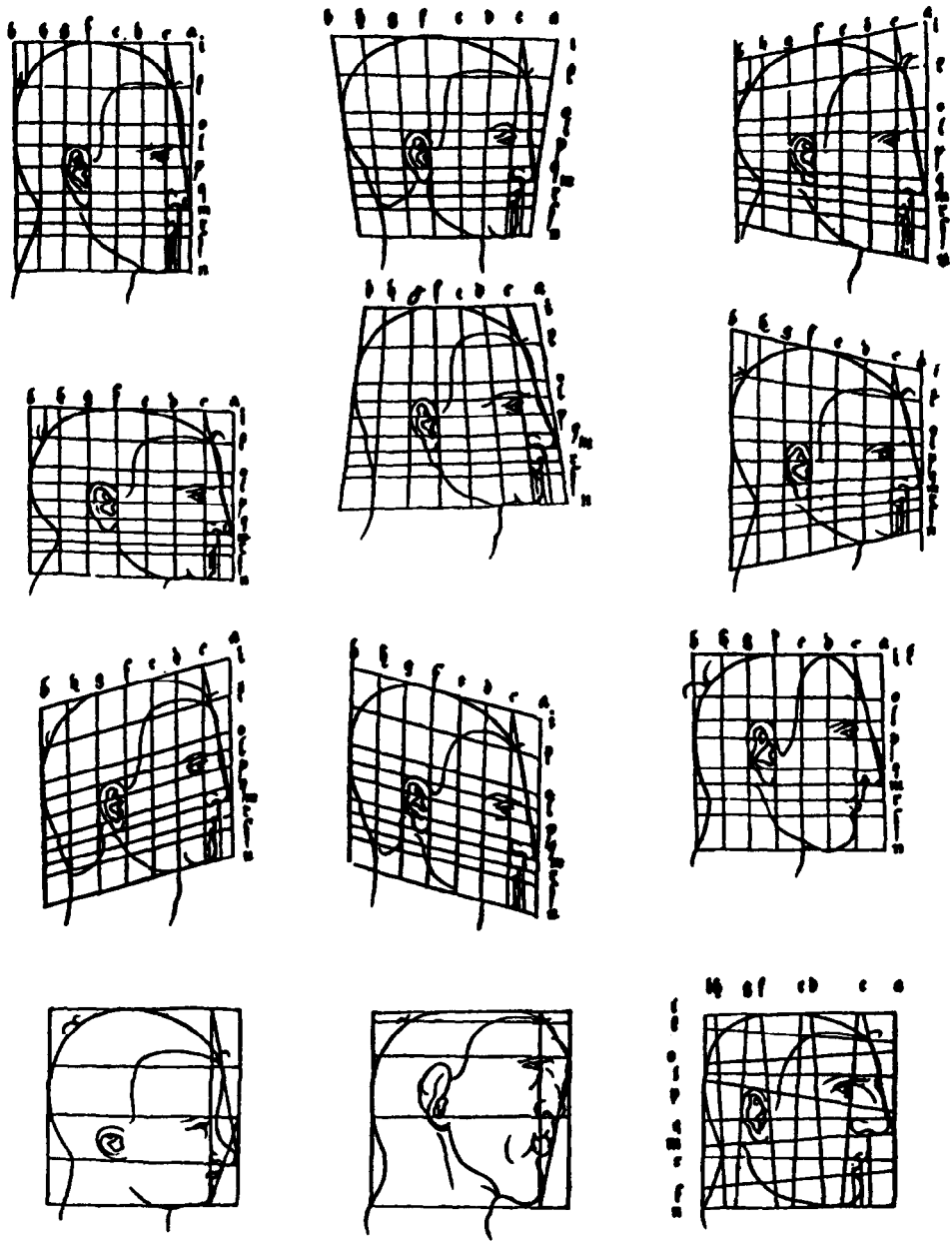

Figure 5. One head with a standard grid and 11 transformations onto other German types. From Dürer (1528), by permission.

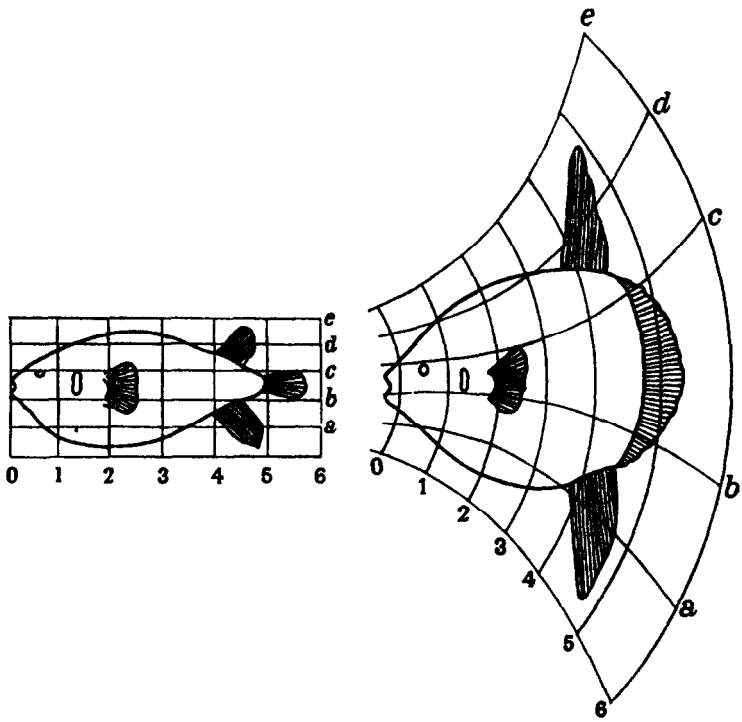

Figure 6. Cartesian transformation, *Diodon* to "Orthagoriscus." From Thompson (1917).

In Thompson's own examples, including the famous exemplar reproduced here as Fig. 6, the Platonic thrust of homogeneity clearly dominates any concern for realism. Thompson's fond hope that these figures would reveal the origins of form in force (an assertion he meant literally) was never realized, and while several later generations of quantitative biologists were tempted by this graphical style, it proved never to lead to quantification in the global mode that Thompson had intended. As a consequence, this once-promising method underwent only "vicissitudes," not development, from its publication in 1917 to its supersession in the middle 1980s by the biometrical graphics of the synthesis. For a historical review, see Chapter 5 of Bookstein (1978).

From the vantage point of 1995, all these earlier attempts at a biometrically quantifiable praxis of transformation grids can be classified by the nature of the compromises they made. The biometric point of view requires that we be capable of representing variation as well as central tendencies, that we be able to cover the full range of potentially meaningful descriptors and patterns in an even and "unbiased" fashion and that graphics are

supplied for those features both separately and in arbitrary composites. Prior to the synthesis, no modification of Thompson's grids met these criteria to any greater extent than Thompson's original suggestion did.

For instance, Sneath and Sokal (1963) argued, by reference to very accurately drawn Cartesian transformations between holotypes of fossil marsupials, that such visualizations did not lead to "features" or to taxonomically valid measures of resemblance. A few years later, Sneath (1967) attempted to express these grids by coefficients of polynomial trend surfaces approximating them, but those coefficients neither corresponded to any natural metric nor sustained any biomathematical interpretations over the figure of the organism. In the context of simple allometry, Huxley's (1932) much earlier method of "growth-gradients" would occasionally lead to suggestive Cartesian transformation diagrams, but was limited to fields of extremely simple structure. My own early method of biorthogonal grids (Bookstein, 1978) provided a canonical coordinate system for Thompson-style grids but did not extend to any visualization of "standard error" or any other biometric aspects of the tensor fields so displayed, and had no vector structure (could not be added, multiplied or averaged). Oxnard's (1973) displays of single principal components of multivariate size measures as grids dealt with biometrically well-characterized shape features but left the semantics of their biomathematical interpretation hopelessly obscure, at the mercy of whatever biomechanical analogies the investigator might imagine (e.g. the "craniolateral twist" of a primate scapula). Yet other methods, such as Lohmann's (1983) "eigenshapes," applied biometric algorithms to valid representations of outline shape but offered no channel by which findings could be interpreted biomathematically, that is, as pertinent to biological homologies.

The earliest applications of tensor analysis in morphometrics, such as that of Richards and Kavanagh (1943), although effectively leading to interpretations of shape change in terms of developmental processes, did not permit group-level operations such as averaging or assessments of variation. The features displayed by later finite-element methods modeled after those early attempts, such as that of Lewis *et al.* (1980) or Bookstein (1984a), proved unacceptably dependent on a priori parcellations of the form into "elements," a step that could not be carried out (and still cannot be carried out) in any biomathematically sensible way. Finally, just prior to the synthesis that is my main theme here, methods of Procrustes analysis were introduced solely in the form of plots like that in Fig. 2, without any semiotics for correlations among the positions of the points fitted severally and without reference to the possibility of orthogonal rotations of this tangent space. Absent that final step, disconnected sets of Procrustes-fit residuals do not conduce gracefully to allometric modeling and cannot be

matched to biomathematical explanations such as regional processes of regulation.

In hindsight it is clear why the morphometric innovations of the 1950s through the early 1980s were so incoherent. There was no agreement about what constituted an appropriate analysis because there was no proper biomathematical understanding of what constituted the data. Oxnard's (1978) review article, for instance, which dealt with data in the form of images, had nothing in common with the approach of Blackith and Reyment (1971), which treated only scalar variables previously extracted by ruler, planimeter or protractor. My first publications on the "method of biorthogonal grids" in the late 1970s incorporated no theory of how landmarks are chosen or how that choice affects interpretation of the grids that were produced. The task would better have been taken as representing the raw data (in this case, whole configurations of landmarks) in a space whose dimensions would *be* transformations, not with the depiction of single changes *as* transformations. On the other hand, my preliminary statistical method for triangles (Bookstein, 1982a, b), lacking only the corresponding distribution theory, never referred to vectors of variables or linear combinations, nor did it hint at any appropriate extension even to pairs of triangles, let alone to landmarks considered without lines connecting them.

In short, none of us acknowledged that multivariate methods would not apply effectively to landmark data until a canonical way was found to make whole landmark configurations into "variables," and none of us thought to pursue the analysis common to alternate visualizations rather than the argument that some visualizations were "better" than others. As an ineluctable consequence of this methodological catholicity, the examples that appeared "successful"—that seemed to lead to biomathematically satisfactory explanations—had little in common. No practical advice emerged for matching techniques to data sets or to biological questions, and although we knew that different analyses applied to the same data set would usually result in incompatible findings, we had no protocol for choosing among them. In this way a good-sized roster of earnest workers—some amateurs, some professionals—circled around the solution that was to come, without ever realizing the crux of our collective problem.

3.2. *The thin-plate spline.* Again, as with Kendall's gift of shape space, what made the morphometric breakthrough possible was help from an unexpected quarter: a relatively esoteric advance in interpolation theory. Computer graphics had long been concerned with the lofting of surfaces

representing physical or sociological quantities measured over "scattered data." Algorithms for smoothing these point observations into surfaces would often begin by promulgating a functional that the resulting interpolation was to optimize under the constraint of according with the given data. Several of us had already tried borrowing from this literature: Sneath's (1967) trend-surface analysis imitated a technique effective in some geological mapping contexts, and my interpolation algorithm of 1978 was modeled on numerical solution of the Helmholtz equation inside a boundary of arbitrary shape.

Even as these unsatisfactory experiments were in progress, two French mathematicians, Duchon and Meinguet, were recasting the interpolation problem in a new form: as the global minimization of a quadratic functional in the constraints of the interpolation. The mathematics of this minimization, it turned out, is identical to that for an unusually old problem in continuum mechanics—the bending of a metal plate subject to physical constraints. The literature of that problem (Timoshenko and Woinowsky-Krieger, 1959) had always expressed the resulting forms in infinite series. Most unexpectedly, Duchon and Meinguet managed to find one version of the problem for which the surface model that emerged can be expressed in very terse closed form. The technique generalizes to a variety of differential operators—interpolants minimizing any of a hierarchy of energy functionals—but it is the first element of the hierarchy which leads to analyses of the greatest interest for biology. For extensions of this specific application of the splines to smoothing problems, see, for instance, Wahba (1990). The same technique has also been extensively modified for application in computerized image analysis per se, well outside the biomathematical context. The following brief summary is taken from Bookstein (1989a).

Let U be the function $U(\mathbf{r}) = r^2 \log r$ and let $P_i = (x_i, y_i)$, $i = 1, \ldots, k$, be k points in the plane. U is a fundamental solution of the *biharmonic equation*: we have $\Delta^2 U \propto \delta_{(0,0)}$, where δ here is Kronecker's function, zero everywhere except at the origin but with integral equal to 1, and Δ^2 is the iterated Laplacian $(\partial^2/\partial x^2 + \partial^2/\partial y^2)^2$. It can be shown (Timoshenko and Woinowsky-Krieger, 1959) that the equation of a thin, uniform metal plate originally flat and now bent by vertical displacements at various points is $\Delta^2 U = 0$ except at points where force is applied. The equation presumes that displacements normal to the rest position of the plate are sufficiently small that strains in the plane of the plate itself can be ignored. The application to morphometrics rests on a scenario with no equivalent in the world of real plates: the displacement of an infinite metal plate at a finite series of discrete points in a world wholly lacking in gravity.

Let the knots of the spline (later to be the landmarks at which we are calibrating the deformation of one biological form into another) be at

points P_i, $i = 1, \ldots, k$, in one single image. Writing $U_{ij} = U(P_i - P_j)$, build up matrices

$$K = \begin{pmatrix} 0 & U_{12} & \cdots & U_{1k} \\ U_{21} & 0 & \cdots & U_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ U_{k1} & U_{k2} & \cdots & 0 \end{pmatrix}, \tag{1}$$

$$Q = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_k & y_k \end{pmatrix} \tag{2}$$

and

$$L = \begin{pmatrix} K & Q \\ Q^t & O \end{pmatrix}, (k+3) \times (k+3), \tag{3}$$

where O is a 3×3 matrix of zeros. The thin-plate spline $f(P)$ having heights (values) h_i at points $P_i = (x_i, y_i)$, $i = 1, \ldots, k$, is the function

$$f(P) = \sum_{i=1}^{k} w_i U(P - P_i) + a_0 + a_x x + a_y y, \tag{4}$$

where

$$W = (w_1 \quad \cdots \quad w_k \quad a_0 \quad a_x \quad a_y)^t = L^{-1} H \tag{5}$$

with

$$H = (h_1 \quad h_2 \quad \cdots \quad h_k \quad 0 \quad 0 \quad 0)^t. \tag{6}$$

Note that the ws multiply copies of the kernel function $U = r^2 \log r$ evaluated with respect to each landmark in turn, while the coefficients a_0, a_x and a_y calibrate the function "at infinity." It is quite important for our morphometric applications that this vector W is linear in the data H of "heights."

Then the function $f(P)$ has three crucial properties (Duchon 1976; Meinguet 1979):

1. $f(P_i) = h_i$, all i. (The function f interpolates the heights h_i at the landmarks P_i.) This is guaranteed by the first k rows of L.

2. The function f has minimum *bending energy* of all functions that interpolate the heights h_i in that way: the minimum of

$$\iint_{\mathbf{R}^2}\left(\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x\,\partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2\right), \tag{7}$$

where the integral is taken over the entire picture plane. This quantity is also called (especially in computer graphics) the *integral quadratic variation*. The integrand is the sum of squares of all the *second* derivatives of the warping function. In that crucial detail these splines contrast greatly with other techniques, such as elastic relaxation, that minimize a functional of *first* derivatives.

3. The value of this bending energy is

$$\frac{1}{8\pi}W^t KW = \frac{1}{8\pi}W^t \cdot H = \frac{1}{8\pi}H_k^t L_k^{-1} H_k, \tag{8}$$

where L_k^{-1}, the *bending energy matrix*, is the $k \times k$ upper left submatrix of L^{-1}, and H_k is the k-vector $(h_1\ h_2\ \cdots\ h_k)$ of "heights."

In the application to two-dimensional landmark data, we compute two of these splined surfaces: one in which the vector H is loaded with the x-coordinates of the landmarks in a second form, and one for the y-coordinates. (The two interpolations use the same matrix L incorporating both Cartesian coordinates of the first form.) The resulting map $(f_x(P), f_y(P))$ is now a deformation of one picture plane onto the other that maps landmarks onto their homologues and has the minimum bending energy of any such interpolant, and its bending energy is now the sum of terms for the x-interpolant and the y-interpolant separately, the form $X^t L_k^{-1} X + Y^t L_k^{-1} Y$, where, temporarily, we write the coordinates of the landmarks in the target form as the pair of k-vectors (X, Y).

This seemingly sterile algebra generates a remarkably cogent biomathematical rhetoric. The bending energy (8) that is a quadratic form in the landmarks of the "target form" is at the same time the integral (7) of quantities that are entirely local. The integrand of (7), although written in terms of second derivatives, is actually a summary of the gradients of the first derivatives of the quantities $\partial f/\partial x$ and $\partial f/\partial y$: a summary description of the local rates of change of the local shape change tensor for "homologous bits of tissue" over the whole image. The quantity is zero when neighboring bits of tissue change by identical shears, and grows as the effect of the transformation on line elements in any direction is graded

more and more sharply in any direction. That the integrals are taken over the whole plane, rather than the interior of the outline of a form, is an inconvenience, but only a minor one. A minimand that arises in classical continuum mechanics as an energy is nevertheless capable of being radically reinterpreted in biomathematical language as a *localization of shape change*. The affine part $a_0 + a_x x + a_y y$ of the spline, which makes no difference for the integrand here, is now an ordinary shear. In the metaphor of the lofted plate, this component can be thought of as an appropriately oriented *shadow* of the original gridded plate after it has been only tilted and rescaled but not bent. The entire formalism is appropriately invariant against similarity transformations of either set of landmarks.

(There is an alternative derivation of the thin-plate spline interpolant that identifies the minimand in (7) with a different quantity, statistical rather than biomathematical: the within-image prediction of a random field observed at one set of landmarks and taking values there corresponding to locations of the other set of landmarks, when prediction is by the variance-minimizing technique of kriging. The equations of the spline can be rederived in identical form from this quite different starting point if one models the prediction between landmark locations as the expression of a Gaussian random field with covariance function $r^2 \log r$, which can be generalized to fractional powers of r. See Kent and Mardia (1994).)

The mapping function from landmarks to landmarks remains linear in the target two-vectors (X, Y). Should those be shape coordinates in the sense of the previous section and the starting form of the spline be a sample average shape, then we have extended the linear machinery of shape space to a system of coefficients (the Ws) and thence to a diagram that visualizes the relation of any specimen shape to the average as a deformation, the coefficients of which are linear in its shape coordinates. In this way Cartesian grids about the average shape have been incorporated into the biometric framework *in toto*, as a direct visualization of one particularly specialized set of linear descriptors.

The analysis of any landmark data set typically begins with scrutiny of single grids of this type. For instance, the summary grid transformation in Fig. 7 visualizes the Procrustes average shape of the landmark configurations for the schizophrenics in Fig. 3 as a deformation of the Procrustes average for the other cases. Its form immediately suggests a meaningful interpretation. The visible strain of the grid lines seems limited to the vicinity of the triangle among landmarks 1, 6 and 13 just right of center. (This is the same region upon which we focused in the analysis of Fig. 4.) The anatomical space indicated by this triangle (recall Fig. 3) is the posterior curve of corpus callosum where it separates the third ventricle from the cistern of the great cerebral vein (Netter, 1989, plate 103), a pool

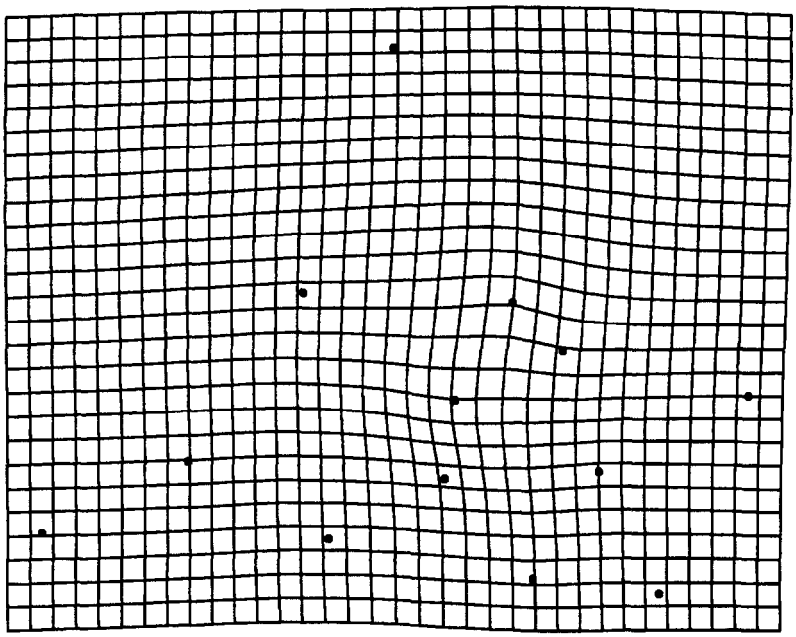

Figure 7. Thin-plate spline for deformation from the average nonschizophrenic shape to the average schizophrenic shape from Fig. 4. The feature to which the eye is drawn—the site of the apparently focal deformation just right of center—spans the splenium of corpus callosum and the cistern of the great cerebral vein in Fig. 3. This region appears expanded, whereas the region to its left, the vicinity of the thalamus, appears compressed.

of cerebrospinal fluid outside the main ventricular system. The organization of this figure permits the biomathematical interpretation at which Fig. 4 could only hint. A rigorous test for the statistical significance of this apparent difference will be introduced in subsequent text.

In terms of the geometry of shape space, what we have done in this example is to follow a very tempting shortcut. The mean difference between the groups of shapes was computed biometrically: as a vector in the tangent space corresponding to Fig. 4, not as a grid. However, its visualization by the method of grids leads us to realize that, by itself, it is already a biomathematically meaningful (i.e. localizable) feature of shape difference: a promising locus for neurophysiological explanation. We shall see in succeeding text that this visual signal (which may be of considerable medical interest) corresponds to the only biometrically significant finding that can be produced by any of the methods of the synthesis. One often finds that a single thin-plate spline display, if carefully crafted, can encompass the import of a whole data set in this way.

3.3. *Eigenvectors of bending energy.* Usually an investigator is not this lucky right at the outset. The more reliable channel by which the spline generates biomathematically sensible reports of shape phenomena is wholly different: a peculiarly useful rotation of Procrustes-fit coordinates its supplies for decomposition of more general shape phenomena into parts each capable of a biomathematical interpretation separately. Specifically, the elements of the most useful basis known for biomathematical interpretation of shape processes are the eigenvectors of the bending-energy matrix L_k^{-1} already introduced. This proposition, the other major formalism underlying the morphometric synthesis, is far from obvious, although it is elementary. It seems unreasonable that a strategy for minimizing a metaphorical "energy" in comparisons of forms two at a time would contribute anything important to descriptive techniques for variation and covariation across samples. To understand how this deep tie comes about, we must inspect the algebra and geometry of spline fits somewhat more closely.

Figure 8 shows a variety of diagrams which, although appearing variously to be surfaces and deformation grids, are actually all thin-plate spline interpolations for four landmarks beginning in the form of a square. For a square of side 1, with landmarks ordered circumferentially, the bending-energy matrix (equation (8)) is

$$L_k^{-1} = \begin{pmatrix} 0.3607 & -0.3607 & 0.3607 & -0.3607 \\ -0.3607 & 0.3607 & -0.3607 & 0.3607 \\ 0.3607 & -0.3607 & 0.3607 & -0.3607 \\ -0.3607 & 0.3607 & -0.3607 & 0.3607 \end{pmatrix}. \tag{9}$$

This matrix has rank 1—it is a multiple of the outer product of the vector $(0.5, -0.5, 0.5, -0.5)$ with itself. Because this vector is an eigenvector of L_k^{-1}, it is immaterial whether we treat it as a set of coefficients or instead, as in the upper left panel of the figure, as a set of heights themselves to be splined: a surface lofted over the landmarks, raised at the endpoints of one diagonal of the square and lowered at the endpoints of the other.

That splined interpolant, which of course looks like a physically reasonable bent sheet (since that was the actual subject of the original physical model), serves in our biomathematical context as one example of a biometrically useful transformation grid. The eye is fooled (as the viewing eye of Cartesian transformation grids is always susceptible to being fooled) by a visually dominant affine component that is quickly, if inappropriately, construed as evidence of tilt in "space." If we reduce the magnitude of that component by rearrangement of the landmarks on the right, we see how the same formalism supplies a solution to D'Arcy Thompson's Cartesian

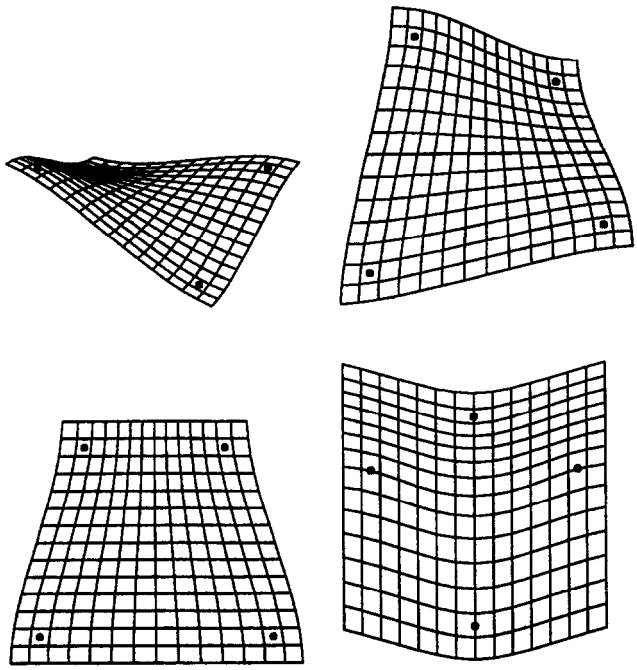

Figure 8. Thin-plate splines of a square. Upper left: The sole *principal warp* for this configuration. This is the form assumed by an idealized (infinite, uniform, infinitely thin, originally perfectly flat) metal plate were it raised over the ends of one diagonal of a square, lowered over the ends of the other and otherwise left free to deform without any pull of gravity. Upper right: A projection of that gridded surface for which "up" and "down" align with the diagonals of the square: deformation of square into kite. Lower left: Another projection, with "up" and "down" now along one set of sides of the square: deformation of square into trapezoid. Lower right: Pointwise, the spline is independent of the orientation of the starting grid. This is the same mapping as that at upper right.

problem: a grid transformation that accords with landmark positions and that uniquely satisfies a well-defined optimal criterion. In this case (upper right), we have projected the two-up–two-down configuration along a diagonal of the square, resulting in the grid for a square-to-kite transformation.

Surfaces like these can be projected down onto the plane of the square in any orientation. When projected along a side of the square instead of along a diagonal, for instance, this same surface leads to the square-to-trapezoid map shown at the lower left. "Up" and "down" are now aligned with a side of the square, not the diagonal. The starting grid can be rotated upon any of these surfaces, as shown at lower right, without altering the interpolation mapping.

Changing the shape of the starting form alters the matrix L_k^{-1}, its eigenstructure and the shapes of all these surfaces. A quadrilateral with three landmarks collinear, for example, generates the principal surface shown lofted at upper left in Fig. 9. Minimum-energy landmark rearrangements can derive from it by projections along or athwart the collinear edge (Fig. 9, lower left, lower right). The noncollinear landmark has nothing to do with this energy; it only affects the affine part of the spline analysis (the apparent "tilt" of the surface, the shear of the homology mapping).

In general, because of the matrix Q appearing in the assembly of L (equation (3)), the bending-energy matrix L_k^{-1} has three eigenvectors of eigenvalue zero, corresponding to the three-dimensional family of planes over any landmark configuration. There remain $k-3$ dimension of nontrivial bending above k landmarks, spanned by the nonzero eigenvectors of the bending-energy matrix. These nonzero eigenvectors, such as the eigenvector $(0.5, -0.5, 0.5, -0.5)$ of the matrix (9), are called *principal warps*. Any splined surface is the superposition of one multiple of each eigenvector, together with a tilt term. If we notate the usual eigendecomposition as $L_k^{-1} = UDU^t$ with U orthogonal and $D = \mathrm{diag}(e_1, \ldots, e_{k-3}, 0, 0, 0)$, the principal warps are the first $k-3$ columns $u_1, \ldots, u_{k-3}$ of U. (This part of U will be denoted U_{-3}.) As in Figs. 11 and 14 to come, each is visualized by loading its coefficients into the slots w_i of equation (4) and then setting a_x and a_y to values that lead to a congenial view.

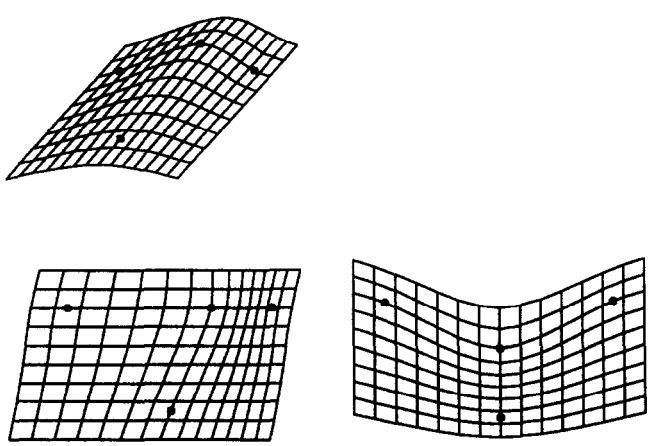

Figure 9. Thin-plate splines of a collinear quadrilateral: isosceles right triangle and the midpoint of the hypotenuse (not shown). Upper left: One view of the sole principal warp. This view also serves as one partial warp. Lower row: Two canonical partial warps (aligned with the principal moments of the starting landmark configuration): left, horizontal; right, vertical. These correspond to views of the surface above from vantage points at suitable azimuths westerly (left) and northerly (right).

To each column u_i of U corresponds not only a three-dimensional surface but also a projection "down" onto the picture plane that captures its role in the spline from the mean to each individual form of a data set. The contribution is a vector of two components: one, $u_i^t X$, for the x-coordinates of the target form; the other, $u_i^t Y$, for the y-coordinates. This paper concerns only the two-vectors that are generated when those coordinates are the Procrustes-fit coordinates F of the shapes of a sample. The partial warp scores are then two-vectors $(u_i^t F_{\text{odd}}, u_i^t F_{\text{even}})$. The whole set can be assembled in the vector $(U_{-3} \otimes I_2)^t F$, where I_2 is the 2×2 identity matrix. The spline map remains the sum of all the partial warps interpreted as mappings (that is, the sum of all the displacements they induce at each landmark, grid intersection, etc.), together with an affine term.

Like the spline itself, the partial warps have a biomathematical interpretation. They are "normal modes" for the *localized description of shape processes*. As eigenvectors of bending energy, they represent component processes that are orthogonal both in terms of shifts in landmark coordinate and in terms of energy. The partial warps, in other words, superpose without interaction. They decompose empirically encountered shape changes in the same way that Fourier components decompose periodic signals, that Bessel functions decompose the shapes of a kettledrum and that the principal warps decompose empirical surfaces lofted over landmarks. Intuitively, metal plates are more difficult to bend as the points at which forces are applied move closer together. Eigenvectors of higher specific bending energy entail discrepant rearrangements in small neighborhoods of landmarks; those of lower specific bending energy, rearrangements widely distributed over the form. It is reasonable to suggest, therefore, that these eigenvectors, ostensibly computed in order of specific bending energy, actually extract a series of shape phenomena in a hierarchy of localization, the same concept we used to interpret the bending-energy formalism, integral (7). The affine term is not localized at all in this sense, and so its bending energy of zero is consistent with the hierarchy.

A quincunx of landmarks (the five-spot of a die) has two principal/partial warps, for instance, as shown in Fig. 10. The upper principal warp, at larger scale, is the two-up–two-down surface—relative displacement of the diagonals—we have already seen. The central one, at smaller scale, represents the displacement of the central landmark over a background of the others, as can be seen directly from its expression in Procrustes-fit coordinates in the right column. Notice that this pair of warps conduces to two different styles of biomathematical interpretation even though they are algebraically entirely analogous. Any rearrangement of the quincunx is the composite of two-vector multiples of these, along with an affine term. For

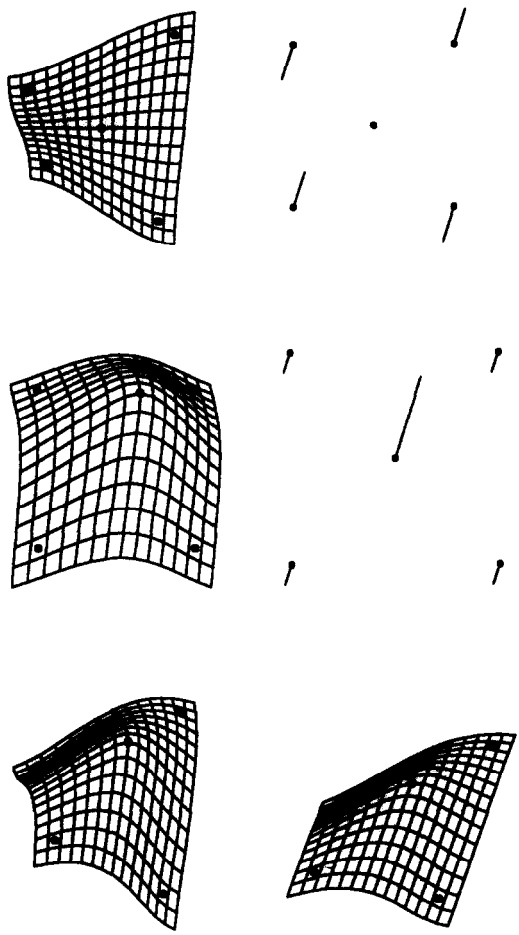

Figure 10. Principal warps of a quincunx. Upper row: the less bent; middle row: the more bent. The equivalent Procrustes displacements (drawn at right) express the principal warps as linear combinations of Procrustes-fit shape coordinates, and so remain approximately at Procrustes superposition. Bottom left: The affine-free transformation for which these figures are the partial warps. Bottom right: The same with an affine term, equivalent to another view of the surface at the left. Both of the upper shape changes are of squared Procrustes length 0.1; hence the warp at lower left has squared Procrustes length 0.2.

instance, the spline at the bottom right is the sum of the two partial warps shown, together with a small shear.

Of the eigenvectors u_i of the bending-energy matrix L_k^{-1}, those that have nonzero eigenvalues are extrema of bending energy for constant sum-of-squares of heights above the base plane. The remaining eigenvectors u_{k-2}, u_{k-1} and u_k, of zero bending, can be taken as the columns of Q, equation (2); they merely translate single coordinates or dilate along

Cartesian directions. Because these dimensions span the similarity transformations in the vicinity of the identity, the nonzero eigenvectors of bending energy lie *within* the tangent structure to shape space we introduced in section 2.2. The nontrivial partial warps, when applied to a Procrustes mean form, generate forms that are still in Procrustes registration with it; that is, the partial warps generate a space of shape variation. In fact, their orthogonality as eigenvectors of L_k^{-1} corresponds to orthogonality in the Procrustes geometry. Then these first $k-3$ principal warps, originally characterized as extrema of bending energy with respect to Cartesian landmark displacements, in fact are extrema of bending energy with respect to Procrustes length in shape space as well: normal modes of shape variation per se. That is,

> the nonzero principal warps of the bending energy matrix in Cartesian coordinates are an orthonormal basis for describing shape variation in the tangent space of shape coordinates.

By this most unexpected connection, the thin-plate spline, via its bending-energy matrix, has supplied us with an orthonormal basis for shape space that makes explicit the hierarchy of localized biomathematical interpretations encoded in the landmark spacings of the average shape.

Partial warps serve as the long-sought bridge between biomathematics and biometrics, the core of the morphometric synthesis: they are at once biomathematically sensible features of shape processes and elements of a biometrically orthonormal basis for shape space. No other orthonormal basis has been suggested for this space. To be useful in this application, furthermore, such a basis would need to have a natural quadratic form like equation (8) and a natural biomathematical interpretation like that in equation (7). In particular, the skew bases, such as two-point shape coordinates (Galton coordinates), permit statistical tests of exogenous associations with shape, but are not consistent with any natural shape metric whenever there are more than three landmarks in general position. (See the discussion in connection with Fig. 16.)

These partial warp scores $(u_i^t F_{\text{odd}}, u_i^t F_{\text{even}})$ total $2k-6$ shape coordinates, but we know that a total of $2k-4$ are required to span the tangent space to shape space at any interesting shape. The two remaining dimensions span a *uniform subspace* of our tangent space, the set of shapes that derive from the average shape by affine transformations leaving parallel lines parallel. Properly speaking, these transformations apply up in the covering Cartesian plane, and must be followed by a Procrustes fit back into alignment with the original mean form. Let the Procrustes mean form, scaled to Centroid Size 1 and oriented with principal axes horizontal and

vertical, have coordinates $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$ and let $\alpha = \Sigma x_i^2$ and $\gamma = \Sigma y_i^2$ be the principal moments along those axes. I verified (Bookstein, 1995a; see also Mardia 1995) that one orthonormal basis for the uniform subspace, which thereby completes a full set of $2k-4$, is the pair of Procrustes-unit vectors

$$Un_1^t = ((\alpha y_1, \gamma x_1), (\alpha y_2, \gamma x_2), \ldots, (\alpha y_k, \gamma x_k))/\sqrt{\alpha\gamma}, \tag{10a}$$

$$Un_2^t = ((-\gamma x_1, \alpha y_1), (-\gamma x_2, \alpha y_2), \ldots, (-\gamma x_k, \alpha y_k))/\sqrt{\alpha\gamma}. \tag{10b}$$

These vectors are of unit length and are orthogonal to one another and to every principal warp of nonzero eigenvalue of the matrix L_k^{-1}. The first of the vectors corresponds to Cartesian shears aligned with the x-axis; the second, to Cartesian dilations along the y-axis. For the formulas to follow it will be convenient to unify these two expressions in one $2k \times 2$ matrix Un. Of course, like every other construction in this tangent space, the orthogonality is exact only in the limit of small shape variation.

In this way a *complete orthonormal basis for shape space* in the vicinity of an average form is produced solely from the Procrustes mean form via the spectrum of its bending-energy matrix. Figure 11 shows this basis for the data set of rodent calvarial landmarks introduced in Fig. 2. The uniform terms, sharing the upper left corner, are those of formulas (10): horizontal shear and vertical dilation. The other terms are represented by lofted surfaces. The $2k-4$ dimensions of shape space are spanned by $k-2$ pairs of these coordinates, as in Fig. 12: one pair for the uniform subspace and another for each partial warp. These are computed as consecutive pairs of the vectors $(U_{-3} \otimes I_2)^t F$, together with the single pair of scores $Un^t F$, as F varies over the Procrustes-fit coordinates of the sample of shapes, Fig. 2. Again the youngest age is coded $+$, the oldest, $\times$. These six pairs of coordinates are a rotation of the eight pairs in the earlier figure, a rotation leaving interspecimen shape relationships invariant. The dimensions annihilated are those corresponding to the (linearized) constraints of the Procrustes fit to the mean. While in any finite sample they appear to have some small variance, those four dimensions could never be used for any multivariate procedure that involved matrix inversion (Bookstein, 1995b). The rotation to the partial-warp basis eliminates any danger of normalizing any of these dimensions by mistake.

For this data set, such a display is enormously more suggestive of a biomathematical interpretation than the unrotated version in Fig. 2. Indeed, the new figure more or less vitiates the need for any further biometrical analysis. The uniform term (upper left) indicates that although the net effect is clearly a relative reduction in height of the braincase, the

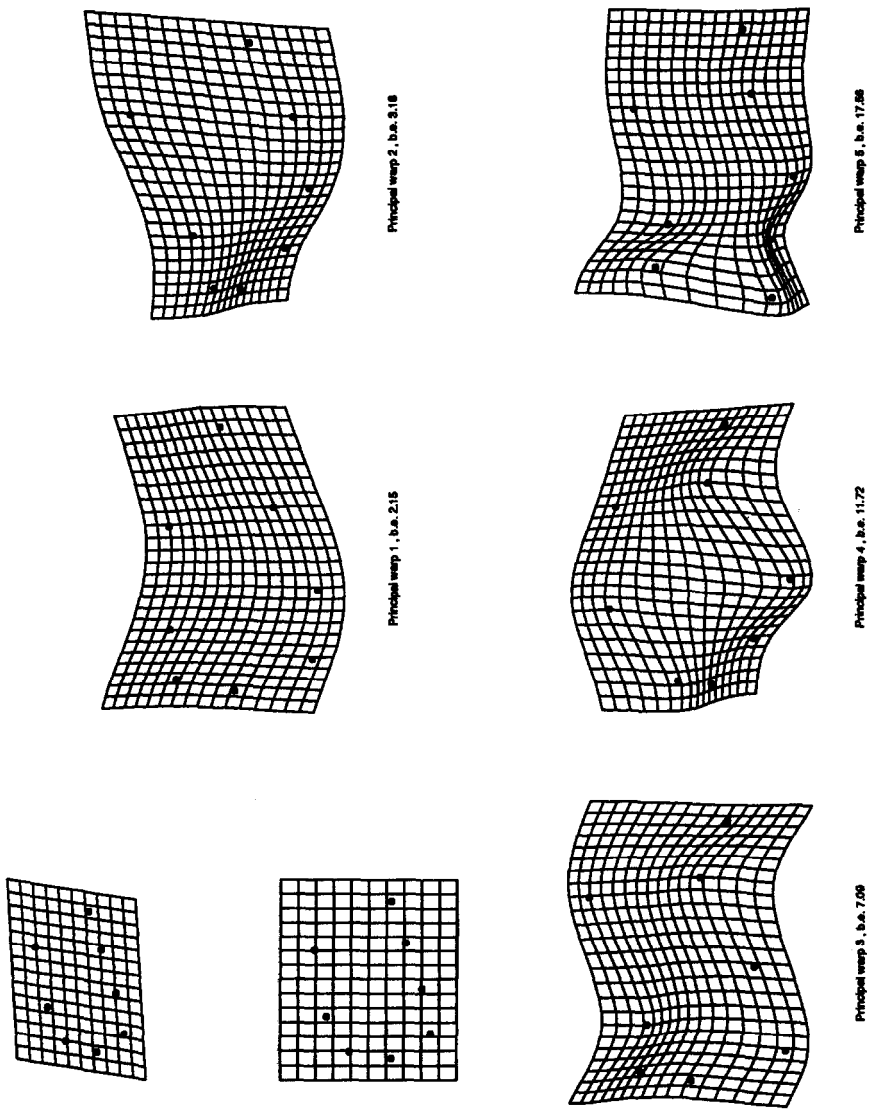

Figure 11. Principal warps for the grand mean rat calvarial form, Fig. 2. Upper left: Uniform components, equations (10). Other panels: Principal warps of nonzero bending, plotted, arbitrarily, at squared Procrustes length 0.036. b.e. denotes specific bending energy.

development of this neurocranium divides cleanly into two epochs. The change of direction occurs at the cessation of brain growth, about 30 days of age. Before that age, growth is aligned with the direction of greatest growth of the brain itself. However, none of the localizable aspects of this growth process (displayed in the other five panels of the figure) participate

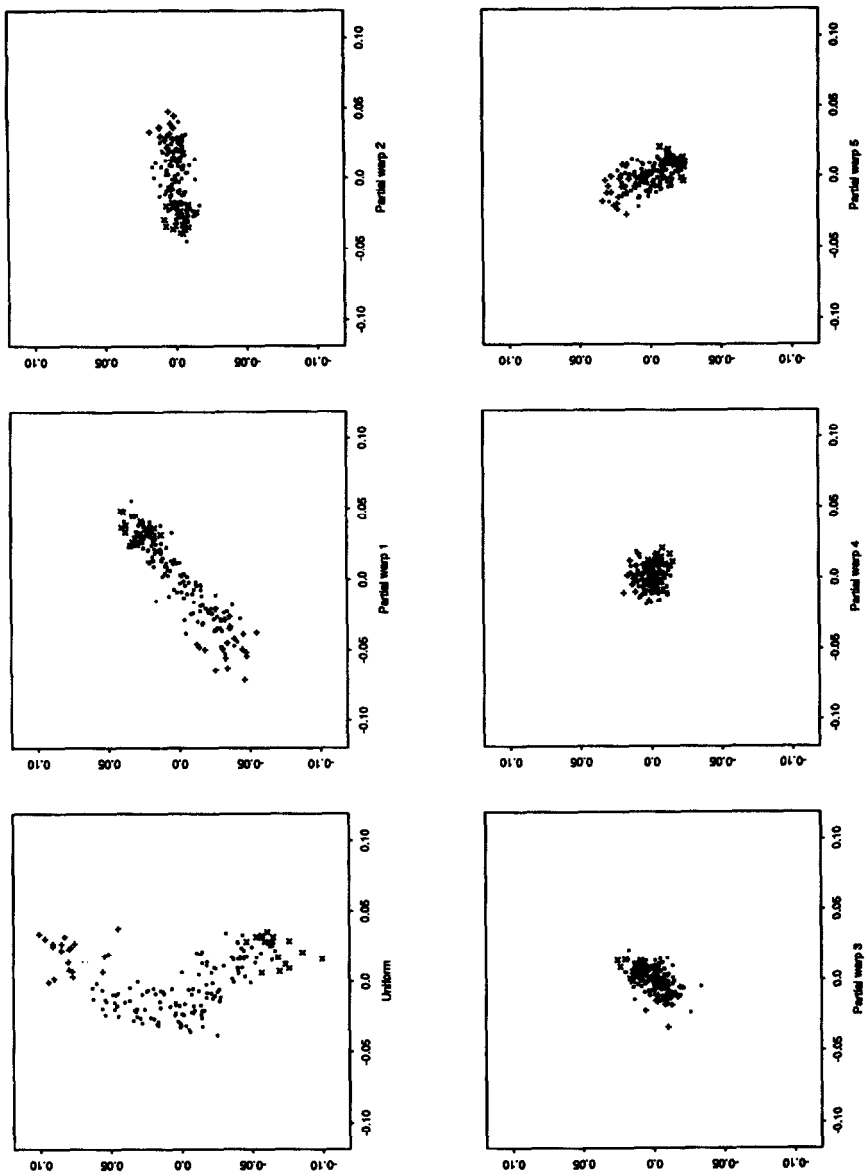

Figure 12. Partial warp scores for the rat data, in the same order of panels as in Fig. 11. These six pairs of shape coordinates are a Procrustes rotation of the eight pairs in Fig. 2. +, 7 = day-old pups; ×, 150-day-old rats.

in this rearrangement mid-development. Except for that uniform term, every partial warp is consistent with regulation of form by one single factor: systematic "orthocephalization" (straightening of the cranial base angle, observed more completely using landmarks further forward) together with a highly localized reconfiguration of the occipital joint. An ordinary (un-

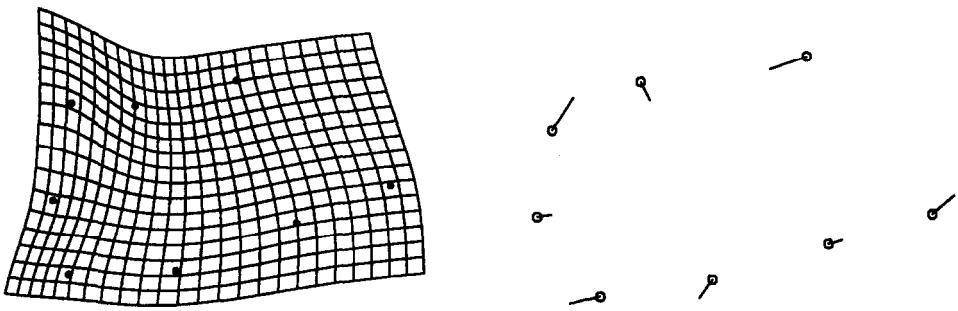

Figure 13. First unscaled principal component or the data in Fig. 12, excluding the uniform component. Left: As a thin-plate spline (equivalently, a superposition of partial warps) for 2 standard deviations' worth of change from the average shape. Right: As a set of displacements Procrustes-fit to the same average (the coordinates of Fig. 2).

scaled) principal component analysis of the last five panels of these shape coordinates confirms that one component explains 83% of all 10 dimensions of Procrustes variation. The other dimensions are effectively spherical error—the second nonuniform principal component explains only one-twentieth as much Procrustes variance as this first one. Drawn as a spline (Fig. 13), the interpretation of this single "growth factor," representing the entire period of growth in the nonlinear subspace, is obvious: a local rearrangement at occiput superposed over a square-to-trapezoid transformation oriented along the couple of cranial base and vault. This visualization improves upon the separate displays of Fig. 12 (again, excluding the panel at upper left) about as much as Fig. 12 already improved upon Fig. 2, the original display of Procrustes-fit coordinates themselves. Notice, for instance, the conversion of the variance at lower right in Fig. 12 into the "pinching" at upper left in the grid, the main feature of the corresponding principal warp (Fig. 11, lower right). This example is discussed in more detail in Bookstein (1991).

The achievement of biomathematical enlightenment does not always dispense with biometric analysis so straightforwardly. Figure 14 presents (as lofted surfaces) the 10 nontrivial principal warps for the shapes in the schizophrenia data set, all drawn to the same Procrustes length. The specific bending energies here vary over more than 1.5 orders of magnitude, from 3.25 to 148. Notice how the first few very smoothly deal with large-scale aspects of the bending of this shape (compare that at the upper left, for instance, to the corresponding panel in Fig. 11). The last few, conversely, are rearrangements of successively smaller features of the form, ending with the smallest triangle of landmarks. The sample scatters of the 28 specimens on the corresponding partial warps, preceded by the uniform

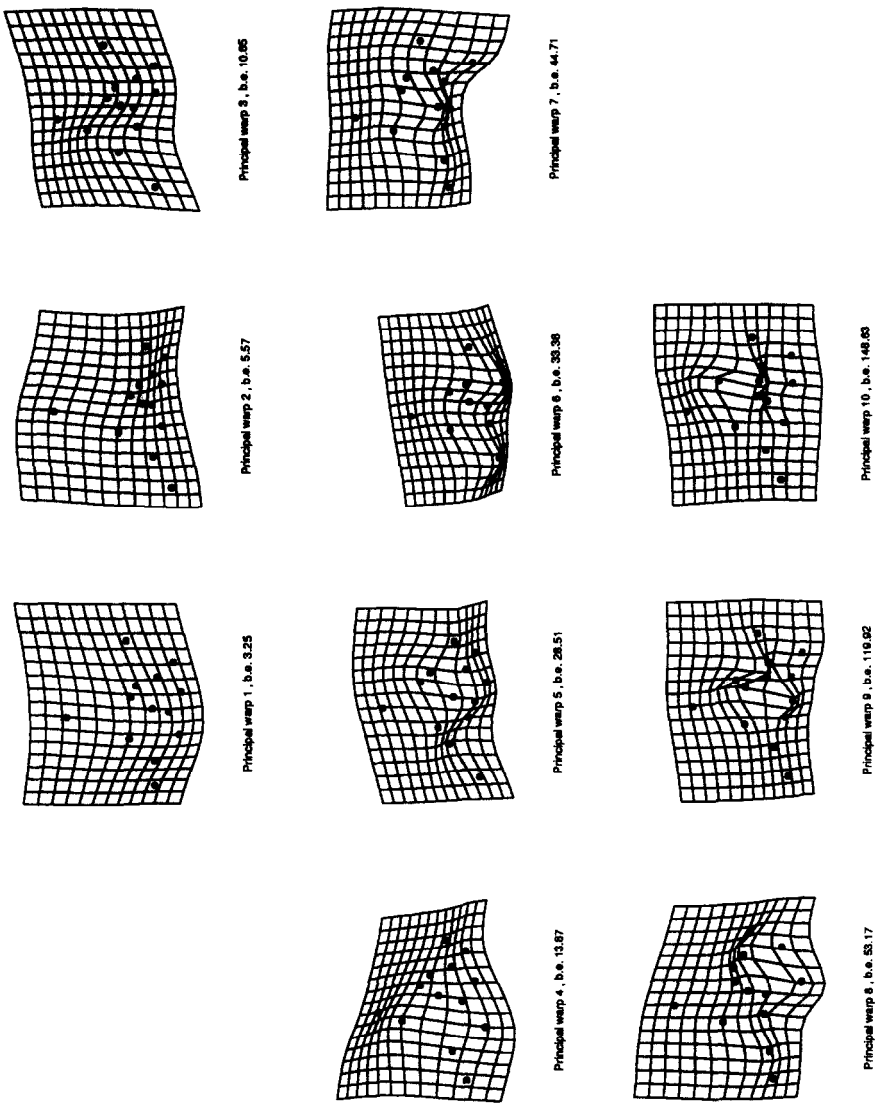

Figure 14. Nonuniform part of the orthonormal shape basis at the mean midsagittal brain landmark configuration, Fig. 4. All 10 principal warp surfaces are drawn at squared Procrustes length 0.036. Both apparent roughness and apparent localization clearly increase with specific bending energy.

component, are shown in Fig. 15. As before, shape coordinates for the schizophrenics are indicated by triangles. Notice the unexpected evenness of Procrustes range across the directions and the scales of these warps. The uniform term accounts for only 26% of the total Procrustes variance, and the first partial warp is not substantially more variable than several others

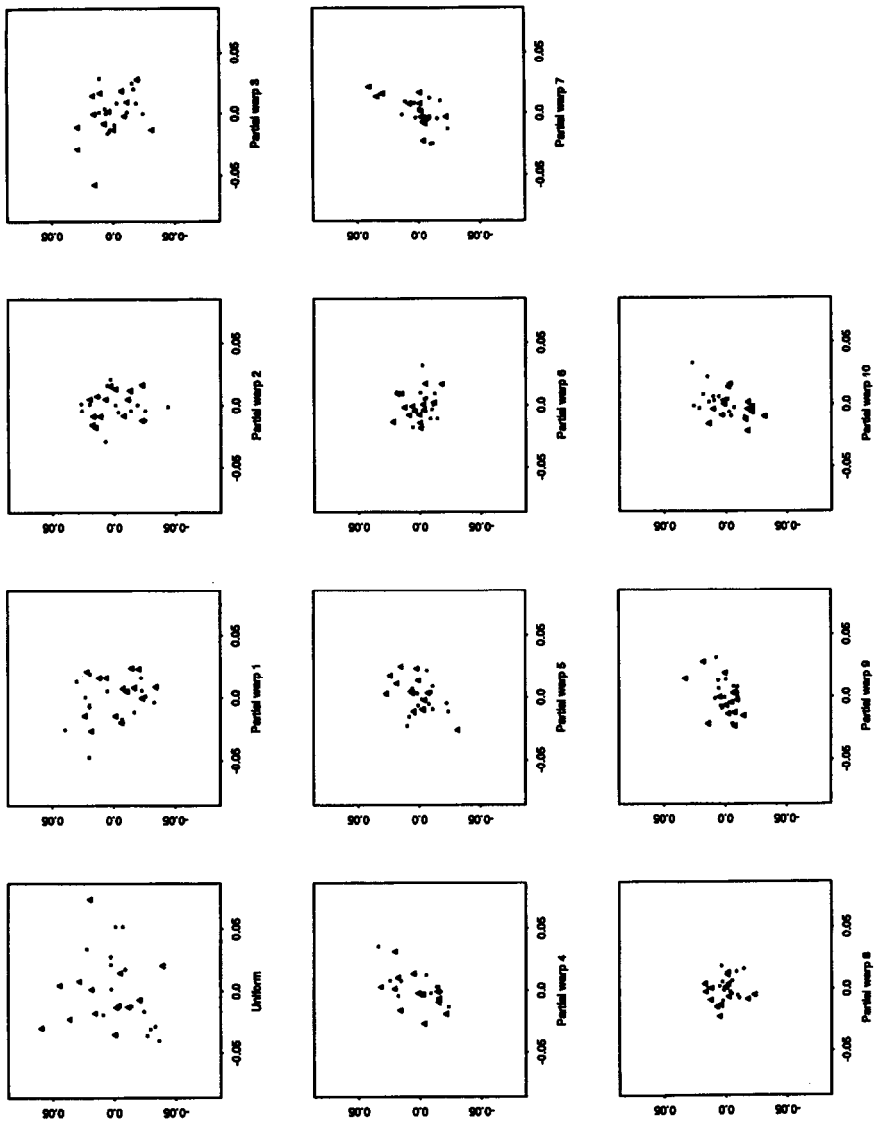

Figure 15. The most useful Procrustes rotation of Fig. 4: the uniform component and the 10 partial warp scores of the brain data set. Triangles, schizophrenics; dots, non-schizophrenics.

at smaller scale. Unlike the situation in Fig. 12, there is not yet any obvious biomathematical finding here.

4. The Biometrics of Landmark Locations: Tests and Diagrams. Those are all the tools necessary for biomathematical-biometrical analysis of

landmark data. The purpose of this paper is to familiarize the reader with this bridge between biomathematics and biometrics, not to explore the neuroanatomy of schizophrenia. Hence discussion here is limited to two further examples: a biometric significance test for the biomathematically suggestive feature visualized in Fig. 7 and a biomathematically sensible version of the familiar biometrical tactic of *ordination* (Reyment, 1991)—a helpful heuristic for the detection of potential patterns at small scale.

For testing linear hypotheses about the independence of shape from exogenous factors (such as Centroid Size or a grouping variable), any full-rank set of $2k-4$ linear combinations of the $2k$ Procrustes-fit coordinates is as good as any other: all yield equally reasonable approximate descriptions of exogenous covariances with shape, whether or not the basis is orthonormal (i.e., whether or not Procrustes distances are preserved). One very convenient set of such skew coordinates is the set of *two-point shape coordinates*, complex affine ratios $(q-p1)/(p2-p1)$ that represent the position of all the other landmarks q, one at a time, after two in particular ($p1$ and $p2$) have been rotated, translated and scaled specimen by specimen to $(0,0)$ and $(1,0)$. I thought I introduced these coordinates in Bookstein (1984b, 1986) for the general testing of linear hypotheses about shape, only to learn belatedly that Francis Galton had published the same construction three-quarters of a century earlier, in 1907 (see Pearson, 1914–1930, Vol. 2, p. 325), in connection with a scheme for telegraphy of facial profiles of criminals. Up through terms of first order in shape variance, any rigid triangulation of a set of landmarks, totalling $k-2$ pairs, supplies an adequate basis for this purpose (Bookstein, 1986).

Figure 16 shows the coordinates of the remaining 11 landmarks of the schizophrenia data set to a baseline from optic chiasm (11) to colliculus (13). In general this is a poorer superposition than that of Fig. 4, as most landmark-specific scatters are larger. At right is an enlargement of the distribution of splenium (1) in this superposition, with the pairs for the schizophrenics marked by triangles. There are clearly distinct mean tendencies for the two groups in this scatter—Hotelling T^2 is significant at 0.0035, somewhat more persuasive than the corresponding findings in Fig. 4. The additional precision arises because the "reference size" against which this shape separation was constructed is the baseline for the shape coordinates, itself fairly localized, instead of the globally estimated Centroid Size (Fig. 1, middle row) that applied to scale the set of Procrustes-fitted coordinates as a unit. The appropriate Bonferroni correction is a factor of approximately 11, not the 78 we needed before, because the 11 shape-coordinate pairs together span the 22-dimensional shape space.

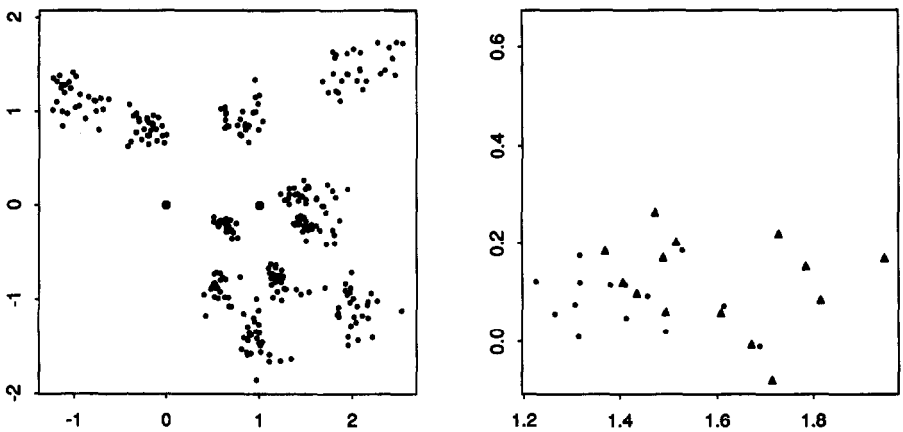

Figure 16. Shape coordinates for the brain data set to a baseline from chiasm to colliculus (landmarks 11 and 13 in Fig. 3). Left: full shape basis; right: enlargement of the coordinates of splenium (landmark 1). The central tendency of the schizophrenics (triangles) is clearly different from that of the others.

When sample sizes are small, as is the case here ("small" means, roughly, "fewer cases than four times the number of landmarks"), it is often helpful to exploit a statistic that involves no such selection or correction. The distributions of these shape coordinates inherit a good deal of symmetry from symmetries in the original landmark distributions. If, for example, the landmarks were generated by independent and identically distributed circular Gaussian perturbations around any mean configuration, then the distribution of derived shapes in shape space is spherically symmetric in Procrustes distance (Mardia and Dryden, 1989; Goodall and Mardia, 1991). In concentrated data, this symmetry extends into the tangent-space projection as well (Mardia, 1995). If the actual distribution of shape coordinates for one's sample looks not too far from spherical—individual shape-coordinate pairs reasonably circular, partial-warp scores not too correlated—this so-called *null model* may be plausible. The orthogonal sections in Fig. 15 are, for this sample size, remarkably close to circularly distributed with the same radius. It may therefore be appropriate to compare the observed Procrustes distance between group means with its expectation on this null model as an F-ratio (Goodall, 1991, p. 314). On this null model, in the limit of small variation, the quantity

$$\frac{N_1 + N_2 - 2}{N_1^{-1} + N_2^{-1}} \frac{\|\bar{F}_1 - \bar{F}_2\|^2}{\Sigma_{\text{groups, cases}} \|F - \bar{F}_i\|^2} \tag{11}$$

is distributed as the statistician's $F_{2k-4,(2k-4)(N_1+N_2-2)}$, where N_1 and N_2 are the two group sample sizes, $\bar{F}_1$ and $\bar{F}_2$ are the group average shapes in the common (pooled) Procrustes registration and $\|\cdot\|^2$ is squared length. The denominator of the expression in the distances is the sum of all squared Procrustes residuals over landmarks and cases.

In the data set here, the squared Procrustes distance between the two group average shapes is 0.00144, and the complete set of 364 squared Procrustes residuals from the appropriate group means totals 0.1389. The corresponding F-ratio is 1.892 on 22 and 572 degrees of freedom; it is significant at $p \sim 0.01$. In this way the biomathematical finding of Fig. 7 is finally issued an appropriate formal biometrical certification of (im)plausibility on a null hypothesis. Other Bayesians and I are aware that this computation is unrelated to any rational approach to scientific reasoning, but editors have come to expect it.

The orthonormal basis for shape space that is a part of the standard morphometric toolkit is the set of partial warps $(U_{-3} \otimes I_2)^t F$, together with a "zeroth" (uniform) component $Un^t F$, for which the entries of L_k^{-1} (and hence U_{-3}) and the coefficients of equation (10) derive from the numerical values of the sample mean shape. Because ordinations of this basis cover the directions of shape space evenly, we can use it to search for directions of excess shape variance in a manner unbiased by differences in the Procrustes lengths and Procrustes angles of the diverse distance ratios, interlandmark angles and so forth, that would otherwise be used. (Recall the discussion of Section 2.) A variety of principal-component strategies, collectively known to the morphometric fraternity as *relative warps analysis* (Rohlf, 1993; Bookstein, 1995b), may be mounted using subsets of this single basis. I have already demonstrated one version of this analysis—the principal component analysis of the full Procrustes-fit coordinate set, which is equivalent to a principal-coordinates analysis of all pairwise Procrustes distances among the specimens of the sample. This initial exploratory tool proved definitive for the growing rat skulls, but unhelpful for the schizophrenia data.

It is often useful to modify this analysis in the direction of greater biomathematical content, by scaling the partial warps by some power α of the bending energy before submitting them to principal components analysis. The new shape basis $(U_{-3}\,\mathrm{diag}(e_i^{-\alpha/2}) \otimes I_2)^t F$ is still orthogonal but is no longer orthonormal in the Procrustes geometry. The modified analysis incorporates specifically biomathematical analytic strategies as follows. By *underweighting* the principal warps in proportion to specific bending energy (the case $\alpha > 0$), we alter the Procrustes ordination to emphasize large-scale aspects of the shape change regardless of the distribution of landmarks at

medium and small scales. Analyses like these are important in studies of growth gradients and of phenomena, such as biomechanical influences, likely to manifest themselves mostly at large scale (see Bookstein, 1991, section 7.6). By *overweighting* the elements of our previously orthonormal basis in proportion to specific bending energy (the case $\alpha < 0$), we alter the Procrustes-orthonormal ordination to more heavily weight discrepancies between landmarks at close spacing (changes on the partial warps of smaller geometric scale). The analysis of Procrustes distance itself uses no such correction: $\alpha = 0$. Other values of α require us to jettison the uniform subspace.

The case $\alpha = -1$, in which squared Procrustes length is scaled by exactly the first power of specific bending energy warp by warp, seems natural for neuroanatomical explorations, inasmuch as small-scale anomalies can have very large effects upon behavior. It also incorporates the collective wisdom of the neuropsychiatric literature: were large-scale anomalies of form to explain the bizarrerie of schizophrenia, they would have been uncovered decades ago. The $\alpha = -1$ computation reduces to a principal coordinate analysis of bending energy in place of Procrustes distance. Figure 17 shows, at the left, the first two relative warps (principal coordinates of bending energy) computed in this way for the schizophrenia data set, and, at the right, the scatter of the first two corresponding scores. The import of this first relative warp is clarified in the plot of Procrustes loadings at upper center, where the localization to the segment between colliculus and splenium is very clear. The eigenvalues of these two warps are 0.697 and 0.440, vs 0.260 for the third. The t-test for the difference between the groups on relative warp 1 (the horizontal in this scatter) is significant at 0.0016. This is not too different from that corresponding to the particular pair of shape coordinates we selected for Fig. 16, but no longer requires any Bonferroni correction for selection, because the decision to look at the first relative warp was made in advance of looking at any data. (The purist might argue that we must correct for the decision to look at the $\alpha = -1$ result instead of those for $\alpha = 0$ or $\alpha = 1$, but the search for small-scale structure is itself a reasonable a priori judgment based in the desirability of small-scale findings, as justified in the foregoing text.) The effect of schizophrenia is found at the smallest scales represented in this data set: a reward for numerosity of landmarks. In fact, this finding is even stronger than the group mean difference of position on the last partial warp (the lower right frame in Fig. 15) owing partly to the concentration of the group difference in one direction rather than two and partly to the combination of more than one partial warp in this first relative warp.

In this way a standard biometric technique, applied to the core structure of the morphometric synthesis, has produced a very promising biomathe-

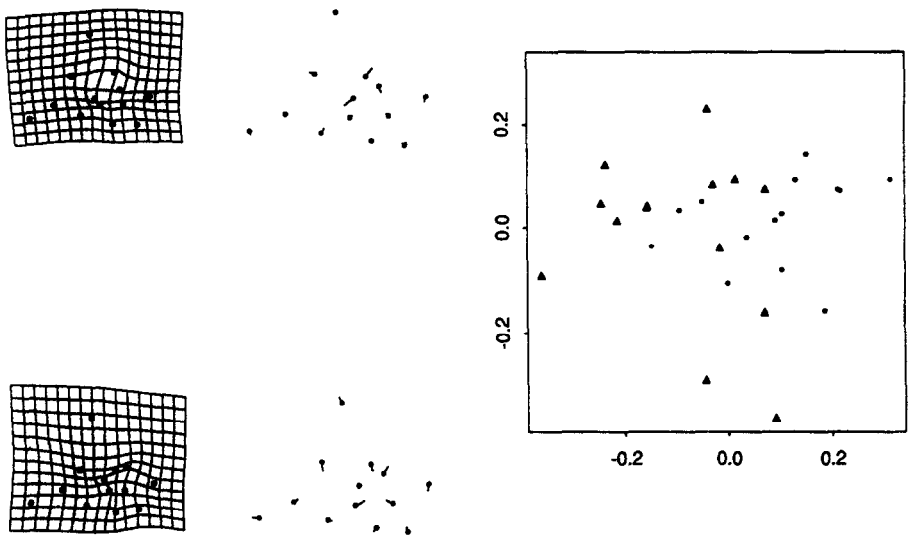

Figure 17. A search for small-scale structure: relative warp analysis of the midsagittal data with $\alpha = -1$ (see text). Left: The first two relative warps of all 28 shapes, shown as partial warps and as Procrustes displacements. Right: Scatter of the first two relative warp scores (triangles, schizophrenics). Compare the analogous scatter for shape coordinates of splenium, Fig. 16.

matical finding by otherwise conventional matrix manipulations of a derived shape basis carefully constructed to be biomathematically reasonable.

5. Biometric Image Warping: "Vertical" and "Horizontal" Morphometrics. The thin-plate spline is, above all, a deformation function. Its original assignment was to warp any picture plane containing landmarks onto any other picture plane containing homologous landmarks. We have exploited the linear structure of those functions to interpret aspects of the biometrics of landmark shape, and we have shown how these deformations affect graph paper, but we have not yet applied those deformations to the original images from which our landmarks arose.

Indexing the pictorial content of the original biomedical images by landmark locations subordinates the original Cartesian coordinate system to the considerably more biomathematical chart given by a rich landmark configuration. While operations of image *processing* (gradient computation, region-growing) are hardly affected by such relabelings, subsequent aspects of image *analysis* are typically far more effective after adjustment of the full biometric space of landmark rearrangement, the shape space of the preceding section. Here "adjustment" is by splined unwarping, not by any procedure linear in the original pixels. The subtler the effect that later analysis is

intended to uncover, the more important is this preliminary shape normalization. The splines serve to align all the pictures of a sample so that, as best the expert morphologist can manage, biological homology obtains across the whole stack pixel by pixel. In such a registration, scientific strategies like pixel averaging or blob detection will have maximum biometric power.

Predictably, the same Francis Galton who produced the first shape coordinates also carried out very early experiments with image averaging (see the discussion in Pearson, 1914–1930, Vol. 2, pp. 283–300). By placing mirrors in the optical path of his enlarger, Galton managed to produce affine corrections by entirely physical means. We are interested in more complicated warps and we would prefer to effect them computationally rather than optically. To manage this operation for landmarks, one splines the sample average shape onto the specimens and then exploits the *inverse* mapping (which is not quite a thin-plate spline itself) pixel by pixel to pull back the pixel values of the individual image to the corresponding loci of the fixed (squared) Cartesian coordinate system of the average. For each group of brains, for instance, there is a Procrustes average landmark shape. For each of the 14 specimens of each group there is a spline that warps the group average onto the picture plane of that image. Each pixel from the plane of the average shape is mapped by this spline to some location in each specimen image. Pixel by pixel, the values encountered in the specimens are copied back to the locations in the original square grid, the "standardized picture." One such pullback is shown in Fig. 18. After they are averaged in place, 14 per group, there result the two frames of Fig. 19. The shape difference of Fig. 7 is still there, but many other contrasts have become visible as well.

These two average images still incorporate different average landmark shapes. We can visualize the group mean difference in one single consistent geometry if we further unwarp each group average image onto the same average shape. This could be a grand mean or, as here, the average for the nonsyndromal group. After this final geometric adjustment, there result the two average images of Fig. 20 and the pixel-by-pixel image difference shown in Fig. 21. Here, at last, are aspects of the original images that are wholly independent of differences in their landmark geometry: the thinning of the schizophrenics' corpus callosum all along its length, for instance, and also the reduction of extent of the thalamus (upon which there happen to be no landmarks). The thinning is particularly marked along a transect from splenium to genu (a "horizontal" diameter of the callosum). The analysis here is a great deal more powerful than that of Andreasen *et al.* (1994), which, by using only an affine image registration (a Galtonian "bounding

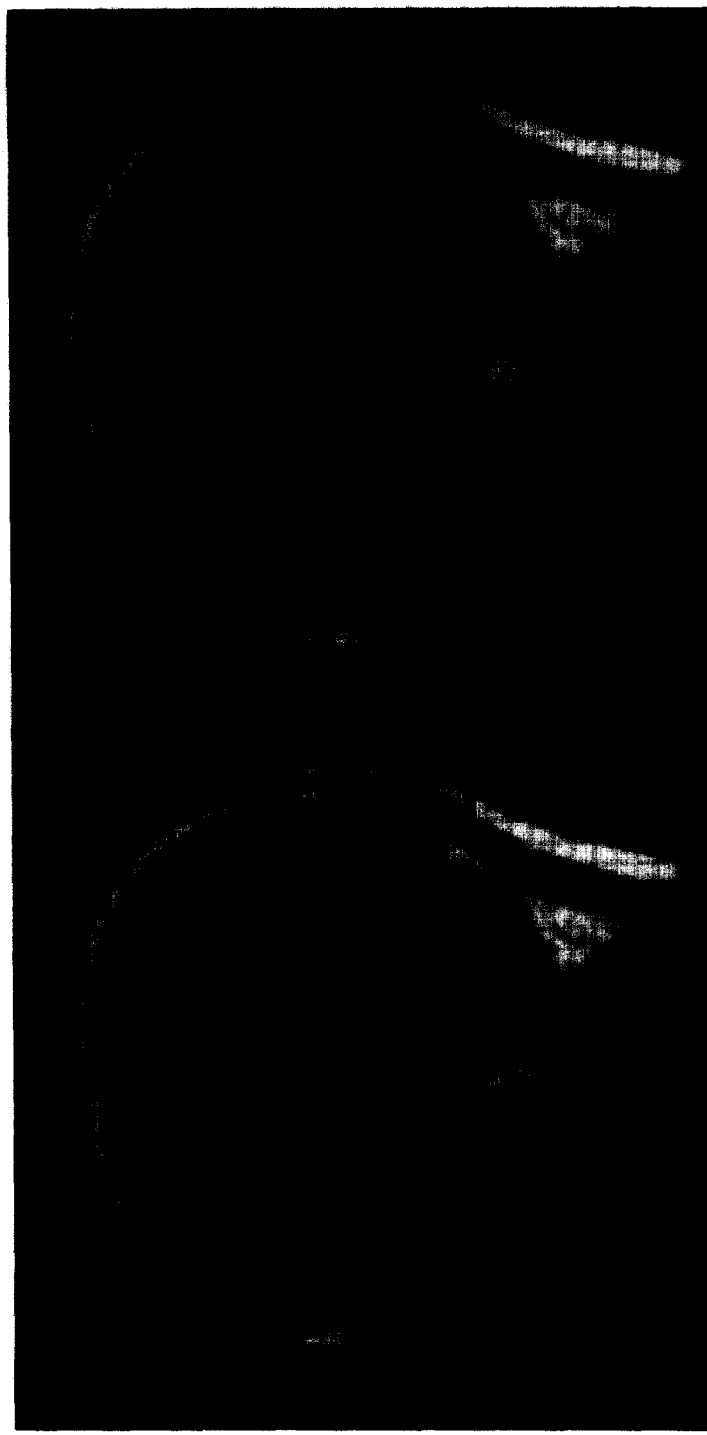

Figure 18. Unwarping of the image in Fig. 3 to the average of the nonschizophrenic group. Right: original image, with original landmarks. Left: Average landmarks, with pixels pulled back in accordance with the splined grid onto the right-hand side. This and the subsequent three figures were produced by William Green's program package Edgewarp (see Bookstein and Green, 1994a, b).

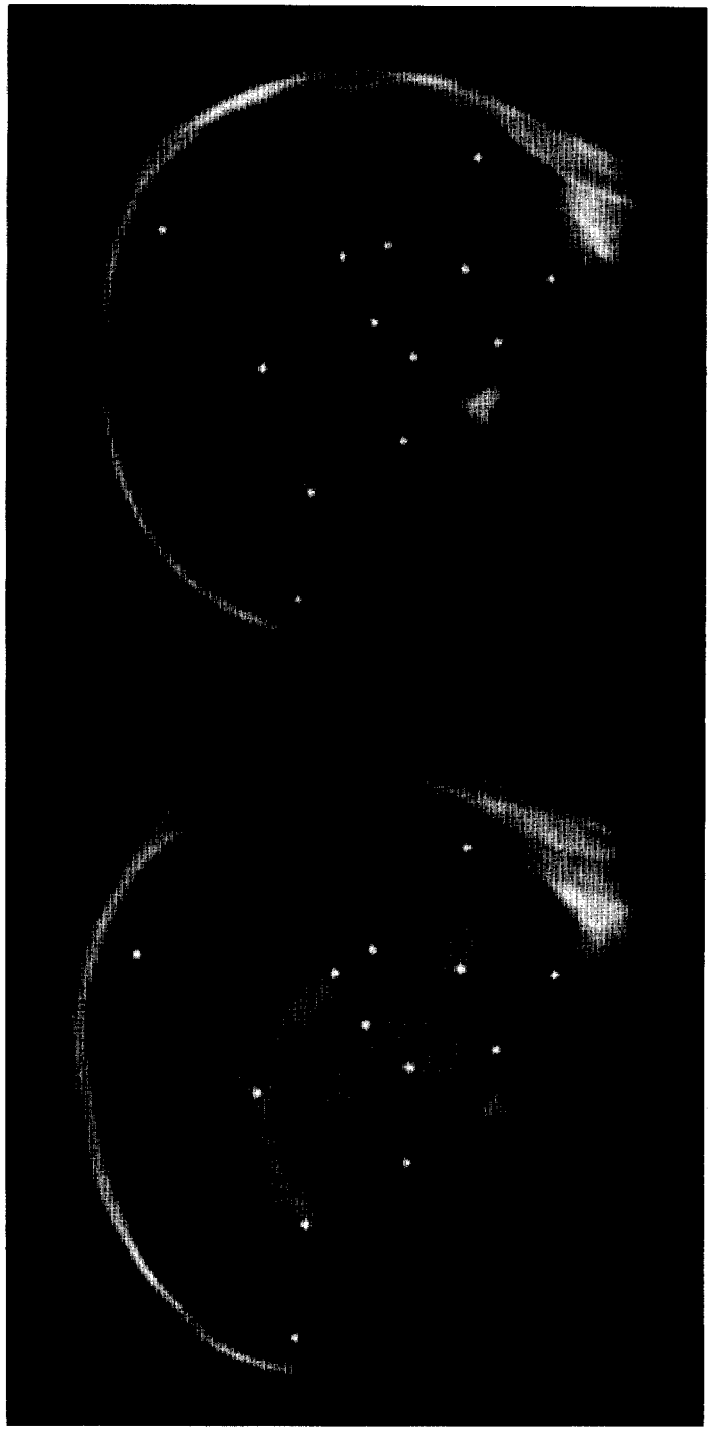

Figure 19. Averaged images of nonschizophrenics and schizophrenics after unwarping each to its group's average shape.

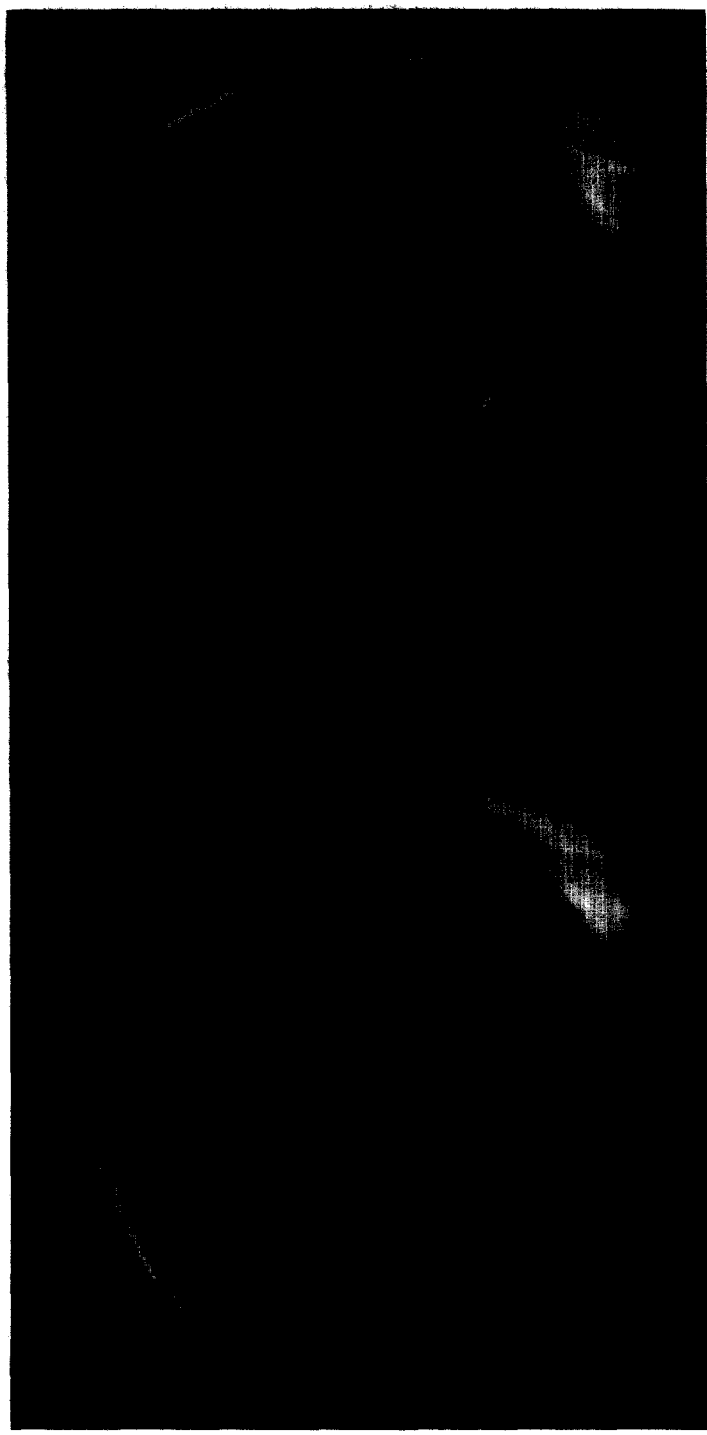

Figure 20. Further unwarping of the right-hand frame in Fig. 19 onto the mean shape for the nonschizophrenics.

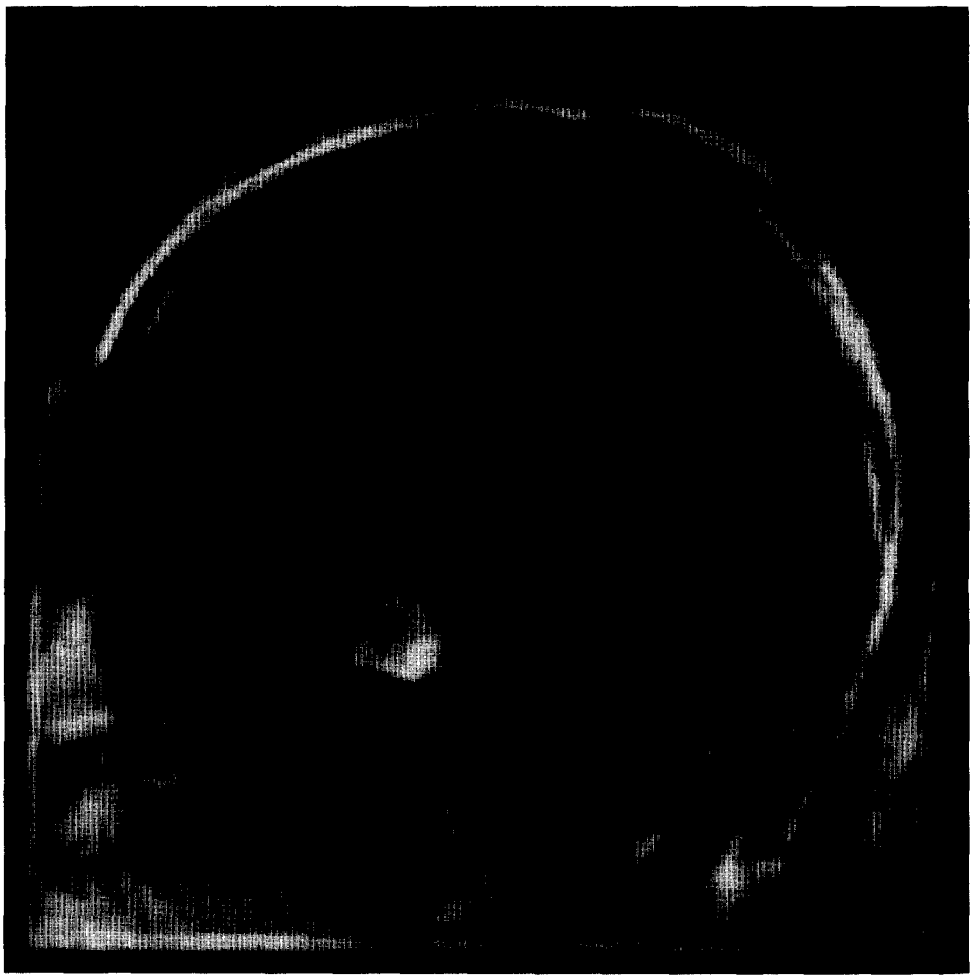

Figure 21. Difference (right minus left) of the frames in Fig. 20: the effect of schizophrenia upon midsagittal anatomy controlling for the position of all 13 landmarks. Zero difference is set to gray level 128. Notice the emergence of additional biomathematical structure, notably along the inner border of corpus callosum, the upper margin of the third ventricle.

box"), systematically blurs the averages and their contrasts in the vicinity of all deep structures.

At the same time, Fig. 21 makes it quite clear that we have not completed the analysis of these images. The residual image here suggests a well-localized spatial structure that overlaps a real anatomical locus—the boundary between corpus callosum and third ventricle—but this feature of form is not yet encoded in the specifically biometric space of landmark

features. We need further landmarks on the inner border to bring this obvious biomathematical signal into the biometrical domain. Using a different sample of brain images, Bookstein (1995c, d) demonstrates how this next step is carried out.

This example hints at a powerful general principle. The images we are conveying for morphometric analysis begin as physical records of interaction with radiation within each cell of a parcellation of an organism's surface or interior. The engineers and physicists who design the recording instruments are tempted to continue representing these often very expensive data sets as scalars or vectors in the Euclidean space of the raw data record, just as they come off the scanners. In that mode, for example, elegant suggestions about geometrically salient features of curving surfaces arise by direct geometric differentiation of surface forms (Koenderink, 1990; Porteous, 1994). However, for our biomathematical context, the temptation to treat Cartesian coordinates of an empirical surface stripped of biological labels as having any possible cogent biological meaning in that form must be strenuously resisted. In particular, conventional image processing algorithms, which invariably compute gradients, convolutions with Gaussian blurs, spectral representations and so forth, pixel by pixel, seem quite incapable of making any biomathematical sense of these records. Because biomathematics cannot make any use of the pixellated coordinate system, it seems pointless to carry out biometrics that way (but see Friston, et al., 1994, for a non-biometrical application). Let us refer to this orientation of linear statistics, that which uses pixel locations for its index set, as *vertical*. Figure 21 is, on its face, a purely vertical presentation: the difference between the two sides of Fig. 20, pixel by pixel.

There is always additional information, then, in the *horizontal* part of this construction: information about where the labeled locations and gradients of the anatomical sketch or textbook diagram actually lie with respect to the pixels, and how their configurations covary with height(s) of the data surface(s) above them. Whenever data are originally visual, and especially if they were originally pixellated, the linear machinery *must* be supplemented by a biomathematical semantics of deformation. The labeled points and directions thereby may move about in their Euclidean domain at the same time that images change above them, leading to multidimensional morphometric patterns that are very interesting both biometrically and biomathematically.

An ultimately vertical analysis uses the geometry of the landmark-labeled image rather as one uses a covariate in a classic experimental design. It is as if the landmark configuration is nuisance variation, instrumental noise, control of which increases the precision with which other phenomena can be addressed. In analyses like these, one wishes to understand picture

gradients or their correlations with physical or biological processes as if they were painted on a shape prototype: we unwarp horizontally only to sharpen this representation. The careers of generations of comparative anatomists testify to the power of this maneuver. At the same time, we must preserve the transformation thus "quotiented out" in some parametric form, with a count of parameters low enough to support some kind of biometrically competent inference about the shape differences being thus "corrected," or else the pictorial findings remain uninterpretable in any biotheoretical context, where shape is just as likely to be signal as to be noise. Methods of fully distributed displacement analysis, such as that recently introduced by Grenander and Miller (1994), do not at present offer such low-dimensional parameterizations. Although those methods often align images for vertical processing better than landmarks can manage, there is no complementary horizontal method and, thus, again, no possibility of arriving at a deeper biomathematical understanding of form.

The morphometric synthesis reviewed here provides the first coherent general praxis for this purpose. Its combination of Procrustes fits, splines and partial warps sustains effective multivariate biometrical analyses of the horizontal at the same time that the vertical is standardized for more subtly theory-laden investigations. Our investigation of brain scans here is intended as a prototype for that broader praxis, the bridging that is my main theme: a vertical summary (Fig. 21) intertwined inextricably with a horizontal one (Fig. 7). The combination leaves to either domain, horizontal or vertical, what is most nearly linear there, and encapsulates each set of crucial nonlinearities (the geometry of shape space, the kernel of the spline, the interplay between landmarks or voids and the function of organs) in a manner conceptually orthogonal to equally necessary manipulations in the other domain.

The triumph of modern multivariate statistical methods in fields arbitrarily far from their biometric origins has seriously distracted us from properly understanding the true meaning of these methods in the quantitative biological sciences. The meaning of statistical methods is inextricably bound up in what a community of scholars believe to be the meaning of their data (cf. Kuhn, 1959; Latour, 1987). The easy availability of matrix manipulations and the ease with which they can lead to publications and tenure is no substitute for an understanding of the nature of the tie between "the data" and the styles of explanation that actually drive the discipline in question. As the example of morphometrics indicates, there need be no prior mathematical model of a phenomenon (for instance, of skull growth), and yet the geometric dissection of the observed patterns of that phenomenon can be effective and suggestive as long as there is a satisfactory quantitative model of the descriptive process itself, the formal-

ism of landmarks and deformations by which the patterns on the scientist's retina are converted into explanations. Figure 7 represents the production of a biomathematical hypothesis by appropriately enlightened biometrical manipulations; likewise Fig. 13; likewise Fig. 19; likewise Fig. 21. Any of these exemplify the discipline at the arch of the bridge that is the morphometric synthesis, but no one of them is complete by itself.

The duality between biomathematics and biometrics embodied in the morphometric synthesis has thus been present just beneath the conventional rhetoric of quantitative studies of form all along, waiting to be unearthed in the course of seemingly unrelated advances in geometrical statistics and image processing. Within Thompson's language of "Cartesian transformations" has always lurked a profound analytic insecurity: how do we know whether we have represented underlying shape phenomena well enough to justify confidence in the explanations suggested by one grid or another? By embracing the biometric language of "adjustment," then applying the absolutely simplest of associated analytic tactics (the construction of the mean difference in Fig. 21), we have come around at last to close the circuit of alternate rhetorics of quantitative biology exposed in my Introduction. The shape space spanned by a set of landmarks can be biomathematically meaningful and reliable to the extent that figures like these, analyses of its residuals, show no further *biomathematical* signal, no further hints of structure. (The evidence in Fig. 21 reminds us that we have not finished the analysis of the schizophrenia data set in this respect.) Similarly, the biomathematical analysis of empirically encountered splines, whether pure partial warps or linear combinations of warps, can be assessed by the extent of the *biometric* spectrum that goes unexplained thereby. This spectrum is not limited to the Procrustes space of landmark shape alone, as in Fig. 15, but extends easily to the additional data supplied by image gradients having their own locations and orientations (Bookstein, 1994; Bookstein and Green, 1993).

The morphometrics of the synthesis bridges grids and vectors, then, by construing each of these abstractions as a partial description of underlying phenomena that must be viewed simultaneously in terms of the geometry of the "labels," those maps and grids, and in terms of the image context, which stands in for the underlying structural processes actually governing. The morphometric synthesis is thus at root a mutual completion of previously competing partial descriptions of the same underlying shape processes. Geometrically, the horizontal–vertical model of biomedical image analysis bridges the schools by tying the landmark location data to the report of a difference. Mathematically, that report is tested in two ways: by whether it conforms to some sensible understanding of how parts of an organism are integrated and, separately, by whether it appears to exhaust

the available information about loci that undergirds the biometric descriptor space on which the investigator has settled.

In the absence of any of these multiple, interlocking verifications—one formalism for reliability of the shape phenomenon, another for biomathematical description of that phenomenon, a third for completeness of that description—I would expect neither quantitative findings about form nor their explanations to be stable across changes in instrumentation or ordinary fluctuations of sample design. The morphometric synthesis of the 1980s, by reminding biometrics of its roots in the observation of organic form, is a fine example of how methodological discipline furthers applied quantitative investigation. In my view, the synthesis is a major methodological triumph, one to be relished by both statistics and mathematical biology, that has for the first time matched descriptive and inferential technique to a powerful classical mode of qualitative biological intuition.

This paper draws on ten years of methodological investigations and collaborations. My colleagues in the synthesis include F. James Rohlf, Colin Goodall and Kanti Mardia, who generated large pieces of its statistical theory or biometric setting, and Richard Reyment, Paul Sampson and Leslie Marcus, fellow teachers and defenders of the faith. William D. K. Green created the Edgewarp package in which the schizophrenia images were processed. Much of the research reported here was underwritten in part by NIH grants DA-09009 and GM-37251 to Fred L. Bookstein. The first of these grants is jointly supported by the National Institute on Drug Abuse, the National Institute of Mental Health and the National Institute on Aging as part of the Human Brain Project.

LITERATURE

Andreasen, N., S. Arndt, V. Swayze, T. Cizadlo, M. Flaum, D. O'Leary, J. Ehrhardt and W. Yuh. 1994. Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* **266**, 294–298.

Blackith, R. and R. Reyment. 1971. *Multivariate Morphometrics*. New York: Academic Press.

Bookstein, F. L. 1978. *The Measurement of Biological Shape and Shape Change. Lecture Notes in Biomathematics*, Vol. 24. New York: Springer.

Bookstein, F. L. 1982a. On the cephalometrics of skeletal change. *Amer. J. Orthodontics* **82**, 177–198.

Bookstein, F. L. 1982b. Foundations of morphometrics. *Ann. Rev. Ecology and Systematics* **13**, 451–470.

Bookstein, F. L. 1984a. A statistical method for biological shape comparisons. *J. Theor. Biol.* **107**, 475–520.

Bookstein, F. L. 1984b. Tensor biometrics for changes in cranial shape. *Ann. Human Biol.* **11**, 413–437.

Bookstein, F. L. 1986. Size and shape spaces for landmark data in two dimensions. *Statis. Sci.* **1**, 181–242.

Bookstein, F. L. 1989a. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Machine Intelligence* **11**, 567–585.

Bookstein, F. L. 1989b. "Size and shape": a comment on semantics. *Systematic Zoology* **38**, 173–180.

Bookstein, F. L. 1991. *Morphometric Tools for Landmark Data*. New York: Cambridge University Press.

Bookstein, F. L. 1994. Landmarks, edges, morphometrics, and the brain atlas problem. In *Functional Neuroimaging: Technical Foundations*, R. Thatcher, M. Hallett, T. Zeffiro, E. John and M. Huerta (Eds), pp. 107–119. New York: Academic Press.

Bookstein, F. L. 1995a. A standard formula for the uniform shape component in landmark data. In *Advances in Morphometrics: Proceedings of the 1993 NATO ASI on Morphometrics*, L. F. Marcus et al. (Eds). New York: Plenum. To appear.

Bookstein, F. L. 1995b. Combining the tools of geometric morphometrics. In *Advances in Morphometrics: Proceedings of the 1993 NATO ASI on Morphometrics*, L. F. Marcus et al. (Eds). New York: Plenum. To appear.

Bookstein, F. L. 1995c. Combining "vertical" and "horizontal" features from medical images. *Computer Vision, Virtual Reality, and Robotics in Medicine. Lecture Notes in Computer Science*, N. Ayache (Ed), Vol. 905, pp. 184–191. Berlin: Springer.

Bookstein, F. L. 1995d. How to produce a landmark point: the statistical geometry of incompletely registered images. In *Vision Geometry IV. S. P. I. E. Proceedings*, (R. A. Melter et al. (Eds), Vol. 2573. Bellingham, WA: SPIE, pp. 266–277.

Bookstein, F. L. 1995e. Utopian skeletons in the biometric closet. Occasional Papers of the Institute for the Humanities, number 2, Institute for the Humanities, University of Michigan.

Bookstein, F. L., B. Chernoff, R. Elder, J. Humphries, G. Smith and R. Strauss. 1985. *Morphometrics in Evolutionary Biology*. Philadelphia: Academy of Natural Sciences of Philadelphia.

Bookstein, F. L. and W. D. K. Green. 1993. A feature space for edgels in images with landmarks. *J. Math. Imaging and Vision* **3**, 231–261.

Bookstein, F. L. and W. D. K. Green. 1994a. Edgewarp: A flexible program package for biometric image warping in two dimensions. In *Visualization in Biomedical Computing 1994. SPIE Proceedings*. R. Robb (Ed), Vol. 2359. Bellingham, WA: SPIE, pp. 135–147.

Bookstein, F. L. and W. D. K. Green. 1994b. Edgewarp: A program for biometric warping of medical images. Videotape, 26 minutes.

Boyd, E. 1980. *Origins of the Study of Human Growth*. University of Oregon Health Sciences Center. Portland, OR.

Burnaby, T. P. 1966. Growth-invariant discriminant functions and generalized distances. *Biometrics* **22**, 96–110.

DeQuardo, J. R., F. L. Bookstein, W. D. K. Green, J. Brumberg and R. Tandon. 1995. Spatial relationships of neuroanatomic landmarks in schizophrenia. *Psychiatry Research: Neuroimaging*.

Duchon, J. 1976. Interpolation des fonctions de deux variables suivant la principe de la flexion des plaques minces. *RAIRO Anal. Numé.* **10**, 5–12.

Duncan, O. D. 1984. *Notes on Social Measurement: Historical and Critical*. New York: Russell Sage Foundation.

Dürer, A. 1528. *Vier Bücher von Menschlicher Proportion*. Dietikon-Zürich: Josef Stocker, 1969.

Friston, K. J. 1994. Statistical parametric mapping. In *Functional Neuroimaging: Technical Foundations*, R. Thatcher, M. Hallett, T. Zeffiro, E. John and M. Heurta (Eds), pp. 79–93. New York: Academic Press.

Goodall, C. R. 1983. The statistical analysis of growth in two dimensions. Doctoral dissertation, Department of Statistics, Harvard University.

Goodall, C. R. 1991. Procrustes methods in the statistical analysis of shape. *J. Roy. Statist. Soc. Ser. B* **53**, 285–339.

Goodall, C. R. and K. V. Mardia. 1991. A geometric derivation of the shape density. *Adv. in Appl. Probab.* **23**, 496–514.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. *Biometrics* **27**, 857–874.

Grenander, U. and M. Miller. 1994. Representations of knowledge in complex systems. *J. Roy. Statist. Soc. Ser. B* **56**, 549–603.

Hopkins, J. W. 1966. Some considerations in multivariate allometry. *Biometrics* **22**, 747–760.

Hotelling, H. 1936. Relations between two sets of variables. *Biometrika* **28**, 321–377.

Huxley, J. 1932. *Principles of Relative Growth*. London: Methuen.

Jolicoeur, P. 1963. The multivariate generalization of the allometry equation. *Biometrics* **19**, 497–499.

Kendall, D. G. 1984. Shape-manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.* **16**, 81–121.

Kent, J. T. 1994. The complex Bingham distribution and shape analysis. *J. Roy. Statist. Soc. Ser. B* **56**, 285–299.

Kent, J. T. and K. V. Mardia. 1994. The link between kriging and thin-plate splines. In *Probability, Statistics, and Optimisation*, F. P. Kelly (Ed), pp. 325–339. New York: Wiley.

Koenderink, J. 1990. *Solid Shape*. Cambridge, MA: M.I.T. Press.

Kuhn, T. S. 1959. The function of measurement in modern physical science. In *Quantification*, H. Woolf (Ed), pp. 31–63. Indianapolis: Bobbs-Merrill.

Latour, B. 1987. *Science in Action*. Cambridge, MA: Harvard University Press.

Lewis, J. L., W. Lew and J. Zimmerman. 1980. A nonhomogeneous anthropometric scaling method based on finite element principles. *J. Biomechanics* **13**, 815–824.

Lohmann, G. P. 1983. Eigenshape analysis of microfossils: a general morphometric procedure for describing changes in shape. *Math. Geology* **15**, 659–672.

Marcus, L. F., E. Bello and A. García-Valdecasas (Eds). 1993. *Contributions to Morphometrics*. Madrid: Monografias, Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Cientificas.

Marcus, L. F., M. Corti, A. Loy, G. Naylor and D. Slice (Eds). 1995. *Advances in Morphometrics: Proceedings of the 1993 NATO ASI on Morphometrics*. New York: Plenum. To appear.

Mardia, K. V. 1995. Shape advances and future perspectives. In *Proceedings in Current Issues in Statistical Shape Analysis*, K. V. Mardia and C. A. Gill (Eds), pp. 57–75. Leeds, U.K.: Leads University Press.

Mardia, K. V. and I. Dryden. 1989. The statistical analysis of shape data. *Biometrika* **76**, 271–281.

Mardia, K. V. and C. A. Gill (Eds). 1995. *Proceedings in Current Issues in Statistical Shape Analysis*. Leeds, U.K.: Leeds University Press.

Meinguet, J. 1979. Multivariate interpolation at arbitrary points made simple. *Z. Angewandte Math. Phys.* **30**, 292–304.

Mosimann, J. E. 1970. Size allometry: size and shape variables with characterizations of the log-normal and generalized gamma distributions. *J. Amer. Statist. Assoc.* **65**, 930–945.

Netter, F. H. 1989. *Atlas of Human Anatomy*. Summit, NJ: Ciba-Geigy Corporation.

Oxnard, C. E. 1973. *Form and Pattern in Human Evolution*. Chicago: University of Chicago Press.

Oxnard, C. E. 1978. One biologist's view of morphometrics. *Ann. Rev. Ecology and Systematics* **9**, 219–241.

Pearson, K. 1914–1930. *The Life, Letters and Labours of Francis Galton* (Three volumes bound as four). Cambridge, U.K.: University Press.

Porteous, I. R. 1994. *Geometric Differentiation for the Intelligence of Curves and Surfaces*. Cambridge, U.K.: Cambridge University Press.

Reyment, R. A. 1991. *Multivariate Palaeobiology*. Oxford: Pergamon.

Richards, O. W. and A. C. Kavanagh. 1943. The analysis of relative growth-gradients and changing form of growing organisms: illustrated by the tobacco leaf. *American Naturalist* **77**, 385–399.

Rohlf, F. J. 1993. Relative warp analysis and an example of its application to mosquito wings. In *Contributions to Morphometrics*, L. F. Marcus et al. (Eds), pp. 131–159. Madrid: Monografias, Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Cientificas.

Rohlf, F. J. and F. L. Bookstein (Eds). 1990. *Proceedings of the Michigan Morphometrics Workshop*. Ann Arbor, MI: University of Michigan Museums.

Rohlf, F. J. and D. Slice, 1990. Extensions of the Procrustes method for the optimal superposition of landmarks. *Systematic Zoology* **39**, 40–59.

Sibson, B. 1978. Studies in the robustness of multidimensional scaling: Procrustes statistics. *J. Roy. Statist. Soc. Ser. B.* **40**, 234–238.

Sneath, P. H. A. 1967. Trend-surface analysis of transformation grids. *J. Zoology* **151**, 65–122.

Sneath, P. H. A. and R. R. Sokal, 1963. *Principles of Numerical Taxonomy*. San Francisco: W. H. Freeman.

Stigler, S. M. 1986. *The History of Statistics: The Measurement of Uncertainty Before 1900*. Cambridge, MA: Harvard University Press.

Thompson, D'A. W. 1917. *On Growth and Form*. London: Macmillan.

Timoshenko, S. and S. Woinowsky-Krieger. 1959. *Theory of Plates and Shells*, 2nd ed. New York: McGraw-Hill.

Wahba, G. 1990. *Spline Models for Observational Data*. Philadelphia: Society for Industrial and Applied Mathematics.

Wright, S. 1968. *Evolution and the Genetics of Populations. Vol. 1: Genetic and Biometric Foundations*. Chicago: University of Chicago Press.

Received 16 July 1995