Allometry for the Twenty-First Century

Fred L. Bookstein

Abstract The current literature that attempts to bridge between geometric morphometrics (GMM) and finite element analyses (FEA) of CT-derived data from bones of living animals and fossils appears to lack a sound biotheoretical foundation. To supply the missing rigor, the present article demonstrates a new rhetoric of quantitative inference across the GMM–FEA bridge—a rhetoric bridging form to function when both have been quantified so stringently. The suggested approach is founded on diverse standard textbook examples of the relation between forms and the way strains in them are produced by stresses imposed upon them. One potentially cogent approach to the explanatory purposes driving studies of this class arises from a close scrutiny of the way in which computations in both domains, shape and strain, can be couched as minimizations of a scalar quantity. For GMM, this is ordinary Procrustes shape distance; in FEA, it is the potential energy that is stored in the deformed configuration of the solid form. A hybrid statistical method is introduced requiring that all forms be subjected to the same detailed loading designs (the same "probes") in a manner careful to accommodate the variations of those same forms before they were stressed. The proper role of GMM is argued to be the construction of regressions for strain energy density on the largest-scale relative warps in order that biological explanations may proceed in terms of the residuals from those regressions: the local residual features of strain energy density. The method, evidently a hierarchical one, might be intuitively apprehended as a geometrical approach to a formal allometric analysis of strain. The essay closes with an exhortation.

Introduction

The subject of this note is a fundamental problem that stands today athwart one current attempt at developing a statistical method for finite element analysis (FEA). That specific vision, the implementation of strain statistics as some sort of extension or complement to geometric morphometrics (GMM) specialized for biomechanics, was the topic of an extended discussion in the course of the NSF-sponsored meeting *Virtual Anthropology Meets Biomechanics* chaired by Gerhard Weber in Vienna, Austria on October 20 and 21, 2010 (see Weber et al. 2011). In my presentation at that meeting and in the ensuing discussion I attempted to establish the proposition that all the currently suggested modes of tying GMM methods to FEA analyses, both published and unpublished, were so severely defective on logical or mathematical grounds as to preclude their valid application in any empirical bioscientific context.

That 2010 critique wove together a long list of features focusing on incompatibilities of differential equation implementation between the domains, in particular the uniform term, along with incompatibilities of graphical semiotics. One specific concern was the differential

F. L. Bookstein (✉)
Faculty of Life Sciences, University of Vienna, Vienna, Austria
e-mail: fred.bookstein@univie.ac.at

F. L. Bookstein
Department of Statistics, University of Washington, Seattle, USA
e-mail: flb@stat.washington.edu

equation itself, the biharmonic equation (iterated Laplacian). In the thin-plate spline approach, the quantity being minimized is the integral of the squared *second* derivatives, and the integral is taken over all of space, both where the biological object is located and where it is not; but the quantity minimized in the course of an elastic analysis derives instead from the sum of squared deviations of the *first* derivatives from unity, and the integral is taken only over the region occupied by actual material. The uniform term in particular, which is so crucial to the interpretation of growth gradients in GMM because it has zero morphometric bending energy, is not relevant to FEA, because within any finite region it does not minimize elastic energy, so that no real elastic transformations of biological material can ever be uniform in the large. The treatment of semilandmarks is entirely different in the two formalisms, as the "sliding" operation in GMM effaces most of the "membrane energy," the within-surface strain tensor, to which elasticity theory pays close attention, whereas the aspect of translation normal to the mean surface, so important to the statistics of GMM, is negligible when viewed as the strain of a shell, the typical point of view of FEA. Regarding semiotics, the diagram styles commonly encountered in the current literature bridging GMM to FEA analyses omit any indication of variability or dependence on models of load, and no current papers provide any rhetoric for claims that two particular images of strain fields are distinguishable or indistinguishable, and, if the former, whether they are nevertheless "similar" enough to merit a positive interpretation of their commonalities. There is more discussion along these lines in (Weber et al. 2011).

One can review most of these fundamentals without direct reference to FEA's actual equations or implementation details. The dilemmas turn out to derive not from the details of FEA software but from discrepancies between the mathematical physics of elasticity theory, which FEA embodies, and the quite different mathematics (*sans* physics) of GMM.

This specific topic fits quite comfortably within the limits of theoretical biology as a special instance of our community concern with methodology for *studies of form and function*. The prediction of function by form, and also the explanation of form by function, are among the standard modes of reasoning cutting across all the modern biological sciences. Their appearance in the quantitative biological literature is traditionally dated to D'Arcy Thompson's (1917) great treatise *On Growth and Form*, but Thompson's reasoning actually derived from a tradition that was already three centuries old as he was writing. Quantitative discussion of biological function under conditions of varying form can be traced all the way back to Galileo's argument about the proportions of load-bearing beams. But the methodology of this topic has not kept up

with the turn to systems description that characterizes many other branches of organismal systems biology today, to the extent that most of the occurrences of the trope "form and function" nowadays arise not in the biological sciences but in the humanities. (For instance, a search on http://www.amazon.com on May 29, 2012 using the retrieval key "form and function" returns 1,421 hits, of which the first is Moussavi (2009), which is a work on architecture and architectonics. The next retrieval deals with typography; the third, with jewelry design.)

To reclaim this phraseology for biology we proceed reductionistically, word by word, inasmuch as each of the three words "form," "function," and "and" involves quantifications of its own. For *form* one modern locus of major innovation is the domain of morphometrics with which I have been associated for many decades (see, e.g., Bookstein 2006). The rhetorical analysis of the notion of *function* is more fugitive, but we are used to the rhetorical opposition of "function" with "evolution" as standing for a difference in time scales (see, for instance, Hanken and Hall 1993; or Lieberman 2011), while the composite phrase "functional morphology" is familiar from the curricula of innumerable academic divisions of the biological sciences. And the notion of "functional explanation" has a philosophical literature all its own (see, e.g., McLaughlin 2007).

There remains the copula, the word "and," whose appearance in the grammatically analogous phrase "size and shape" I have already discussed in this journal (Bookstein 2009). As applied to quantitative studies of "form and function," this essay will construe the word "and" as referring to the possibility of a successful *joint quantification* comprising (1) the measurement of form by one instrument or modality, (2) the assessment of function by a physically separate instrument or modality, and (3) the existence of a contingent relationship, such as a near-identity or a regression equation, relating the quantifications (1) and (2). While the wording here is prolix, the underlying concept is familiar. Galileo was talking about form and function, for instance, when he pointed out the exponents for beam strength as a non-dimensionless algebraic expression in measures of form.

My topic falls under this heading of "and" analysis, mode (3), exploration of a practical quantitative notation for contingent relationships (numerical regularities) between measures of form and measures of biomechanical function. More specifically, these would be measures of form variation as juxtaposed to measures of the variation of one particular type of biomechanical function. Our assignment is made feasible by the prior development of quantitative methods for these domains separately. For describing variations of form under a logic of biological homology, this is the methodology usually referred to as GMM (see, e.g., Weber and Bookstein 2011, Chap. 4).

For describing variations of biomechanical function, there are as many methodologies as there are domains of the measurement of function. In this essay I will concentrate on the measurements that involve deformations of elastic biological organs or tissues. The parent methodology, then, is *elasticity theory*, which has constituted a branch of applied physics for hundreds of years (cf. Todhunter 1886). The subject supplies a major component of today's undergraduate engineering curricula: see, for instance, Sadd (2009).

The argument to follow falls into three parts. It begins with some relatively simple paradoxes generated by inconsistencies between the mathematical foundations of today's GMM and biomechanics when considered separately (without the "and" of "form and function"). In other words, if you do morphometrics the way you were taught, and also do biomechanics the way you were taught, you have no access to any semantics by which to talk about them at the same time. In the central section, I explore the existing literature of GMM more closely in search of a possible formalism that circumvents these paradoxes. And there is one: a modification of the method of relative warps (principal components of form-variation) that optimizes the prediction of biomechanical deformation instead. A closing discussion indicates the function of this new construal as bringing forward into the twenty-first century one canonical form of form-function studies, *allometry*, that has in fact been with us nearly as long as D'Arcy Thompson has.

Some Elementary Contradictions and Paradoxes

Begin by focusing on the correspondence between GMM and FEA when both are construed as relying on a single summary scalar. Such a strategy is an innocuous pedagogical choice, not an essential aspect of the mathematical development. For GMM that central scalar must be Procrustes distance (although there will be comments on morphometric bending energy from time to time in the exegesis here). For FEA, it is the analogously positive quantity that is strain energy, the net potential energy stored in a solid that has been deformed from its resting form by the application of reversible forces. For several years now we have been teaching the multivariate methods of GMM as derivable from solely the information encoded in Procrustes distance. See, for instance, the introduction to morphometrics in Chap. 4 of (Weber and Bookstein 2011), or the more technical survey in Part III of (Bookstein 2013). For FEA, one can revert to the classic literature of linear elasticity theory, subject of any of the elementary textbooks (e.g., Sadd 2009; or Boresi et al. 2010), wherein any observed deformation must minimize the potential energy cost incurred over the course of the deformation.

This minimand, though capable of driving the entire analytic or computational setup all by itself, is nevertheless encountered only rarely in reports of empirical analyses. Papers such as Strait et al. (2009); O'Higgins et al. (2011); or Cox et al. (2011) show simulated strains of a skull under load in remarkable detail but nowhere mention the quantity of work that must have been simulated in order to impose those particular simulated strains. If the subject is a stress analysis of chewing, for instance, the necessary quantification would be the work done by the muscles of mastication, work that has to be accounted in the organism's energy budget and that has itself presumably been optimized over the mechanisms (muscle mass, muscle insertions, tooth crown geometry) that are responsible for its delivery.

In this initial simple setting, where each of GMM and FEA is represented mainly by one nonnegative scalar (Procrustes distance or strain energy, respectively), we can learn a great deal merely by pursuing the following naive-seeming question:

> What is the statistical relationship, if any, between Procrustes distance and strain energy for loads at the end of a cantilevered bar?

The choice of a cantilever is suggested not only by its position near the front of most textbooks of mechanical engineering but because of its similarity to the role of the mandible in chewing.

Corresponding to this simple question are results that follow immediately from formulas in the standard textbooks. I consider two settings: one (Fig. 1) in which the bar is of uniform cross-section, and a second (Figs. 2, 3, 4, 7, 8, 9, 10,11) in which it is instead tapered from one end to the other. The tapers of the tapered beams will vary—that gives rise to one Procrustes distance archive—and in addition each beam will be subject to a load, giving rise to a separate Procrustes distance. It makes no sense to combine these two distances, because, it will be shown, the analysis of the effect of a constant system of loads on a varying initial form is a function of only a few relevant parameters, not the huge mass of degrees of freedom typically agglomerated in any net Procrustes quantity such as a set of relative warp scores.

The Uniform Bar

Assume a horizontal cantilever as shown in Fig. 1, with a point load on the unsupported end. Then for a load directed *vertically* (e.g., the weight W in the figure), the well-known solution of Leonhard Euler (cf. Segel 1977) shows that the beam must bend in a profile that is a suitable multiple k of the unique polynomial $-x^2(3 - x)$, where x, now rescaled to range between 0 and 1, is the positional coordinate along
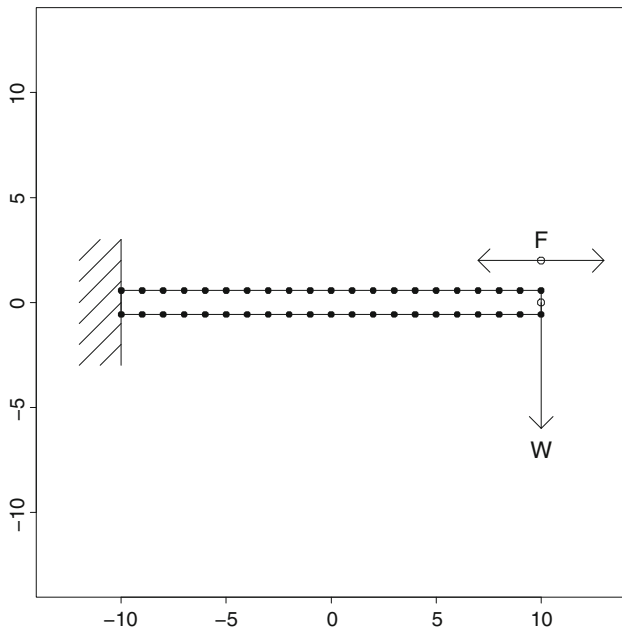
Fig. 1 Scenario A: classical scenario of a cantilevered beam loaded at its free end. The beam is intended as a structure in three dimensions with a thickness into the page equal to its height as shown on the page. As the magnitude of the force F or W is varied, strain energy for forces like W perpendicular to the axis of the beam is proportional to Procrustes length of the corresponding shape change. But for forces like F parallel to the axis, strain energy is proportional instead to the *square* of that Procrustes distance. The *dots* indicate the locations of the 4 landmarks and 38 semilandmarks used to represent these beams for Procrustes purposes
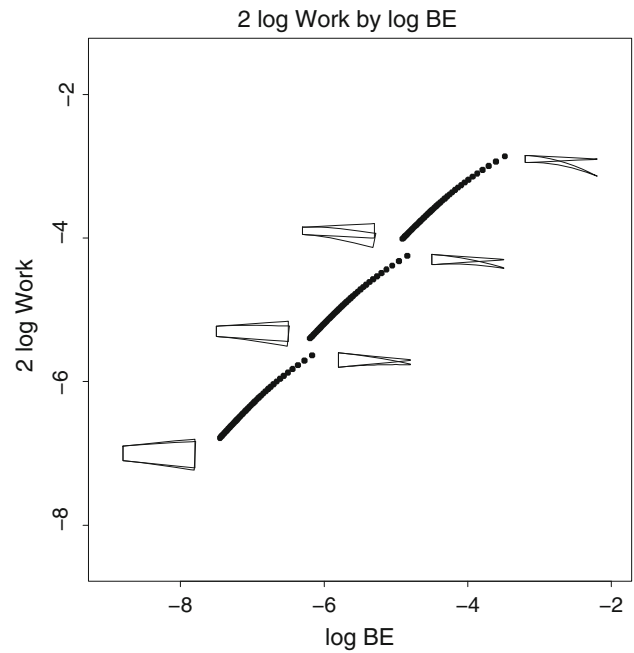

Fig. 2 Scenario B: refinement of the preceding demonstration into a two-parameter system of tapering beams. There are three groups of beams, as the height on the left (fixed) end is 0.05, 0.07, or 0.10 of the length of the beam. Our analysis lines them up approximately along one single dimension corresponding to both morphometric bending and biomechanical strain increase, which are loglinear with respect to one another. Forces like the F in Fig. 1 are no longer considered, and the force W is presumed constant over all analyses. The beam deformations have all been computed according to the derivation by virtual work that was set out in Nguyen (2007). The geometric design here is plotted as an enhancement of a scatterplot showing the nearly perfect proportionality of GMM bending energy to the square of the FEA strain energy, as explained in the text. Icons, from lower left to upper right, show unstressed and stressed positions of the following exemplars: fixed end 10 % of beam length, free end double that; fixed end height 10 % of beam length, free end a cusp; fixed end height equal to 7 % of beam length, free end double that; fixed end 7 % of beam length, free end a cusp; fixed end 5 % of beam length, free end double that; fixed end 5 % of beam length, free end a cusp. The thickness of the beam in the third dimension (into the page) is presumed constant. *BE* bending energy. All logs are to base e. For the derivation of the scaling of this plot, with its slope of ~ 1, see the text

the beam. The coefficient k prefixing this polynomial is a function of Young's modulus together with the moment of the beam in cross-section.

Let us restrict the discussion to realistically small bendings of an original form that has a straight centerline. In that setting these deformations all manifest Procrustes distances proportional to the constant k, which serves as an "amplitude" of this specific convex-downward profile both before and after a suitable Procrustes registration. But, also, strain energy is proportion to the drop of the weight attached to the end of the beam. This drop is by a distance $k \cdot 1^2 \cdot (3 - 1) = 2k$, likewise proportional to k. Hence:

> For the case of a simple cantilever beam loaded at the free end *normal* to its axis, Procrustes distance of the strained form from the initial form and strain energy of the associated deformation change precisely in proportion to one another as the applied load is varied.

That was for forces perpendicular to the axis of the beam. For forces *along* the beam axis (forces like F in the figure), there will result a material expansion or compression of the beam by some fraction ε_x along its axis. To a first approximation, Procrustes distance will be proportional to this same ε_x, but strain energy, which goes as

Hooke's Law, is proportional to ε_x^2 instead. This is worth setting off in its own indented paragraph:

> For the case of a simple cantilever beam loaded at the free end *parallel* to its axis, strain energy of the associated deformation is proportional to the *square* of Procrustes distance as the applied load is varied.

Dependence on the Starting Form

Anatomical forms vary in geometry. In this particularly simple setting, such variation could involve the thickness of the bar. Assume that this is a three-dimensional structure, with thickness into the page identical to thickness

upon the page. Then, according to the formulas, the response to the *bending* load, the W in Fig. 1, will scale inversely as the second moment of the cross-section of the bar. That moment is proportional to the product of thickness by the square of height, and so goes as $1/h^3$, where h is that thickness (which is 0.1 in the figure—call this h_0). But the response to an *axial* load like F in the figure is simply distributed over cross-sectional area, and so will scale (in this example) as $1/h^2$ only. In all of these changes, the corresponding Procrustes distance between the equilibrium states of the beams will be proportional to $|h - h_0|$. For moderate ranges of h, corresponding perhaps to empirical variation of homologous anatomical structures in contemporary or Pleistocene anthropoid/hominoid species, the dependence of formulas like $1/h^3$ on $h - h_0$ is not captured at all well by the first term of the corresponding Taylor series. Hence:

> Shape changes of constant load applied perpendicularly to the axis of beams that differ in anatomically realistic ways do not relate linearly to any Procrustes analysis of that same variation of the starting forms.

(And furthermore there remains the pesky matter of that squared term for axial loads; these surely do not relate "linearly" to ordinations of shape space either—and the dependence of *this* term on starting shape has a different coefficient from the dependence of the effect in the normal direction.)

The situation is thus *extremely troublesome*, troublesome enough that somebody should have pointed this out a while ago. The Procrustes toolkit was explicitly designed to be invariant against rotations of the coordinate system; so were all the manipulations associated with thin-plate splines. Yet the interrelations between GMM and FEA arising in the course of this very simple textbook example—perhaps the simplest possible—violate that rotational invariance. *Analyses of strain are not compatible with the symmetries of the Procrustes toolkit.* Phrased even more pointedly, *our standard analyses of form are not compatible with the simplest analyses of biomechanical function*, a circumstance that must be regarded as terribly embarrassing. The statistics of loads on a beam (such as the mandibular corpus) will be functions of angulation between load and beam in a way that wholly confounds any linear modeling of strain in terms of Procrustes distance, because of the variation of the order of dependence (linear, vs. quadratic) as a function of the direction of load, which is, in turn, a function of anatomical variation in the form of the mandible. A correct analysis will have to separate the relation of strain energy to shape distance into these two separate components, analyzing them first separately and then together. The symmetries of the Procrustes toolkit no longer apply to any joint statistical approach bridging

GMM to FEA, nor may we expect any linear modeling to work sensibly when loads have an axial component (to which the strain response is quadratic instead). Nor may we expect responses to similar loads to be dominated by effects linear in the first principal component of shape. The dependence must be studied with far more subtlety.

A Tapered Bar

Figure 2 shows the situation for a cantilevered bar that is now no longer uniform in cross-section but instead tapers linearly from one end to the other. In the computational design here, this taper varies according to one of three settings of thickness at the left (fixed) end along with one of 41 values ranging from 0 to 200 % of the initial height at the free end. The 1st and 41st of each of these little series have been sketched on the graph next to the points at one end or the other of the incorporated curves that they delimit. To simplify the algebra, I have not tapered the third dimension (into the page or the screen) correspondingly, but held it constant. Then the equation according to which these beams deform under an end load normal to the centerline remains a cubic polynomial (except for the beams that come to a sharp edge, for which the equation is quadratic instead of cubic), and the deformation incurs a *physical* strain energy that is not hard to compute following the examples of Nguyen's (2007) formulas. As Fig. 3 implies, the analyst may relate this physical strain energy to either of two different GMM quantities. One is the Procrustes distance between the two forms of the beam, deformed and undeformed. The other, the variant actually used in Fig. 2, is the morphometric bending energy of the same comparison, computed with respect to the undeformed beam.

As Fig. 3 shows, both relationships are loglinear except at the (unrealistically) highest degrees of bending. But the scaling is different in a way that adds additional insight to the relation between GMM and FEA for more general configurations. For all these beams, it remains the case that strain energy is proportional to the fall of the weight inducing the end load. For all the deformations, this is the evaluation of the corresponding polynomial. It will always be proportional to the form $-x^2(M - x)$, where M is a function of the taper, except that for the beams that taper to zero height, M goes to infinity, resulting in the simpler profile that is just $-x^2$, a parabola downward. To the extent that $-Mx^2$ is the dominant term in all these deformations, Procrustes distance will be proportional to M, which is proportional to the Mx^2 evaluated at the endpoint $x = 1$ of the beam. Hence, once again, Procrustes distance should be expected to be proportional to strain energy except at the highest values of strain, which bend the beam so much that

Fig. 3 When bending energy (in the GMM sense) is replaced by simple Procrustes length of the shape change, the slope of the relationship between their logarithms reverts from 2 back to 1 but the errors of fit (the effect of all the small algebraic incompatibilities of formula) are somewhat larger by eye. *BE* bending energy of the thin-plate spline, from GMM (missing for the forms having a cusp), *PD* Procrustes distance, *Work* net work done by the falling weight *W*, which is also the strain energy of the deformation
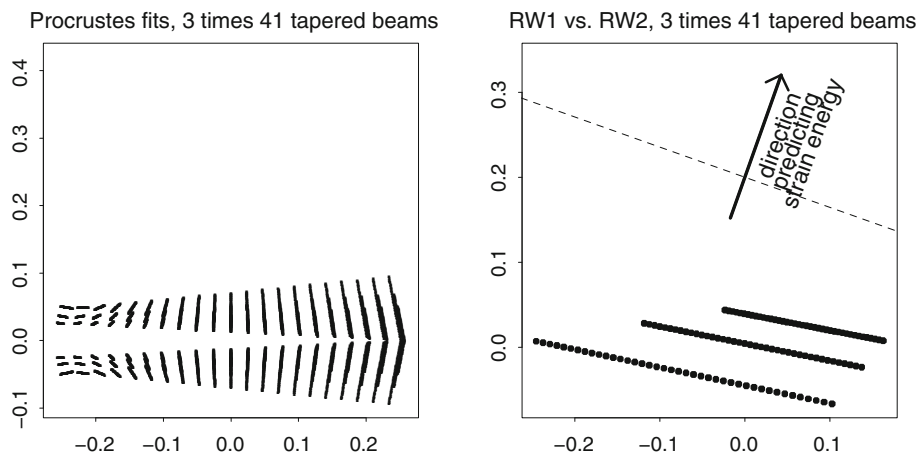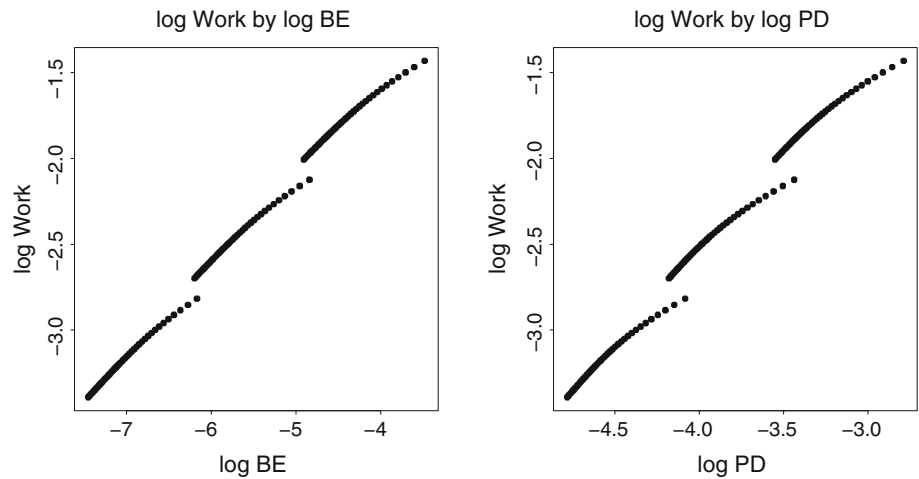

Fig. 4 Conventional GMM analysis of the 123 tapered beams before deformation. *Left* Procrustes shape coordinates. The presence of the three subsamples (Fig. 2) according to height at the fixed end is clear at the left margin of this figure, along with the systematic variation in height at the right (free) end. *Right* conventional relative warp analysis of these 123 beams shows no obvious linear ordination resembling either horizontal axis in Fig. 3. The origin of the data in three subgroups is now considerably clearer. We will see later that

something useful arises as a projection along the direction shown in the heavy arrow floating at the top of the diagram. The regression in the *left-hand side panel* of Fig. 9 projects out the component of the shape space lying along the *dashed line* up there, which appears visually to connect the right end of either of the lower two subscatters to the left end of the subscatter immediately above. *RW1*, *RW2* relative warp 1, relative warp 2

the affine (vertical-only) approximation to Procrustes distance threatens to become invalid.

On the other hand, we might be more interested in a different quantity likewise copied over from GMM, the formal *morphometric bending energy*, defined as the integral over all space of the summed squared second derivatives of the deformation. In this case of bending that is negligible in the *x*-direction, this is the single term $(d^2y'/dy^2)^2$ (for a notation in which the point at ordinate y is moved downward to new location y'). To the extent that these curves are nearly parabolic, the net fall of the weight is proportional to M, whereas the term $(d^2y'/dy^2)^2$ over space will be proportional to its values along the interval from 0 to 1 (the beam itself), where it evaluates to M^2. Hence:

For the case of a tapered cantilever beam loaded at the free end normal to its axis, Procrustes distance of the strained form from the initial form and strain energy of the associated deformation continue to be proportional, now as the taper is varied, whereas GMM's bending energy is now proportional to the square of Procrustes distance, which is to say, proportional to the *square* of strain energy. This will be referred to below as the *scaling of biomechanical strain energy to the 1/2power of morphometric bending energy*.

Further examining Fig. 3, we notice that these beams align on the graph, both in their strain energy and in their GMM parameters, along a single nearly straight line. In other words, the Procrustes variation among these initial

shapes, which is considerable, is quite irrelevant to the evaluation of their response to strain *except for* the one single dimension along the arc here. But this is not any convenient principal component of these forms, as one can see from the RW1–RW2 scatter in Fig. 4. *Nothing about this figure indicates the simplicity of the finding at which we will eventually arrive*—the interpretation in terms of the relation to biomechanics is *not* a natural component of the GMM toolkit.

> In a realistic setting of systemic anatomical variability, the parameter of starting form that best predicts response to an unchanging elastic load normal to the axis of a tapered beam is not equivalent to any self-evident principal-component-based summary of the distribution of those starting shapes.

You simply have to do the hard work for your forms that I did for the beams here, in other words, a scaling analysis the way a good engineer would have been trained to do it. You must search explicitly for the parameter that answers the question the scientific context actually requires (how do I relate the extent of a strain to the details of the form undergoing the strain and the details of the registration of the stresses upon the form) rather than the question you wish you had been asked instead (how do the results of the analysis of form I did yesterday predict this strain that somebody told me to concern myself with today).

Bending of a Cylindrical Shell

The preceding two examples involved no curving of the starting form, no semilandmarks; the next example involves both. Once again, although the application will be to computations carried out as finite element analyses, the example is one for which the elasticity analysis is known in closed form. This is the topic of the bending of an initially straight circular cylindrical shell, one having a radius that is small in proportion to its length and a thickness that is small in proportion to that radius. The problem has been understood for approximately a hundred years, since early work by Saint-Venant. The formula below is from Brazier (1927), and the diagram in Fig. 5 is redrawn from

Fig. 5 Scenario C: the bending of an initially straight thin-walled circular cylinder. *Top* bending induces a compression strain along half the generators of the cylinder and a corresponding tension along the other half of the generators. As these compressions or tensions are proportional to distance from the centerline of the cylinder, they are somewhat mitigated by an *ovalization* that converts the cross-section of the cylinder from circle to ellipse. After Karamanos (2002). *Bottom* the squared Procrustes length of such a shape change is the sum of two components: one at large scale, proportional to the square of the curvature of the cylinder axis, and one at small scale, corresponding to the shift from circle to ellipse and proportional to the fourth power of the same curvature. The smaller-scale term adds to the Procrustes length of the deformation but subtracts from the strain energy. *PD* Procrustes distance
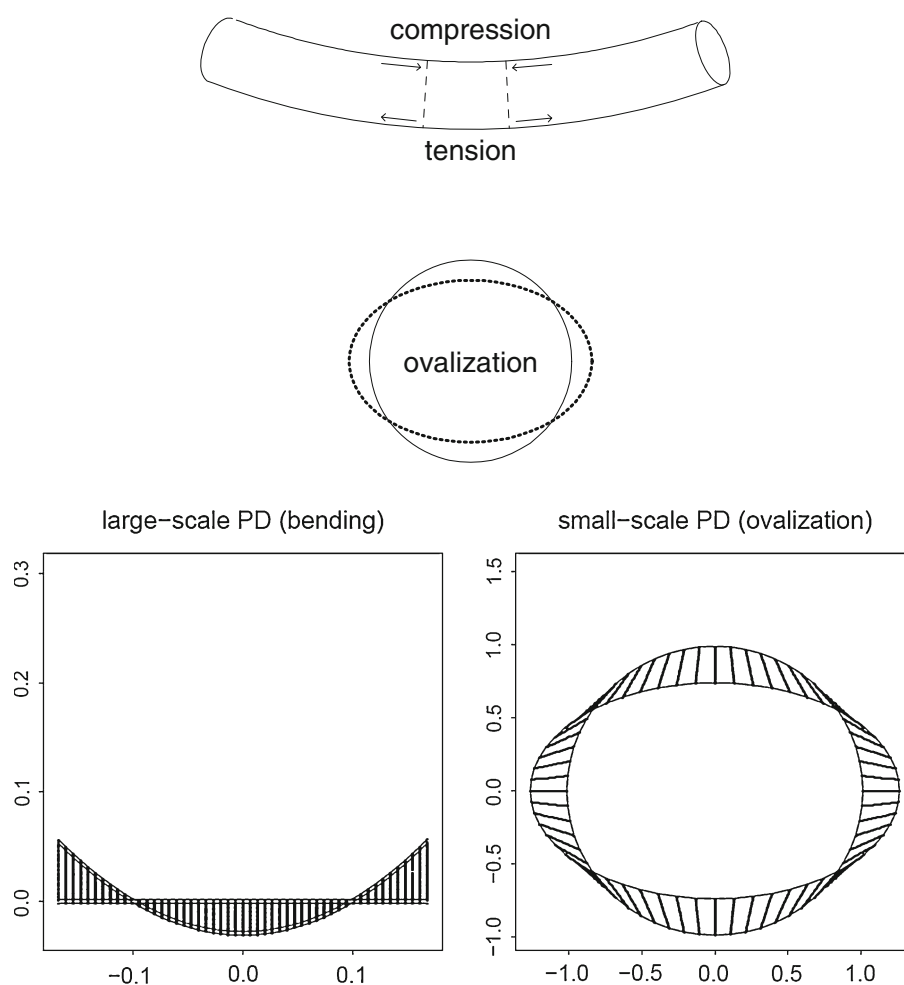
Karamanos ([2002]). Think of this, if you want, as a model for what the zygoma is doing during chewing, or of stresses on the supraorbital ridge.

As the Karamanos figure makes clear, the effect of bending on a cylindrical shell is the composite of two effects at quite different scales. At larger scale there is a couple of forces on opposite sides of the axis of the cylinder: a compression along the generators that are shortening, balancing a tension along the generators that are lengthening. Distinct from these, and at much smaller physical scale, is the so-called *ovalization* of the cross-section: the diameter in the plane of the bending is shortened, and the diameter perpendicular to that plane lengthened. (This makes intuitive sense: if the diameter in question is shortened, both the compression and the tension in Fig. [5] are by smaller fractions, and thus require less strain energy.)

Brazier ([1927]) derived the formulas that quantify this intuition. In his notation, where r is the radius of tube in cross-section (before deformation), c is the curvature of the torus into which the cylinder is bent, t is the thickness of the material shell, and E and σ are Young's modulus and Poisson's ratio of the material from which the cylinder is made, then the strain energy per unit length goes as

$$\frac{E}{2} \pi r^3 t c^2 \left(1 - \frac{3}{4} \frac{r^4 c^2 (1 - \sigma^2)}{t^2} \right).$$

The extent of ovalization turns out to go as $c^2 r^5/t^2$. For the bending of the originally straight cylinder into a parabola, Procrustes distance is proportional to the curvature c and thus squared Procrustes distance to c^2, whereas at the smaller scale, the change of circular section into elliptical section involves a Procrustes distance proportional to c^2 (in its shape subspace, which can be presumed orthogonal enough to that of the cylinder axis) and thus a contribution to squared distance proportional to c^4.

In a more compact notation,

Strain energy $= A(r,t)c^2 - B(r,t)c^4$

(where c is curvature) and

Procrustes distance squared $= A'(r,t)c^2 + B'(r,t)c^4.$

These are the same two expressions in powers of c^2, but with contrasting signs. The minus sign in Brazier's formula corresponds to the relaxation that ovalization embodies. (At the larger geometric scale, the larger the Procrustes distance, the larger the strain; but at the smaller geometric scale, positive Procrustes distance goes with *smaller* strain energy.) This joint structure can certainly be retrieved by appropriate polynomial regression analysis, *if we know in advance what sort of signal it is that we are attempting to retrieve*: in this instance, dependence of strains on both the square and the fourth power of Procrustes distance, the one

at large spatial scale and the other, with a minus sign, at much smaller scale. Back in Fig. [2], that would be the requirement of the equivalently strong prior knowledge that the profile of centerline bending was an exact cubic polynomial. But if we knew features like these in advance, we would have completed our biomechanical analysis already—there is no need for any finite-element software at all when the analysis of form versus function has been completed a priori.

The coefficient multiplying squared curvature in the dominant term of Brazier's formula for net strain energy per unit length is r^3t up to constants. Of the three powers of r (radius of the circular sections of the torus), one conveys the proportionality of strain to net area element of the infinitesimal hoop involved, while the remaining product r^2t represents the second-order scaling of moment with respect to the longitudinal axis, analogous to the same term encountered in the analysis of the bending of the cantilever. Because the term arises from a compression or tension, it corresponds to a maximum directional strain on the outer surface that is proportional to cr, and would be graphed as such on color maps of FEA simulations. There is a version of the formula for cylinders that are curved to begin with (see, again, Karamanos [2002]). It would be instructive to see how this remarkably simple formula matches either computed strains along the long axis of ridges under simulated stress or actual strain-gauge measurements at corresponding points of a real biological structure.

Sketch of a Way Forward

To cut through all of these inconveniently counterintuitive mathematical facts a principled reformulation of the problem is required. This section sketches one possible approach consistent with this intention.

There are three degrees of complexity at play here. Most complicated is the representation of the solid bony skeleton, via a discretization of its CT scan or μCT scan. When data come from the most modern machines this involves from several million to several hundred million voxels, each with material properties that might be scalar (bone density) or tensor (the stress-strain relationship). A radical simplification of this data resource arises from the separation of space into bone vs. nonbone, followed by the representations of the bony compartment that we call GMM. Points are set on the surfaces of this separatrix, on its edges, and on its characterizable discrete points (landmark points) when they might be available. The mathematical space of these configurations, quotiented by the group of similarity transformations, makes up a familiar Riemannian manifold, *Kendall's shape space*, that has Procrustes distance for its metric. Many familiar forms of

multivariate statistical analysis can go forward for these representations once they have been projected onto a space of *shape coordinates*, which live in the tangent space to the manifold that touches at the sample average form. Usually (though not in the examples here) one supplements this shape space by one additional dimension conveying the original geometric scale of the configuration. There results a representation of up to a few thousand geometrical degrees of freedom (df), far fewer than were embodied in the original CT scans but not just a selection from those millions of df insofar as we have added the information borne in the labeling scheme (the "template") to whatever was produced in obedience to the physics and engineering of the scanner.

A second aspect of complexity is the sample of *forms* involved in the comparative analysis. These might arise as a haphazard collection of individual fossils, a systematic sample from some medical imaging resource, a growth study of selected individuals followed through time, an evolutionary series recovered from sediments, or any combination of these designs (see, for instance, the discussion in Bookstein and Ward 2012). To each sample provenance corresponds a conventional interpretation of the principal components of form that will be carrying the suggested new biometric statistic. In the central example of this paper, for example, the principal components (Fig. 4) are rotated away from an alignment with within-group allometry in order to accommodate a between-group allometry (that variation of the overall thickness of the beam) that might correspond to a subspecies-level taxonomic distinction.

Of a dimensionality lower than either of these is the third essential feature of our situation, the set of *loads* or *probes* by which we pretend to stress our CT scans in order to calculate the resulting strains. Each load is applied to every specimen in the sample of forms, under the oversight of the GMM scheme specifying exactly how points of application of force or constraint correspond across the specimens. Every load is presumed to be free of net vector resultant and net torque, so that the deformed body may be presumed at rest (we are not engaged in dynamics here, only in statics). From well-known theorems of linear elasticity, the effect of superimposed loads is additive in the effects of those loads separately. From the paradoxes reviewed earlier in this note, though, it appears that the effect of loads is nonlinear in many of the parameters of the GMM representations that attach to our original CT imagery. Once the loads are realized in the space of each individual specimen, it is the effect of those loads that is modeled by the black boxes we call *finite-element software*; but please keep in mind that the correspondence of those loads from form to form is governed not by finite-element algebra but by the geometry of shape space and the

requirement that forces and torques allow for an equilibrium solution.

GMM begins with minimization of a single summary scalar, the Procrustes distance metric, and the preceding discussion has suggested that it might be useful to make explicit the corresponding minimand governing the finite-element computations, the net strain energy of the deformed configuration. The problem to be posed below involves the approximation of one deformation by another based on different information, and for that to make mathematical sense, there needs to be a metric for the approximation. In keeping with the existence of the governing minimand for the FEA, this metric might be taken as the root-mean-square (RMS) discrepancy between the strain tensors of the various deformations.

For this to work, we need a metric for 3×3 strain tensors. The squared distance between two strain tensors (each taken in positive definite form—the effect of the strain on lengths and dot products along the cardinal axes)[1] can conveniently be taken as the canonical natural metric $\sqrt{\Sigma_{i=1}^{3} \log^2 \lambda_i}$, where the λ_i are the *relative eigenvalues* relating the two matrices, the eigenvalues of either in a coordinate system in which the other has been linearly transformed into a sphere. See Mitteroecker and Bookstein (2009), for an informal review of the mathematical reasoning underlying this claim. The "length" of a single such 3×3 tensor can be taken as its distance from the identity matrix of order 3, which stands for the state of no deformation and thus no strain energy. The formula assigns a value to this squared length that is the sum of squares of the logarithms of its principal strain ratios. From the approximation $\log(1 + \epsilon) \sim \epsilon$ we can approximate this squared length as just the sum $\Sigma \epsilon_i^2$ of squares of the deviations of its eigenvalues from unity. But, by Hooke's Law, that is precisely the formula for the strain energy density of the little element being strained in this manner (as long as the underlying material is isotropic). In other words, in the limit of small strains, *squared geometric length of the strain tensor of a finite element is equivalent to its strain energy density*.

It is an important technical point that this biomechanical quantity, the strain energy, does not correspond to any of the standard Procrustes formulas—one cannot substitute Procrustes distance for strain tensor distance $\Sigma \epsilon_i^2$ and still hope to make biomechanical sense. Rather, neither elasticity theory nor the FEA software has knowledge of Procrustes formulas or any interest in them. For instance, Procrustes distance is a function of the shape of the finite element; it takes on a different value for long and skinny elements or

---

[1] If any strain matrix has singular-value decomposition UDV^t, U, V orthonormal and D diagonal, each 3×3, then the representation is $(UDV^t)(UDV^t)^t = UDV^tVDU^t = UD^2U^t$.

pancake-shaped elements than for elements that are nearly cubical in shape. (Among these inconveniently anisotropic elements are the elements right at the surface of the bone, those whose strain tensors are typically the ones converted to colors and plotted in the published figures.) Even deep inside a form, where actual finite elements might be nearly cubical, the corresponding Procrustes formula is $\Sigma(\epsilon_i - \overline{\epsilon})^2$, the sum of squares of the ϵ's around their average, not around zero. The two quantities are functionally equivalent over a mesh only under special circumstances that are extremely unlikely in practice. Hence just as FEA analysis has no use for Procrustes parameters, so the general Procrustes shape analysis pays no attention to the most important physical parameter of a strain experiment, nor can it help formalize it for any kind of subsequent analysis. (And shifting to form space does not help resolve our dilemma, since strains change volume and Centroid Size in different ratios, and the shift from shape space to form space is not the same as "deciding not to divide out Centroid Size.") Most important of all, strain energy density follows material properties (for instance, it is zero for the regions of air inside fossils), whereas Procrustes distance knows nothing about the difference between air and tissue and accrues everywhere over the form, indifferent to the segmentation of tissue from air and, within the tissue, indifferent to the difference between bone and softer tissues. If the beams in Fig. 2 were all hollow, the strain analysis would be totally different, but the Procrustes analysis would be exactly the same.

Therefore, inasmuch as Procrustes distance and strain energy are not interchangeable, our strategy will be as diagrammed in Fig. 6. In the upper portion of this diagram, under the heading "Virtual reality," is the situation that is accurate to the best of the ability of the biomechanicist. A biologically interesting tissue (the concentric structure at left) is represented by way of geometric extents and associated material properties that are inferred from its CT scan. The representation is deformed according to a simulated load (the heavy solid arrows) with the aid of an expensive black box (the FEA software, heavy dashed arrow). (An anonymous reviewer of an earlier draft of this manuscript helpfully observed that as the price of this software diminishes with calendar time, the problem I am highlighting is actually becoming more serious rather than less.) There results a deformed solid (right) with strains inside all of its finite elements; it is these strains that will sustain the figure of merit in real examples for which strain energy is not available a priori in closed form. In the lower part of the diagram is what we must do in addition if we are to proceed with the extension of these analyses to some sort of scientific explanation. Now the loads come from a space of their own, representing multiple probes of the same interesting tissue over multiple specimens. The starting
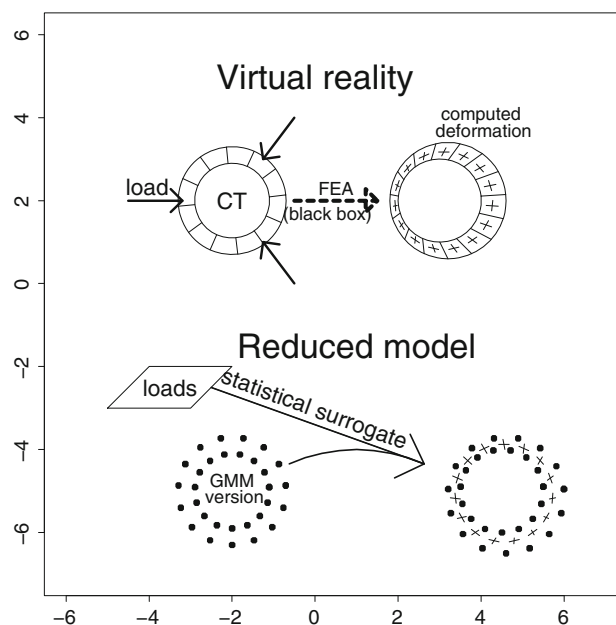

Fig. 6 Summary diagram of the statistical strategy for building a multiscale nonlinear bridge between GMM and FEA. The actual finite-element analyses model the effect of loads on a parcellated CT scan (*upper left*) using FEA software as a "black box" (*upper center*). The usual graphical summary is a representation of the achieved strain tensors (*little crosses*) within each finite element in turn (*upper right*). In contrast, our modeling (*lower center*) combines a representation of one or more patterns of load with a GMM version of the same CT data as corresponding points rather than volumes (*lower left*) in an approximation to the output of that black box (*lower right*), with fitted values but also prediction errors in the locations of the semilandmarks and in the predictions of the strain tensor

form is now represented in a GMM version that involves no parcellations of space, only point locations (lower center), and in place of the expensive FEA software we use only statistical summaries of the experiences with virtual reality, the earlier rounds of computation from the top of the page. There results an estimated deformation, lower right, that is not quite the same as what would have been simulated by that FEA software. The rms difference between the two strain fields calibrates the inaccuracy of the approximation, and our job is to build the appropriate statistical machine, the fusion of the information from GMM and from the earlier experience with explicit loads, for minimizing this rms. According to the paradoxes reviewed early in this essay, this cannot be a linear pattern engine, but instead must be something a good deal subtler and more complex to install.

The machinery we need is a composite of two approaches that have hitherto appeared in different literatures. One, the study of *quadratic shape trends*, was treated in Chap. 7.3 of Bookstein (1991), where it was introduced as a biometrical method for examining the details of growth-gradients and other aspects of continuous allometry in
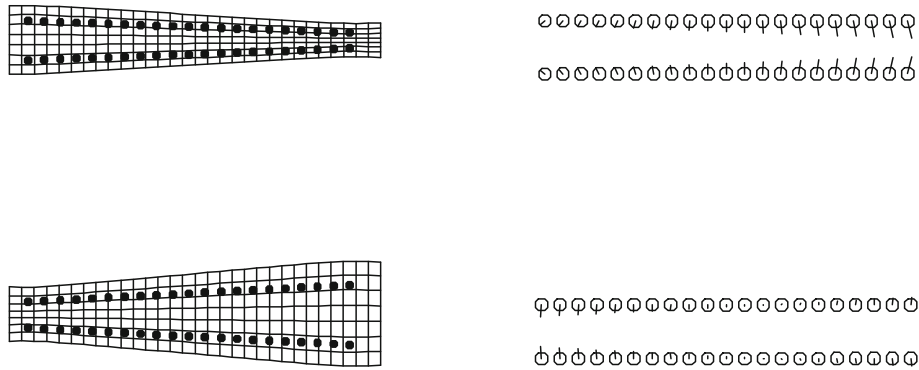
Fig. 7 The only two relative warps for the 123 tapered beams capture the correct two parameters in two tapered combinations. *Left* the warps as grids; *right* as displacements of Procrustes shape coordinates. Upper row, RW1; lower row, RW2. The two grids appear to be mirror-images, but these forms are not symmetric around a vertical midline, and so this mirroring is not a Procrustes operation—the two RW's are not "inverses"

shape studies. The other contributory stream comes from modern applied systems theory: it is the toolkit of *model order reduction* (Antoulas 2005; Schilders et al. 2008) that is used to reduce the dimension of systems simulations from a few million to up to a few dozen degrees of freedom, the tractable range for explanations or design of further experiments.

Formally, the task is to understand the dependence of strains on form for a sample of forms observed in CT or μCT and subjected to a range of stress analyses by state-of-the-art FEA. This will be presumed a fully crossed design, each specimen subjected to each of the same list of simulated stresses (with points of application of the stresses supplied by GMM software so as to guarantee consistency with the steps to follow below). Let there be L load pattern probes, K specimens, N_G shape coordinates of (semi)landmarks in the GMM representation of form (plus 1, for size), and N_F strain tensors, one per finite element of the finite-element analyses. K is of the order of tens, L likewise (and perhaps far fewer), N_G of the order of a thousand, and N_F in the millions.

The first part of the report is a computation based on the $K \times K$ matrix of Procrustes distances of the specimen forms as they were encountered in nature before any deformation. Each squared distance is a sum of squared differences from one configuration to another over the N_G shape coordinates of the template. Subsequent analysis proceeds step by step as follows.

1. Convert the Procrustes distance matrix of unstressed forms to a matrix of relative warp scores by the standard principal-coordinates double-centering procedure (Torgerson 1958; Gower 1966; Weber and Bookstein 2011, Chap. 4). The computation is an eigenanalysis, $K \times K$.

There results the pair of relative warps in Fig. 7, which, for this particularly simple exemplar, exhausts all the information in the Procrustes distance matrix. The scatterplot of their scores has already been displayed in the right-hand panel of Fig. 4. In the general case, these should be the principal coordinates of the Procrustes form distance matrix, the version that incorporates a term for log Centroid Size squared, because strain energy is indeed a function of change in size (although the dependence is quadratic, not linear). In the tapered-beam example here, the effect of the strains on Centroid Size is negligible up to very high levels of strain, and so it is sufficient to work with the distances in the space of lower dimension, the shape space only.

2. For each of the L load probes, analyze the strain energy of the bent configuration under this condition by ordinary polynomial regression. Strain energy will be a sum over N_F components, unless it is supplied in advance by a textbook formula such as in the three examples reviewed in connection with Figs. 1, 2, 5. In this example, $L = 1$.

The obvious univariate regressions, Fig. 8, show no meaningful sharing of information at all. *Neither relative warp, by itself, affords any sensible insight into the relationship between unloaded form and response to the load.* This is unfortunate, as it is precisely this technique that corresponds to the custom of the earlier publications I have been criticizing. Rather, what is needed is a *multiple* regression. It will prove sufficient to use terms through second degree, Fig. 9. This regression, on a total of five terms (RW_1 and RW_2 as extracted by the standard Procrustes machinery, augmented by RW_1^2, RW_1RW_2, and RW_2^2) accounts for 98.9 % of the variation of the strain energy. (See also Fig. 10.)

Fig. 8 The prediction of strain energy by the two relative warp scores separately looks unpromising. The near-symmetry here is intriguing but, in the end, unhelpful
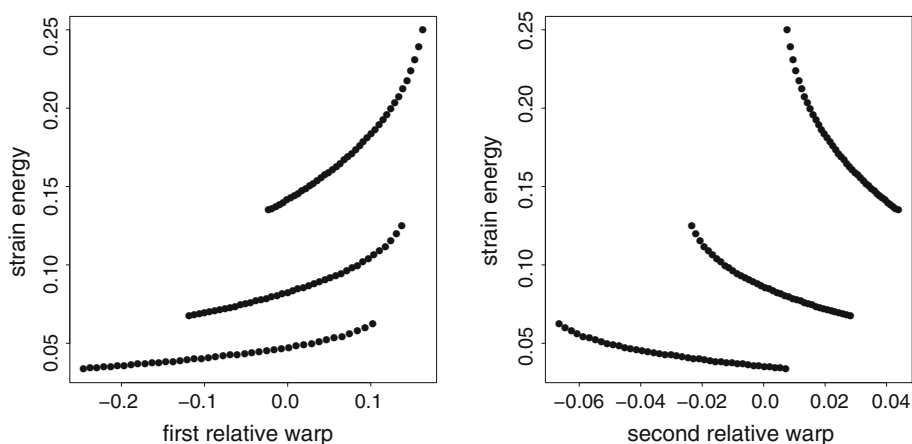

Fig. 9 Multiple regression proves an effective tool for bridging the GMM data base to the FEA computations. *Left* the best linear predictor of strain energy from the two relative warp scores. The horizontal axis here is the projection sketched already in the *right-hand panel* of Fig. 4. *Right* improved prediction when using the three quadratic terms in RW_1 and RW_2 as well. *SE* strain energy
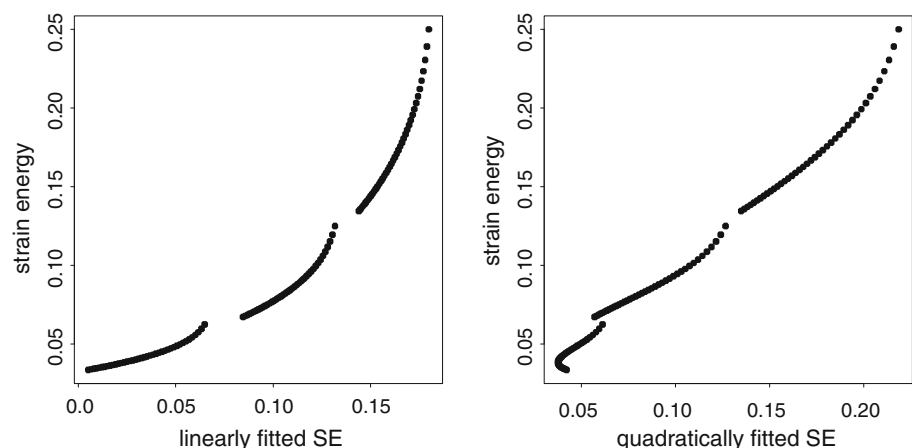
3. For each of the L load probes, analyze the deformed shape of the semilandmark configuration and the strain tensors at every semilandmark by the analogous regressions. The regression here is for each of N_G shape coordinates.

The second-order polynomial regression with five predictors that accounts for 98.9 % of the variation of strain energy accounts for 99.69 % of the Procrustes shape variance in the same configurations (versus only 95.4 % of the same variance, for the regression on the first-order terms only). The expansion of the prediction machinery to incorporate quadratic effects thus reduces the extent of shape variation left unexplained by 93 %, confirming the essential nonlinearity of these accountings but also their relatively gentle character. It would appear that we have analyzed only the net strain energy, not the strain tensors per se; but *in this specific example*, where only the vertical coordinate is shifting, not thickness, the information in a strain tensor is functionally equivalent to the information in the shape coordinates themselves—taking advantage of that coincidence, we simply ignore the Mitteroecker–Bookstein formula.

4. For large ranges of shape, deformation, or function, the analysis may be more comprehensible when strain energy and its density are converted to logarithms.

If we take the logarithm of strain energy before doing the regressions in Fig. 9, we obtain the analyses in Fig. 10, from which undesirable curvature has been wholly eliminated. The geometry of the predictor shape patterns is hardly affected at all—the linear term, for instance, is still the direction indicated upon the scatterplot at the right in Fig. 4.

5. The relative warps are, by design, features at the largest scales. Interesting biomechanical interpretations become possible just when the residuals from these large-scale explanations of strain or of displacement appear to be concentrated in particular regions of the form. The overall analysis may thus be considered a *hierarchical regression approach*, with findings at different scales interpreted using different vocabularies of quantification.

In the tapered-beam example here, that is indeed the case. As Fig. 11 shows, the residuals from the large-scale

Fig. 10 Recomputations after a log-transformation. Most of the prediction error in the multiple regressions of Fig. 9 is circumvented when the dependent variable is replaced by its logarithm. This is the same variable that runs along the vertical axis in Fig. 3
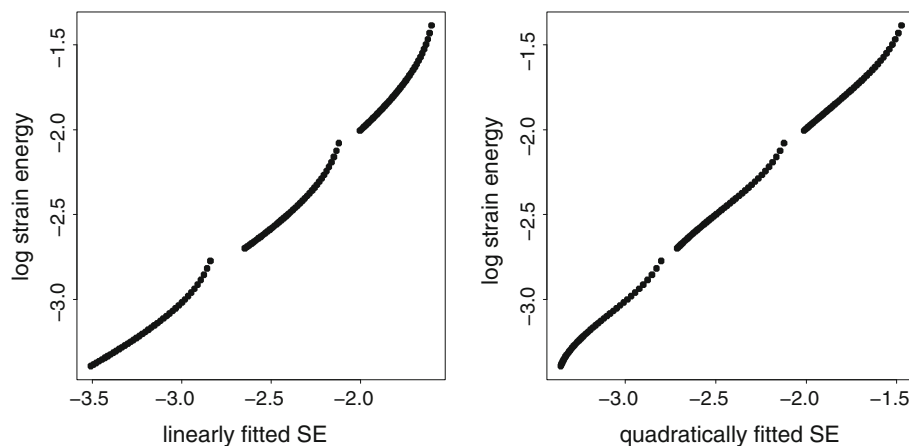
analysis that are largest in mean square are those right at the end of the beam, where the effect of the tapering parameter is the strongest (see Fig. 2). In the present setting, this is a mathematical artifact of truncating the Taylor series for the true model, which is a ratio of polynomials in the taper parameters, at just the site upon the models where the effect of parameter range (the length of those segments) is largest. In settings more driven by data than by that sort of pedagogical artifice, these would be domains in which to look for special explanations in terms of selectively plausible morphogenetic or biomechanical factors of far more local scope.

6. When there are multiple load probes—when $L > 1$—one can continue the procedure via extraction of the principal components of those regression formulas themselves, followed by a similar analysis of the residual strain fields.

If the strains were represented in a linear vector space, this would just be an ordinary partial least squares (PLS) analysis of the covariance structure of the strains against the shape coordinates; but the strains are not represented in such a space, so the PLS possibility is blocked. Nevertheless, the proposal is a close cousin to the *method of snapshots* that is often invoked to import proper orthogonal components into computational fluid dynamics studies. The L separate load-specific matrices of Procrustes distance after deformation will each have a principal coordinates analysis, and for each the principal coordinates can be represented as a rotation of the principal coordinates for the Procrustes distances of the same K forms before deformation. The model reduction would proceed by a standard method of multidimensional scaling, Carroll and Chang's (1970) INDSCAL, and will not be discussed further here.

By this means, we have established a model of reduced dimensionality, of greatest use if it spans from two to four dimensions, that, all at the same time—

- covers the principal dimensions of variability of the undeformed specimens,
- spans the spaces of effects of one or more load prototypes,
- permits the extension of linear regressions of strain on load to accomodate the reality of nonlinear strain formulas, and
- has a computable error as a predictor of strain. In the examples here, for which no explicit strain tensor needed computing, we see this in the fit of that energy (and, even better, of the logarithm of that energy) to the analogous regressions on the relative warps of the undeformed beams. In more realistic cases, where strain energy is computed *post hoc* from the actual strain tensors returned by the FEA software, an analogous error would be computed as the rms difference between the reported strains and those arising from the GMM-mediated reconstructions. I do not demonstrate this second version here.

Discussion

To the extent that these regressions exhaust most of the available variance of outcomes, our methodological goal has been achieved (and if the assumptions do not obtain, there is no alternative logic that would make any sense in its place anyway). In that happy circumstance, the connections identified between features of the form and smaller-scale features of the strain map responding to the specific loading patterns of the probe set can be extracted from this biometric context and treated as novel explanations of either function or form pursuant to the study context. Here in 2012 it is irresponsible to content ourselves with images of strains that lack all information about uncertainties and thereby all possibility for numerical or biotheoretical inference. If there is to be a bridge between

GMM and FEA analyses, it must be based on valid statistical approaches to the information they share, information which is realized only in the context of particular models of biomechanical load. The approach here exploits the parallel construal of each of the two techniques as a minimization of a scalar quantity (Procrustes distance or strain energy, respectively) and implements the biomathematical style of explanation by observing the alignment (or not) of principal coordinates of form with systematic aspects of the variation of strain as the probes of a study cycle through all of the load configurations that were considered important that day.

The logic of this load-dependent aspect of the phenomenon is complementary to that of Ulrich Witzel (see, e.g., Witzel and Preuschoft 2005), who, in effect, treats shape coordinates as functions of load in a particularly high-dimensional nonlinear computational setting. (Witzel's, in turn, can be viewed as the most recent in a series of investigations, beginning with D'Arcy Thompson's, of "the origins of form in force.") The approach of this paper instead examines the dependence of strains on shape coordinates, which may be a more nearly linear (or at least more gently nonlinear) explanatory domain more suited to eventual explanations in terms of selection gradients or similar phenomena that can be formalized as regressions.

There is a cost to the hierarchical regressions of strain fields on form recommended here, a cost paid mainly "in silico." The method requires the computation of all KL of those finite-element analyses: the development of a workflow for applying "the same" loads to each of a whole sample of CT scans, and its realization over a weekend, or a month, on some university's high-performance grid. The reward for this care is the apparent mastery of the informatics by which FEA and GMM have drawn on the same geometric data, as they must have done. If you achieve a reduction of the prediction error to some 1 % or so of the variance of shape or of strain energy, as in the example of the tapered beam, you are entitled to claim a corresponding degree of understanding of the associated allometric explanations; if you can further rationalize the emergence of spatial foci of the residuals from those regressions, perhaps speculative explanations of a biological sort may emerge as well. Conversely, if a paper does not report the extent to which starting form together with load model does in fact explain the variation of the outcome strains, then no matter how colorful the presentations of those strains, no authority should be granted the adjacent diagrams as regards biomechanical or evolutionary explanations.

I put forward this partition of strain phenomena in terms of dependence on the large-scale relative warps as the appropriate GMM version of the scaling analyses that have been a familiar style of biomechanical explanation of form ever since the time of Galileo. Explanations associated with the *regressions* (the predictions of strain by large-scale features of shape) should be seen not as functional or evolutionary but simply as allometric. They express mainly a previously hidden dimensional analysis of the various geometric coefficients in the laws of elasticity. For biological explanations to emerge from FEA analyses they must emerge via the strategy demonstrated in Fig. 11: analyses of the information left over after the effect of scaling is regressed out via these same first few relative warps of shape. The proper role of GMM analysis in interpreting multiple FEA computations, then, may well be this service of pulling out the part of the observed strain patterns that arises on purely geometrical grounds, so that there is a chance to see biological explananda in the residuals.
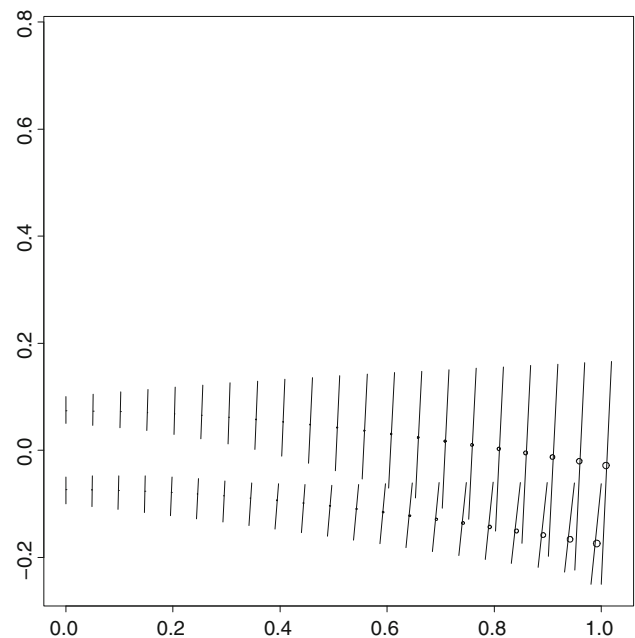

Fig. 11 Prototype for a multiscale analysis: semilandmark-specific (or strain-specific) residual variances from semilandmark-by-semilandmark regressions on the shape features of largest scale. The radius of each circle (plotted at the mean location of the corresponding semilandmark over the 123 deformations) is proportional to the standard deviation of the residual of the corresponding semilandmark after the quadratic regressions of the right-hand panel in the preceding figure. The segments through these circles connect the extremes of variation of the corresponding points over the full sample of 123 bent beams. In this artificial example, the concentration of unexplained variation at the free end of the tapered bars arises as truncation error in the Taylor series that the regressions are approximating, a truncation error that grows as the size of the effect (the displacement of the beam's end, the net strain energy of the deformation) increases. Any focus of the residual signal—any local maximum of residual amplitude interior to the domain of analysis—would evince a potentially biologically interesting phenomenon at smaller geometric scale. In real applications the dependent variable would not be position, as in this simplistic example, but instead the computed strain tensor, using the strain metric given in the text

The division of the space of summary descriptions of strain into those of geometrically "large scale" versus those of "small scale" shares this rhetoric with several other multivariate application domains. In GMM itself, there is a variant of the relative warps analysis whereby patterns emerge in order of a specific quantity balancing spatial scale against variance accounted for (Bookstein 1991). The parameter governing this balance, notated α, is at the discretion of the user. For standard relative warps analysis it takes the value zero, but the GMM–FEA match might proceed better if it were set at a value of 0.5, corresponding to the scaling of biomechanical strain energy with respect to morphometric bending energy, or at some empirically derived value instead. In factor analysis, the issue of how many factors to extract is always present, and likewise in partial least squares the issue of how many singular vector pairs to investigate further. There are other domains in which the examination of residuals proceeds separately from the examination of predicted effects, for instance, the separation of an observed group difference of form into its allometric and its non-allometric components (Schaefer et al. 2004). In the world of econometrics a multivariate time series may be explicitly partitioned into an autoregressive component versus a shock or perturbation—in this setting it is the order of the autoregression, the number of past observations of the series to take into account, that is uncertain—and in dynamical systems theory the Kalman filter likewise separates any suitable series observed incrementally into a predictive component updated at every observation together with an "innovation" that applies only to the latest observation (but that represents a true resetting of the state of the process). In all these domains the setting of the boundary between the scope of the contrasting explanatory styles is a matter of craftsmanship, not theorems. That is not the case in the example of the tapered beam only because the count of relative warps (dimensions of geometrically feasible prediction) was capped at 2 by the design of the computational experiment.

I encourage proponents of any biostatistical analytic style in this domain to become far more critical of their own methods than is the custom at present, and urge their readers—and mine—to become far more skeptical of previously published pattern claims. Only from extended new data sets analyzed by the methods of the heightened rigor advocated here can arise the persuasive new worked examples that lead to novel biomechanical or evolutionary insights. Those who pursue analyses on this particular interdisciplinary interface must be much more demanding of the biomathematical protocols on which they rely, and our community should insist that data sets be rich enough to sustain these more stringent methods of analysis rather than serving mainly to illustrate hypotheses that were already embraced qualitatively a priori. For example, there are clear implications of the methods here for the design of biomechanical studies of empirical strain: applications of loads must be on homologous principles from form to form, loading must apply to all the forms of a data set (not just forms chosen as extreme according to the relative warps analysis), and it is best if there are multiple loading patterns applied to the sample of forms, not just a single simulated "function." It will be interesting to see to what extent the relatively simple results of the toy example here are emulated in biological samples. But any such application needs to be cognizant at all times of the statistical structure of the *unstrained* shapes—the presence or not of multiple taxa, the presence of allometry or sexual dimorphism within taxon—and, as already noted, the simulation of loads that drives the biomechanics must be in accord with the geometric scaling (of relative beam size, relative shell thickness, relative position of components) that is already captured in the morphometrics. Due attention must also be paid to the discrepancy of scaling dimensions described in connection with our first example: strain energy goes as the first power of a Procrustes distance for loads in some directions but as the second power for loads in other directions. Interpretation of regressions or other linear machinery is likely to be defeated by these considerations quite early in any analysis that relies on conventional multivariate statistical simplifications if aspects of scaling are not explicitly formalized as they were in connection with our examples here.

In short, it is time for some principled skepticism about the bridge between GMM and FEA: skepticism about the pro-forma applications of standard GMM, and principles borrowed from the applied mathematics of continuum mechanics, which is so much deeper than the mathematics of GMM. The methods of GMM can serve to differentiate between two subspaces of the strain-related reports emerging from FEA analysis: the component that has the nature of a scaling, locked to the large-scale features of the variation of form; and the residual component, which is the proper domain for the biological and biomechanical styles of explanation. It may be that this discrimination between two forms of logic applied to the same FEA computations is the highest and best use of GMM in connection with FEA: as a quantitative rhetoric for separating the more obvious from the less obvious in those beautiful color pictures of strain.

University of Vienna, for the original invitation to argue an earlier version of these themes at the 2010 meeting he chaired, and to Paul O'Higgins for helping me to clarify the assumptions governing the conventional approach that I am criticizing here. The presentation here has benefitted greatly from comments by Philipp Mitteroecker, Department of Theoretical Biology, University of Vienna, and Jim Rohlf, Stony Brook University, as well as two anonymous reviewers for *Biological Theory*.

References

Antoulas AC (2005) Approximation of large-scale dynamical systems. Society for Industrial and Applied Mathematics, Philadelphia

Bookstein FL (1991) Morphometric tools for landmark data: geometry and biology. Cambridge University Press, Cambridge

Bookstein FL (2006) My unexpected journey in applied biomathematics. Biol Theory 1:67–77

Bookstein FL (2009) Measurement, explanation, and biology: lessons from a long century. Biol Theory 4:6–20

Bookstein FL (2013) Reasoning and measuring: numerical inference in the Sciences. Cambridge University Press, Cambridge (in press)

Bookstein FL, Ward PD (2012) A modified Procrustes analysis for bilaterally symmetrical outlines, with an application to micro-evolution in *Baculites*. Paleobiology (submitted for publication)

Boresi AP, Chong K, Lee JD (2010) Elasticity in engineering mechanics, 3rd edn. Wiley, Hoboken

Brazier LG (1927) On the flexure of thin cylindrical shells and other "thin" sections. Proc R Soc Lond A 116:104–114

Carroll JD, Chang J-J (1970) Analysis of individual differences in multidimensional scaling via an n-way generalization of the "Eckart-Young" decomposition. Psychometrika 35:283–319

Cox PG, Fagan MJ, Rayfield EJ, Jeffery N (2011) Finite element modeling of squirrel, guinea pig and rat skulls: using geometric morphometrics to assess sensitivity. J Anat 219:696–709

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338

Hanken J, Hall BK (1993) The skull, vol 3, functional and evolutionary mechanisms. University of Chicago, Chicago

Karamanos SA (2002) Bending instabilities of elastic tubes. Int J Solids Struct 39:2059–2085

Lieberman DE (2011) The evolution of the human head. The Belknap Press of Harvard University Press, Cambridge

McLaughlin P (2007) What functions explain: functional explanation and self-reproducing systems. Cambridge studies in philosophy and biology. Cambridge University Press, Cambridge

Mitteroecker PM, Bookstein FL (2009) The ontogenetic trajectory of the phenotypic covariance matrix, with examples from craniofacial shape in rats and humans. Evolution 63:727–737

Moussavi F (2009) The function of form. Actar and the Harvard University Graduate School of Design, New York

Nguyen CT-C (2007) EE C245 – ME C218, Introduction to MEMS design, Fall 2007, University of California at Berkeley. Lecture 16, Energy Methods. http://inst.eecs.berkeley.edu/~ee245/fa07/lectures/Lec16.EnergyMethods.pdf. Accessed 22 Feb 2012

O'Higgins P, Cobb SN, Fitton LC, Gröning F, Phillips R, Liu J, Fagan MJ (2011) Combining geometric morphometrics and functional simulation: an emerging toolkit for virtual functional analyses. J Anat 218:3–15

Sadd MH (2009) Elasticity: theory, applications, and numerics, 2nd edn. Academic Press, St Louis

Schaefer K, Mitteroecker P, Gunz P, Bernhard M, Bookstein FL (2004) Craniofacial sexual dimorphism patterns and allometry among extant hominids. Ann Anat 186:471–478

Schilders WHA, Vorst HA, Rommes J, (eds) (2008) Model order reduction: theory, research aspects and applications. Springer, Berlin

Segel LA (1977) Mathematics applied to continuum mechanics. Macmillan, New York

Strait DS, Weber GW et al (2009) The feeding biomechanics and dietary ecology of *Australopithecus africanus*. Proc Natl Acad Sci USA 106:2124–2129

Thompson D'AW (1917) On growth and form. Macmillan, New York

Todhunter I (1886) A history of the theory of elasticity and of the strength of materials. Cambridge University Press, Cambridge

Torgerson WS (1958) Theory and methods of scaling. Wiley, New York

Weber GW, Bookstein FL (2011) Virtual anthropology: a guide to a new interdisciplinary field. Springer, Vienna

Weber GW, Bookstein FL, Strait DS (2011) Virtual anthropology meets biomechanics. J Biomech 44:1429–1432

Witzel U, Preuschoft H (2005) Finite-element model construction for the virtual synthesis of the skulls in vertebrates: case study of *Diplodocus*. Anat Rec 283A:391–401