



# Reflections on a Biometrics of Organismal Form

Fred L. Bookstein<sup>1,2</sup>

Received: 11 June 2018 / Accepted: 1 March 2019 / Published online: 29 April 2019  
© The Author(s) 2019

## Abstract

Back in 1987 the physicist/theoretical biologist Walter Elsasser reviewed a range of philosophical issues at the foundation of organismal biology above the molecular level. Two of these are particularly relevant to quantifications of form: the concept of *ordered heterogeneity* and the principle of *nonstructural memory*, the truism that typically the forms of organisms substantially resemble the forms of their ancestors. This essay attempts to weave Elsasser's principles together with morphometrics (the biometrics of organismal form) for one prominent data type, the representation of animal forms by positions of landmark points. I introduce a new spectrum of pattern descriptors of the shape variation of landmark configurations, ranging from the most homogeneous (uniform shears) through growth gradients and focal features to a model of totally disordered heterogeneity incompatible with the rhetoric of functional explanation. An investigation may end up reporting its findings by one of these descriptors or by several. These descriptors all derive from one shared diagrammatic device: a log–log plot of sample variance against one standard formalism of today's geometric morphometrics, the bending energies of the principal warps that represent all the scales of variability around the sample average shape. The new descriptors, which I demonstrate over a variety of contemporary morphometric examples, may help build the bridge we urgently need between the arithmetic of today's burgeoning image-based data resources and the rhetoric of biological explanations aligned with the principles of Elsasser along with an even earlier philosopher of biology, the Viennese visionary Hans Przibram.

**Keywords** Bending energy–partial warp variance (BE–PwV) plots · Walter Elsasser · Morphometrics · Nonstructural memory · Ordered heterogeneity · Philosophy of biology · Hans Przibram · Thin-plate splines

## Epigraphs

*Living matter is the repetitive production of ordered heterogeneity.*

— Rollin Hotchkiss, in Gerard, ed., 1958, p. 145

*The essentially new property of living matter is the transfer of information over finite intervals of time without an intermediate message.*

— Walter Elsasser, 1998, p. 120

*Perhaps the very success of the biologists in dissecting cells into their constituent organelles and molecules,*

*along with the ability to synthesize some of the latter, has overcome their rightful sense of awe at the existence of living matter.*

— Frederick Seitz, 1993, pp. 106–107

*In sciences other than biology, it is commonly recognized that a fixed range of forms is inherent to every type of matter and that variability, where it exists, is only expressed within that range.*

— Stuart Newman, 2018, p. 2

## Introduction

*How should biometrics relate to the theoretical biology of organismal forms? One can imagine two different structures for this interdisciplinary negotiation. In the first, which characterizes most of today's textbook treatments, the formalisms are those of probability theory or statistics, as applied to the distillation of experimental or nonexperimental*

✉ Fred L. Bookstein  
flb@stat.washington.edu

<sup>1</sup> Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria

<sup>2</sup> Department of Statistics, University of Washington, Seattle, WA, USA

observations into schemes of quantities. Simultaneous with any academic selection from this domain (see, for example, my recent offering, Bookstein 2018) one can generate a series of steadily more biting critiques that jointly delimit the domains outside of which these adaptations or metaphors are very likely invalid. In general, while the formalisms for summarizing sets of measures of length, area, or volume in terms of averages, variance, and covariance have been stable for decades, analyses of repetitive spatial patterning have lagged well behind. Quantitative analysis in this more complicated domain has barely emerged from its infancy (Grenander 1994; Grenander and Miller 2007).

But the theoretical biologist might approach this task in a second way instead. That biologist, aware of Hotchkiss's mantra that life is "the repetitive production of ordered heterogeneity," is much more concerned with the converse intellectual flow, starting from the theoretical biology of that heterogeneity and ending up more constructively with a tool kit that arguably might be responsive to the biologist's axioms instead of flouting them the way almost every statistical maneuver we teach today's students does. The present article focuses on one particular suggestion that might serve the purposes of this rhetorical reversal particularly well. The data analyses in Bookstein (2015) were the first to note a novel regularity in patterns of morphometric data—a frequently encountered pattern of loglinearity of partial warp variance (PwV) as a function of bending energy (BE) over a range of application domains (growth, evolution, birth defects)—that might help circumvent the critiques I was publishing at the same time. Bookstein (2017b) extended this discovery by a new form of factor analysis, customized for applications to analyses of organismal form, that applies to the residuals from the loglinearity.

But it turns out that that loglinearity is not just a convenience of morphometric data analysis. It articulates with principles of theoretical biology set down by a variety of disciplines at midcentury but mainly overlooked since then. So this might be a good time to turn the attention of theoretical biologists and evolutionary biologists to a deeper exploitation of this new pattern engine, one much more closely aligned with those old theoretical concerns than anything flowing from the "data science" to which our current image resources are so often unfortunately consigned. Hence this article, which explicitly combines two domains of discourse that have hitherto been pursued separately:

the themes of *multivariate morphometrics*, adaptation of the classic 20th-century tool kit of covariance matrix analysis to accommodate the newly automated data resources of measures of the size and shape of organisms living, dead, or extinct. For the purposes of this essay that tool kit will be represented mostly by publications of my own;

and

the themes of *theoretical and evolutionary biology*, articulations of the principles that lie at the basis of modern biological explanations and in particular that differentiate those explanations from the not always differently phrased explanations apposite to inanimate nature, on the one hand, or to social organization, on the other. For the purposes of this essay that list of themes will be represented by three of its classic embodiments in book form: Gerard (1958), Waddington (1968, 1969, 1970, 1972c), and Elsasser (1998). According to this literature, the essence of living matter is its embodiment of "repetitively organized heterogeneity" and "memory without storage," whereupon its quantitative descriptors need to arise as a meaningful selection of simple and stable descriptors from out of this complexity.

To interweave these two streams, some clearing of biotheoretical underbrush is needed, in order to formalize some aspects of the information that the algebra is variously managing or disregarding. This essay is a first step in that direction, focusing on a new language for the reporting of that "repetitive production of ordered heterogeneity" in one of its aspects, its geometric scaling. The language to be introduced below attends equally to a potentially important new quantification pertaining to samples of sufficiently similar organismal forms and the persistence of a fundamental characterization of living matter that bears implications for the rhetoric of pattern reports pertaining to those forms.

Elsasser's own contribution to modern biology is, in retrospect, an ironic one. A recent review (Gatherer 2008) summarizes the situation:

During his lifetime, Elsasser's ideas were the subject of considerable comment, but very little in the way of detailed critique. The revival of interest in his work has produced several short introductions but no full reassessment of his ideas. This article attempts to fill this gap, with the emphasis on his relevance to systems biology, a field which Elsasser foresaw. ...Until recently, it seemed that Elsasser's biological work might vanish into obscurity, with references to it in the literature becoming increasingly rare from the mid-70s onwards, even though Elsasser remained scientifically active in biology until the late 1980s. Under such circumstances, the present article would scarcely be required, but his star is once again on the ascendant, with the realisation that Elsasser's themes are important to modern biology in a way that they were not in an era of simpler model systems studied using simpler methods. Now that the technical difficulties of systems biology have become apparent, anyone with original thoughts on the meaning

of complexity is well worth a reassessment. (Gatherer 2008, pp. 1, 9; citations omitted)

The present article represents my preliminary attempt to fill an adjacent gap, the relevance of his ideas to current trends in the biometrics of organismal form. My basic argument, indeed, echoes Gatherer's here. Let me paraphrase, then:

Elsasser's themes are important to modern morphometrics in a way that they were not in an era of simpler statistical models studied using simpler methods. Now that the technical difficulties of measuring organismal form have become apparent, anyone with original thoughts on the meaning of its complexity is well worth a reassessment.

## Ordered Heterogeneity, Nonstructural Memory

### *On Growth and Form*

A hundred years ago the British natural philosopher D'Arcy Thompson produced this celebrated book, which explicitly broached one of the arguments I am trying to fuse in this essay. Thompson wrote,

When the morphologist compares one animal with another, point by point or character by character, these are too often the mere outcome of artificial dissection and analysis. Rather is the living body one integral and indivisible whole, in which ...aspects of the organism are seen to be conjoined which only our mental analysis had put asunder. The co-ordinate diagram throws into relief the integral solidarity of the organism, and enables us to see how simple a certain kind of *correlation* is which had been apt to seem a subtle and a complex thing.

But if, on the other hand, diverse and dissimilar fishes can be referred as a whole to identical functions of very different co-ordinate systems, this fact will of itself constitute a proof that variation had 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. It will not only show how real and deep-seated is the phenomenon of 'correlation,' in regard to form, but it will also demonstrate the fact that a correlation which had seemed too complex for analysis or comprehension is, in many cases, capable of very simple graphical expression. (Thompson 1961, pp. 275–276; emphasis Thompson's)

Thompson's key idea was to work with pairs of drawings of organisms rather than with measurements of individuals.

To show a comparison, Thompson would draw a grid of vertical and horizontal lines over the picture of one organism in some orientation, then deform these verticals and horizontals over the picture of a second organism so that grid blocks that corresponded graphically would roughly correspond morphologically as well. He liked to interpret these graphical patterns via a metaphor, the explanation of form-change as the result of a "system of forces."

Thompson thereby foreshadowed all three of my main themes: production of a graphical representation of "non-structural memory," the resemblance of juvenile forms to their later configurations or of ancestral forms to their descendants; simplification of those comparisons by careful selection of the features pertinent to the claimed similarity, disregarding innumerable others; and attribution of the particulars of this similarity to explanations (in Thompson's examples, "forces").

Left for future development was the technical task of operationalizing those "Cartesian transformations," the grid diagrams: realizing Thompson's intuitions objectively via instrument-derived data objectively manipulated in groups. This project proved intractable. As Peter Medawar put it some 40 years later, "The reason why D'Arcy's method has been so little used in practice (only I and one or two others have tried to develop it at all) is because it is analytically unwieldy" (Medawar 1958, p. 231). Another three decades had to pass before a methodology emerged that replaced the examination of Thompson's "grid lines" by arithmetical operations upon actual point locations. (See my brief analysis in Bookstein 2018, at the end of Sect. 5.1.4.) But the emergence of that tool kit, usually called "geometric morphometrics," is not the subject of my present argument. Instead the issue is the epistemology of the deeper claim here, that "the living body" is "one integral and indivisible whole," the "integral solidarity" of which a coordinate diagram "throws into relief" for the purpose of furthering specific explanations and the theories they illuminate.

It is *this* part of Thompson's argument that has proven the most difficult to operationalize for realistic data. We need to archive and aggregate whole samples of instances of organisms (not just single textbook illustrations), compare them by reliable algorithms comprehensible to the nonmathematician, interpolate or extrapolate those comparisons, and ultimately wield them rhetorically in a manner coherent with the ultimate explanations that justify calling these enterprises "quantitative bioscience." The relevant literature, as I will review it in a moment, did not show the growth pattern usual for scientific controversies, but instead arose as a series of three bursts of intellectual effort, the first two 10 years apart and the third a full 20 years later, followed, in effect, by silence as the profession changed the focus of its attention to the discrete information channels of the -omic domains. Thus the state of the discussion as of 1987 remains

in general its state today. My specialty of morphometrics, on the other hand, has made nearly all of its remarkable progress over the years since 1987. Perhaps it has lessons to impart, then, for the articulation of theory with numerical praxis in this domain of organismal form.

### Hans Przibram and the Vivarium

Even before this effort of Thompson's, the centenary of which was observed in 2017 with much pomp, there had been a remarkably sophisticated attempt to transform biology into a theoretical science that respected mathematization, wherever feasible, along lines similar to those of physics. This was the research program of the *Biologische Versuchsanstalt* (Biological Experimental Institute)—the “Vivarium”—established in 1902 in Vienna, which endured until it was closed by the Nazis. Its underappreciated effort to quantify developmental biology was recalled to our attention in the same year as Thompson's centenary via the retrospective symposium chaired by Gerd Müller and then the proceedings volume he edited (Müller 2017).

The Vivarium not only carried out strikingly original, technically masterful research on the environmental determinants of developmental variation and regeneration (Coen 2006) but also produced many publications in domains we would now classify under the rubric of theoretical biology. A particularly clarion statement of their program can be found in Chap. 2, “Organometrie,” of Hans Przibram's short pamphlet *Aufbau mathematischer Biologie* (“The Structure of Mathematical Biology”) of 1923. In just seven pages, unleavened by any figures, this brief chapter reviews or anticipates an entire half-century's investigations into organismal development by analogy with physics (a vision superseded only 30 years later by the discovery of the role of a science even more fundamental than physics, to wit, the science of information as it applies to the molecular biology of DNA).

Przibram argues that form, when studied in appropriately controlled experiments, must be regarded as the integral of continuous morphogenesis or development as observed over varying conditions. (Here I am not translating, merely interpreting in today's much-altered language.) But development takes place inside the organism more than on its surface—we must have access to an appropriate parametrization of the “repetitively ordered heterogeneity” before we attempt to distill its statistics via external measurements of the organism expressing it. Hence the measurement of internal components is essential for comparisons at all taxonomic ranks (thus the chapter's title, “organometry”). In analogy with mechanics, which divided (back then) into statics, kinematics, and dynamics, biological studies ought to be divided into studies of single organismal forms one by one (“Statik”), then what we now call growth trajectories (“Kinematik”),

and, ultimately, once we have identified the effective factors, what we would now call studies of ecophenotypy (“Dynamik”). Comparisons of species must likewise be considered dynamically, so as to accommodate differences in their developmental conditions, “although these data do not yet afford of a mathematical treatment” (Przibram 1923). In short, reports of organometric findings must be under the control of functional or ecophenotypic explanation *at the appropriate experimental level*. In this context of principled reductionism, where form is an epiphenomenon of developmental physiology, if mathematically precise descriptions aligned with the laws of physics are to apply to organismal phenomena, then the biological descriptors will need to be the descriptors of physics.<sup>1</sup> Then the next important technical issue must be the distribution of that reductionist modeling over the interior of the organism in extenso.

To the proponent of such a methodology the system of Cartesian transformations latterly proposed by D'Arcy Thompson could only have been seen as deeply misguided. In Przibram's critique of 1923, Thompson's method, which essentially reduces to comparisons of proportions among organs, is purely inductive, and does not conduce to explanations. Indeed, he notes, Thompson gives us no insight into how an organism made up of diverse organs could submit to analysis via a holistic (“einheitliche”) assessment in the first place, let alone change into another species while remaining an integrated, functioning entity. Przibram concludes this frustratingly brief chapter with an insight astonishingly ahead of its time:

Thompson's holistic deformations can be made comprehensible if we can visualize a space lattice upon the living form, so as to assess how each little piece changes its shape under conditions that vary by species. Here lies open a rich, nearly undeveloped field that invites a mathematization, one whose erection we hope will begin very soon. At the conclusion, measurement of the physical space grid will have been unified with stereochemical organometry.<sup>2</sup> (Przibram 1923, p. 14)

<sup>1</sup> This program would be explicitly refuted 30 years later by Williams (1956), who points out somewhat iconoclastically that most biochemical titers vary by factors of upwards of 10 within a taxon or within an organism over time. I return to this point at the end of this section.

<sup>2</sup> “Thompsons einheitliche Deformation verständlich würden, falls wir in der lebenden Form ein Raumgitter erblicken dürfen, bei dem die Anordnung kleiner Teilchen unter bestimmten, mit der Veränderung der Art aber wechselnden Winkeln vor sich geht. Auch hier steht noch ein reiches, fast unbebautes Feld offen, das zur Errichtung einer mathematischen Theorie einladet, und dessen Bebauung hoffentlich in allernächster Zeit wird angefangen werden. Den Abschluß sollte dereinst die Vereinigung der physikalischen Raumgitter- mit der stereochemischen Organometrie sein” (author's translation).

He had been somewhat more explicit the year before, in Contribution 20, “Das Winkelmaß der lebenden Formen” (Angle Measurement of Living Forms), from his collection *Form und Formel im Tierreiche* (“Form and Formula in the Animal Kingdom,” Przibram 1922). On the very last page of this little essay, he argues that Thompson missed the point of his own grids—he talks too much about lengths and only occasionally about changes of angle (= proportion). Przibram concludes, in a sentence that I find simply astounding:

It is a promising and enticing future task to design the diagrams for this organismal space grid so as to represent animal forms in formulas.<sup>3</sup> (Przibram 1922, p. 158)

Przibram’s “enticing future task” in effect foretells the formally multiscale analysis propounded in the present essay. It is as if the Hans Przibram of 1922–1923 had explicitly anticipated a description of the features of “space grids” in the appropriate statistical language that would appear 95 years later. (It must be conceded that the physicists of his time did not yet know how to go about the description of variation in lattices—random field analysis was not to emerge within probability theory for another 50 years.) Yet Przibram and his program go unmentioned by all three of the later sources I am reviewing in this section. Had his work been memorialized appropriately—had our field turned to the description of those “space lattices” in the middle of the 20th century instead of in the present decade of the 21st—the story I am telling in this article might have been considerably different, or might have been told much earlier.

### Concepts of Biology

Vienna’s Vivarium notwithstanding, a review of the modern approach to quantification in theoretical biology might begin with a slender volume of 1958, edited by the neurophysiologist Ralph Gerard, that reports a synthesis scripted by the embryologist Paul Weiss (who happens to have briefly been an assistant of Przibram’s at the Vivarium) and polished at a conference organized by Gerard in 1956. The bulk of this 124-page publication is the “condensed transcript of the conference” from its rapporteur, Russell Stevens. Of particular interest is the tabular summary of the symposiasts’ suggested “principles,” page 148. I reproduce this scheme

**Table 1** “Principles,” from Gerard (1958)

---

|  |
|--|
| I Hierarchization of biological entities   |
| A. Entities are populations; consequently statistical rather than typological                  |
| II Order   |
| A. Heterogeneity in interaction arrangements   |
| B. In energy supply and enzymes  |
| C. Detected by ordered observer—specificity  |
| III Stability (homeostasis, equilibrium, steady state)   |
| A. Open system   |
| B. Equifinality  |
| 1. Efficiency—power  |
| 2. Rate of return to stability   |
| IV Adaptation  |
| A. Adaptive amplification  |
| B. Information storage—learning  |
| V Replication of similar individuals   |
| A. Pre-existing individuals interact with random factors                                       |
| VI Evolution (genetic alteration of molecular structure under selective impact of environment) |
| A. Recombination of individuals and species—adaptive diversity                                 |

---

in Table 1 in its original format, a transcription of text scribbled on a blackboard “during a coffee break.”

Following which, Ernst Mayr (Gerard 1958, p. 149) commented shortly afterwards, “from II to VI might possibly be simplified into three headings: the creation of order; the maintenance of order; and the changing of order.” To which Weiss (Gerard 1958, p. 149) appends a friendly Vivarium-flavored amendment: “When you speak of order you have evolution in mind, but I think the change of order is equally characteristic of development.” The crucial concept, then, is this trope of *order*, at all levels, especially those higher than the organelle or the cell.

This scheme, which persisted to the end of the symposium, could itself be summarized by Rollin Hotchkiss in the words I have already quoted as my first epigraph: “Living matter is the repetitive production of ordered heterogeneity” (Gerard 1958, p. 145). Elsasser indeed cites this as the explicit anticipation of his “theory of organisms” to be published 30 years later (Elsasser 1987, p. 39; see below). The principle was restated vividly by the meeting’s organizer, Paul Weiss, in the form of an aphorism: “Identical twins are much more similar than any microscopic sections from corresponding sites you can lay through either of them” (Gerard 1958, p. 140). We know that to be good biologists we must ignore irrelevant details, but we have no a priori specification as to what makes details irrelevant. As in statistical thermodynamics, it takes a long time for a community to come to some agreement on what those criteria might be.

<sup>3</sup> “Es ist eine vielversprechende, lockende Aufgabe der Zukunft, die Diagramme für die organischen Raumgitter zu entwerfen, die tierischen Formen in Formeln darzustellen.” How can it be that I first encountered this man’s thinking only in 2018, 41 years after my thesis on precisely this topic?

## Towards a Theoretical Biology

A decade later, the British geneticist C. H. Waddington chaired a series of symposia under the auspices of the International Union of Biological Sciences resulting in a splendid four-volume compendium (Waddington 1968, 1969, 1970, 1972c) covering the full range of speculations across the branches of biology of its time. That is not to say that it represented any sort of canonical synthesis. The title of the compendium, after all, begins with the word “Towards,” and, as the preface to the first volume notes, “Theoretical Physics is a well-recognized discipline. ...In strong contrast to this situation, Theoretical Biology can hardly be said to exist as yet as an academic discipline. There is even little agreement as to what topics it should deal with or in what manner it should proceed” (Waddington 1968, Preface). This position was echoed by several of the other symposiasts, for instance, Martin Garstens, who noted (1969, p. 285), “Nothing in the history of biological research has appeared and been found acceptable as a formulation of theoretical biology.”

Four years later, in his “Epilogue” following volume 4, Waddington (1972b, p. 284) concludes that the main difficulty is aligned with Elsasser’s concern for how “things with a certain global simplicity — a transfer of gaze, a tissue cell of a certain type, a four- or five-jointed leg — rise from a set of microstates of extreme complexity.” And in a trope that seems to have anticipated this essay of mine, he guesses that issues like these will have to be characterized “with the help of the analogy of language,” beginning with his own work (Waddington 1940) on the wing veins of *Drosophila*. He means this almost literally: insofar as the theory of general biology is concerned with “algorithm and programme,” it will need to cooperate with other aspects of the human uses of language, including (though Waddington overlooks this one) their use in quantitative descriptions of patterns in nature such as his own exploration of fruit fly wing patterns in this same volume. While there is no mention of Wittgenstein (nor anybody else from the Vienna Circle) in Waddington’s volumes, his conclusion falls not too far from Wittgenstein’s orphic summary of the role of physics in philosophy, “Wovon man nicht sprechen kann, darüber muss man schweigen” (Wittgenstein 1922, Proposition 7).

The only contributor to Waddington’s symposia who cites the earlier Gerard publication is Walter Elsasser, who notes (1970, p. 158) that his own 1958 announcement of ideas that would culminate in the 1987 essay was, “unknown to me at the time,” verbally very similar to the Gerard group’s ideas being floated at the same time: “my own approach [both of 1970 and of 1987] is no more nor less than a rather abstract formulation of what these scientists [of the Gerard volume] have been expressing.”

## Reflections on a Theory of Organisms

The brief eruption of biotheoretical explorations I am reviewing here culminated, well after the deaths of both Waddington and Gerard, in a late-life essay by the physicist-turned-theoretical-biologist Walter Elsasser that used the rhetoric of theoretical physics itself to distinguish what was unique about theoretical biology. (I have modified the title of Elsasser’s 1987 essay for use as my own.) I have already quoted his pointed summary regarding the centrality of memory without structure as the second epigraph at the top of the present essay. In an alternate phrasing (Elsasser 1998, p. 139), “it becomes indispensable to introduce the observed stability of organic form and function, over many millions of years in some cases, as a basic postulate of biology.”<sup>4</sup> (Earlier, on p. 96, he had referred to the then-novel concept of punctuated equilibrium in precisely this connection.) Together this pair of aphorisms captures the essence of Elsasser’s “four principles,” the topic of his Chap. 3, which are worth reviewing in greater detail. I will argue that the new biometrics I will be introducing in later sections constitutes the first systematic translation of these principles into the domain of morphometrics, indeed, into the domain of organismal biology *sensu lato*. Although in many respects Elsasser’s thinking is strikingly modern (i.e., his reference to “the competitive struggle for scientific results,” p. 104), his principal argument is grounded entirely in comparative philosophy of science, to wit, to investigate “whether and to what extent the idea that theory imparts structure to the empirical material can be applied to biology” (1998, p. 16).

At several junctures Elsasser implies that organismal form might serve as the best testbed for his principles. If biological complexity is the combination of “structural complexity together with temporal variability” (1998, p. 25), then for introducing oneself to this complexity “the best practical means is the perusal of an atlas of human anatomy,” a type of text “singularly suited to arouse that sense of marvel which is the nourishing ground of every scientific endeavor.” (In which respect see Bookstein 2017c.) But the “Cartesian method,” the resolution of complex phenomena into “smaller components,” must be abandoned in the course of the transition from theoretical physics to theoretical biology (1998, pp. 26–27), because, unlike the situation in statistical mechanics, variation among internal microstates cannot be depended on to average out.

Following this prolegomena Elsasser introduces his four “principles” of bioscience. Again he raises the horizon of his view for a moment of striking modernism:

<sup>4</sup> Notice how Elsasser here conflates the holistic invariance of a species with the precise, quantum-level invariance of inanimate categories such as minerals or atoms and their components. I thank Norm MacLeod for this point.

Since science is for all practical purposes an institutionalized endeavor, the scientists themselves may not even be aware of existing prejudices if these are the outcome of long continued training. ...In this book I am trying to face squarely the fact that the passage from mechanistic to holistic biology goes across a discontinuity. (Elsasser 1998, p. 36)

Then the four principles—the “holistic principles”—are enumerated: “ordered heterogeneity,” “creative selection,” “holistic memory,” and “operative symbolism.” These rules will have a surprising relevance to the new biometric method that I will eventually introduce here.

### Ordered Heterogeneity

By this phrase Elsasser means essentially the same rule of method that Hotchkiss had already adumbrated 30 years earlier in the Gerard symposium: an order (e.g., the pattern of boundaries of biological macrocompartments, such as individual bones) that is invariant relative to irregularities at the molecular or cellular levels. This heterogeneity is one expression of “a concept of major biological importance, *individuality*” (Elsasser 1998, p. 40), meaning the same thing Paul Weiss meant in his bon mot about the twins that I have already quoted. “Individuality is nonexistent in the physical sciences,” Elsasser goes on to note, or, phrased more operationally, “the laws of physics do not preclude unbounded repetition of an experiment, the regularities of biology (morphology) do” (1998, p. 41). The concept of homogeneous classes is limited to physics and chemistry—“everybody knows what is meant by a bottle of chemically pure water, and everybody understands that the molecules in this bottle are exactly alike.” But “organisms, on the other hand, show an observationally quite different behavior” (1998, p. 60). Earlier he had put this more magisterially: “We shall conceive of biological processes as an inextricable mixture of mechanisms with individualities” (Elsasser 1970, p. 140), meaning, processes that cannot be explained solely by repetitively controllable, measurable physical properties of the entities involved. The difference between a live and a dead cat is that the living cat is continually producing or regenerating organized inhomogeneity, whereas the dead cat isn’t.

### Creative Selection, Holistic Memory

These two principles may be reviewed jointly. “The regularities of morphology (so far as they cannot be derived mechanically) are the result of a selection made by nature among the immense multitude of atomic-molecular states,” and this selection, which “is a primary expression of

biological order,” is via a criterion of information stability.<sup>5</sup> “The organism uses its freedom to create a pattern which resembles earlier patterns either in the same organism or in earlier (parental) ones” (Elsasser 1998, p. 43). There is nothing illogical in this concept of “memory without storage,” Elsasser continues; it is the core of his theoretical scheme. He distinguishes further between the genetics of *replication*, which is “homogeneous” (meaning, constituent parts of DNA molecules can be identical), and the biology of *reproduction*, which is “heterogeneous” (meaning, holistic, creative). The new morphometrics I am proposing will be pertinent to the heterogeneity pole of this antinomy. In a nutshell: growth can sometimes (at least in principle) be reduced to biophysics—in effect this is what Przibram was pointing out (cf. Huxley’s 1932 formalism of “relative growth”); reproduction, however, cannot.

It gratifies me that Elsasser turns here to two of my own main favorite compendia, Anson’s (1963) atlas of anatomy and Williams’s (1956) review of human biochemistry, to illustrate his points about individuality and creative selection. Of particular importance to Elsasser is the fact that whether in the anatomical domain or in the biochemical domain, variation of measures of proportion as well as organ size across samples of normal human adults can range upwards of factors of ten. But, importantly in light of the next principle, “we expect to find the most striking evidence for biochemical individuality when we look at details, rather than at crude summations” (1998, p. 67). (Waddington 1972a, p. 109, makes the same point about branching anatomy, in a different language, but does not cite either of these sources.)

<sup>5</sup> This usage of the word “selection” is not the same as Darwin’s, of course. As Newman (2018), among others, points out, there are severe restrictions on the forms of animals—“tissue layers, lumens, segments, appendages”—that can be “readily generated by physical organizing forces acting on tissue masses with minimal programming by the genome. With respect to evolution, it means that body plans and organ forms will inescapably be characterized by these motifs despite their not having arisen by multiple cycles of selection for improved fitness.” His point of view suggests that the appropriate domain for a biometrics of selectable variability might well be restricted to quantifications *within* these relatively unvarying motifs. Elsasser (1975, p. 203) had already noted a similar point, that “the description of a living object ...implies an abstract selection. ...One goes from theoretical physics to theoretical biology not by the method of superadding new, formal constructs ...but by omitting an overwhelming fraction of an immense number of available descriptions.” In other words, the domain of biometrics is, in a sense, the complement of the domain(s) of inherency. This biometric domain might sustain studies of selection even when the processes that account for inherency cannot.

## Operative Symbolism

Elsasser's final principle states the compatibility of heterogeneous reproduction with the second law of thermodynamics, but goes beyond it. Nothing in inanimate nature can preserve an information-rich structure over millions of years the way a biological species can. "Ordinary chemical stability," he notes, "can hardly be sufficient for the continued maintenance of all information required to keep a species similar to itself over several million years" (1998, p. 81). The second law deals only with molecular disorder; memory without structure, a type of biological order, is, to use the modern jargon, supervenient over Boltzmann's level. "The order of inorganic science can be described mathematically but the order of biological classes which are heterogeneous cannot conveniently be represented by mathematics. ...In biology we then propose to speak of *regularities*. ...That organisms maintain their peculiarities, anatomical, physiological, and behavioral, over their lifetime is a result of the most ordinary human experience. It is far more satisfactorily understood in terms of a dynamic process of heterogeneous reproduction than by a purely static model of replication" (1998, pp. 72–73). (In this way Elsasser implicitly extends his dismissal of the multivariate Gaussian models to incorporate the Ornstein–Uhlenbeck class of recurrent stochastic models as well.)

By including the word "behavioral" in the preceding list Elsasser (1998) means to include brain processes too: "cerebral memory is mainly a matter of heterogeneous reproduction, ...a property of all higher organisms independent of their degree of consciousness." That memory is a matter of heterogeneous reproduction seems a requisite, for instance, for the persistence of memories across the extended human lifetime. (These notions intriguingly anticipate one current theory of brain science, the thermodynamic view of Clark 2016 and Friston 2010.) In short, "for the first time in history the biologist has to face squarely his own epistemological problems," Elsasser (1998) states on page 89 in boldface type. "The time has come to face and digest the concept of 'memory without storage,' " not only as regards ordinary "cerebral memory" but also "all information stability in the organism." Even more tersely: "the ability to reproduce structural organization without intervening storage [is] a basic property of living matter" (1998, p. 89).

Taking the four principles together, we arrive at the "decisive hypothesis" (Elsasser 1998, p. 115) that the living world "has its own internal order which can only be described as radically different from anything that physicists and chemists have ever conceived ...The reproduction or maintenance of information is exactly the point where the organism is different from any mechanism that engineers may have constructed." This holistic memory is "a primary

phenomenon of nature" (1998, p. 118) that will prove to align well with the new approach to morphometrics I will be sketching in sections below. First, however, it is appropriate to review what is wrong with the existing statistical approaches.

## Today's Multivariate Statistical Tool Kit is Inadequate for Our Purposes

If we agree (however tentatively) that the basic postulate of biological science is "the observed stability of organic form and function" (Elsasser 1998, p. 139), or, equivalently, Hotchkiss's repetitively produced ordered heterogeneity, then any coherent biometric methodology must accord with that principle. However, the great majority of contemporary approaches to data analysis have not yet accommodated this fundamental aspect of biological organization; they may, indeed, have no way to do so.

One can trace much of the difficulty here to a tacit decision by the statisticians of the previous century to set their discipline in a formalism of "variables" across "cases"—the idea of data as a *matrix*; see, for example, Mardia et al. (1979). The corresponding mathematical model (the linear finite-dimensional vector space) and statistical model (the multivariate Gaussian distribution) are unsuited to organismal biology in all of the following ways. In a nutshell, the Gaussian models all rely on one particular model of ignorance, the model of maximum entropy (cf. Jaynes and Bretthorst 2003), that simply fails to match what we biologists most urgently need to learn. In Elsasser's phrasing, this would be the heterogeneities of creative selection that drive the living kingdoms, not the homogeneities that bind them to the physical world.

## Difficulties in the Passage from Instruments to Information

Just as the essence of biological science is ordered heterogeneity, the essence of the corresponding quantifications must be the ordering of information; and for adapting studies of organismal form to that principle, the model of the finite-dimensional vector space is just not adequate. Despite early attempts by numerical taxonomists such as Rohlf and Sokal (1967), the conversion of images into data for purposes of understanding (as distinct from classification) has proven intractable by any algorithms short of neural nets (which are, after all, imitations of human perception). The information we need for our descriptions of organismal form—the decompositions often referred to in toto as "computational anatomy"—are *highly* nonlinear in the actual pixel intensities captured by imaging instruments, so that the manipulations (which are strenuous) required to reduce an image's

raw data into landmarks, curves, surfaces, centerlines, organ volumes, and so on, must precede any subsequent quantitative manipulations. This is equally true of the spacetime approaches (cf. Cressie and Wikle 2011) that follow images through time, as in analyses of morphogenesis or its converse, the progress of diseases characterized by lesions. Lengths and proportions (Bookstein 2018, Sect. 5.1), volumes, average pixel values over organs—all quantities like these are nonlinear in the image data, so that all require the prior enunciation of bioscientific principles that justify their actual empirical formulation. Even the concept of a “pixel value” requires attention, after all, based as it usually is on unspecified biophysical aspects of the interaction of radiation with living matter. One usually circumvents the associated logical problems by presuming that an anatomical boundary is a locus that can be detected redundantly by a wide range of these biophysical models. But how is such an axiom to be justified, and what are the limits of its validity?

In this connection, even the basic idea of a “coordinate system,” such as the Cartesian framework of a digitizing tablet, is fraught with problems. The analyst’s task is to report the underlying biological phenomena independent of artifacts of instrumentation. Thompson, for instance, would begin with a square grid on some “starting” form. The orientation of that grid may derive from functional considerations (gravity, symmetry) but in practice (e.g., for the anthropoid skull viewed mediolaterally) it is typically pure artifact (cf. Bookstein 2016b). Less obviously, so is the straightness of its axes, inasmuch as the associations among loci given by the gridded graphical links have nothing to do with biological reality. We need descriptors of shape change that do not rely on accidents of coordinate system choice, and the difficult problem of constructing such sets of descriptors for focal phenomena has not yet been solved. One early attempt, the biorthogonal grid of Bookstein (1978), properly applies only to uniform transformations (Bookstein 2018, Sect. 5.4); I will discuss this option in a later section.

After all, the role of a coordinate system is to “co-ordinate,” to link continua of spatial location that are related in some potentially meaningful manner. Hilbert and Cohn-Vossen (1932/1952) summarize the classic nineteenth-century understanding of the manifold ways in which “space,” including our laboratories, can be organized. To select one of these systems, the Cartesian system or any other, requires reference to biological postulates; but the machinery for that articulation is simply lacking. Our task must instead be the construction of descriptive systems that do not depend on a priori choices like these, but that instead, if co-ordination is indeed a finding, *produce* those coordinates by objective analytic manipulations. For a discussion of this crucial point, see Bookstein (1981, 1985).

One way to approach this problem is René Thom’s formalism of catastrophe theory (see his chapters in

Waddington (1968, 1970), or his well-regarded book *Structural Stability and Morphogenesis* of 1975). The fundamental observation here is not specific to biological systems, but applies more broadly: a dynamical system can sometimes make irreversible commitments to one path over another, *bifurcations*, that can be reconstructed from observations at a later time—a version of Elsasser’s “nonstructural memory” in the sense that the conditions at the time of the bifurcation are not measurable at later stages, but must be inferred from a valid model. Furthermore, such a system is not “aware” of the bifurcation at the moment it arises—the bifurcation is, in the jargon of systems analysis, an *emergent* property of the dynamics, not an explicit quantity measurable at a single instant. For applications of this approach at the level of organelles, see, for example, Tabony (2006). One application of Thom’s catastrophe theory to morphometrics is the breakdown of the map between Cartesian grids and principal strains at *umbilic points* (see Bookstein 1981). Another, more speculative at this time, is my suggestion of 2000 that descriptions of spatially varying grids might concentrate on the points where there is a *transition-to-lips catastrophe*, an extremum of directional derivative that affords a local “horizontal” and “vertical” based on details of the shape change not at single points but over extended regions (Bookstein 2000, or see under “crease” in Bookstein 2004 or Bookstein 2018).

### Difficulties with the Notion of “Linear Combination”

The time is past, if ever there was such a time, when you can just compute a linear combination and turn it loose in the world and assume that you have done good.

— modified from Berry 2000, p. 145

The point I am hinting at in my parody of Berry’s pithy advice (the original had the phrase “discover knowledge” in place of my “compute a linear combination”) is of fundamental importance: *the main theoretical construct of today’s multivariate biometrics, the linear combination, seems not to articulate with any of the principles of theoretical biology.* For linear multivariate quantifications to be a reliable aspect of the biologist’s tool kit—for the goal to be understanding rather than classification—there *must* be such an articulation. This section, a brief survey of the available options, concludes that no such articulation is possible within the standard statistical geometry of “data vectors” absent an a priori theoretical structure for the roster of dimensions involved. The penultimate main section will instead exploit a much more general structure, for which the geometry of distance between patterns is

different for patterns corresponding to essentially different styles of biological explanation.

In the interest of brevity I confine my discussion in this subsection to the truism that to interpret an “unconstrained” linear composite, such as a multiple regression formula, is not defensible in our biosciences (nor in the other domains of applied statistics, either). Two principles must intervene. First, statistical computations must be referred to invariant reference systems, the way the astronomer uses positions of the stars with respect to the ecliptic, not the horizon. In mathematicians’ language, to be biologically interpretable a descriptor needs to have the epistemology of an *equivalence class* of equally meaningful quantifications under changes of arbitrary aspects of the instrumentation. (The first of these to enter into morphometrics was invariance with respect to the scale of the ruler: Mosimann 1970.) It is usually the case, however, that equivalence classes have a geometry totally distinct from the original instrumentation, just as the geometry of the surface of a sphere is entirely different from the geometry of the points in space whose “directions out of an origin” the sphere represents for some theory or other (Fisher et al. 1993). Second, explanations will always involve *comparison* of observed data summaries with other data-derived points or with points according with particular theories. (I refer here not to significance tests of nonzero versus zero mean differences or correlations, of course, but to comparisons of theories that are roughly equally well-supported prior to the current investigation.) We will need a methodology for such comparisons; it will, in general, be entirely distinct from whatever methodology was involved in generating the separate individuals or samples being compared.

Today’s applications of multivariate statistics can be reviewed under two headings, as the approach is either a priori or data-constrained.

The *a priori approaches* are easily dismissed. The standard model of allometric growth (Jolicœur and Mosimann 1960) analyzes log-transformed extents with respect to a vector of all-1’s. But this requires data that are all nonnegative numbers in the same units of length, for example, all in centimeters, or all in  $\text{cm}^2$ . Such a restriction on the information that we are permitted to extract from raw imagery would be intolerable (cf. the many examples of inconstant dimensionality in Blackith and Reyment 1971). In some fields, such as mathematical geology (Koch and Link 1980) or medical image analysis (Pennec et al. 2016), there is a second approach that frees coefficients not to be identical but just to have the same sign—a pattern having the format  $\sum a_i x_i$  with all  $a_i > 0$ . Such a formula is some multiple of a *weighted average* of the data, a construction that in other fields is sometimes called a *barycenter*. But for the purposes of theoretical biology, such a formalism requires that the

entities  $x_i$  being thus averaged cannot themselves vary in sign—in practice this means they must already be separate theoretically meaningful patterns themselves.

The *data-dependent approaches* fall into a few clusters. In the simplest of these approaches, from the larger context of signal processing the bioscientist borrows the formalism of a *hierarchically organized universal orthogonal basis*, such as Fourier analysis (for a 1-D signal or a 2-D curve) or spherical harmonics (for, among other applications, star-convex surfaces in space). The justification of formalisms like these is in terms of their energetics (the connection of the Fourier coefficients with vibrating strings, that of the spherical harmonics with entities governed by a Laplace equation, e.g., electron clouds); neither seems appropriate to the comparative description of organisms.

I have deconstructed the *principal component analysis* (PCA) approaches in a recent essay (Bookstein 2017b) that does not need recapitulation here. PCA has far too many shortcomings to be of use in any theoretically coherent context except when there is an exogenous quadratic form (e.g., energy) that is in control. Otherwise, the algebraic principles that govern a PCA express mainly the list of measurements included, a list whose construction is usually subjective, thus without biological content. Jolliffe (2002, p. 297) puts it in a nutshell: “If the [phenomena to be explained] are not expected to maximize variance and/or to be uncorrelated, PCA should not be used to look for them in the first place.” Neither of these goals has any relationship to biological explanation, and thus neither articulates with the biotheoretical foundation for quantifying organismal form at which this essay is aimed. (The standard literature of these methods completely ignores the difficulty here by tacitly restricting the domain to specialized atheoretical branches of the biosciences such as numerical taxonomy.) Related modern statistical techniques—factor analysis (Bookstein 2017b), structural equations modeling (Spirtes et al. 1993), and the like—are likewise disarticulated from quantitative biological explanation without strenuous modifications. Factor analysis, for example, applies mainly to lists of scalar measurements that lack any further parametric structuring principle (but see Bookstein 2017b for a version compatible with the pattern language I’m about to introduce here); structural equations modeling is based on a naive physicalist model (see the critique in Bookstein 2018) or probabilistic model (Pearl and Mackenzie 2018) of causation lacking any relationship to Elsasser’s dicta.

A third class of data-driven approaches builds linear combinations by the method of *multivariate calibration* (Martens and Næs 1989): formulas  $\sum a_i x_i$  where each coefficient  $a_i$  corresponds to a covariance or correlation describing how each of the  $x$ ’s of the list is predicted by some exogenous scalar (possibly under side-conditions). Specific algorithms under this heading include multiple

regression, the “lasso” method, partial least squares and its regression version, and others. In general the problem of any such composite was already identified by Sewall Wright nearly a century ago (see, e.g., the review in Bookstein 2018, Chap. 3). For the pattern of these coefficients to serve in *explanations*, as distinct from merely using the arithmetically derived values  $\Sigma a_i x_i$  casewise in a forecast or prediction, the list of ancillary criteria that must obtain increases in length as the square of the number of predictors. Furthermore, the formalisms of linear regression, and even the coefficients of covariance on which regressions are based, are profoundly problematic when applied to lists of variables for which there is not already a prior reductionist theory controlling the elementary aspects of addition and multiplication that those formulas for covariance and regression rely on. (See Bookstein 2016a.)

Even deeper problems arise when the scientific model of independent finite-dimensional samples driving all these computations is replaced by another that is equally plausible as bioscience but leads to wholly different algebra. For data in the form of time series, for instance, both regression and PCA will fail completely unless radically modified (see Bookstein 2013a); for data of any dimension that is high with respect to the sample size, *all* the likelihood-based approaches fail for any of several reasons (for random samples, the problem of excess dimensionality is distilled in the Marchenko–Pastur theorem, see Bookstein 2017a or Cardini et al. in preparation; for problems of evolvability, the dominance of edge-effects over bulk effects in multivariate analyses of selection, Gavrillets 2004). The typical Gaussian formalism likewise excludes Boolean models, networks, and models with chreods, among others. In my judgment this is too much insight to give up for mere reasons of software availability or the comfort of revisiting the ideas one was taught in graduate school.

In short, **none of today’s standard machinery of biometric statistics, whether univariate or multivariate, offers any reassurance to bioscientists that their data analyses accord with the fundamental principles of bioscientific inference from quantitative data gathered at the level of whole organisms and their organs.** In Przibram’s language, we still do not have sturdy methods for assessing the variability of those spatial lattices (*Raumgitter*) at the multiple scales required to make sense in whatever theoretical context is driving the organometry. We must start afresh with a different methodological assignment: *within a restricted domain of biological phenomena and their instrumentation*, to construct a geometry of measurement and explanation for morphology consistent with biological theory. In the rest of this essay I offer such a novel approach, a fusion of biological geometry and descriptive statistics governed by a newly discovered principle of morphometrics

that explicitly embodies Elsasser’s insights: pattern analysis by a variety of views of one fundamental new graphic, the *BE–PwV plot*.

## BE–PwV Plots and the Quantitative Summaries that Interpret them

This section reviews the methodology of the proposed new method and the insights that it affords. The next section will sketch a wide range of its applications, under four subheadings, and the corresponding four types of graphical reports that it can generate.

### Measures of the Similarity of Pattern Findings Must Explicitly Refer to the Consequent Biological Interpretations

As a first attempt at a synthesis I replace the task of “searching for a biological explanation” by a radical simplification: the construction of a biotheoretically sensible quantity expressing the *similarity* that is the subject of Elsasser’s notion of “reproducible heterogeneity”—the similarity of a pattern to one member of a family of possible alternate explanations erected *in advance* in conformity with the nature of the data abstracted from the original biological imagery. Following a practice that has become standard in contemporary morphometrics (Bookstein 2018, Chap. 5), reports take the form of sentences that say either “this comparison looks like the following theoretically interesting phenomenon” or “this comparison looks like a combination of the following two (or three, or four) theoretically simple phenomena at different levels of spatial detail.”

In other words, I am operationalizing the task of “quantifying organismal form” as the construction of a *suite* of pattern metrics for summarizing observed comparisons, across a range of spatial scales, by the extent to which they align with one of a set of prior quantitative explanatory patterns *at that scale*. Such a replacement echoes the protocol recommended by Platt (1964) in his method of “strong inference,” which he suggested be copied from molecular biology as a guide throughout all the biosciences. (See Bookstein 2014, Sect. 3.3; and then Anderson et al. 2008.)

In its computational geometry, the approach I am suggesting would couch *all* the descriptors, both empirical and explanatory, in a space of their own within which quantitative summaries must align with canons of biological explanation per se, prior to any consideration of patterns in the particular data set at hand. The reporting rhetoric exemplified in the various examples of the next section here will reduce observed quantitative summaries to sets of graphical distances that should almost all be small, or almost all be large. But the “points” involved in these descriptions pertain

to highly processed versions of the raw data. For instance, the relevant quantities of the central graphic upon which the examples of the next section will rely, the BE–PwV plot (Bookstein 2015), are residuals from a regression of the log variance of linear-trend or quadratic-trend partial warps on log specific bending energy, a very highly derived quantification indeed. Yet considered as a whole, the method closely matches one of the four principles set out by Elsasser (1998) that underlie a coherent “theory of organisms.”

The embodiment of Elsasser’s principles within this new organismal biometrics is via the following dictum: to convey, even approximately, what it is that we have discovered—to turn multivariate arithmetic into understanding—within each of a specified class of findings, patterns must be considered as locations within a neighborhood characterized by a particular sum-of-squares formulation, a selection from one of many dissimilarities (distance formulas) that vary from neighborhood to neighborhood so as to match an a priori class of simple potential explanations. What we’re approximating is the *meaning* of a linear combination or some other pattern descriptor: a metric among patterns, not among forms.

The explanation chosen to report any empirically derived pattern will be an assignment to one of a small number of interpretable pattern families, each of much lower dimension than the dimensionality of the data themselves. Explanations, in other words, are not summaries of the data; they are simple statements (or simple combinations of simple statements) that are not too far from data-derived points of the same pattern geometry. Then the methodologist’s assignment is to design reasonable and plausible metrics for pattern similarity.

Here I take as an axiom the demonstration in Bookstein (2016a) that, principally for biological reasons, Procrustes distance is inappropriate as a candidate metric for biometric shape spaces describing organisms when explanations involve any sort of experimentally reproducible mechanism—when they deal with any level of process except the phylogenetic and with any theory other than neutral evolution.<sup>6</sup> This is because the Procrustes formulation precludes all the other principles of biological order accrued over centuries as natural history slowly transformed itself into organismal biology. Procrustes distance, with or without its accompanying size measure, is simply too symmetrical—too homogeneous; furthermore, the model of variation that accommodates it (the model of total lack of integration) is wholly incompatible with the “ordered heterogeneity”

driving all the morphometric methodologies of the present century. In the course of building sensible substitutes for Procrustes distance, the corpus of established bioscientific knowledge must play the dominant role, the extant lore of multivariate statistics a secondary role only. A pattern analysis of BE–PwV plots substantially aids our search for an appropriate methodology of organismal form.

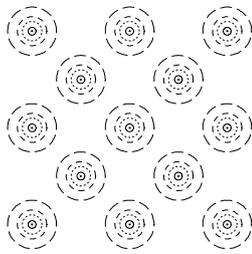
### The Algebraic Machinery of BE–PwV Plots

“A BE–PwV plot” is the concise name for a new type of scatterplot, the plot of log partial warp variance (PwV) against log bending energy (BE) over any sample of sufficiently similar homologous landmark configurations. (For the meaning of “sufficiently similar,” see the Discussion.) The radical reformulation of organismal biometrics that this article puts forward for data sets of landmark configurations rests on the understanding of these BE–PwV plots. If this approach is new to the reader, there is no substitute for studying the original presentation (Bookstein 2015) or the somewhat lengthier didactic summary in Chap. 5 of Bookstein (2018) prior to the further decomposition of this pattern language to be pursued here. (Earlier pedagogies, such as Chap. 7 of Bookstein 2014 or Chap. 4 of Weber and Bookstein 2011, were written prior to the construction of the BE–PwV machinery.)

We need some vocabulary. A sample of organismal forms will be represented for present purposes as a configuration of named *landmark points*, of count  $p$  in total, that are considered to be homologous from specimen to specimen on grounds that do not concern us here. At left in Fig. 1 is a possible statistical model for such data sets in some laboratory coordinate system. It is assumed that the position and orientation of these specimens are irrelevant (or else the positioning and orienting features, such as the elevation above ground or the orientation with respect to a fluid flow, would be encoded as additional points in the data configuration). *Shape*, which further ignores scale, is represented for these purposes by  $2p$  Procrustes shape coordinates (or, in 3-D,  $3p$  of them) that embody constraints upon the three aspects (totalling four dimensions, for 2-D data; seven, in 3-D) of the isometry group we are choosing to ignore. (Size may very well matter for explanations such as allometry or biomechanics; the methods here extend easily to incorporate it if the shape data set is extended by a single log size measure, an augmentation that will not concern us here.) The plot at right in Fig. 1 shows a sample of the shapes generated by the statistical model at its left when standardized in these three ways.

Combining the spirit of Hans Prizbram with some elegant formulas from the computational vision literature of the 1980s, in order to most effectively generate potential biological explanations, we display interesting shape

<sup>6</sup> This is a different issue from the “homology” of landmark points per se, an issue I have raised in Sect. 5.1 of my 2018 textbook. This essay takes up the issue in the Discussion where I dig deeper into Prizbram’s metaphor of “grids.”

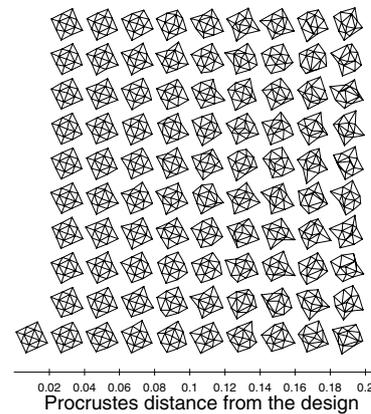


**Fig. 1** The isotropic Mardia–Dryden model for variation of landmark configurations in their original Cartesian coordinates (left). Samples of empirical configurations arise as perturbations of a holotype (13 points from a square grid) by identical circular Gaussian variations at every landmark independently. Shown here are four possible settings of double the standard deviation of those Gaussians. That null model

comparisons (differences of group averages, or patterns of association with causes or effects) by *thin-plate splines*, deformations of Cartesian grids squared on one form that minimize the total of local information (sum of squared second derivatives of the map) over all smooth interpolants consistent with whatever landmark rearrangement is under investigation. The total of these squared second derivatives is not a function of the definition of horizontal or vertical. And as all these derivatives are zero for transformations that are not “bent”—that take parallel lines to parallel lines—the physicist would recognize the integral over the picture plane of a total like this as the *bending energy* (BE) that such a transformation would have if it pertained to a real metal plate instead of a picture. This formalism is now familiar in detail from several morphometrics textbooks, including mine.

It is a remarkable mathematical fact that the integral of that sum of squared second derivatives over the whole plane can be computed as a simple quadratic form applied to the shape coordinates. The formula for that quadratic form, the *bending energy matrix*, is a function of the average shape of the landmark configurations. For a data set of  $p$  landmarks it has  $p - 3$  nonzero eigenvalues, the *specific bending energies* corresponding to shifts of shape coordinates totaling 1.0 in summed squared Procrustes distance, and  $p - 3$  corresponding eigenvectors, the *principal warps*, each of which specifies a pattern of coordinated landmark translations in any shared direction. When the principal warps are examined in order of increasing specific BE, their drawings show a series of steadily more and more spatially specific focused transformations. The series goes from border-to-border growth

### assorted examples of Procrustes distance



is absurd for applications in organismal biology (right). This chart assembles a series of simulations at varying distance from the template at lower left, showing how unlikely it is that totally uncorrelated morphogenetic patterns like these would lead to biological explanations of any cogency

gradients at large scale, having the smallest specific BEs, to the relative displacements of closely spaced landmark pairs or triples that correspond to the principal warps of smallest spatial scale (greatest specific BE). We can rotate any sample from the representation by the Procrustes shape coordinates to this new basis, the *partial warp scores*, if we carry along a representation of the subspace of no bending, the *uniform component*, as well. There arises a new and biotheoretically crucial auxiliary plot, the *BE–PwV plot*, which scatters the log of specific bending energy against the log of partial warp score variance over the partial warps of any data set.

But “bending energy” is not an energy in any physical or physiological sense; it is only a useful metaphor. Here the concept applies to our *descriptions* of form — our language for *reporting* patterns of change of Przibram’s grids: not the physiology by which the changes are actually produced, but their decomposition for descriptive purposes into a range of spatial scales suggestive (in many examples) of separate developmental processes. In this context one can think of the partial warps as analogous to the sines and cosines of the more familiar *Fourier analysis* of observed periodic processes via their amplitude spectra. The analogy is actually quantitative: the two components of a partial warp for two-dimensional data correspond to the two coefficients, sine and cosine or amplitude and phase, for each frequency in a Fourier analysis, and the slope of  $-1$  that plays a central role in the taxonomy the next section introduces is equivalent to the slope of “ $1/f$  noise,” the “pink noise” observed in the physical spec-

tra apposite to many biological systems. (In pink noise, each octave carries the same amount of noise energy — of course, the net span of the underlying spectrum must be truncated for the total energy to be finite.) But whereas ordinary Fourier analysis diagrams the amplitudes *per se* of these components, the BE–PwV plot diagrams the *variance* of these amplitudes: a strategy that proves more helpful for biological interpretation in most applications.

It may be helpful at this point to clarify a distinction originally introduced by Bookstein (1989) between two distinct uses of this thin-plate spline. Algebraically, the technique takes the form of a pair of interpolation functions, one for the  $x$ -coordinates of a target form and one for its  $y$ -coordinates, that together generate the deformation formalizing D’Arcy Thompson’s “Cartesian transformation” between anatomies as an exact interpolation between configurations of landmark points. (This is the situation for morphometric data in two dimensions. If the data come in three dimensions, there is another term for the  $z$ -coordinates as well.) Each of these coordinate-specific interpolations represents the corresponding target coordinate as a smooth surface (see Diagram 5.62 in Bookstein 2018)—the algebra of this surface interpolation had been introduced a few years earlier by Terzopoulos (1983) for the quite different application domain of computer graphics. (His formulas, in turn, derived from theoretical advances in applied mathematics of the preceding decade, advances that I exploited, too.) In either context, surface interpolation or grid deformation, the thin-plate spline is the *unique* mapping that minimizes a certain nonlinear expression exactly matching a physical quantity that had been in use in engineering for over a century: the BE of a thin flat metal plate under normal deflection. The “thin-plate” portion of its name represents this physical analogy, and the term “spline” refers to the energy-minimization property. The terminology is the same for data in three dimensions except that the “surfaces” are now *hypersurfaces*, the four-dimensional analogue.

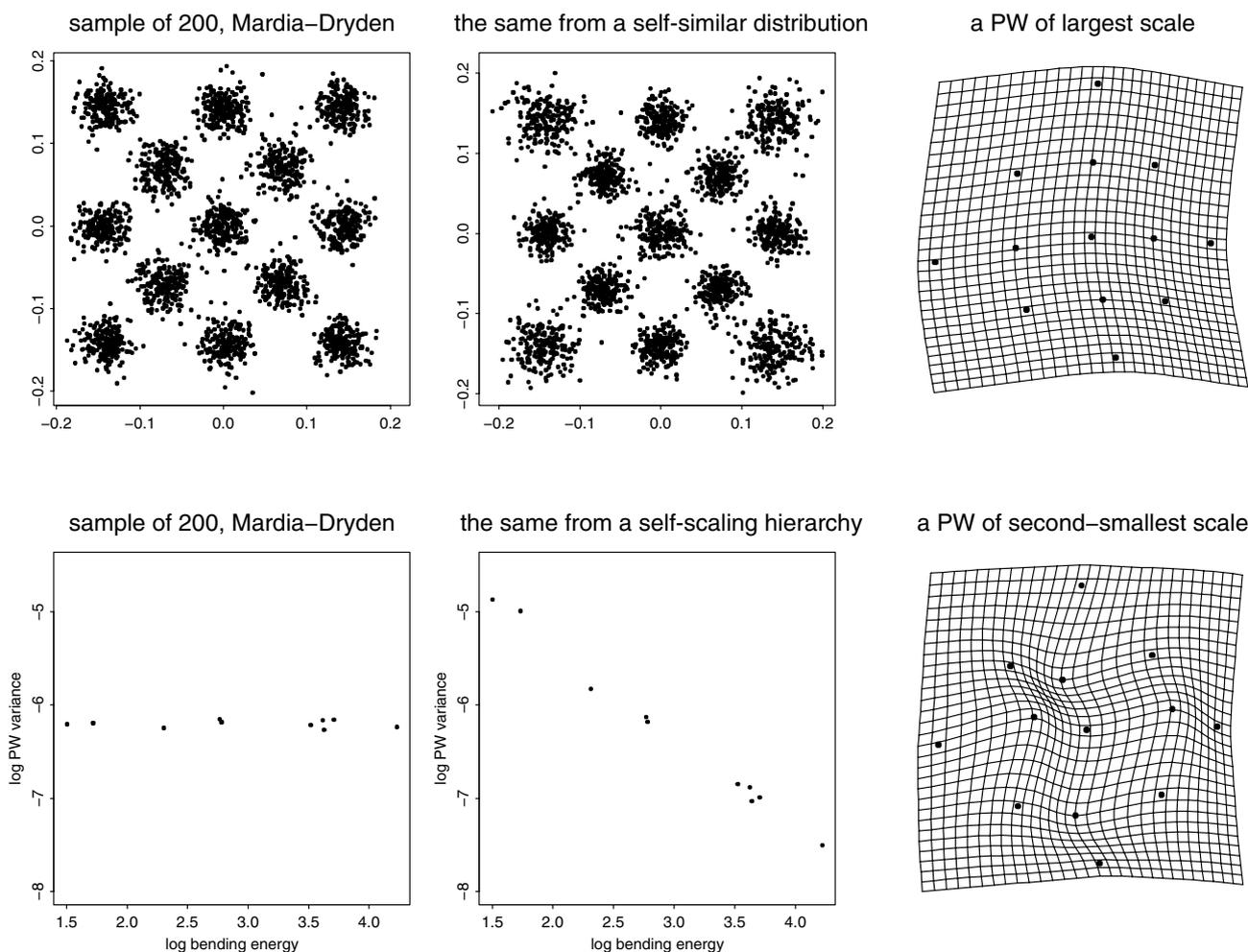
My 1989 paper seems to have been the first to notice that because this BE happens to be a quadratic form in the Cartesian coordinates of the target form, it can be eigenanalyzed; and furthermore that this eigenanalysis generates a surprisingly useful decomposition of the interpolation function itself into a hierarchy of components, each one a 2-vector multiple of the corresponding eigenvector. (For data in three dimensions, this will be a 3-vector.) The jargon of “principal warps” and “partial warps” was introduced in this original paper. Each partial warp can itself be drawn as a thin-plate spline, and in this sense the configuration of landmark displacements driving any deformation grid is itself the sum of the landmark displacements whose grids visualize its partial warps one by one (displacements of

all landmarks by multiples of one single 2-vector corresponding to the elements of the principal warp). Because any particular deformation grid can be decomposed into its partial warps, each of which is itself a deformation grid on the same starting configuration of landmarks, any sample of specimens (landmark configurations) can be considered a sample of these deformations, *and also of the hierarchy of their components*, where each grid interprets each specimen or each component partial warp as a deformation of their joint Procrustes average configuration. As I have already mentioned, that hierarchy is taken in a conventional order, from eigenvalue zero (the “uniform term”) through the low eigenvalues (for the large-scale deformations, those less bent per unit summed squared Cartesian displacements) right up to those of the highest specific BE, which usually involve shifts of just two or three landmarks at close spacing.

The present essay, then, will consider two separate biological interpretations of such a sample of stacks of deformations. Interpreting the components individually, especially when we are considering only one single biological phenomenon (a growth pattern, a two-group comparison), we can speak of the amplitudes of individual components (lengths of the corresponding 2-vectors) as larger or smaller. For instance, if one of them is substantially larger than all the others (Example 7), it can be called “dominant,” and if the first three are all larger than all the others, the phenomenon can be identified as a quadratic gradient, as in Example 2. But for studying the specific phenomenon of *integration* of variations in organismal form, a different language is appropriate in which the amplitudes of the partial warps of *an entire sample* are considered in terms of trends in their scale-specific variances. These often take the form of deviations from a linear relationship between  $\log$  BE and  $\log$  PwV, and it is a useful null model for integration to posit that this regression is a tight fit around a slope of  $-1$  (the case of self-similarity; Bookstein 2015). In that setting, the components that are “relevant” to the description of integration are those that deviate upwards from the regression. If there are no such components, then none of them qualify as “relevant” to the task of describing integration even if at the low-BE end their amplitudes happen to be large. In this text, I will reserve the word “relevant,” in quantitative contexts, for this integration interpretation (cf. Example 6), and refer to individual components or narrow ranges of components as “dominant” when the discussion is about individual partial warp amplitudes rather than their integration *per se*.

Note that sample summaries of amplitudes can be approximated from the BE–PwV plot. Since the quantities here are all nonnegative (squared lengths of those vector multiples of the principal warps), the variance of each serves as a lower bound on its own mean square.

For the isotropic Mardia–Dryden shape distribution itself, the points of a BE–PwV plot vary around a horizontal



**Fig. 2** Replacement of the isotropic Mardia–Dryden by a biologically much more appropriate null. Distributions of conventional Procrustes shape coordinates: upper left, Procrustes-distributed shape coordinates from the distribution in Fig. 1; upper center, a sample from a self-similar distribution, derived by deflation of the Procrustes distribution by the formula in the text. BE–PwV plots (log partial warp variance against log bending energy) for the ten partial warps of this

mean configuration: lower left, for the isotropic Mardia–Dryden, variance is the same for every partial warp; lower center, for the deflation recommended here, where log PwV is linear in log BE with slope  $-1$ . Upper right, an instance of the first partial warp with loadings  $(0.1, 0.1)$ . Lower right, the second-last partial warp, showing eight times as much bending for the same loading vector

line—the same variance for each partial warp score regardless of its scale (Fig. 2, lower left). This does not correspond to any biological data set I have ever observed. (In Przibram’s language, it would represent a *Raumgitter*, or, better, a *Flachsgitter*, that had no explicable features whatever, at any scale, and thus no possibility of any “measurement” except for one single parameter, the positional variance (perhaps mere instrument error) common to all the landmark coordinates.) To make sense of real data sets, we need more realistic models. One easy way to generate other reference shape distributions than the Mardia–Dryden of Fig. 1 is by scaling the principal warps so as to manipulate the slope of these BE–PwV plots in order to correct for this highly inappropriate constancy of geometrical signal strength over the

range of possible geometrical scales at which explanatory features might be detected.

Figure 2 shows the most useful of these transformations, in which the variance of each partial warp score is “deflated” by exactly the corresponding BE. (NB: One does not deflate one’s *data*, only the Mardia–Dryden model of pure Brownian noise.) In symbols (Bookstein 2015), the deflation of any shape *Pdist* is the adjusted shape  $defl = Pmean + \alpha \sum_{k=1}^{p-3} \sqrt{\frac{E_1}{E_k}} (W_k \cdot Pdist) W_k$  where *Pmean* is the Procrustes mean shape of the simulated sample,  $\alpha$  is any scaling factor, the *E*’s are the nonzero BEs of the  $p - 3$  partial warps  $W_k$ ,  $k = 1, \dots, p - 3$ , and the quantities  $W_k \cdot Pdist$ , taken as complex numbers, are the corresponding  $p - 3$  partial warp scores. In words, one begins by sampling

from an isotropic Mardia–Dryden distribution over some mean configuration, as in Fig. 1, rotates to the basis of the partial warps (ignoring the uniform component), deflates each nonuniform partial warp by the square root of its specific BE, and finally sums up all the components after that deflation.

This maneuver is demonstrated in Fig. 2 for the same 13-gon template used in Fig. 1. (The clustering of BEs along the horizontal axis of either BE–PwV plot is owed to the symmetries of this artificial quincuncial design.) Notice how much more the diagrams in the lower row contrast than those in the upper row, even though the underlying geometric resources are identical. The panels at lower left and center deal explicitly with the concern of *integration* by rotating the Procrustes shape coordinates of the top row to a far more informative orientation (namely, the partial warps), and the right-hand column exemplifies the scale-dependence of BE (upper panel, partial warp 1 with loadings (0.1, 0.1); lower panel, partial warp 9 at the same loading, but entailing roughly eight times as much bending). BE is proportional to the integral of the squared second derivatives of an interpolating map matching the average landmark configuration to each specimen of a data set, and these second derivatives can be intuited in these diagrams via the contrasts of shape of adjacent grid cells. In their algebra, the partial warps *are* eigenvectors of something, but that something is a matrix that treats discrepant shifts of nearby landmarks as standing for much greater shape changes than the same discrepancies among landmarks at greater distance. This BE matrix is not a covariance matrix, so the partial warps are not principal components.

It is another remarkable mathematical fact that the slope of  $-1$  in the lower center panel of Fig. 2 corresponds on its own to a powerful biometric model, *self-similarity*, meaning equivalent expected amplitude of shape phenomena regardless of spatial scale. Figure 3 demonstrates the truth of this assertion for the 13-gons of shape coordinates being modeled in Fig. 2, but the proposition is actually true in general (Kent and Mardia 1994). Shapes simulated using this model (Fig. 4) present a much more realistic panoply than Fig. 1 does of apparent features of form that much more easily suggest explanations in terms of the spectrum that will be introduced in the next section. I have shown examples of empirical distributions with this slope in Bookstein (2015): one for the callosal midcurve of the human brain under prenatal alcohol challenge, another for larger-scale aspects (such as “neuroglobularity”) of the anthropoid skull under hominization. For shape distributions according with the self-similar model, patterns ostensibly emerging at any scale (e.g., an end-to-end gradient) should be considered to have been just as “uncaused” (that is, just as inexplicable) as the analogous “patterns” emerging from examination of random walks, a mathematical structure affording the same

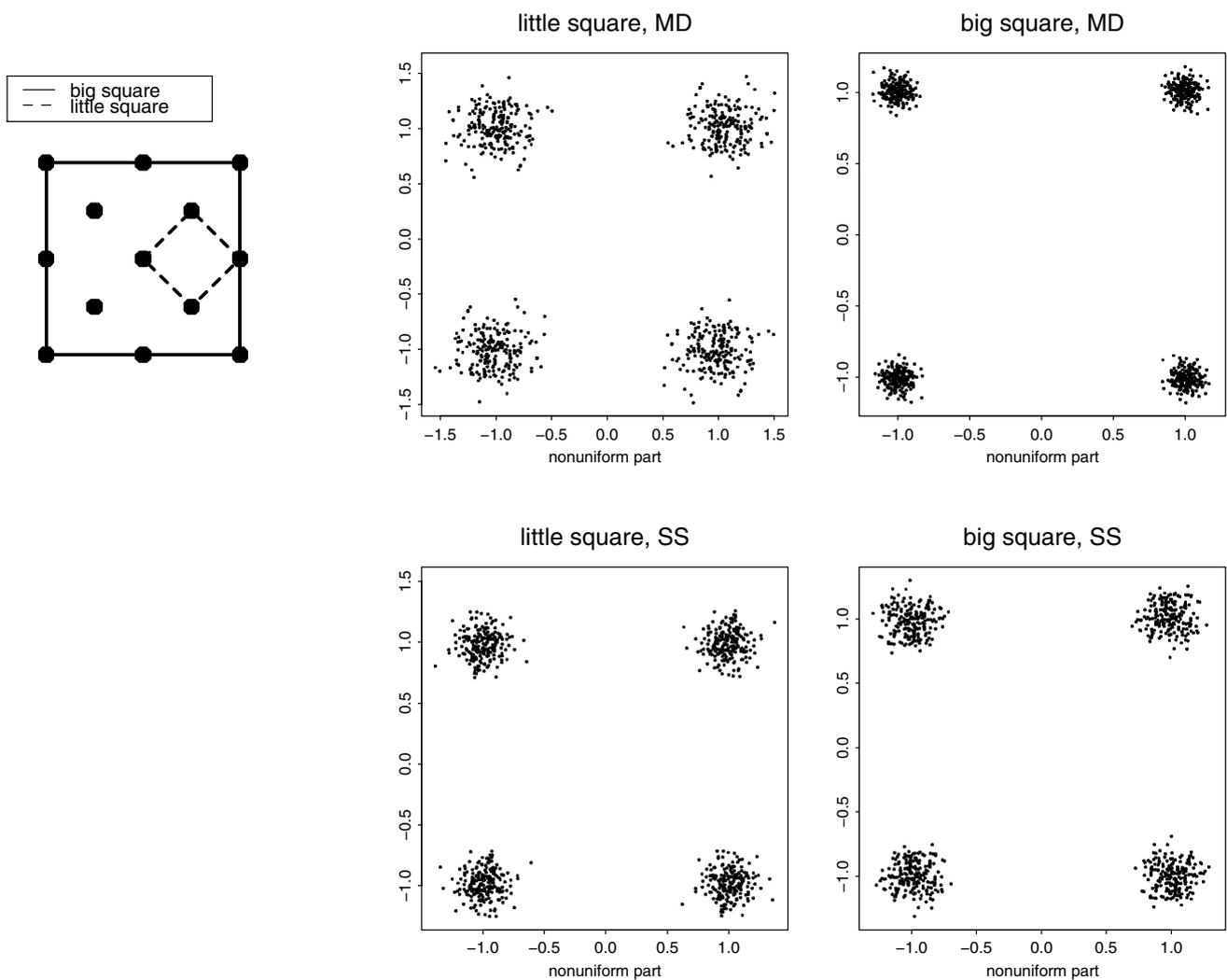
self-similarity but ordinated now in one dimension (usually, time) instead of two or three.

## A Tentative Pattern Language for BE–PwV Plots: Four Regimes

This section will suggest a variety of different pattern quantifications for shape comparisons corresponding to different configurations of the BE–PwV plot or its relation to the uniform component. These subclasses of pattern analysis have somewhat the character of a *spectrum*. Again Elsasser anticipated this circumstance, when he noted (1970, p. 137) that “individuality is the endpoint of a spectrum” with homogeneity at the other end. In terms of the classification I am about to introduce, Elsasser would call the separate landmark-by-landmark displacements (discussed below) “individuated,” while the uniform transformations and the growth gradients in the next two subsections appear homogeneous both algebraically and visually. Patterns at this level are to be submitted to holistic functional explanations instead of regionalized ones. In-between are the transformations with linear BE–PwV plots, which may be explained by regional functional arguments or instead by Elsasser’s “creative selection.”

In all the following types of explanatory findings, pattern similarity will be coded graphically rather than numerically, as adjacency or linearity on the paper or the computer screens bearing these data displays. Numerically, squared distance between patterns (such as the assessment of residuals, Figs. 14, 16) would be a sum of squares with a formula different from type to type. This dependence of report rhetoric on the actual data is, of course, not a flaw. The biologist must argue *against* any single canonical distance formula—the lists of dominant features are different in different parts of the shape space. (An analogy: here on Earth squared map distance is the sum of squared latitude difference and squared longitude difference near the equator but not near the poles.) Explanations associated with points of any of these metric structures need to be exogeneous to morphometrics: explanations that involve, for example, energy (in its various incarnations), information (of which BE itself is an instance), or consequences of some evolutionary or developmental process that might be functional or instead neutral. The variety of these explanations will roughly correspond to the variety of genetic, epigenetic, or ecological parameters that describe or differentiate them.

Before launching into this typology of descriptions, readers may benefit from a schematic overview. The subject is the classification and corresponding reportage of diagrams like those in the bottom row of Fig. 2, which sort the variation of a multidimensional shape change signal according to a range of components of different spatial



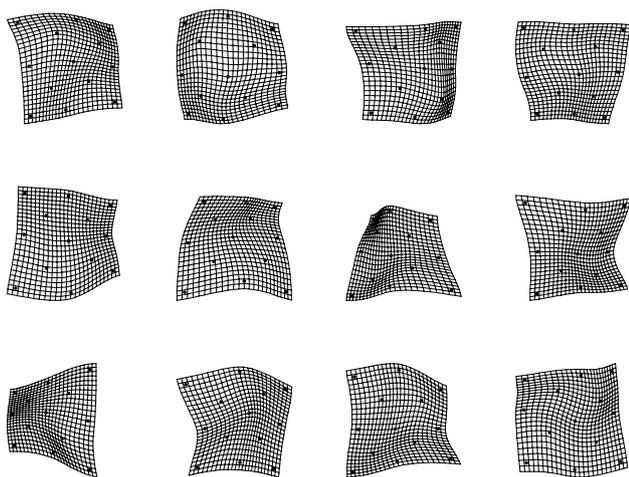
**Fig. 3** Confirmation of self-similarity for the distribution in the middle column of Fig. 2. The two quadrilaterals selected from the template for this demonstration (far left). In the Mardia–Dryden distribution, squares of different sizes have different shape variances, making it difficult to scan data analyses for meaningful patterns in any intu-

itively accessible way (upper row, center and right). In the deflated distribution (central column in Fig. 2), the shape variance of subconfigurations is the same regardless of their scale (lower row). (MD: the Mardia–Dryden distribution; SS: the self-similar distribution generated by deflation of the MD)

scales. The next subsection is concerned with the uniform transformations, those for which the single component that is *omitted* from this diagram (because it can be assigned no horizontal coordinate in the plot) accounts for most of the sample shape variation. Typically, uniform transformations express classical biomechanical factors; the example given in the subsection is the change of rodent neurocranial shape during the period of fastest neural growth. The following subsection considers the situation where the points at the upper left of these charts are substantially higher than all the others. These comprise the “lowpass” domain, illustrated three different times: by the pattern of adaptive radiation of the mammalian skull in midsagittal section, by an analysis of variability in the human disease

called “deviated septum,” and by an analysis of the difference between the geometries of two different human smiles.

The “Loglinearity” subsection deals with BE–PwV plots that take the form of straight lines sloping downward toward the left. These slopes can be steeper than  $-1$ , more gradual than  $-1$ , or very nearly equal to  $-1$ . The particular value of  $-1$  plays an important theoretical meaning similar to the role of 0 (no correlation or no mean difference) in classical statistical inference. As shown in Figs. 3 and 4, data for which the BE–PwV plot takes on this slope serve as an appropriate null model (self-similarity) for the task of describing shape variability. At a slope of  $-1$  there is no preferred scale, and so data approximating this situation should be described rather



**Fig. 4** Twelve forms from the distribution in the upper center panel of Fig. 2. The variability here consists principally of visually discernible separate features which emerge at distinct scales, thereby much more suited to explanations of organismal form

as random walks should be described: only by a variance, not by any list of specific features. (In other words, when the BE–PwV slope is  $-1$ , none of the partial warps, or any other feature of the sample shape variation, can be assumed meaningful—any of these dimensions may instead be expressing Elsasser’s third principle, “creative selection.”) The steeper this slope, the faster PwV drops with decreasing scale, i.e., the more integrated the transformations. One example is offered from the rodent neurocranial data set, now modeled over a longer epoch of its growth—this example proves to incorporate three *separable* biological developmental processes—and another from a study of correlations between human facial form and its covariates (body mass index, salivary cortisol) or cognitive effects (ratings of “health”).

Finally, the last subsection before the Discussion describes those BE–PwV plots that present, in whole or in part, segments appearing to show a horizontal trend. Two examples are on offer here, one (the residuals of the mammal cranial example already used in the “Lowpass Filters” section) that is resistant to biological explanation (a “null model” of complete lack of integration, never seen in biological materials at the organismal level) and the other extracted from Jim Rohlf’s public data set of wing venation landmarks for the genera of mosquitos by restricting the range of spatial scales considered. Taken together, these domains span all the types of examples I have encountered in my explorations of this diagram style.

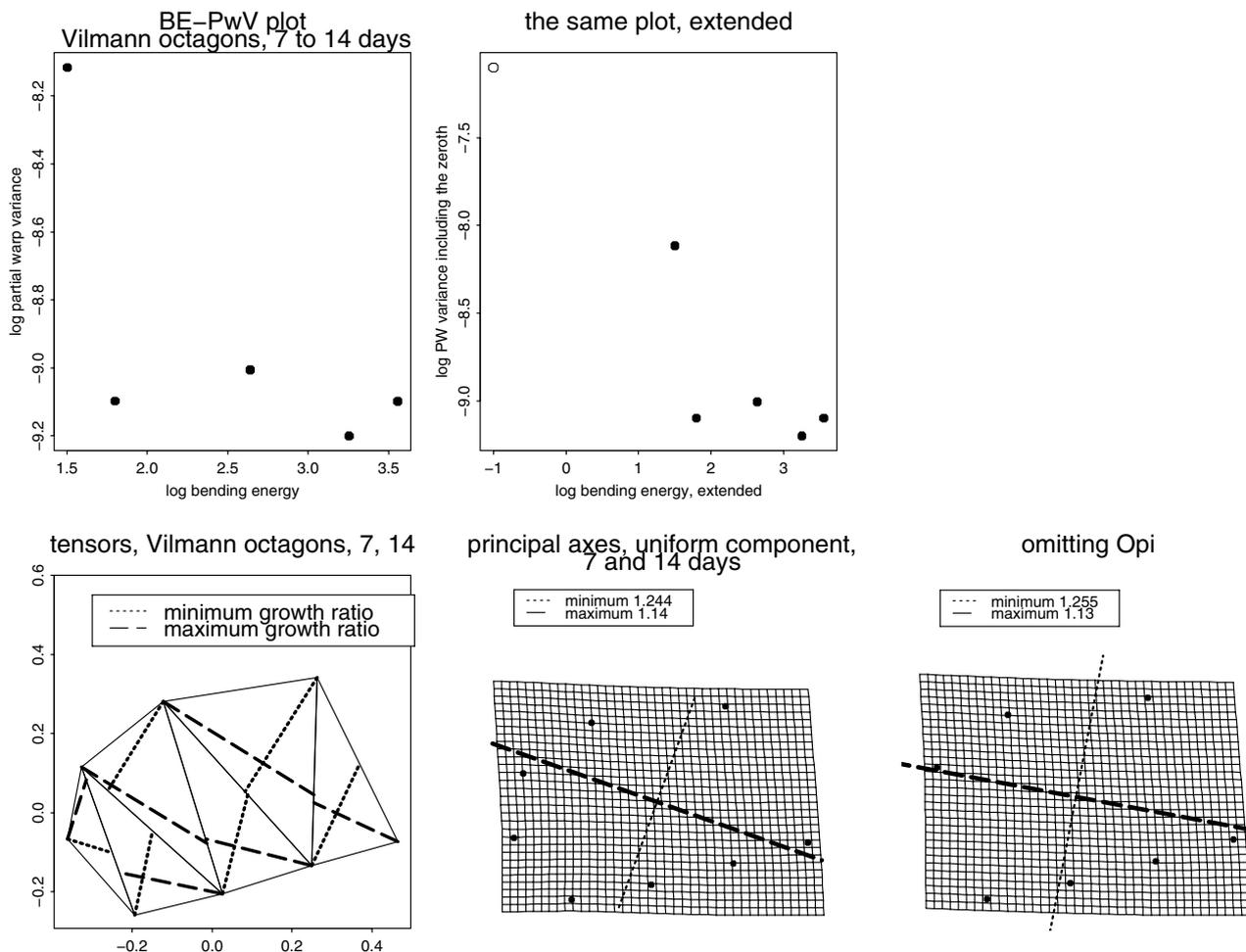
### Uniform Transformations and Their Explanations

The simplest class of geometric changes comprises the **uniform transformations**, those that take parallel lines into parallel lines. The description of these can exploit a

particularly simple mathematical substrate, 19th-century tensor algebra. The corresponding quantitative biological rhetoric affords such well-known reporting formats as summaries of the distribution of angles or proportions of lengths (Bookstein 1991, pp. 218–221)—reports hardly different from what Rudolf Martin, pioneer of anthropometrics, exploited more than a century ago (see Bookstein 2018, Sect. 5.1). The natural metric here would be the simplest ( $2 \times 2$ ) case of the general formula  $\sum_i \log^2 \lambda_i$  where the  $\lambda$ ’s are the *relative eigenvalues* of one positive-definite matrix with respect to another. In the original announcement (Mitteroecker and Bookstein 2009) that this metric had biological relevance, these were covariance matrices. In the present application, they are the *uniform components* that help to represent a sample of forms as deformations of their average. For general landmark configurations, the formula for the uniform component is rows 5 and 6 of the  $J$  matrix that, for samples distributed Gaussianly without too much variance of position, orientation, or size, approximates the Procrustes superposition by a linear projection operation applied to the  $2p$  original Cartesian coordinates. (See Bookstein 2018, pp. 409–410.)

**Example 1** *Vilmann’s rodents, 7 to 14 days*. The classic example of a uniform component is one aspect of the earliest example of shape coordinate statistics: the analysis of the often-exposed *Vilmann data set* of coordinates of eight midsagittal neurocranial landmarks for 21 male laboratory rats digitized in the early 1980s by Henning Vilmann from serial cephalograms at ages 7, 14, 21, 30, 40, 60, 90, and 150 days. The data are listed in an Appendix to Bookstein (1991) and also in the online resource associated with Bookstein (2018). Clockwise from left in the lower row of Fig. 5, the landmarks are Opisthion, IPP (intraparietal suture), Lambda, Bregma, SES (sphenothmoid synchondrosis), Intersphenoidal suture, SOS (sphenoöccipital synchondrosis), and Basion. I will draw two examples from this data resource, corresponding to two different types of explanation: here, in Example 1, a uniform transformation over this restricted age range; later, in Example 6, which considers a longer ontogenetic epoch, a curvilinear growth gradient together with a focal deviation from it.

Considering now only the first pair of ages of these animals, 7 days and 14 days, there result the representations in Fig. 5. At upper left is the standard BE–PwV plot whose algebra I reviewed above. While the plot appears to have some structure, note how limited is the range of the vertical coordinate for the rightmost four points. To put this variation in context, the upper central panel augments this plot by an artificial point for the log variance of the uniform component (which, of course, has no log BE). This variance clearly dominates the growth pattern, while the last four warps show the null pattern of nondecreasing variability



**Fig. 5** Shape change of the Vilmann rodent neurocranial octagons between ages 7 and 14 days is very nearly uniform. The BE–PwV plot restricted to these two ages only (upper left). Note the restricted vertical scale. Augmentation by the uniform variance, *x*-coordinate

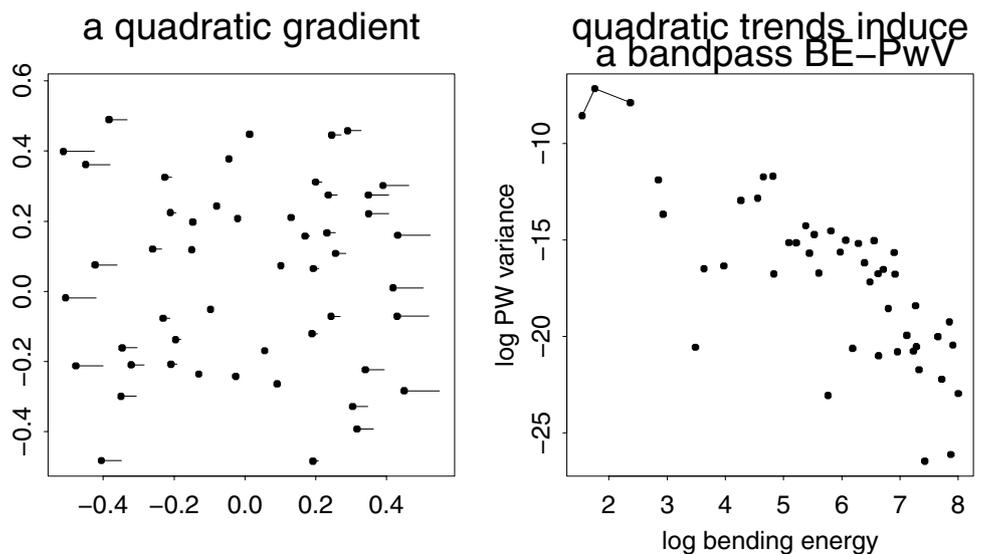
arbitrary (upper center). A convenient triangulation (lower left). Thin-plate spline grids and the corresponding uniform components for configurations of eight and seven landmarks, respectively (lower center, lower right). See text

already exemplified at lower left in Fig. 2. I focus here on the representation of that uniform term alone, the simplest possible description of growth over this earliest of the available age intervals. (There is a speculation on the meaning of that higher variance for the first partial warp at the end of Example 6.)

From a triangulation of these octagons we extract triangle-by-triangle representations of average growth across the form (Fig. 5, lower left panel) suggesting uniformity except in the vicinity of Opisthion. The least-squares fit to the full transformation (Bookstein 2018, Eq. 5.7), diagrammed in the lower center panel along with the unconstrained thin-plate spline, shows an anisotropy of about 9% in directional growth rates (1.244/1.14). When the posteriormost landmark, Opisthion, is omitted (lower right panel)—in other words, dropping the triangle on Opisthion—the anisotropy

increases to about 11% for a ratio of distances that evidently align with the directions SES–IPP and Bregma–SOS. (The corresponding values of  $\sum \log^2 \lambda$  for shape are about 0.004 and 0.006, respectively.) The first of these directions, Sphe-noethmoid to Intraparietal, might be a presumptive growth axis of the neural contents that could be hypothesized to account for the shape change of this functional matrix at this growth epoch. (Such a sentence exemplifies the match of metric to explanation suggested in the introduction to this section.) Example 6 returns to this data set to extract two additional descriptors, a growth gradient and a local perturbation, from the full temporal record.

**Fig. 6** Example of a shape phenomenon at large scale and its interpretation. A random 50-gon of “landmarks” with a shift that is precisely quadratic in form (landmarks all translated by a multiple of  $x^2$ , their original horizontal coordinate) (left). The corresponding BE–PwV plot, showing the dominance of the first three partial warps by a factor of up to  $e^5 \sim 150$  over the variance of every subsequent term (right). See text



## Lowpass Filters

BE–PwV plots dominated by the few partial warps of the largest scale convey the explanations that fall closest to biophysics in their combination of simple differential equations and randomness. (Recall that quote from Elsasser about the inextricable mixture of mechanisms and individuality. To Elsasser, mechanisms are governed by differential equations, individuality by randomness. But this is not *stochastic* randomness—individual differences surely arise from their own embedded mechanisms. Rather, the “randomness” here refers to those aspects of the developmental process that cannot be reliably replicated in experiments; it is their *modeling*, their *reporting* that is stochastic.) The lowpass part of a BE–PwV analysis is the simplified signal that suits explanations that are experimentally replicable mechanisms, meaning, selections that (in Elsasser’s sense) are not “creative.”

**Growth Gradients: A Simulation.** The largest-scale morphometric phenomenon after the uniform component that bears a fixed parameterization is the *quadratic trend*, suggested by Sneath as long ago as 1967, for which I offered a canonical form in Section 7.4 of my treatise of 1991. The general form of a quadratic growth gradient supplements the uniform component by additional features (for 2-D data, there are three of these: multiples of  $x^2$ ,  $xy$ , and  $y^2$ ) each multiplied by a vector analogous to a partial warp weighting. In the example of Fig. 6, based on a random configuration of 50 “landmarks,” there is no shape change except for a pure quadratic phenomenon precisely along the horizontal axis. This has nonzero projections on all 47 of the partial warps, but note (1) the clear separation of the variance of the first three partial warps against that of all the others, along with (2) the chaotic decomposition of this same pattern against

the remaining 44 dimensions of bending. The slope of this residual misspecification here, about  $-2$ , corresponds to what I published in Bookstein (2015) as the “true” self-similarity dimension for the full Vilmann data set to come in Example 6. The vertical difference between the low-pass cluster of PwVs and the others, which here has the magnitude of almost 5.0 on this natural log scale (a factor of about 150 before the log-transform), serves as a figure of merit for the claim that the low-pass description is adequate to the phenomenon.

This slope of  $-2$ , which I have never seen exceeded in practice, seems to be characteristic of nonuniform phenomena of largest scale. Besides an explanation in terms of an actual growth gradient or similar allometric process, a finding at this scale might be explained as the rotation of one anatomical compartment with respect to another (such as is corrected by the “block registration” exploited by Senck and Coquerelle 2015 or Currie 2018), the relative expansion or contraction of one compartment with respect to others, or a biomechanical process parameterized by a torque or moment. Any of these might have a uniform part as well, with which it might or might not share a report. Perhaps the simplest of all these options is the familiar phenomenon of bilateral symmetry of a paired compartment as modified by asymmetry of positioning (see in general Bookstein and Mardia 2003).

**Example 2** *The Marcus mammalian skulls.* Another way to handle this misspecification of quadratic gradients over partial warps is to construct and then correct for a quadratic term explicitly. My second example carries out such a procedure for the pedagogically useful data set compiled by Leslie Marcus and colleagues for analysis in Marcus et al. (2000):

calvarial landmarks for nearly all the major taxonomic subgroups of the mammals. The analysis by Marcus et al. was of 37 landmarks in three dimensions; the analysis here concerns only the 13 landmarks on the midsagittal plane (Fig. 7, top panel). The standard Procrustes representation, in the panel at middle left, shows huge variation at nearly every landmark, particularly inion and frontal-parietal (the outliers at upper left in the plot are those for *Homo sapiens*). The corresponding BE–PwV plot, lower left, shows a clear dominance of the first two partial warp variances over all the others.

To accommodate the growth gradient suggested there I replaced the standard thin-plate spline, with kernel  $r^2 \log r$ , by the quadratic version with kernel  $r^4 \log r$  (see Wahba 1990 or Bookstein 2004) along with three additional trend terms  $x^2$ ,  $xy$ ,  $y^2$  alongside the usual  $x$ ,  $y$ , and 1. When the standard plots are regenerated after this adjustment, the resulting scatter of residuals, center right, is far more tractable than the standard version to its left, and the corresponding BE–PwV plot, lower right, is indistinguishable from the flat model discussed below. (Note the difference of vertical scales between the two plots in the bottom row.) Any attempt at a numerical taxonomy of these skulls, or any attempt to correlate genetic distance against morphological distance, should be carried out using the sums of squares of the displacements in the right-hand panel of the central row, not those from the panel at its left. Such a “correction” could well grossly alter the conclusions of nearly every published article (e.g., Marcus et al. 2000) that has attempted analyses of correlations like these, with or without a phylogenetic correction, over a wide range of taxonomic groupings.

When this quadratic residual is taken as homogeneous at the mesoscale in the sense of Elsasser and this article, there remains the task of describing the larger-scale components of this sample. For this purpose I divided the quadratic spline’s total of eight trend terms into two for the uniform component (see previous subsection)—the conventional pair (Bookstein 2018), vertical dilation along the ordinate and horizontal shear along the abscissa—along with six for the coefficients of  $x^2$ ,  $xy$ , and  $y^2$  in the splines of  $x$ - and  $y$ -coordinates of the Procrustes-fitted landmarks separately. In view of the gross undersampling of both specimens and landmarks it is enough to resort to an ordinary PCA of this latter sextet of coefficients. The eigenvalues of this analysis are 4.7, 3.9, 1.4, 1.1, ..., so (according to the stepdown test of Bookstein 2014, pp. 322–324 or 2018, pp. 295–296) we need only consider the first two. There results the joint analysis set out in Figs. 8, 9, and 10 and their captions.

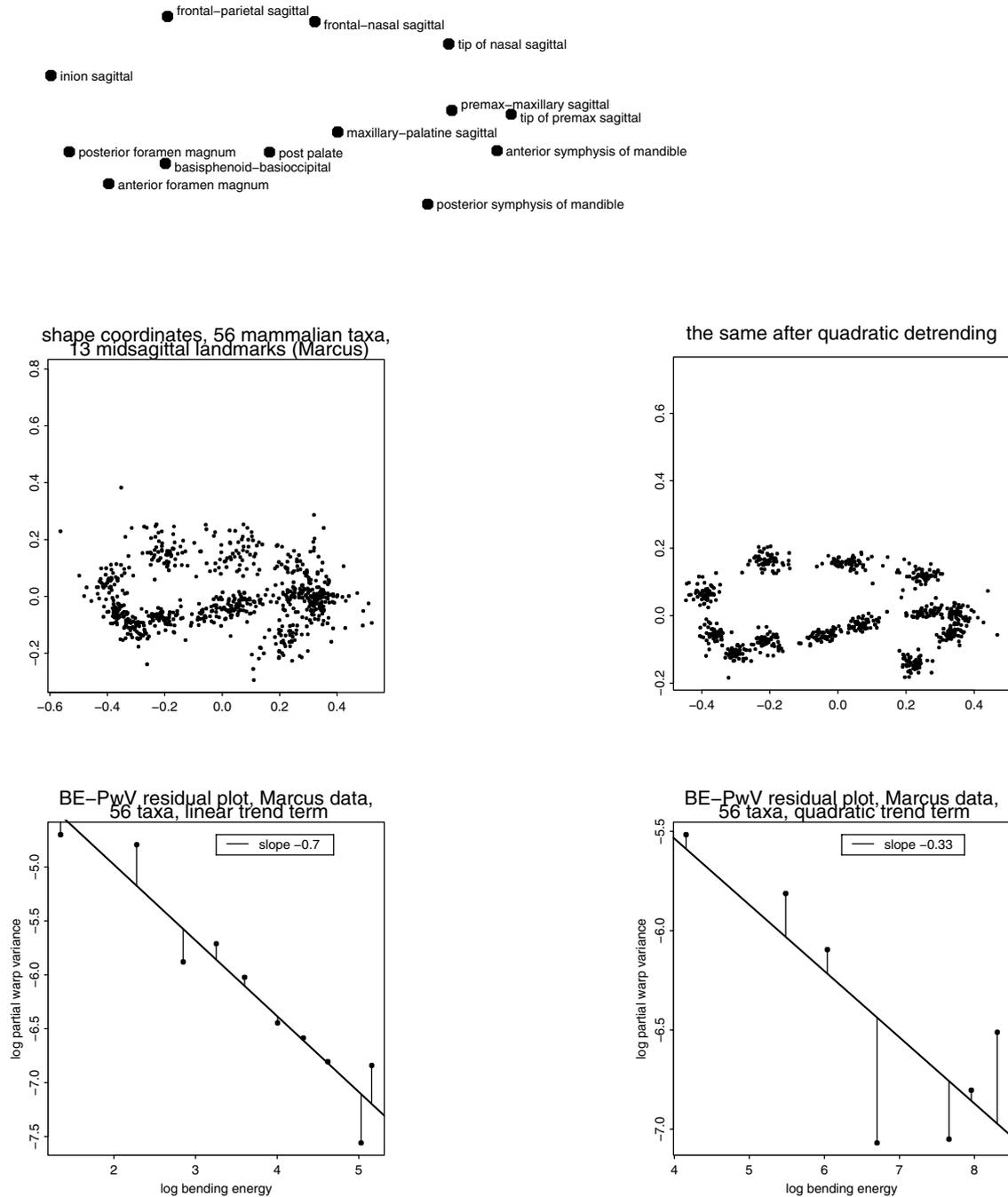
Perhaps the gradient diagrammed in the central column of Fig. 9 corresponds to exo-devo findings at earlier stages of morphogenesis in this clade. Or, instead, this pattern might prove an aspect of mammalian variation exemplifying Elsasser’s “creative selection,” the phylogeny’s choice of aspects that could well have been otherwise. Either way, this

approach to close scrutiny of the BE–PwV plot provides a much more sturdy bridge to theoretical biology of this compartment than ordinary Procrustes analysis (Marcus et al. 2000) could manage. The residual at middle right in Fig. 7 is factored further in Bookstein (2017b); the *Ondatra* grid emerges there as one specialized varimax factor, the dimension of contrast from tapir to hare as another.

**Example 3** *Net roughness*. Here the metric is just the simplest sum-of-squares representation extractable from the partial warp analysis: the sum of the variances of partial warps 1 through  $k$ , where  $k$  is a descriptive parameter capped at the total count  $p - 3$  of those warps. The corresponding explanations would likely incorporate actual biophysical arguments aligned with such mechanisms as wear (Bravo Morante et al. 2018) or turbulence (Currie 2018). Figure 11 is an example from the latter project, the main concern of which is the clinical problem of nasal septal surgery (correction of “deviated septum”). The task was to operationalize today’s most common clinical classification, the seven “Mladina types” (Mladina 1987), by an objective analysis of the form of the nasal septum as represented in computed tomography.

From a sample of septal forms for 49 clinically diagnosable cases of deviated septum, digitized from their CT scans using the 132-point two-sided template at upper left in the figure, the variances of the partial warps’ mediolateral coordinates only (with respect to a posterior cranial midline) were extracted and two BE–PwV plots produced, left and right, as shown. Both slopes are near  $-1$ . The obvious suggestion of self-similarity after the first two warps permits us to just ignore the subsequent details of these plots, instead summing over all the separate contributions (because at a slope near  $-1$  they don’t *deserve* any individual interpretations) to arrive at a total BE per septum, a net deviation from the traditional atlas’s ideal of “flatness” (Bookstein 2017c). The lower left panel of Fig. 11 ordines this “net bending” against the seven Mladina classes. Such a diagram is not promising, but the physiology of the system suggests a *physical* scaling dimension corresponding to airflow turbulence. We explore a quantification, therefore, limited to just a finite number of these warps, and one of these lowpass versions, limited to the first 19 partial warps out of the available 73, resulted in the splendid discrimination set down in the panel at lower center. (The threshold was taken at the short vertical segments in the BE–PwV plots of the top row.) This criterion of relatively large-scale nonflatness segregates all five instances of the most serious clinical category along with four other instances that appear to have been misclassified beforehand into one of the milder categories.

Currie’s dissertation argues persuasively that the analysis here, which confirms Takahashi’s (1987) explicitly evolutionary theories of how the hominid nasal septum got to

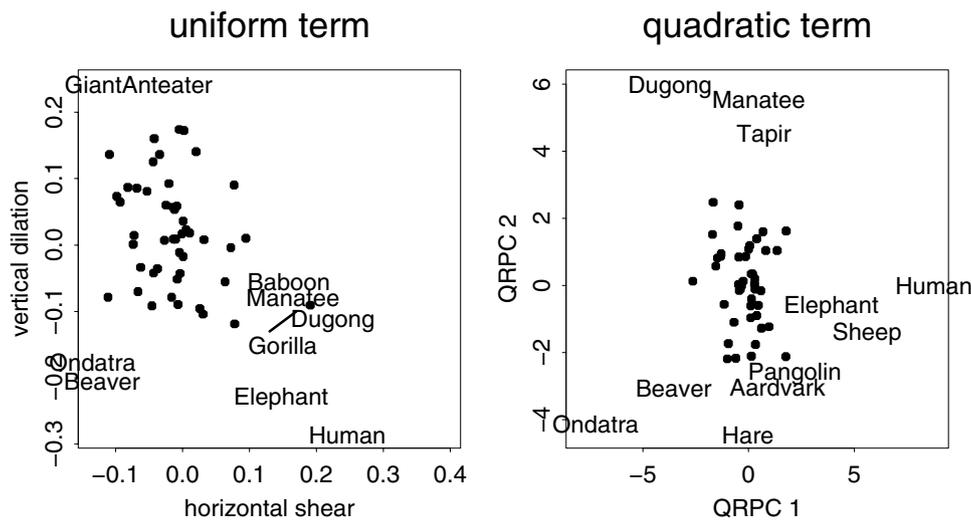


**Fig. 7** BE–PwV analysis of the 55-taxon version of the midsagittal landmarks from Marcus et al. (2000). Upper row: the 13 landmarks at their Procrustes average locations. Middle row: standard Procrustes shape coordinate scatter (left) and residuals from a quadratic detrend-

ing (right). The corresponding BE–PwV plots (lower row): at left, for the standard thin-plate spline, suggesting loglinearity of the plot; at right, the corresponding plot for a quadratic thin-plate spline, consistent with an absence of residual structure

be this way, should urgently be packaged for use by head-and-neck surgeons along with other craniofacial biologists, and should also be adapted for use by the biophysicists who are currently modeling nasopharyngeal airflow with the aim

of accounting for surgical failures in this context. A low-pass filter of BE corresponds to the clinical goal here, the quantification of Mladina's clinical hunches, far better than either Procrustes distance or total BE. This interpretation



**Fig. 8** Tentative division of the variability of the midsagittal mammalian 13-gon data into two subspaces by spatial scale. Scatterplot for the 55 conventional uniform terms (left). The same for the first two principal components of the corresponding 55 lists of coefficients of the quadratic trend term (right). Selected outlying points are printed as centered names. (The name “Gorilla” pertains to the indicated

point between Manatee and Dugong.) Note how frequently the specimens extreme on one of the plots are also extreme on the other—these two aspects of large-scale variation are deeply entangled—and that neither scatter is Gaussian. QRPC: principal component of the quadratic trend of the quadratic thin-plate spline

corresponds to a dynamics of *surface roughness* down to a specific scale without any homologous geometry. Takahashi refers to this as a morphogenesis via *crumpling*, nonreplacable across cases in its geometrical details but functionally significant nevertheless. Thirty years later, this means that clinical classification in terms of form would best be checked via the corresponding models of computational fluid dynamics. This example comes closest of those in this article to Przibram’s proposal that morphometric data come in the format of *Gitter* (grids).

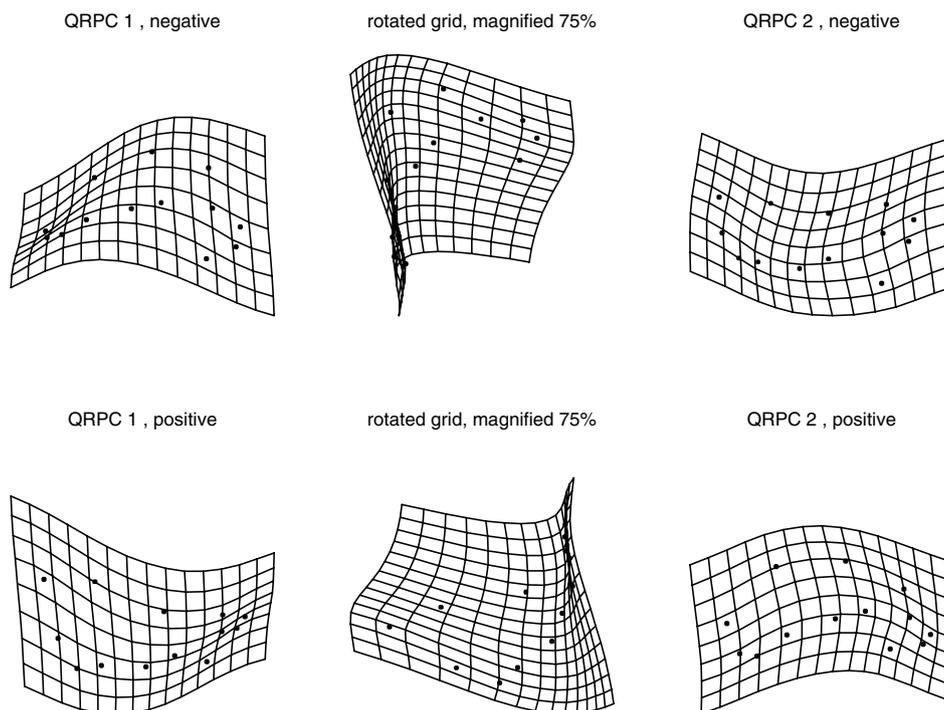
**Example 4** *Explanation by two muscles, one per partial warp.* In this analysis (Mardia et al. 2018), a decagon of landmarks around the human vermilion border (lip line), Fig. 12 upper left, was followed as 14 subjects who had been taught to keep their teeth in occlusion performed two calibrated smiles, open-lip and closed-lip, that lasted about one second each. (The long-term goal of this project is to apply this analysis to the smiles of patients with surgically corrected clefts.) These were recorded in frontal images at 60 Hz, submitted to a semiautomatic digitization, and analyzed as trajectories. As the middle row of the figure indicates, the only partial warps to show substantial loadings for the phenomenon under study were PW1 and (for the open-lip smile) PW0 (the uniform component), corresponding to the action of two distinct facial muscles (the zygomaticus and the orbicularis oris, respectively). The center row of the figure shows the variations of these components in comparison

to that of all the other partial warps, and the lower row shows how very strikingly different these two classes of trajectory are. Because these smiles were symmetric, we needed only one coordinate from each partial warp instead of the usual two. The measure of pattern dissimilarity I am using here is just ordinary Euclidean separation upon this display plane, and in the proffered explanation, each of the two dimensions represents the action of one muscle (the orbicularis) or a symmetric pair (the two zygomaticus). The correspondence of the two largest-scale signals to the actions of two individual muscles is consistent with Przibram’s (1923) approach to the relation between “organometry” and the deformation geometry of grids.

An interesting transitional case between uniform and lowpass analysis arises in the context of linear biophysical strain distributed over a topological sphere of orientations that varies at large scale. The strain might be actual isovolumic change, as with, for instance, the contraction of heart muscle within the thick shell of the ventricle (Varano et al. 2018), or perhaps an abstracted version constrained to align a principal strain with the normal direction to some anatomical bounding surface, as for the analysis of cerebral atrophy in Alzheimer’s disease (Tang et al. 2015).

**Loglinearity**

For shape changes that are more complex than any of these lowpass possibilities, surprisingly often we find a different



**Fig. 9** Visualization of the two dimensions of quadratic patterns by thin-plate spline as applied to the grand average mammalian 13-gon (Fig. 7 top). The uniform term of the Procrustes analysis has been discarded from all panels. Both components involve a bending of the craniocaudal axis of this form, but the first component incorporates a further oblique organization, a gradient along an axis from dentition through braincase, the derivative of which is aligned perpendicular to this axis. The first QRPC appears to precisely match the classic qualitative understanding of the hominization gradient (reduced maxilla,

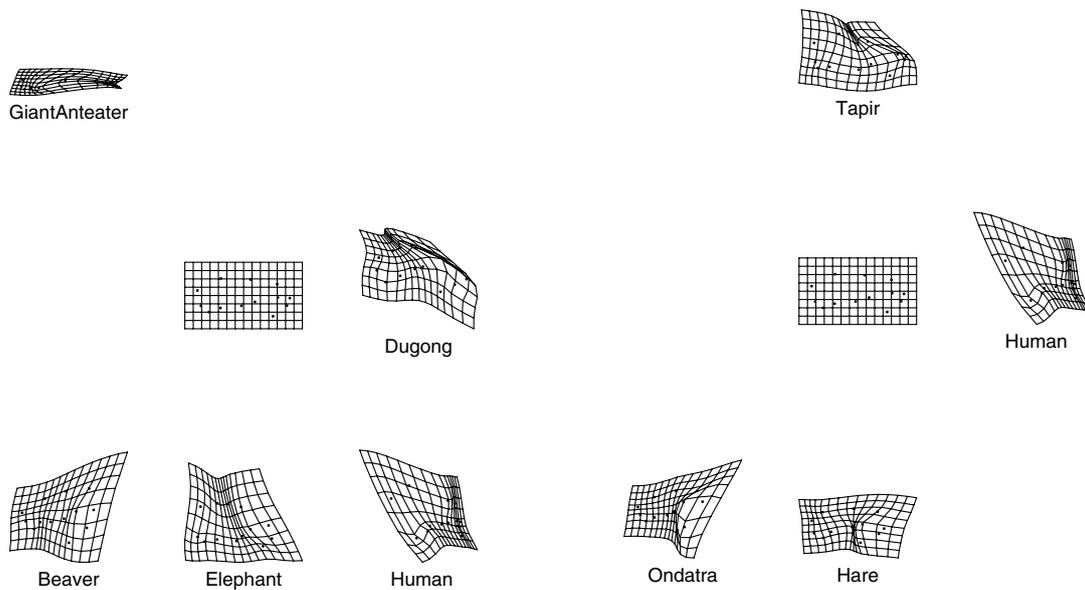
enlarged braincase). Splines for the horizontal axis at right in Fig. 8, drawn on a grid squared to the coordinate system at top in Fig. 7 (left). Extrapolation by 75%, drawn to a grid rotated 45° clockwise from the grid for the panels at left (center). This representation corresponds to the visualization as a “crease” in Bookstein (2000). Notice that the axis of the crease (here, vertical on the page) is perpendicular to the axis of the growth gradient here, which, after the rotation, now lies horizontal on the page. Grids, again in the coordinate system as above in Fig. 7, for the vertical axis in the plot at right in Fig. 8 (right)

parameterization to govern the BE–PwV curve: fit to a nearly straight line, sometimes accompanied by one or two substantial and meaningful residuals therefrom. The slope of the fit typically falls between  $-0.5$  and  $-1.5$ . (Intercepts are less important for analyses that, like this one, center on residuals.) The metric for the residuals is the topic of the alternative to factor analysis that I suggested recently in Bookstein (2017b). Before rotation, its geometry can be conveyed via a simplistic but not unrealistic example involving just six landmarks (hence a total of three partial warps) for which the specific BEs are in even geometrical progression (e.g., 1, 2, 4). In that case the sum of squared residuals from the loglinear regression of  $\log$  PwV (or anything else) on that sequence is proportional to just the square of the residual at the middle term, which is the deviation of the value there from the average of the values at the first and third. In Cartesian coordinates it has the formula  $(x - 2y + z)^2$ . This function is drawn in Fig. 13 as five sections of the corresponding solid distribution, each of which takes the form of a paraboloid on an axis corresponding to the intersection of

the sectioning plane with the plane  $y = (x + z)/2$  of perfect regression (summed squared residuals zero).

Figure 13 sketched the general case. The next two examples proffer two simpler approaches, one for which the explanations center on the slope parameter and the other for which that explanation combines with a focal phenomenon at much smaller scale.

**Example 5 Human social perception.** In this example, from Windhager et al. (2017), the explanatory goal deals explicitly with the slopes of the BE–PwV plots in the right column of the figure. The data comprise configurations of 71 landmarks and semilandmarks on frontal photographs of young women’s faces that were correlated to two biological measurements, body mass index (weight divided by height squared) and salivary cortisol, along with a “health rating.” The vectors submitted to the BE–PwV analysis are actually the regressions of these landmark locations on the three measurements, so that the slope of the BE–PwV plot is based on a regression over regression slopes themselves.



**Fig. 10** Combined visualization of the two trend subspaces by splines from the grand average 13-gon to a selection of specimens extreme on one or another axis in Fig. 8. A selection based on the uniform

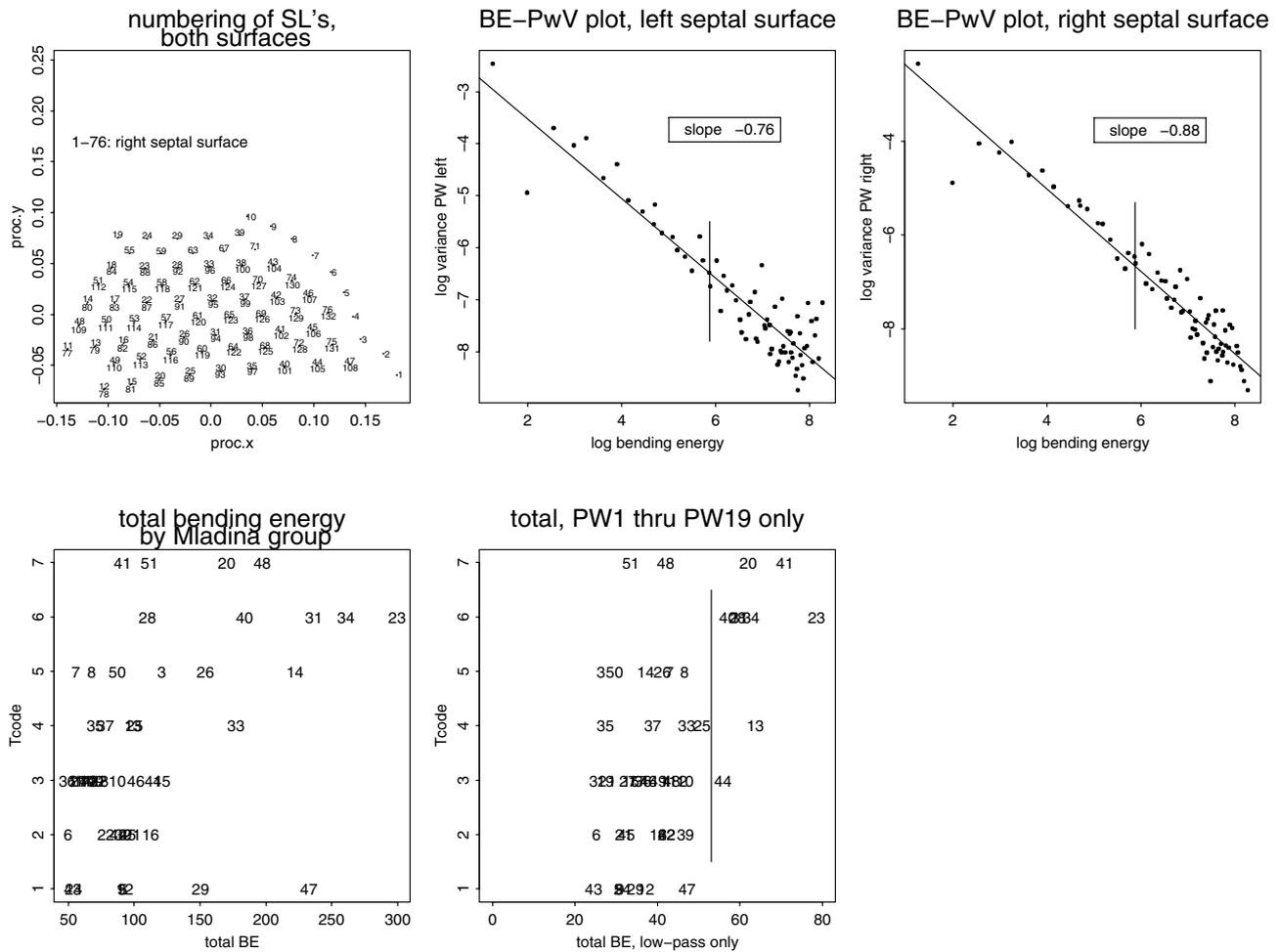
term (left). On the quadratic term (right). The *Ondatra*, an aquatic mammal, is better known as the muskrat. In these grids the uniform terms have been restored

Both regressions are displayed in Fig. 14: the slopes of the landmark shifts in the first two columns, the slope of the loglinear BE–PwV fit in the third column. These BE–PwV regression slopes differ significantly from top to bottom of the figure, but their numerical comparison is less effective than the simpler representations to their left, which segregate the contributions of just the largest three scales of partial warps. It is plain that the visible weight of red and blue (or thick black lines; see electronic version for color figure) decreases strikingly from upper left to lower left here (i.e., from the description of the association of BMI with shape to the association of perceived health with shape). The finding is reducible to a single sentence: “patterns of correlation of facial shape with physiological measurements are more integrated than patterns of correlation with ratings.” That was the title of the publication itself.

**Example 6** *The rodents again.* I return to the Vilmann rodent neurocranial octagons to show a combination of multiple explanations applying to the same data set, this time the growth trajectory of these animals over the full duration of the experiment, 7 days of age to 150. (The previous uniform finding, not shown here, bifurcated over this interval. While the early drop of aspect ratio of this octagon continued over growth, the shearing of the upper surface backward over the

period to 30 days reversed to a forward shear from 30 days on. For details, see Bookstein 2018, Sect. 5.6.)

Figure 15 compiles the diagrams most relevant for this summary. The first five panels simply show the scatters of the five partial warps one by one. Their variances decline from the first through the fourth but then rise again for the fifth. At lower left is the corresponding BE–PwV plot, indicating that those first four warps lie on a line of slope about  $-1.4$  whereas the fifth warp falls well above that line, with a deviation much greater than those of any of the warps preceding it. Such an analysis suggests that the net change from average age-7 form to average age-150 form, shown doubled at center right, can be seen as the superposed effect of three distinct processes: a component for uniform change of aspect ratio (the obvious widening in the net deformation at center right); a growth gradient at largest finite scale, the quadratic term displayed at lower center; and the specific residual for partial warp 5 (the only component that is relevant to the description of integration), graphed by itself at lower right. The second and third of these components are further submitted to a factor analysis in Bookstein (2017b). The import of this example is that one must sometimes combine analyses over multiple geometric scales in order to arrive at a coherent statement of a finding, a statement that might lead to a coherent suite of corresponding holistic explanations. The quadratic growth gradient here probably accounts for



**Fig. 11** Analysis of nasoseptal dysplasia by lowpass filtering. (After Currie 2018.) Template for digitizing the two septal surfaces after block registration on posterior cranial structures (upper left). Landmarks: 1, anterior nasal spine; 10, foramen caecum. The face is facing toward the right. BE–PwV plots for left and right septal surfaces (upper center, upper right). The short vertical segments show the threshold (the 19th partial warp) used for the plot at lower right. Total

BE (of the mediolateral coordinate only) against Mladina class (lower left). Printed numbers in this panel and the next are case numbers. Lowpass BE against Mladina class, affording a far better operationalization of the severity of the patients in Class 6 along with diverse detected misclassifications to the “less serious” categories (lower center)

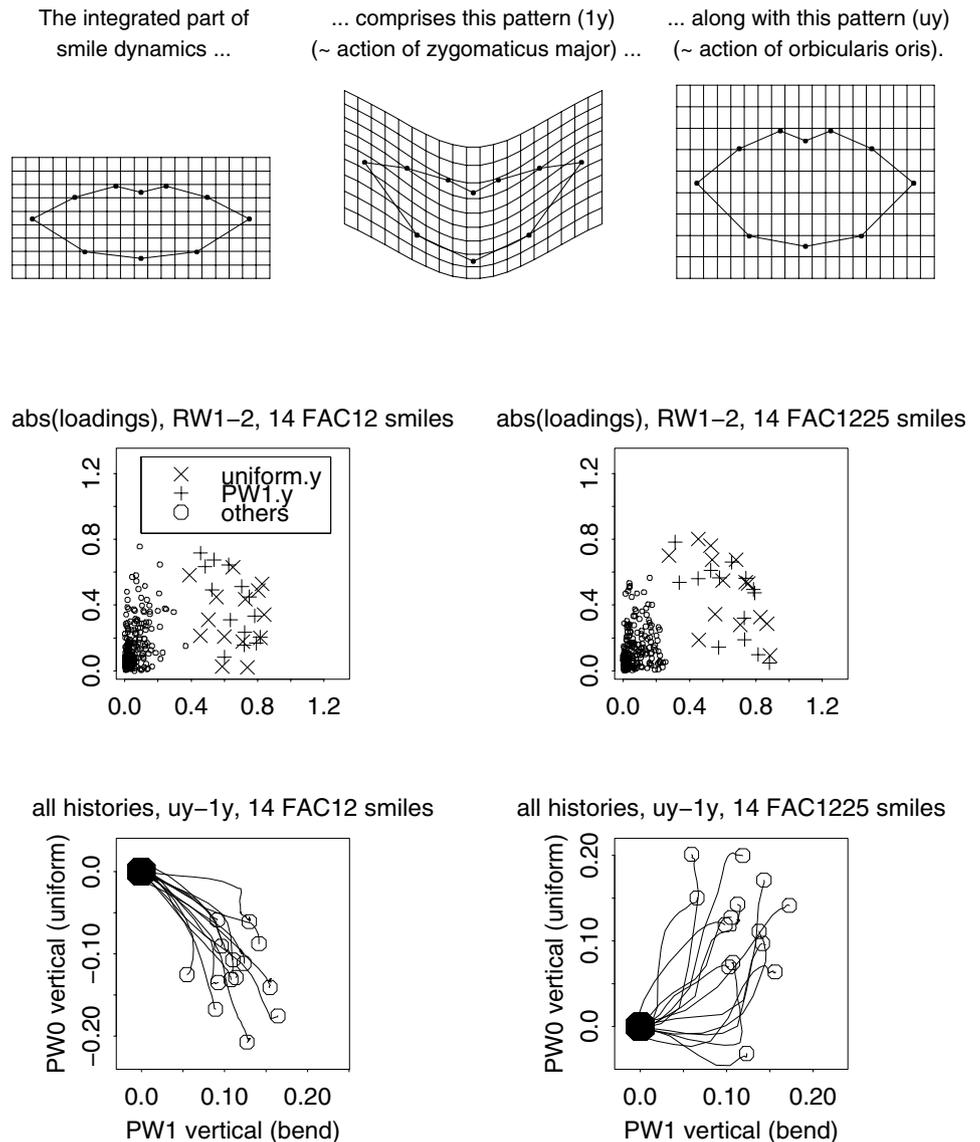
the highest point in the upper left panel of Fig. 5, implying individual differences in the implied rate of maturation over this sample of animals even at this young age.

**Flat BE–PwV Plots or Subplots**

In contrast to the self-similar models of slope near  $-1$ , flat BE–PwV plots are good examples of Elsasser’s “homogeneity,” the domain of microstates that *do* average out like thermodynamic quantities do. Explanations, if any, must be idiopathic—heterogeneous over instances (if clinical, lives; if evolutionary, autapomorphies and synapomorphies). At left

in Fig. 2 we already saw that the isotropic Mardia–Dryden model produces such a flat BE–PwV plot. Actually, it does so as a superposition of one such plot for each landmark coordinate, Fig. 16 upper left, with the corresponding BE–PwV plot for this single pair at lower left. The variance around this regression line is larger than if multiple such perturbations are summed in one single configuration of meaningless variation. Morphometricians once referred to such an explanation of a shape change, focused on the shift at just one landmark, as a “Pinocchio effect.” This article instead proposes referring to it as a *quasiwedgel*, meaning, a single isolated shift of exactly one landmark: an empirical (as opposed to mathematical, infinitesimal) version of the

**Fig. 12** Summary of the analysis of pairs of smiles for 14 subjects (Mardia et al. 2018). Top row: left, the landmark decagon; center and right, actions of the two muscles well-known to be responsible for the largest-scale aspects of the human smile. Of the ten points here, four (the points nearest the corners of the mouth along both borders) are semilandmarks, and the other six proper landmark points. Representation of the BE–PwV plot by scatters of partial warp loadings instead of just their variances (middle row). There is an obvious discrimination of the two types of smiles by the vertical element of the uniform component, without much contrast vis-à-vis action of zygomaticus (bottom row). Black dots: standardized starting configuration. Open circles: configurations at the moment of maximum smile, rendered as displacements of the corresponding starting forms. 1y: vertical loading of partial warp 1. uy: vertical loading (dilation) of the uniform component. FAC12: category 12 of Paul Ekman’s Facial Analysis System (Ekman et al. 2002), the “closed-lip smile,” FAC1225: combination of action 12 with action 25 of the system, corresponding to the open-lip smile

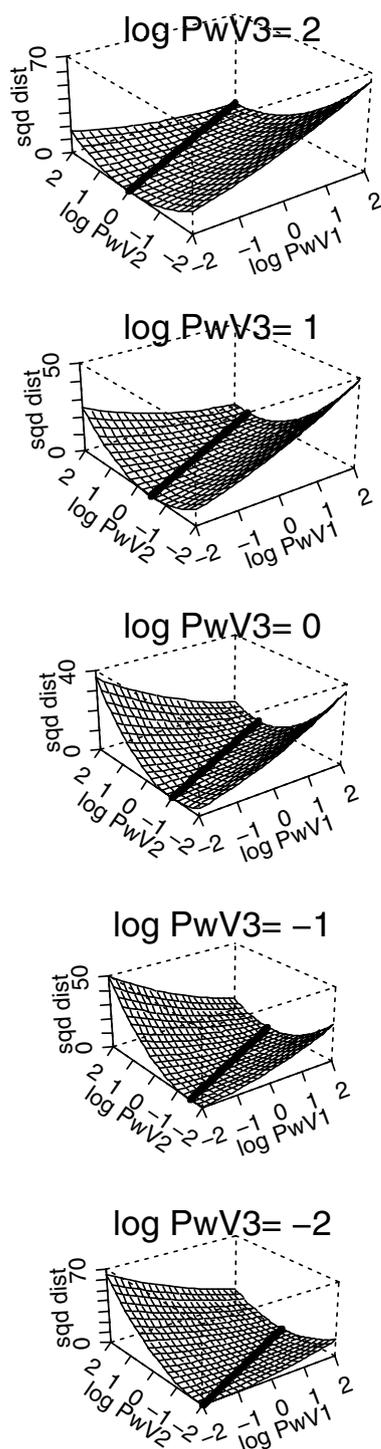


edges that were the subject of my 1993 paper with William D. K. Green. A quasiedgel like this could also correspond to a tightly focused volume change, such as a tumor, although if such a phenomenon is suspected on empirical grounds there are better ways to search than by a BE–PwV plot.

I suspect Elsasser might refer to the isotropic Procrustes model of Fig. 1 as one unique macrostate superimposing a myriad of equivalent microstates without additional biological structure. But such a circumstance (the absence of all additional structure), however inhomogeneous in its actual appearance (compare Fig. 1), would not meet the Hotchkiss aphorism’s requirement that its heterogeneity actually be ordered.

**Example 7** *Rohlf’s mosquito wings*. Nevertheless one can search the literature of morphometrics for data sets that partly show this degree of homogeneity. Jim Rohlf discovered one of these in 2015 and kindly permitted me to publish

on it in Bookstein (2016a) and again here. His data set comprises 18 landmarks on the wings of 127 species of mosquito spanning all the genera. The analysis is laid out in the center and right columns of this same Fig. 16. The Procrustes shape coordinate scatter, at upper center, is unremarkable, but the BE–PwV plot, lower center, has an unusual structure: a dominant first partial warp, then a run of eleven warps whose variances show no particular trend, and finally a sloping run-out that I suspect has something to do with the vortex sheet such a wing leaves behind its trailing edge in flight. (I would hazard a speculation that the wing’s net longitudinal stiffness probably matters for the efficiency of the insect’s flight, and the shape of the wing’s trailing edge is tightly canalized, but apparently not the detailed geometry of those venation branch points in the central region.) The dominance of that first warp variance, together with a uniform term (not shown), suggests a custom residualization by regressing out



**Fig. 13** Residual metric from the model of a linear BE–PwV. (What is being diagrammed here is the geometry of deviation from the model, not the loglinear model itself.) The figure shows one example of that change of analytic focus from summed squared log PwVs (i.e., properties of individual warps) to summed squared *residuals* from *loglinearity* of those variances (a plausible measure of net feature strength for deviations from self-similarity). The rendering here is of the value of the surface  $(x - 2y + z)^2$  for  $x$  equal to  $\log \text{PwV1}$  (log variance of the first partial warp),  $y$  equal to  $\log \text{PwV2}$ , and  $z$  equal to  $\log \text{PwV3}$ . The sections are taken at equally spaced levels of  $z$ . The heavy straight line in each frame is the locus of points on a perfect regression—points in the section also having  $x - 2y + z = 0$ , where the fit to this simplistic loglinear model for the BE–PwV plot would be perfect

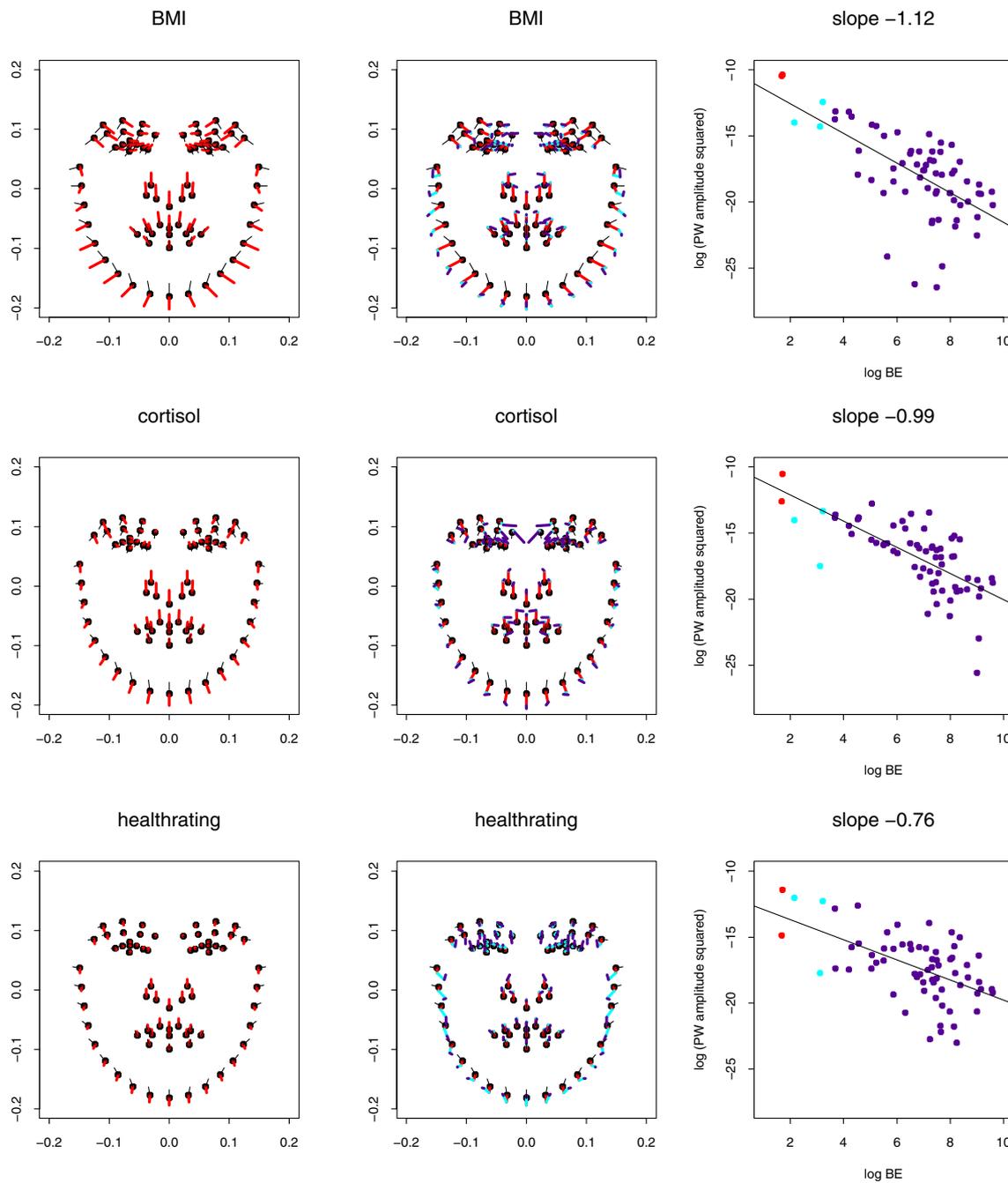
just those two pairs of shape dimensions. The result, upper right, indicates that indeed that run of eleven equivalent partial warps in the center panel could have arisen from isotropic (i.e., biologically meaningless) variation at the seven landmarks of greatest variance in this final panel. As the final figure panel shows, that variance is directional, much more along the wing’s axis than along the direction of flight.

The corresponding explanatory theory, at last, embraces the hypothesis of homogeneity that both Przibram and Elsasser would otherwise view so skeptically at this organismal level. Or, rather, there really *isn’t* any explanation of why these veins go where they do—no functional aspects within this central range of geometric scales. Waddington himself guessed that this might be true (see Waddington 1972a, his essay in his final symposium volume of 1972, especially his Figs. 5, 6). The *actual* location of these vein branch points from species to species is thus an excellent example of what Elsasser called “creative selection,” his principle 2. Przibram’s program presumably would find it to have no meaningful physiological correlates.

## Discussion

The development of models for organismal form-comparisons pioneered by D’Arcy Thompson (for a concise history, see Bookstein 1978) was cut short by the—in retrospect inappropriate—irruption of too superficial and imitative a multivariate analysis into biometrics in the 1970s and 1980s and then the twin distractions of “industrial biometrics” and the -omics beginning in the 1990s. (The present author contributed to the first of these detours to some extent.) I think it is now possible to repair this misdirection. The previous section reviewed a range of examples that, taken together, cover the full span of the dimension Elsasser was surveying, from homogeneity to individuation, by diverse biometric reporting styles all applying to a particular class of data about organismal form, namely, homologous landmark configurations.

The analysis here expands a bit on the distinctions I raised in Bookstein (2015). The typology there divided the domain of these plots into just three classes: the “integrated” plots, those with BE–PwV slopes negative and greater than 1.0 in absolute value; the “self-similar” plots having slope at or close to  $-1.0$ , such as the deflated Procrustes distribution itself at right in Fig. 2; and the “disintegrated” deformations with a slope between  $-1$  and  $0$ , such as that for the major mammalian taxa at lower left in Fig. 7. Figure 16 includes an example of a *real* data set of BE–PwV slope surprisingly close to zero: the quadratic-trend residuals of the Rohlf mosquito wing-vein data. This supplements not only the simulated example of an actual Procrustes data set, Fig. 2

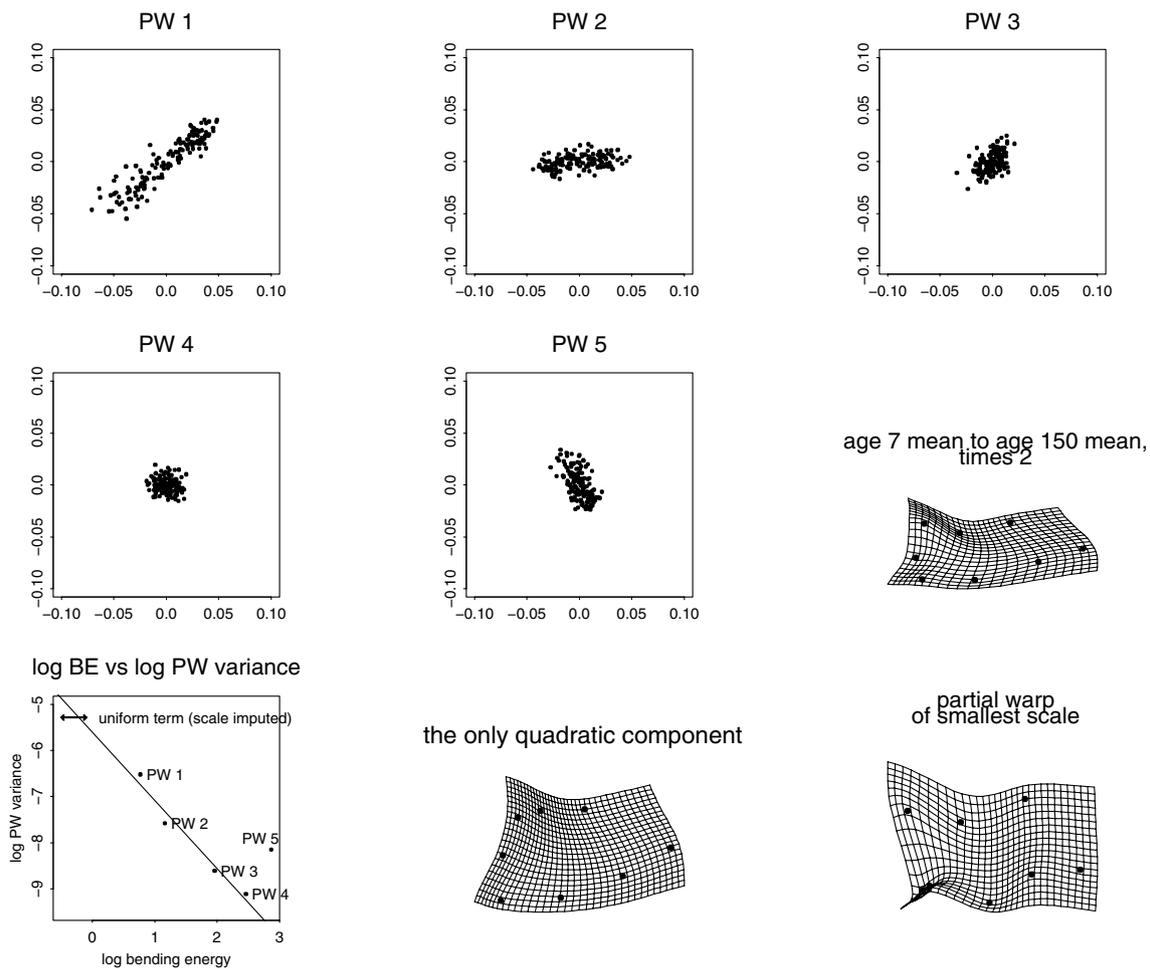


**Fig. 14** BE–PwV plots and their visualization for the Windhager et al. (2017) study of correlates of facial shape. Top row, for the conventional body mass index; middle row, for salivary cortisol; bottom row, for a “health rating” averaged over 101 raters of the same photographs. The pattern displayed is the standard way of regressing a shape on a predictor: the displacements of all the landmarks at once corresponding to their separate regressions on the named predictor. Left column, largest-scale variation: uniform component (thin black lines) and partial warps 1 and 2 (in red or thick black lines; color

figure in electronic version) plotted separately. In the middle column, the other partial warps, representing small-scale variation, are added (the next largest three in cyan (or thick black lines), the others in purple (omitted in grayscale version)). Right column, the standard BE–PwV plots. In the panel at upper right, no red dot (or big open circle) is missing; rather, the first two partial warps have nearly equal specific BEs and variances and hence print in indistinguishable locations. The slopes of the regression lines shown indicate the level of integration pattern by pattern

(left), but also the class of “flat” transformations exemplified at the left in Fig. 16 or at the lower right in Fig. 15, the shift of one landmark or interlandmark segment only. The

highly integrated transformations, with slope more negative than  $-1$ , are now not only those with a slope between  $-1$  and  $-2$ , such as the BE–PwV plot in Fig. 15, but also plots



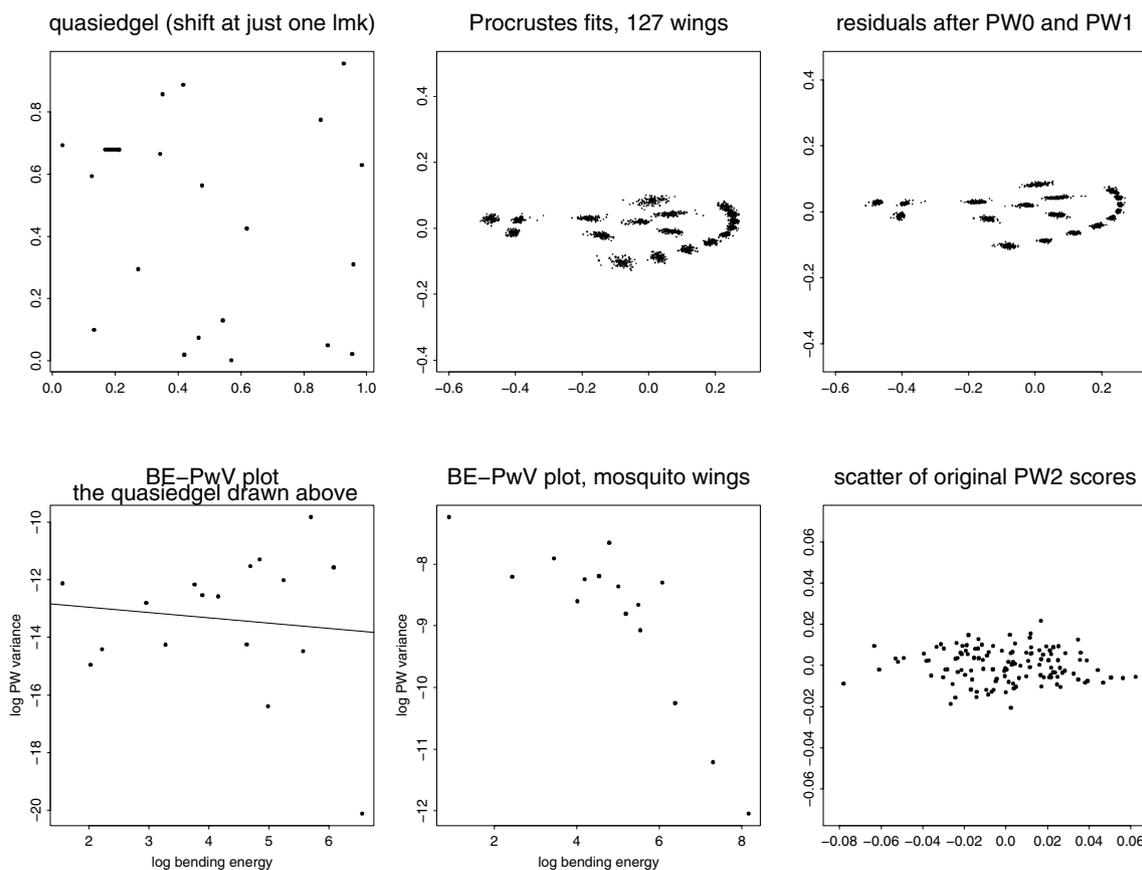
**Fig. 15** The Vilmann analysis, continued: full displays of all five partial warp scores for the complete data set of eight images per animal, along with the total warp from age 7 days to age 150 days (the grid at center right, doubled for legibility), the BE–PwV plot (lower left) and the two aspects of bending (the largest-scale quadratic term, lower center, and the most focal change, between Basion and Opisthion,

lower right) that, together with the obvious drop in relative height of this octagon, sum to the net deformation from age 7 to age 150 drawn in twofold extrapolation at center right. Axes for the partial warp plots are the standard axes assigned to the Procrustes coordinate plots, the axes for which the principal moments are horizontal and vertical

like that in Fig. 6 that, in effect, represent a regression that incorporates a step function downward: up to three large-scale partial warps of high variance (here, a growth gradient; elsewhere, adjustments of the sizes or relative positions of compartments) trailed by an indefinitely large number of meaningless components of lesser spatial scale at an incomparably smaller scale of variation. The self-similar deformations, those with slope approximately  $-1$ , remain resistant to explanatory speculations, for the same reason that patterns in random walks should not be interpreted featurewise (Feller 1957; Bookstein 1987, 2013b).

Perhaps the most important gap I have left in this preliminary report is the restriction of the data domain (except in Example 3) to landmark *points*, meaning, detached individual loci. Today’s geometric morphometricians, especially those concerned with 3-D data, are turning rapidly to focus

instead on *semilandmarks*, “sliding landmarks” that are not homologous from form to form as points but that instead represent a choice of points from curves (or, in 3-D, from surfaces) by optimizing some figure of merit, often BE itself. For that reason, the variance of BE and its spectral components is reduced below what it would have been if these points on curves had been proper landmarks instead (for the distinction, see Bookstein 2018, Sect. 5.1, and for the sliding formulas, Sect. 5.5.4 there). For curves in 2-D, the more abstract mathematical approach of the *square-root velocity representation* (Zhang et al. 2015; Srivastava and Klassen 2016) is interesting for general applications but unnecessary, in my view, for the carefully controlled samples of limited variance that arise in studies of whole organisms aimed at explanation rather than, say, classification.



**Fig. 16** BE–PwV plots can be horizontal. A simulation: arbitrary shift of one single arbitrary “landmark” (left). Above, the shift; below, the corresponding BE–PwV plot. Note the slope of zero for the BE–PwV regression, albeit with a substantial scatter of PwVs around it. Rohlf’s data set of 18 landmarks on 127 mosquito wings (center, right). Procrustes shape coordinates of the full data set (above

center). The corresponding BE–PwV plot (below center). Substitute shape coordinate scatterplot after the residualization by four more terms as described in the text (above right). Original scatter on partial warp 2, indicating strong constraint along the direction of airflow over the wing (vertical in this plot) in spite of the obvious horizontal homogeneity of perturbations (below right)

What could Przibram have meant by those “Gitter” (grids), his preferred (because experimentally validated) alternative to Thompson’s “Cartesian transformations”? Przibram apparently meant a persistent pattern of material particles that neither appear nor disappear over a study, but only *move*, thereby representing how the material in which they are embedded is deforming in some way analogous to a physical strain, such as would be characterized by the derived pattern of relative changes of interpoint distance at small and large scales. (This applies to landmarks just as pertinently as to the anonymous lattice intersections Przibram had in mind.) But such data sets are *very* difficult to compile. The first practical context seems to have been clinical cardiology, where the technology of myocardial *tagging* was developed starting in 1988 (Zerhouni et al. 1988; for a review see Jeung et al. 2012). In a gentle irony, the physics of this technologically clever imaging echoes D’Arcy Thompson in invoking not a grid of *points*, which

is what Przibram actually imagined, but instead a grid of *lines* projected onto the magnetic tissue structure of a beating heart, where it lasts long enough to be imaged (by magnetic resonance) repeatedly over the contractile phase of a single heartbeat, whereupon the myocardial strains can be computed directly. In other words, the grid is *manufactured experimentally* in order to be innocuously embedded in the tissue whose deformation it is intended to describe.

A corresponding analysis of biological processes at longer time scales seems at present to be restricted to interfaces that can be characterized by their inherent markings (e.g., the surfaces of embryos, following the pioneering imaging of Jacobson and Gordon 1976). Applications to solid tissues, understandably, are lagging, but see the work of, for example, Mitteroecker and his colleagues (Mayer et al. 2014). Any such tagged data set would be susceptible to the manipulations of the BE–PwV plot; the limiting factor here is not the silicon of the computational approach but the sheer

technical demands of producing the data. A contemporary Przibram would probably also take into account the concept of inherency embedded in contemporary evo-devo science (see Newman 2018 and in particular footnote 5 *supra*): the most efficient approach to selectable quantifications requires a prior awareness of these inherent interfaces. To measure their geometry, one must already have fixed their topology. The closest this article manages to come to such a Przibram-pleasing data structure is the system of semilandmarks for the nasal septum exploited in Example 3. That data collection, though computer-assisted, nevertheless required many hundreds of hours of graduate and postdoctoral student time.

A more pragmatic series of caveats is also appropriate here. Samples of organisms to which these considerations apply will need to be “sufficiently similar”—limited in their geometric range. The class of Vertebrata *sensu lato* is much too wide; Mammalia (Marcus et al. 2000) probably likewise too wide, even with the cetaceans excluded (although I have attempted exactly such an analysis here—Figs. 7, 8, 9, 10—by substantially altering the definition of Procrustes shape coordinates). Primates? Higher primates? Anthropoids? One single animal or genetically homogeneous family followed through time? See the brief discussion in Bookstein (2018); but this notion of “similarity” deserves a much more probing deconstruction than I have space for here. Also, I am aware that several thrusts of modern bioscience are excluded here: not just the -omics, which center at the molecular level, but also “industrial biometrics,” cell cycle analysis, and most of the domain of macroevolution. The subset of bioscientific styles to which my suggestion applies is a small fraction of the full range covered in Gerard’s introduction to the volume he edited. Nevertheless it is my hope that it might serve as a useful “existence proof” regarding a feasible bridge between morphometrics and theoretical biology—the possible import of these biometric advances for the broader range of quantitative biosciences having a spatial or geometrical component.

A few “quick and dirty” accommodations might be in order. The practitioner should not mix proper landmarks with semilandmarks; that was the case, for example, in the callosal midcurve example of Bookstein (2015), which found the effect of prenatal alcohol exposure on this structure to be self-similar and therefore, so to speak, free of anatomically explicable features. Or, in the course of computing the variance of any single partial warp score, one might simply discard the observed variance of any semilandmark along the tangent to its curve, as I have done for the semilandmark-by-semilandmark variance displays in Fig. 9 of Bookstein et al. (2001), the paper that first announced the focus of the fetal alcohol signal at one specific location along the border between isthmus and splenium (see Bookstein 2019). Finally, since the maximum effect of suppressing one dimension out of two is  $\log 2 = 0.693$  ( $\log 3 = 1.1$  for 3D

data), one might simply ignore the problem whenever the vertical profile of the BE–PwV plot is of greater amplitude than this. The contrast of variance between partial warp 1 and the average of PW’s 2 through 11 in Fig. 16 (upper center) is substantially greater than this threshold, for example, as is the contrast between the variances of PW’s 0 and 1 and the others in the middle row of Fig. 12. The comparisons of slopes in Fig. 14 are likewise unaffected by this complication.

It is no accident that the rhetoric for reporting analyses of this type centers on diagrams, not decimal numbers. The subject, after all, is a spatial one, not a matter of genetic encoding or instrument readings of inner physiological states. It is diagrams, not numbers, that convey dissimilarity of data-derived patterns of shape from their theory-derived ideal types. (The fundamental metaphor here, dissimilarity as Euclidean distance, is built deep into the statistics of complex organized systems.) For us to interpret any composite expression—to turn its arithmetic back into biological understanding—we must be able to picture it in a geometry that translates into statements biologically meaningful at the level of the organism.

That geometry may be theoretical (“a priori”), as with the BE–PwV plot exploited so intensively here. Thompson, it turns out, had the right idea after all, but only for one end of our spectrum, and Przibram the right idea for the other end. Geometric morphometrics, in its premature adoption of the Procrustes tool kit, in effect started at the wrong end of the spectrum, the “homogeneous” end of unrelated quasiedgels, rather than the level of integrated phenomena that Thompson had always intended to couple directly to the level of explanation *per se*. It has taken many decades for quantification to back out of the culs-de-sac of principal components and significance levels—to catch up and produce a classification of transformation grids (or, better, whole *samples* of transformation grids) for which those susceptible to Thompson’s preferred type of explanation emerge as a detectable subtype. This point is tenable regardless of the type of contrast borne in the deformation, whether growth, allometry, sexual dimorphism, ecophenotypy, or taxonomic contrast at any level for which the underlying model of deformation remains itself plausible.

What we have arrived at via the BE–PwV spectrum thus is actually the nucleus of a genuinely multiscale biometrics compatible with theoretical biology, just what the combination of Elsasser’s principles and Przibram’s metaphor of the space lattice calls for. Holistic memory of the “ordered heterogeneity” is what characterizes living form—to supply useful explanations, instead of mere technology, any biometrics will need to recognize the distinctions implied by this summary of Elsasser’s approach. Even Thompson himself had (inadvertently) hinted as much, when on page 274 of the 1961 edition, he notes his assumption that, for explanations

of form by force to obtain, “the form of the entire structure under investigation should be found to vary in a more or less uniform manner, after the fashion of an approximately homogeneous and isotropic body, ...and that our structure vary in its entirety, or at least that ‘independent variants’ should be relatively few.”

We now know that even if such descriptors obtain, they are only partial descriptors, conveying only the large-scale aspects of deformations and thereby hinting at the homogeneous causes that might be expressed thereby. Such claims would have to be investigated, then, by the rigorous experimental methods of Przibram’s Vivarium put in place while Elsasser himself was still a boy: the explicit investigation of how these finitely measurable quantities vary from point to point, organ to organ, by developmental conditions like temperature or nutrients. In their physiology these might reduce to biophysics; but in their variation from point to point of a tissue they would reify the much earlier concerns of Przibram about conditions that could only be studied experimentally, not theoretically.

I never met Elsasser, who died in 1991, but I like to think that he would have been delighted with this new applied context for his deepest ideas. And Przibram died at the hands of the Nazis in 1944. The *Aufbau* from which I quoted in Sect. 2 has never been translated into English, and so, unlike Waddington’s symposium, it played no role in the development of my own thoughts. But the essays on theoretical biology of these two pioneers, along with the contributors to Waddington’s series, deserve the attention of every sophisticated quantitative biologist right now, as we work to inject biological theory—the implications of being alive—back into the current turn toward data science that so greatly increases our collective information load without any commensurate increase in explanatory power for the patterns that actually describe the natural variation of organisms. When Darwin wrote of the grandeur in “this view of life,” it was these aspects of the natural world he was calling to our attention, not just the levels at which data come in the discrete elements our instruments are trained to call molecules. But valid organismal explanations do not share much of their explanatory rhetoric with molecules.

**Acknowledgements** Open access funding provided by the University of Vienna. Many of the ideas here were brainstormed or polished in conversations with Philipp Mitteroecker, University of Vienna. Gerd Müller, University of Vienna and KLI, directed my attention to Hans Przibram’s seminal work of the 1920s and enthusiastically helped to shape my discussion of the Vivarium Project. Norman MacLeod, of the Natural History Museum, London, commented trenchantly on many points of an earlier draft; and Harald Martens, Norwegian University of Science and Technology, supplied a crucial insight about calibration. This version of the essay has benefited greatly from the efforts of two anonymous reviewers of a previous version. My thinking about the contemporary relevance of D’Arcy Thompson’s work arose in the course of discussions with Kanti Mardia, University of Leeds, as did

my understanding of the self-similarity of the deflated Procrustes distribution. I thank Peter Currie for permission to use excerpts from his dissertation of 2018 and Guillermo Bravo Morante for access to the abstract he submitted to ESHE for its 2018 Meeting. Several of the examples were computed with partial support from National Science Foundation Grant DEB1019583, “Integrating quantitative characters, discrete characters, fossils, and morphometrics,” to Joseph Felsenstein and me, 2010–2013.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Anderson DR (2008) Model based inference in the life sciences: a primer on evidence. Springer, New York
- Anson BJ (1963) An atlas of human anatomy, 2nd edn. W. B. Saunders Company, Philadelphia
- Berry W (2000) Life is a miracle: an essay against modern superstition. Counterpoint, Washington, DC
- Blackith RE, Reyment RA (1971) Multivariate morphometrics. Academic, London
- Bookstein FL (1978) The measurement of biological shape and shape change. Lecture notes in biomathematics, vol 24. Springer, Berlin
- Bookstein FL (1981) Coordinate systems and morphogenesis. In: Connelly TG, Brinkley L, Carlson B (eds) Morphogenesis and pattern formation. Raven Press, New York, pp 265–282
- Bookstein FL (1985) Transformations of quadrilaterals, tensor fields, and morphogenesis. In: Antonelli PL (ed) Mathematical essays on growth and the emergence of form. University of Alberta Press, Edmonton, pp 221–265
- Bookstein FL (1987) Random walk and the existence of evolutionary rates. *Paleobiology* 13:446–464
- Bookstein FL (1989) Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans Pattern Anal Mach Intell* 11:567–585
- Bookstein FL (1991) Morphometric tools for landmark data: geometry and biology. Cambridge University Press, New York
- Bookstein FL, Green WDK (1993) A feature space for edgels in images with landmarks. *J Math Imaging Vis* 3:231–261
- Bookstein FL (2000) Creases as local features of deformation grids. *Med Image Anal* 4:93–110
- Bookstein FL, Sampson PD, Streissguth AP, Connor PD (2001) Geometric morphometrics of corpus callosum and subcortical structures in the fetal-alcohol-affected brain. *Teratology* 64:4–32
- Bookstein FL, Mardia KV (2003) The five components of directional asymmetry. In: Aykroyd RG, Mardia KV, Langdon MJ (eds) Stochastic geometry, biological structure, and images. LASR2003 proceedings. Leeds University Press, Leeds, pp 35–40
- Bookstein FL (2004) After landmarks. In: Slice DE (ed) Modern morphometrics in physical anthropology. Kluwer Academic Publishers, New York, pp 49–71
- Bookstein FL (2013a) Allometry for the twenty-first century. *Biol Theory* 7:10–25
- Bookstein FL (2013b) Random walk as a null model for geometric morphometrics of fossil series. *Paleobiology* 39:52–74
- Bookstein FL (2014) Measuring and reasoning: numerical inference in the sciences. Cambridge University Press, New York

- Bookstein FL (2015) Integration, disintegration, and self-similarity: characterizing the scales of shape variation in landmark data. *Evol Biol* 42:395–426
- Bookstein FL (2016a) The inappropriate symmetries of multivariate statistical analysis in geometric morphometrics. *Evol Biol* 43:277–313
- Bookstein FL (2016b) Reconsidering, “The inappropriateness of conventional cephalometrics”. *Am J Orthod* 149:784–797
- Bookstein FL (2017a) A newly noticed formula enforces fundamental limits on geometric morphometric analyses. *Evol Biol* 44:522–541
- Bookstein FL (2017b) A method for factor analysis of shape coordinates. *Am J Phys Anthropol* 64:221–245
- Bookstein FL (2017c) Review of Thieme’s atlas of anatomy, third edition, Latin nomenclature. *J Anat.* <https://doi.org/10.1111/joa.12684>
- Bookstein FL (2018) A course of morphometrics for biologists. Cambridge University Press, New York
- Bookstein FL (2019) The mathematician in criminal court (in preparation)
- Bravo Morante G, Bookstein FL, Schaefer K, Slice DE, Aguilera A, Botella López M (2018) Predicting age at death from the shape of the human pubic symphysis by bandpass filtering of bending energy. Submitted abstract. ESHE
- Cardini A, Rohlf FJ, Bookstein FL et al (in preparation)
- Clark A (2016) Surfing uncertainty: prediction, action and the embodied mind. Oxford University Press, Oxford
- Coen DR (2006) Living precisely in fin-de-siècle Vienna. *J Hist Biol* 39:493–523
- Cressie N, Wikle CK (2011) Statistics for spatio-temporal data. John, Hoboken
- Currie P (2018) From geometric morphometrics to clinical classification: using bending energy to assess the functional significance of nasoseptal deformity in a nonsurgical Austrian sample. Doctoral Dissertation, Department of Anthropology, University of Vienna
- Ekman P, Friesen WV, Hager JC (2002) Facial action coding system: the manual (CD ROM). A Human Face, Inc., Salt Lake City
- Elsasser WM (1970) The role of individuality in biological theory. In: Waddington CH (ed) Towards a theoretical biology. Drafts, vol 3. Edinburgh University Press, Edinburgh, pp 137–166
- Elsasser WM (1975) The chief abstractions of biology. North-Holland, Amsterdam
- Elsasser WM (1998) Reflections on a theory of organisms, 1987. Republished posthumously with new introductions by Harry Rubin and Frederick Seitz. Johns Hopkins University Press, Baltimore
- Feller W (1957) An introduction to probability theory and its applications, vol 1, 2nd edn. Wiley, New York
- Fisher NI, Lewis T, Embleton BJJ (1993) Statistical analysis of spherical data. Cambridge University Press, Cambridge
- Friston K (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138
- Garstens MA (1969) Statistical mechanics and theoretical biology. In: Waddington CH (ed) Towards a theoretical biology. Sketches, vol 2. Edinburgh University Press, Edinburgh, pp 285–292
- Gatherer D (2008) Finite universe of discourse: the systems biology of Walter Elsasser (1904–1991). *Open Biol* 1:1:9–20
- Gavrilets S (2004) Fitness landscapes and the origin of species. Princeton University Press, Princeton
- Gerard RW (ed) (1958) Concepts of biology. Publication 560. National Academy of Sciences, National Research Council
- Grenander U (1994) General pattern theory. Oxford University Press, Oxford
- Grenander U, Miller MI (2007) Pattern theory: from representation to inference. Oxford University Press, Oxford
- Hilbert D, Cohn-Vossen S (1952) Anschauliche Geometrie. Springer, Berlin 1932. Translated as Geometry and the imagination. Chelsea, New York
- Huxley J (1932) Patterns of relative growth. Dial Press, New York
- Jacobson AG, Gordon R (1976) Changes in the shape of the developing vertebrate nervous system analyzed experimentally, mathematically, and by computer simulation. *J Exp Zool* 197:191–246
- Jaynes ET, Bretthorst GL (2003) Probability theory: the logic of science. Cambridge University Press, New York
- Jeung MJ, Germain P, Croisille P, El-ghannudi S, Roy C, Gangi A (2012) Myocardial tagging with MR imaging: overview of normal and pathologic findings. *RadioGraphics* 32:1381–1398
- Jolicoeur P, Mosimann JE (1960) Size and shape variation in the painted turtle. A principal component analysis. *Growth* 24:339–354
- Jolliffe IT (2002) Principal component analysis, 2nd edn. Springer, New York
- Kent JT, Mardia KV (1994) The link between kriging and thin-plate splines. In: Kelly FP (ed) Probability, statistics, and optimization. Wiley, New York, pp 325–339
- Koch GS, Link RF (1980) Statistical analysis of geological data. Dover Publications, New York
- Marcus LF, Hingst-Zaher E, Zaher H (2000) Application of landmark morphometrics to skulls representing the orders of living mammals. *Hystrix* 11:27–47
- Mardia KV, Kent JT, Bibby J (1979) Multivariate analysis. Academic, London
- Mardia KV, Bookstein FL, Khambay BS, Kent JT (2018) Deformations and smile: 100 years of D’Arcy Thompson’s on growth and form. *Significance*, June 2018, 20–25
- Martens H, Næs T (1989) Multivariate calibration. Wiley, Chichester
- Mayer C, Metscher BD, Müller G, Mitteroecker PM (2014) Studying developmental variation with geometric morphometric image analysis (GMIA). *PLoS ONE* 9(12):e115076
- Medawar PB (1958) Postscript. In: Thompson RDA (ed) D’Arcy Wentworth Thompson. Oxford University Press, Oxford
- 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
- Mladina R (1987) The role of maxillary morphology in the development of pathological septal deformities. *Rhinology* 25:199–205
- Mosimann JE (1970) Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. *J Am Stat Assoc* 65:930–945
- Müller GB (ed) (2017) Vivarium: experimental, quantitative, and theoretical biology at Vienna’s Biologische Versuchsanstalt. Vienna series in theoretical biology. MIT Press, Cambridge
- Newman SA (2018) Inherency. In: Nuño de la Rosa L, Müller GB (eds) Evolutionary developmental biology. Springer, New York
- Pearl J, Mackenzie D (2018) The book of why. Basic Books, New York
- Pennec X, Fillard P, Ayache N (2016) A Riemannian framework for tensor computing. *Int J Comput Vis* 66:41–66
- Platt JR (1964) Strong inference. *Science* 146:347–353
- Przibram H (1922) Form und Formel im Tierreiche. Beiträge zu einer quantitativen Biologie, I–XX. Franz Deuticke, Leipzig
- Przibram H (1923) Aufbau mathematischer Biologie. Abhandlungen zur theoretischen Biologie, vol 18. Verlag von Gebrüder Borntraeger, Berlin
- Rohlf FJ, Sokal RR (1967) Taxonomic structure from randomly and systematically scanned biological images. *Syst Zool* 16:246–260
- Seitz F (1993) The science matrix: the journey, travails, triumphs. Springer, New York
- Senck S, Coquerelle M (2015) Morphological integration and variation in facial orientation in *Pongo pygmaeus pygmaeus*: a geometric morphometric approach via partial least squares. *Int J Primatol* 36:489–512

- Sneath PHA (1967) Trend-surface analysis of transformation grids. *J Zool Lond* 151:65–122
- Spirtes P, Glymour C, Scheines R (1993) Causation, prediction, and search. Springer, New York
- Srivastava A, Klassen EP (2016) Functional and shape data analysis. Springer, New York
- Tabony J (2006) Historical and conceptual background of self-organization by reactive processes. *Biol Cell* 98:589–602
- Takahashi R (1987) The formation of the nasal septum and the etiology of septal deformity: the concept of evolutionary paradox. *Acta Otolaryngol Suppl* 443:1–155
- Tang X, Holland D, Dale AM, Younes L, Miller MI, The Alzheimer's Disease Neuroimaging Initiative (2015) The diffeomorphometry of regional shape change rates and its relevance to cognitive deterioration in mild cognitive impairment and Alzheimer's disease. *Hum Brain Mapp* 36:2093–2117
- Terzopoulos D (1983) Multilevel computational processes for visual surface reconstruction. *Comput Vis Graph Image Process* 24:52–96
- Thom R (1975) Structural stability and morphogenesis. Benjamin, Reading
- Thompson DAW (1961) On growth and form, abridged and edited by J. T. Bonner. Cambridge University Press, Cambridge
- Varano V, Piras P, Gabriele S, Teresi L, Nardinocchi P, Dryden IL et al (2018) The decomposition of deformation: new metrics to enhance shape analysis in medical imaging. *Med Image Anal* 46:35–56
- Waddington CH (1940) The genetic control of wing development in *Drosophila*. *J Genet* 41:75–139
- Waddington CH (ed) (1968) Towards a theoretical biology. Prolegomena, vol 1. University of Edinburgh Press, Edinburgh
- Waddington CH (ed) (1969) Towards a theoretical biology. Sketches, vol 2. University of Edinburgh Press, Edinburgh
- Waddington CH (ed) (1970) Towards a theoretical biology. Drafts, vol 3. University of Edinburgh Press, Edinburgh
- Waddington CH (1972a) Form and information. In: Waddington CH (ed) Towards a theoretical biology, Essays, vol 4. Edinburgh University Press, Edinburgh, pp 109–140
- Waddington CH (1972b) Epilogue. In: Waddington CH (ed) Towards a theoretical biology, vol 4. Edinburgh University Press, Edinburgh, pp 283–289
- Waddington CH (ed) (1972c) Towards a theoretical biology. Essays, vol 4. University of Edinburgh Press, Edinburgh
- Wahba G (1990) Spline models for observational data. CBMS-NSF regional conference series in applied mathematics, vol 59. Society for Industrial and Applied Mathematics, Philadelphia
- Weber GW, Bookstein FL (2011) Virtual anthropology: a guide to a new interdisciplinary field. Springer, Vienna
- Williams RJ (1956) Biochemical individuality: the basis for the genotrophic concept. Wiley, New York
- Windhager S, Bookstein FL, Millesi E, Wallner B, Schaefer K (2017) Patterns of correlation of facial shape with physiological measurements are more integrated than patterns of correlation with ratings. *Sci Rep* 7:45340
- Wittgenstein L (1922) Tractatus logico-philosophicus. Harcourt Brace, New York
- Zerhouni EA, Parish DM, Rogers WJ, Yang A, Shapiro EP (1988) Human heart: tagging with MR imaging—a method for noninvasive assessment of myocardial motion. *Radiology* 169:59–63
- Zhang A, Xie Q, Srivastava A (2015) Elastic registration and shape analysis of functional objects. In: Dryden IL, Kent JT (eds) Geometry-driven statistics. Wiley, Chichester, pp 218–235