The Concept of Morphospaces in Evolutionary and Developmental Biology: Mathematics and Metaphors

Philipp Mitteroecker
Department of Theoretical Biology
University of Vienna
Vienna, Austria
philipp.mitteroecker@univie.ac.at

Simon M. Huttegger
Department of Logic and Philosophy of Science
University of California, Irvine
Irvine, CA, USA
shuttegg@uci.edu

Abstract

Formal spaces have become commonplace conceptual and computational tools in a large array of scientific disciplines, including both the natural and the social sciences. Morphological spaces (morphospaces) are spaces describing and relating organismal phenotypes. They play a central role in morphometrics, the statistical description of biological forms, but also underlie the notion of adaptive landscapes that drives many theoretical considerations in evolutionary biology. We briefly review the topological and geometrical properties of the most common morphospaces in the biological literature. In contemporary geometric morphometrics, the notion of a morphospace is based on the Euclidean tangent space to Kendall's shape space, which is a Riemannian manifold. Many more classical morphospaces, such as Raup's space of coiled shells, lack these metric properties, e.g., due to incommensurably scaled variables, so that these morphospaces typically are affine vector spaces. Other notions of a morphospace, like Thomas and Reif's (1993) skeleton space, may not give rise to a quantitative measure of similarity at all. Such spaces can often be characterized in terms of topological or pretopological spaces.

The typical language of theoretical and evolutionary biology, comprising statements about the "distance" among phenotypes in an according space or about different "directions" of evolution, is not warranted for all types of morphospaces. Graphical visualizations of morphospaces or adaptive landscapes may tempt the reader to apply "Euclidean intuitions" to a morphospace, whatever its actual topology might be. We discuss the limits of metaphors such as the developmental hourglass and adaptive landscapes that ensue from the geometric properties of the underlying morphospace.

Keywords

adaptive landscapes, affine space, developmental hourglass, morphometrics, phenetic space, sequence space, shape space, skeleton space, theoretical morphology, topology

January 30, 2009; accepted September 5, 2009
Biological Theory 4(1) 2009, 54–67.

In his celebrated 1917 book, D'Arcy Thompson stressed the significance of construing the physical configuration of an organism as a mathematical object—represented in terms of Cartesian coordinates—for establishing morphology as a rigorous scientific discipline (Thompson 1917). However, already from the late 19th century on, initial developments in statistics such as correlation and regression were applied to morphological measurements, yet without explicitly considering the mathematical spaces induced by the measured data. For example, in 1888 Frances Galton introduced the correlation coefficient and applied it to a variety of measurements on humans such as body height, head length and width, or arm length (Galton 1888). He further produced a scatter plot of two variables, stature versus cubit length, and assessed the geometry of the data points across the two-dimensional space of the plot. In 1907 Galton invented a method to quantify facial shape that has later been termed as two-point shape coordinates (Galton 1907; see also Bookstein 1991). In an attempt in 1934 to compare a mummified head, claimed to be Oliver Cromwell's, to a range of portraits and busts of Cromwell, Galton's student Karl Pearson and the anthropologist G. M. Morant (Pearson and Morant 1934) computed all the ratios of distances among pairs of morphological measurement points and compared them across the heads. Without referring to the space of shape variables itself, Galton as well as Pearson and Morant already had the numbers in hand that would place the configurations as points in the underlying shape space.

Only after the foundations of multivariate statistics had been laid down in the first half of the 20th century did the study of mathematical spaces and analytic geometry become an integral part of statistics. Basically, two sorts of spaces are used in applied multivariate statistics to represent data for p measurements on n cases: The p-dimensional space spanned by the variables, often called Q-space, in which each of the n measured specimens is represented by a single point, and the n-dimensional R-space (or dual space) with p points representing the variables (e.g., Mardia et al. 1979; see also Figure 1).

The correlation between two variables can be represented in R-space as the cosine of the angle between the two vectors connecting the corresponding points (the two variables) with the origin. Thus, phenomena like morphological integration and modularity, which often are defined in terms of statistical associations among morphological variables, are best studied in this space (Mitteroecker and Bookstein 2007). The distance between points in Q-space, in contrast, reflects similarities among the measured objects. Classically, distances between points are computed as the Euclidean distance, thereby taking the measured variables as Cartesian coordinates of an underlying Euclidean vector space. However, modern literature on mathematical statistics comprises numerous (relatively abstract) spaces that do not exhibit a Euclidean structure, but are based, for example, on Riemannian geometry (e.g., Amari 1985; Mitteroecker and Bookstein 2009; see also below).

Morphological spaces, or *morphospaces*, are mathematical spaces describing and relating the phenotypic configuration of biological organisms and are central tools in nowadays theoretical and mathematical biology. They have been used both in a merely metaphorical sense and in the context of actual mathematical and statistical computations. In a typical morphospace, the morphological configuration of an organism is represented by a single point, and the dimensionality of the space is determined by the number of measured variables (Q-space).

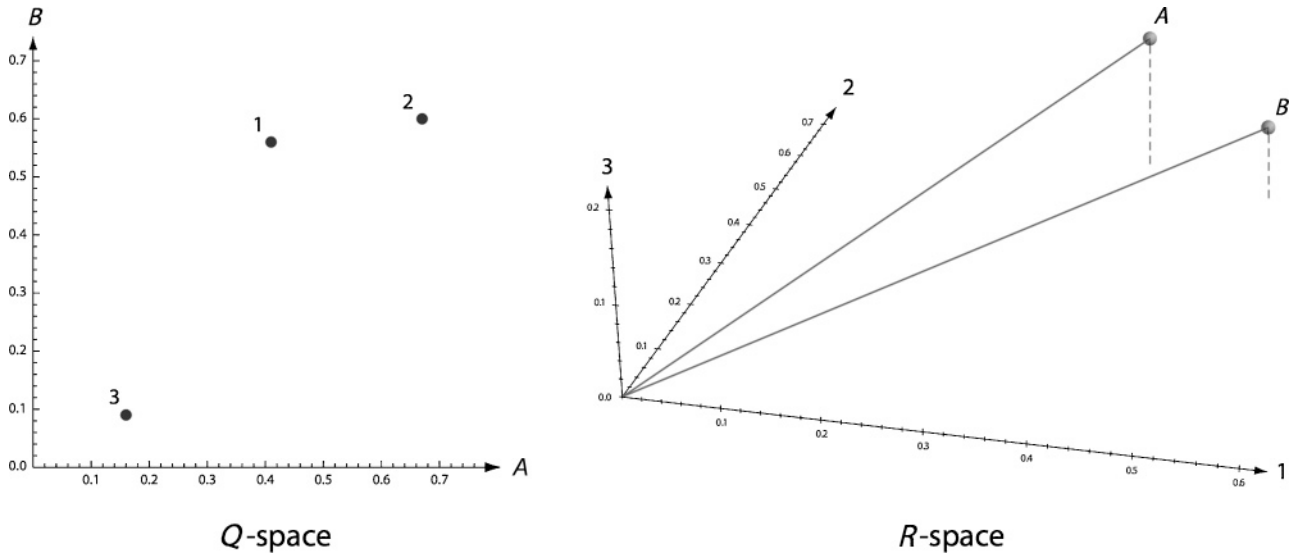

Figure 1.
A simple data matrix, comprising two measurements A and B for the three cases 1–3, is visualized in two different ways. In Q-Space (left) the two variables are taken as axes and the three cases are single points each. R-space (right) is spanned by the three cases, and the variables are two points in this space. When the data is standardized to a mean of zero, the cosine of the angle between the two vectors (the continuous gray lines) connecting the origin with the points A and B is equal to the correlation between A and B (for these data 0.98).

However, in many studies the typically high-dimensional morphospaces are represented by low-dimensional summaries or projections.

The geometric relationship among points in a morphospace should reflect biologically meaningful relationships among the corresponding morphologies. For example, we expect that the distance between points represents morphological similarity: "closer" morphologies are more similar than more "distant" morphologies. Furthermore, a simple geometric structure in morphospace, such as a straight trajectory of several morphologies, should correspond to a simple underlying cause or explanation, like a single developmental process or evolutionary transformation. Two nearly parallel linear trajectories are expected to indicate similar underlying processes, whereas diverging trajectories should be due to different processes.

Our intuitive concept of "a space" is usually dominated by Euclidean spaces, such as the three-dimensional physical space or the two-dimensional surface of a sheet of paper, even though many other spaces with different geometric properties are conceivable. The geometric properties will in general depend on the empirical structure under consideration. A Euclidean vector space is a space with a range of quite stringent relations and properties, allowing many algebraic and geometric operations. The algebraic structure underlying Euclidean geometry is given by a vector space in which there is a well-defined notion of *addition* and (scalar) *multiplication*. A vector (e.g., representing some phenotypic change) can be extended or truncated (multiplied) while maintaining its direction, and it is possible to add two vectors or decompose a vector into a sum of two or more others. For instance, one can predict the phenotype that lies "in-between" two other phenotypes, or calculate a phenotype that results from several different transformations.

Apart from this algebraic structure, Euclidean space has various other properties. First, it is a *metric space*, i.e., there exists a metric distance d that meets the following three conditions for any two points x and y in the space:

(1) $d(x, y) \geq 0$, where equality holds if and only if $y = x$;
(2) $d(x, y) = d(y, x)$; and
(3) $d(x, y) + d(y, z) \geq d(x, z)$.

That is, in a metric morphospace two phenotypes inhabit the same position only if they are identical, and the distance between a phenotype x and a phenotype y is identical to the distance between y and x. Furthermore, the distance between two phenotypes is smaller or equal to the sum of the distances between these two phenotypes and a third one.

In addition, Euclidian space is characterized by the existence of an *inner product* and a *norm*. The standard Euclidean norm of a vector is defined as the square root of its summed squared elements and is equal to the "length" of the vector. The concepts of inner product and norm allow us to define angles and distances between vectors. In Euclidean space, the distance between two points x and y is equal to the norm of the vector $x–y$, and is a metric measure of the distance.

All these geometric relations appear very natural and almost trivial, as we are used to them from the physical space or the real plane R^2. These notions also are deeply entrenched in the rhetoric of modern biology: Two species may exhibit different "directions" of evolution; two forms may be "closer" (phenotypically more similar) than two other ones; we may "extend" or "truncate" an evolutionary or developmental trajectory, or may express a phenotype as the "sum" of two different transformations. Whether a morphospace is used as a metaphor or for computational purposes, it must possess the mathematical properties listed above to allow a language of that style.

Even though phenotypic spaces are often implicitly assumed to have a Euclidean structure, we will show that in most cases this assumption appears to be too strong. The actual geometric properties of a space depend on the data that induce the space, and in particular on the nature of the measured variables. When formalized precisely many morphospaces are weaker than Euclidean spaces (and sometimes they are much weaker); in other words, they exhibit different groups of symmetries or invariances. This implies that some of the usual concepts and properties that we associate with our informal concept of "space" are not meaningful in certain morphospaces.

Affine Morphospaces

In the late 1950s and 1960s, three more or less different scientific communities adopted similar and relatively informal notions of a morphospace. In *numerical taxonomy*, aimed at classifying organisms by phenetic data, "taxonomic distances" were computed based on a range of observed characters (Sokal 1961; Rohlf and Sokal 1965; Sneath and Sokal 1973). For the purpose of visualizability, most researchers preferred to express taxonomic distance by the Euclidian distance among individuals or taxa, so that they can be plotted as points in a Euclidean "phenetic space." Several other metrics (such as Manhattan distance or Canberra distance) have also been suggested that do not constitute a Euclidean space, but ordination methods like principal coordinate analysis were often used to compute a space in which the Euclidean distances among individuals approximate the original distances.

The application of the multivariate statistical methods that emerged in the first half of the 20th century to a wide array of phenotypic measurements has been termed as *multivariate morphometrics* and is nowadays frequently referred to as traditional morphometrics (Blackith and Reyment 1971; Marcus 1990; Reyment 1991). In a variety of biological and paleontological contexts, scatter plots of selected morphological

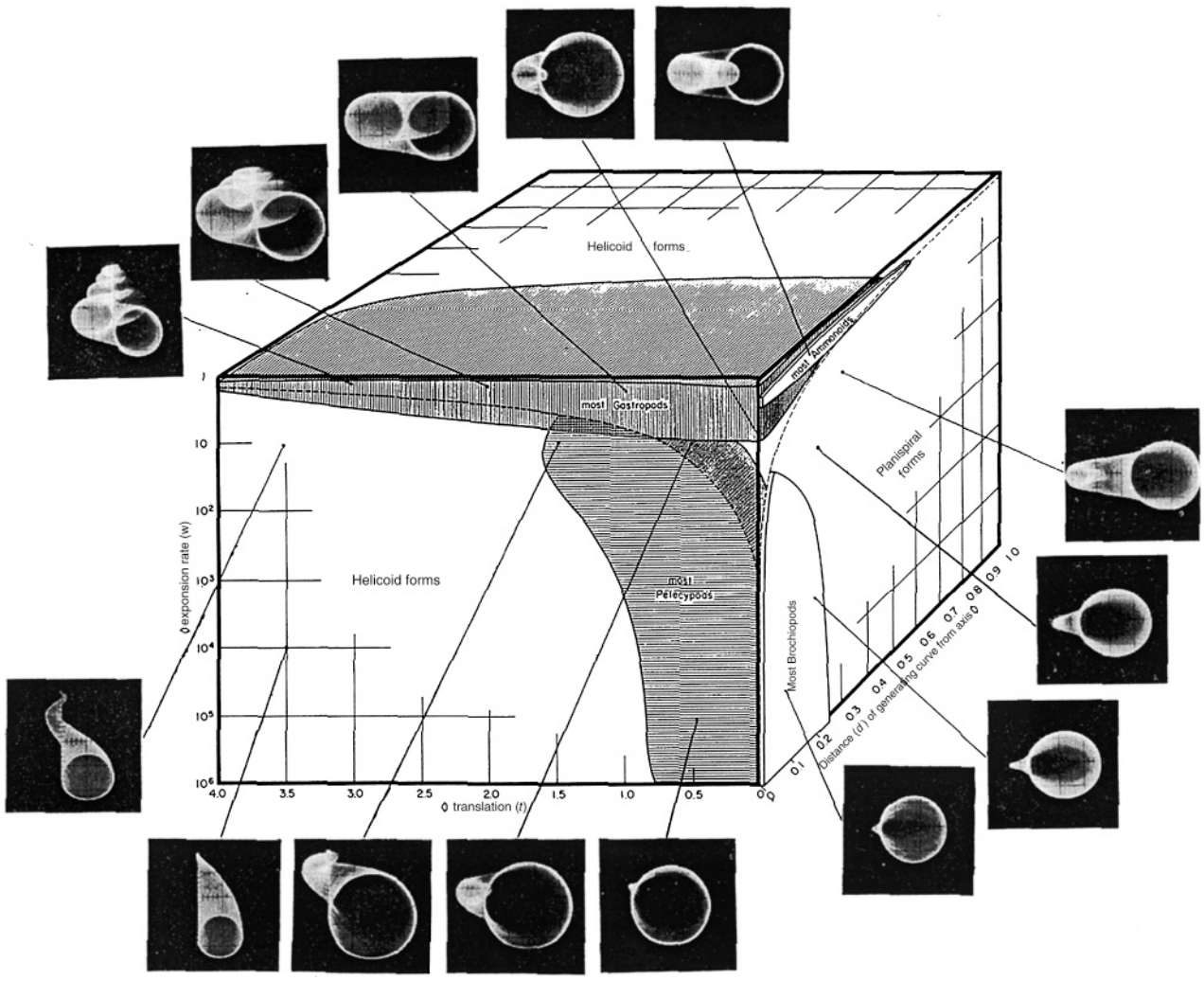

Figure 2.
Raup's (1966) morphospace of coiled shells. He used three out of several parameters in a geometric model simulating shell geometry to define the axes of his space. These parameters are the rate of increase in the size of the generated shell cross section per revolution, the distance between the cross section and the coiling axis, and the rate of translation of the cross section along the axis per revolution. Many existing taxa can be placed within this space, but most parts of the space remain "empty" (the white regions).

variables or ordinations such as principal component analysis and canonical variate analysis were used to produce two- or three-dimensional spaces in which the relationship among different phenotypes can be studied.

Somewhat independently, David Raup developed a simple geometrical model of coiled shells, essentially based on three different parameters, and used early analog computers to plot the modeled geometry (Raup and Michelson 1965; Raup 1966). In his influential approach, he took these three parameters as orthogonal axes of a morphospace, in which he compared the areas occupied by different taxa to those regions of the space that have not been explored by Nature (Figure 2). Attempts in that style are sometimes referred to as *theoretical morphology*; an extensive review can be found in McGhee (1999).

In most of these, meanwhile classical, approaches the measured variables constitute a vector space. But in contrast to Euclidean space, the axes do not necessarily posses commensurate units. For example, in traditional morphometrics, different measures, such as distances, angles, or volumes, sharing no common scales, are often analyzed jointly. Niklas and Kerchner (1984) constructed a three-dimensional morphospace of branching plant forms based on measures as different as "probability of branching termination," "bifurcation angle," and "rotation angle." Two of Raup's three parameters are of the same unit, but they serve fundamentally different roles in his geometric model and there is no "natural" relation among their scales. In some cases the axes may also lack a meaningful origin. Furthermore, some variables may be geometrically dependent (such as the distances and angles of a triangle) so that it might be misleading to take them as orthogonal axes of a morphospace. (Also, Raup's three parameters are not independent; a change in "whorl expansion rate" automatically leads to changes in the other two variables; see Schindel 1990.)

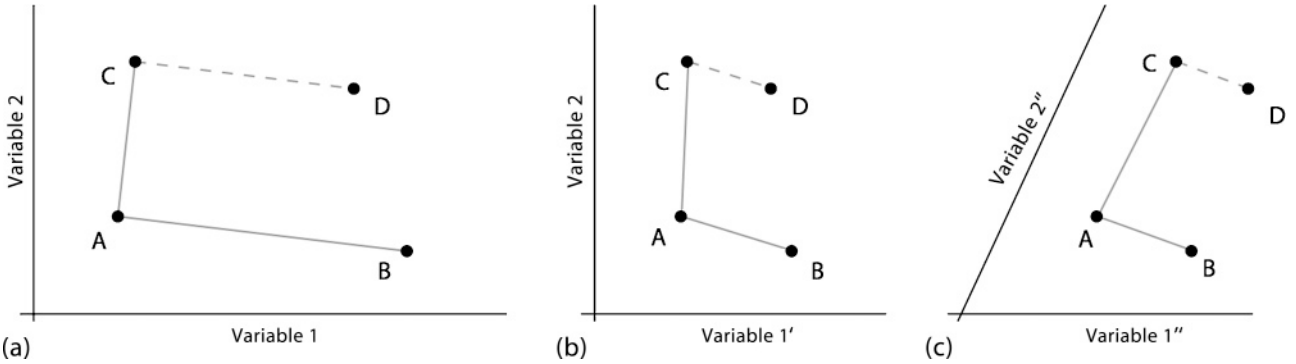

Figure 3.
(a) Four points (phenotypes) in a two-dimensional morphospace. The configuration in (b) is the same as in (a), except that the scale of variable 1 is linearly compressed. Configuration (c) is both compressed and sheared. In an affine space only the properties that are invariant to affine transformations (translation, rotation, scaling, shearing) are meaningful. While in (a) the distance *AB* is larger than *AC*, the reverse is true in (b). Also the angle *CAB* depends on the scale and orientation of the axes. But the vectors *AB* and *CD* are parallel in all three spaces and the ratio *AB*/*CD* remains constant.

In a space based on such variables, geometric findings are only meaningful when they do not depend on the scaling and the orientation of the axes as well as on the origin of the space. A geometry that is invariant (unchanging) relative to such *affine transformations* is called affine geometry and the underlying space is an *affine space*. Affine spaces are weaker than Euclidean spaces; i.e., a Euclidean space is also an affine space, but not every affine space is Euclidean (Suppes et al. 1989). In a vector space (such as in Figure 3), affine transformations correspond to the scaling of the coordinate axes and the shearing of the space, i.e., a rotation of the axes. As distances and angles are not preserved under such transformations, they are not meaningful in an affine space.

However, certain geometric properties are invariant relative to affine transformations. For example, they preserve collinearity among points, i.e., three points lying on a line continue to be collinear after an affine transformation (parallel lines are transformed into parallel lines, and parallelograms into parallelograms). As a consequence, affine transformations preserve incidence relationships (geometric statements such as "a point X lies on a line l" or "line l_1 intersects line l_2") and ratios of distances along a line.

Furthermore, affine transformations leave barycentric combinations unchanged (linear combinations with weights that sum up to 1). That is, the idea of a mid-point or a centroid is meaningful in an affine space.

A familiar measure of distance in statistics that is invariant to affine transformations is Mahalanobis distance, which expresses the distance between points *relative* to the statistical distribution of a subset of the points. Intuitively, affine transformations affect the distances in the same way as the reference distribution and hence leave the relative distances unchanged.

Many classical morphospaces, such as Raup's space or the ones produced in early morphometrics, are affine spaces. Biological inferences from an empirical distribution in such a morphospace necessarily are constrained by the underlying affine geometry. Phenotypes that occupy the same position in an affine morphospace are identical, and there is a well-defined notion of parallel vectors. That is, two phenotypic transformations can be identified as "the same" even if they are applied to different template phenotypes. Vectors also can be added and multiplied by scalars, which is to say that a certain phenotypic change can be extended or truncated. Yet, the distance between two phenotypes cannot be compared with the distance between two other phenotypes (except when the two distances are along parallel directions). Angles also cannot be related directly in an affine space. Different nonoverlapping regions of a morphospace stay nonoverlapping under affine transformations, but the "size" of these regions is not meaningful per se, only ratios of volumes are invariant. In theoretical and comparative biology, several researchers are concerned with the study of "disparity," i.e., the area of a morphospace occupied by different taxa (e.g., Foote 1993). In a (locally) Euclidean space, a range of different measures of disparity may be applied (such as the trace of the population covariance matrix), but in an affine space, disparities of different taxa can only be compared as ratios of the generalized variance (determinant of the population covariance matrix). The generalized variance relates to the "volume" of the distribution of phenotypes in the underlying morphospace, so that ratios of such measures are invariant to affine transformations. (For a proof and details on the statements of this section, see Huttegger and Mitteroecker in preparation.)

Kendall's Shape Space and Riemannian Geometry

Modern morphometrics—the branch of statistics concerned with the measurement of biological form—is based on measurement points, so-called *landmarks*, and is usually referred to as *geometric morphometrics*. Unlike Galton's and Pearson's early attempts as well as traditional multivariate

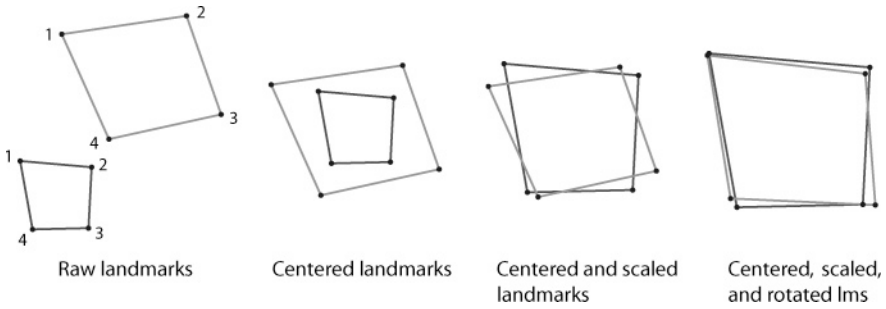

Figure 4.
The three steps of Procrustes superimposition. All landmark configurations are translated to a common origin, scaled to unit centroid size, and rotated to minimize the sum of squared Euclidean distances among the homologous landmarks. The resulting landmark coordinates are called Procrustes shape coordinates; they are standardized for location, size, and orientation to reflect only shape information. The Euclidean distance between two configurations of shape coordinates is called (partial) *Procrustes distance* and is a metric measure of shape difference.

morphometrics, landmark-based approaches can express statistical results in terms of landmark coordinates and thus as actual morphologies or deformations of morphologies (Bookstein 1991; Slice 2005; Mitteroecker and Gunz 2009; for historical reviews see, e.g., Rohlf and Marcus 1993; Bookstein 1998).

The mathematical theory of shape for such landmark configurations was developed somewhat independently by several researchers in the 1980s (Kendall 1981, 1984; Bookstein 1991; Small 1996; Dryden and Mardia 1998). Kendall's shape space, named after the Scottish mathematician David Kendall, is the mathematical space for the shapes of different point configurations. Take a set of n objects in k dimensions and describe their forms by p landmarks each. The shape of an object equals the geometric properties that are invariant under translation, rotation, and scaling of the object. In his pioneering work, Kendall (1981, 1984) showed that a metric that takes into account only shape differences between two landmark configurations induces a $pk-k-k(k-1)/2-1$-dimensional shape space. But this space is not a "flat" Euclidean space; rather it constitutes a "curved" Riemannian manifold. As a simple example of a Riemannian manifold one may consider smooth two-dimensional surfaces in three-dimensional space. The distance between any two points on a manifold is defined as the length of the shortest curve connecting these two points. At any point on the manifold there is a Euclidean space tangent to the manifold (e.g., a plane tangent to a curved surface). Thus, a Riemannian manifold by itself cannot be regarded as a Euclidean space (just as a metric space); but in a sufficiently small neighborhood of any point it is approximately Euclidean. The metric on Kendall's shape space is usually referred to as Procrustes distance and is estimated in most applications by a Generalized Procrustes Analysis (Figure 4; Rohlf and Slice 1990). For the simplest two-dimensional object, a triangle, shape space is a two-dimensional manifold with the form of the surface of a (hemi)sphere (Figure 5; see also Slice 2001). Kendall's shape space for more than three landmarks is more complex and cannot be visualized easily.

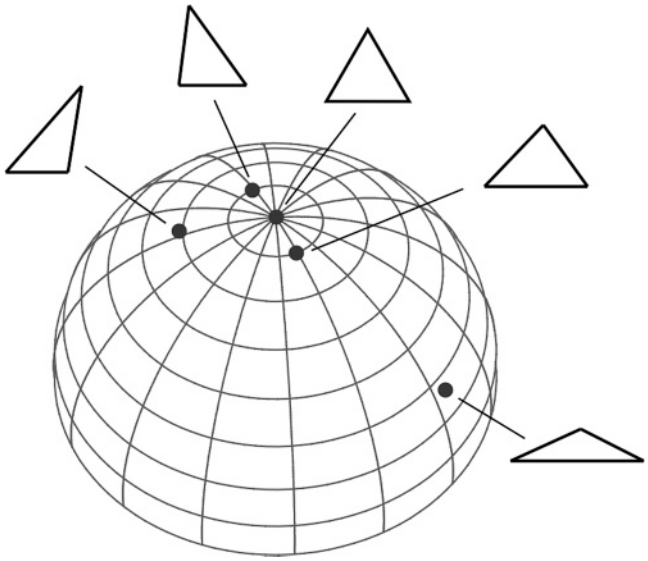

Figure 5.
Kendall's shape space for triangles is a two-dimensional manifold with the form of the surface of a hemisphere. Any point in this space represents the shape of one triangle. The figure shows one representation of this space along with five different shapes, where the "north pole" is an equilateral triangle.

For small-scale variation, the nonlinear shape space can be approximated by a Euclidean tangent space (Rohlf 1999). In an analysis of cranial shape, Marcus et al. (2000) demonstrated that this linear approximation provides a good representation of shape variation for a wide range of different mammalian taxa. As most statistical tools available are based on linear models, the majority of modern morphometric analyses are performed in this tangent space rather than in Kendall's shape space.

In almost all practical applications, the shape space underlying the set of measured landmarks is of very high dimension and cannot be visualized without the loss of information (for instance, shape space for 10 three-dimensional landmarks has 23 dimensions). In statistics, low-dimensional representations of high-dimensional data spaces are called *ordination analyses* (see, e.g., Mardia et al. 1979; Johnson and Wichern 1998). The

most classical and widespread method is principal component analysis (PCA), which provides a low-dimensional orthogonal projection of the full space so that most of the variance in the data is preserved by the projection. Other ordination methods, such as canonical variate analysis, multidimensional scaling, and correspondence analysis, are not simple orthogonal projections and hence distort the geometry of the space; these methods were used mainly in non-landmark-based approaches. The full shape space can also be partitioned into different subspaces based on statistical and geometrical criteria such as symmetric versus asymmetric shape variation, functional significance, integrated versus modular variation, or the geometric scale of shape variation (Bookstein 1989; Rasskin-Gutman and Buscalioni 1996; Klingenberg and McIntyre 1998; Mardia et al. 2000; Manfreda et al. 2006; Mitteroecker and Bookstein 2008).[1] By augmenting the shape coordinates with the natural logarithm of centroid size—the measure of overall size in geometric morphometrics—shape space can be extended to Procrustes form space (also called size-shape space; Mitteroecker et al. 2004).

Kendall's shape space is an example of a morphospace that is locally Euclidean and thus warrants (locally) all the notions of distance, length, and angle. However, this geometric morphometric approach requires the identification of the same set of biologically homologous landmarks at every measured individual and thus constrains the diversity of organisms that can be compared at once in such a space. Due to the curvature of shape space, this methodology further imposes limits on the range of variability that can be assessed for these landmarks, but in most real-world applications this range is not exceeded.

Metric Spaces

Euclidean spaces and affine spaces have the algebraic structure of a vector space, and Kendall's shape space is locally Euclidean, i.e., the tangent space at each point in Kendall's shape space has the structure of a vector space. Metric spaces do not need to be vector spaces, however. Important biological examples are given by the space of nucleotide sequences or amino acid sequences of finite length i (e.g., Stadler et al. 2001; Stadler and Stadler 2004). Clearly, these spaces are not vector spaces over the field of real numbers: multiplying a sequence with a real number or adding two sequences does not necessarily yield another valid sequence. Yet, a metric can be defined by the so-called Hamming distance. The Hamming distance between two sequences s and s' is given by the number of positions where s and s' differ. For example, s is one step away from s' if there exists exactly one position i at which s' has a value different from the value of s at i (Figure 6a).

The Hamming distance is meaningful for point mutations in nucleotide sequences. There are four possible entries at each position, and at each position these possibilities are the same

Figure 6.
(a) Three DNA sequences with seven nucleotides each. The Hamming distance between two sequences is given by the number of positions at which the sequences differ. The first sequence in the figure has a Hamming distance of 1 to the second sequence and of 2 to the third sequence. Sequences second and third differ at three positions. (b) Three sequences with five positions each, but in contrast to the nucleotide sequences, the positions here are not comparable among each other. It is unclear if a metric distance is warranted in this case. Yet, a qualitative notion of neighborhood or similarity is possible. For example, the first sequence differs at position three from the second sequence, and the third sequence differs at the same position from the first one plus in one further position (position two). The third sequence hence is farther away from the first sequence than is the second sequence.

(A, G, C, and T or U). If we assume that any mutation of one nucleotide into another has the same probability (which is, of course, an idealization), then changing position i in the sequence can be assigned a real number $r > 0$ regardless of the incumbent and the mutant nucleotide. Since this holds for any position in the sequence, the resulting metric distance between two sequences x and y will be r times the Hamming distance of x and y. Hence, the Hamming distance is unique up to multiplication by a positive real number.

The space of nucleotide sequences together with the Hamming distance constitutes a metric space. The space of nucleotide sequences is not a morphospace, of course (although it is related to certain interesting morphospaces such as the space of RNA's secondary structure; see below). But we can think of a morphospace having a similar metric structure as the space of nucleotide sequences. Take a set of k nominal phenotypic variables that can take on p different values, e.g., the k body segments of an arthropod, each having one of the p possible types of appendages. If it would be biologically reasonable to assume that (evolutionary) changes of these character states are equally likely (e.g., by comparable homeotic mutations), a meaningful distance function would be the number of elements in which two phenotypes differ. This distance would correspond to the probability of an evolutionary transition from one phenotype to another and induces a metric space. The notion of a distance between two phenotypes is thus meaningful in such a space, even though no vector structure exists (vectors cannot be added or multiplied) and the concepts of direction and angle are undefined (there is no inner product). A Euclidean graphical representation of a metric space without a vector structure can be misleading, as it would indicate geometric relationships that actually do not exist. A more correct representation is possible by a Hamming graph connecting all states differing by a

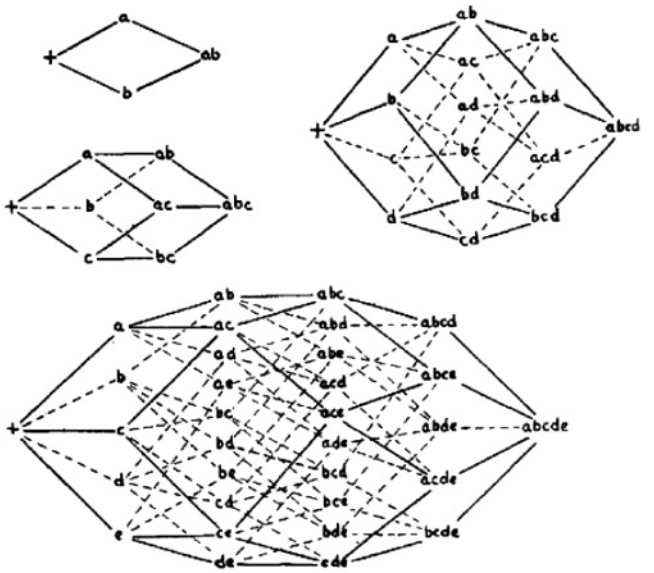

Figure 7.
Representation of two to five "combinations of paired allelo-morphs" as Hamming graphs in Sewall Wright's famous 1932 paper, in which he introduced the concept of fitness landscapes.

unit Hamming distance. Figure 7 shows five such graphs with which Sewall Wright (1932) aimed to represent the distances among different "combinations of paired allelomorphs."

Topological Spaces and Weaker Structures

Thomas and Reif (1993) introduced the "skeleton space" and applied it to categorize morphological variability in hard skeletons of Burgess Shale animals during the Middle Cambrian (Thomas et al. 2000). This attempt is another example of a morphospace where phenotypes are distinguished by having different states for certain morphological characteristics. The elements of the skeleton set consist of seven variables or positions, which can be represented as a 7-tupel $(x_1, \ldots, x_7)$. There are two to four possible states at each position, e.g., one position describes whether the skeleton of the specimen is internal or external; another one describes the growth pattern of the skeleton (accretionary; addition of serial units; molting and replacement; remodeling).

The skeleton space, as described by Thomas and Reif (1993), is in fact a set of certain modes for the design of animal skeletons aimed at capturing the phenomenon of convergent evolution. In their enumerative approach, the authors did not provide any relations such as a metric among the possible modes (and hence it is not a "space" in a strict sense). It is tempting, though, to apply the Hamming distance to the skeleton set, i.e., to say that specimen a is k steps from specimen a' if they differ at k positions. But the seven variables are not comparable in their developmental or phylogenetic meaning, nor are the possible values comparable within one position. For example, the third variable in Thomas et al. (2000), denoting the number of skeletal elements, can take on the values "one," "two," or "many." It is quite easy to see that using the Hamming distance cannot be justified as in the space of nucleotide sequences. At each position we may assign different numbers to a change in that position. Moreover, numbers representing differences may vary across the positions in the sequence. Changes within as well as across positions are too unrelated to establish meaningful relations between the numbers representing distances in a straightforward way.

Even if it turns out that no meaningful *quantitative* measure of distance can be assigned to such a set, some *qualitative* notion of "nearness" may be warranted. For example, take a phenotype B differing in k positions from A. A phenotype C that differs in the same k positions from A plus in one further position can be regarded as farther away from A than is B (Figure 6b). Alternatively, it might be possible to formulate "neighborhood relations" in terms of the accessibility among phenotypes: Two phenotypes are neighbors if a direct (evolutionary) transition from one phenotype to the other is possible (but this relation does not need to be symmetric). Sets equipped with such a qualitative nearness or neighborhood relationship may give rise to topological or pretopological spaces.[2]

There are several other interesting pretopological spaces in biology. For example, the space of nucleotide sequences—which is metric when considering point mutations only—is incompatible with a metric distance measure when unequal crossover is allowed, but satisfies the conditions for a pretopology (Stadler et al. 2002). Fontana and Schuster (1998a, 1998b) introduced an RNA shape space describing the folding of RNA sequences into secondary structures. The topology of this shape space is determined by the "genetic accessibility" among the folding structures: Several RNA genotypes fold into the same secondary structure and hence constitute one "neutral network" of RNA sequences (Figure 8). The nearness of an RNA phenotype to another phenotype is determined by the size of the joint boundary between the two corresponding neutral networks relative to the total size of the network, and thus relates to the probability of a transition from one phenotype to the other through random point mutations. As neutral networks for different RNA shapes may be of different sizes, the nearness relation may be asymmetric and hence does not support a metric geometry but can be described as a pretopology. Perhaps this elegant approach may never be applied to actual organismal morphologies (where the genotype–phenotype map is considerably more complex). But it allows several inferences that are likely to apply to the realm of macroscopic anatomy as well (see also Stadler et al. 2001).

Morphospaces as Metaphors

Morphospaces are frequently used in quantitative fields of research like morphometrics, theoretical morphology, and

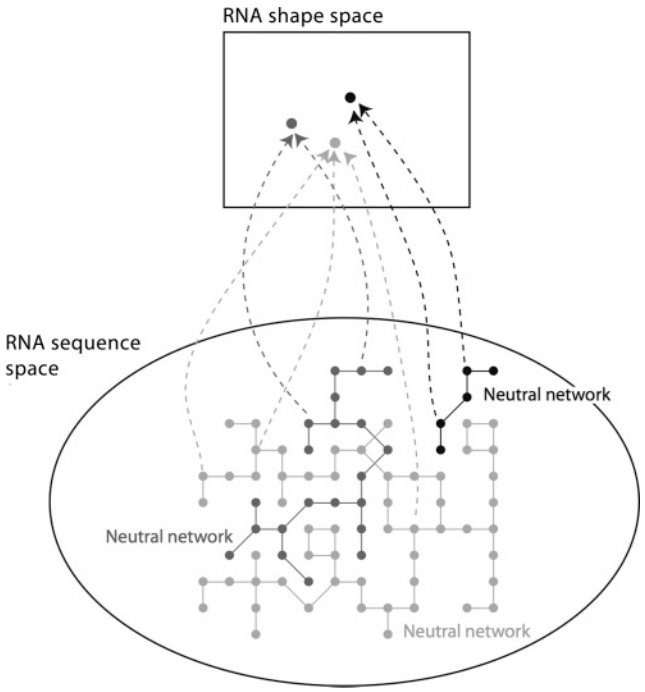

Figure 8.
RNA shape space equipped with an accessibility topology. Each "neutral network" of RNA sequences (black, gray, and light gray) gives rise to one specific secondary RNA shape. The nearness between two phenotypes is determined by the relative size of the joint boundary between the two corresponding neutral networks and reflects the genetic accessibility based on random point mutations. For example, a random step off the black network is likely to end up in the light gray network, but not vice versa. Thus, the black RNA shape is near the light gray one, but the light gray shape is not near the black RNA shape. (Modified from Stadler et al. (2001))

population genetics. As demonstrated in the preceding sections, the biological meaningfulness of empirical or theoretical results obtained from such computational studies strongly depends on the geometrical and topological properties of the employed morphospace. However, phenotypic spaces are also used in a more metaphorical sense in developmental and evolutionary biology. Below we will discuss two widely used metaphors, the developmental hourglass and the fitness landscape. We show that even in a less quantitative context, the utility of these concepts is partly determined by the geometric properties of the underlying morphospace.

The Developmental Hourglass

In 19th-century embryology, the idea of recapitulation—in particular Ernst Haeckel's biogenetic law—was the source of continued dispute. Basically, Haeckel emphasized that evolution proceeds by appending new developmental events to the end of the descendent's ontogeny; hence the claim that ontogeny recapitulates phylogeny. One prominent opponent of this view was Karl Ernst von Baer, who argued that embryos develop from general and uniform early stages to more and more specific and individualized older stages—ontogeny is not recapitulation but differentiation. In his *Entwicklungsgeschichte der Thiere*, he wrote, "The further we go back in the development of vertebrates, the more similar we find the embryos both in general and in their individual parts." (von Baer 1828: 221).

No *adult* stage of another animal is ever recapitulated during ontogeny, yet different taxa may share common *embryonic* stages in their development while differing in later ontogeny. Milne-Edwards (1844: 72, cited after Gould 1977) expressed von Baer's arguments in a very figurative way:

> The metamorphoses of embryonic organization, considered in the entire animal kingdom, do not constitute a single, linear series of zoological phenomena. There are a multitude of these series . . . They are united in a bundle at their base and separate from each other in secondary, tertiary, and quaternary bundles, since in rising to approach the end of embryonic life, they depart from each other and assume distinctive characteristics.

The biogenetic law was eventually rejected with the rise of experimental methods in embryology in the late 19th century and finally with the rediscovery of Mendelian genetics in the 1910s and 1920s (Gould 1977). In contrast, von Baer's view of comparative embryology—as tree-like divergence from common undifferentiated early stages—is still maintained in modern developmental biology, but has been extended by the observation that very early development is subject to substantial interspecific variation too. In contemporary EvoDevo literature, this extended view is often illustrated by the developmental hourglass (Raff 1996) or the phylotypic egg timer (Duboule 1994; see Figure 9). The ontogenetic trajectories of different species may be clearly distinct at the earliest stages (right after conception), but converge to a *phylotypic stage* at mid-development, and diverge again thereafter in a "von Baerian" style. Duboule speculated that the phylotypic stage is due to the highly conserved and linearly organized HOX gene expression along the early body axis. Raff, instead, claimed that the phylotypic stage is a developmental period characterized by strong interactions and developmental dependencies across the whole embryo; thus, evolutionary modifications of development at that stage are likely to be lethal (Galis and Metz 2001). The relatively simple mechanisms of axis formation during earliest development, in contrast, are more prone to evolutionary modification. During late development, when the organs have started to develop, the organism is more "modular" and hence easier to modify locally (Wagner and Altenberg 1996).

The developmental hourglass represents a sort of morphospace in which morphological characteristics of different organisms are plotted in the course of their development. The so-defined trajectories differ initially, converge at the phylotypic stage, and eventually diverge again. More specifically, each horizontal cross section through the hourglass is

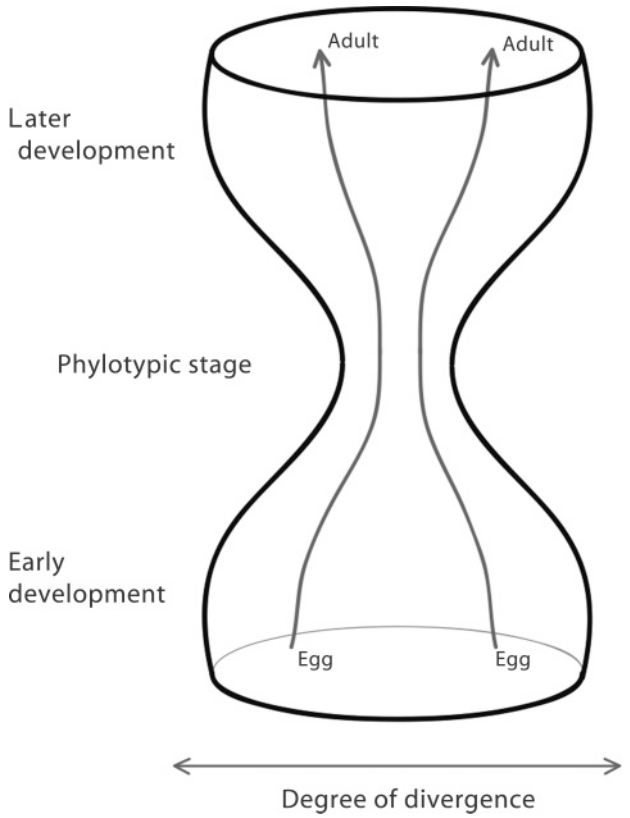

Figure 9.
The developmental hourglass. (After Raff 1996)

a morphospace encompassing the phenotypic variability of different organisms at the same developmental stage. Morphological distances among specimens or species at early and late stages are supposed to be larger than distances during mid-embryonic stages. But it is unclear what the geometrical structure of such a morphospace would be and if a consistent notion of distance is warranted within each "cross-section," let alone across all developmental stages. Certainly, adult or even late embryonic characteristics are not present at early developmental stages. Limb morphology, as an example, cannot be determined at the time of gastrulation, when the limb buds have not started to develop yet. Similarly, at later stages the phenotypes are already differentiated in a complex way so that quantitative comparisons across species may be difficult. How to compare the beak of a chicken to the human lips? Casting a specimen into a morphospace with a vector structure requires one to specify all coordinates, i.e., the same phenotypic measures need to be defined for all specimens being compared. Consequently, a morphospace encompassing all this evolutionary and developmental diversity is unlikely to be Euclidean or affine. There is also no straightforward way to establish a metric when different variables are undefined for different specimens. Furthermore, as pointed out above, metrics like the Hamming distance are based on characteristics that are meaningfully comparable within and across the individuals (see Figure 6). Perhaps a morphospace underlying the developmental hourglass may be topological or pretopological, if a reasonable classification can be devised at all.

However, for such a weak geometrical structure, no single quantitative notion of distance may be applicable across the full range of phenotypic variability that is typically covered by the developmental hourglass. One cannot quantitatively relate morphological similarities among organisms at the blastula stage to the similarities at the time of somite formation, let alone to the diversity among fetuses. What biologically meaningful measure of similarity or distance could be applied at all these levels of complexity? In fact, even a qualitative comparison of variability across these stages would be vague. The developmental hourglass thus cannot be regarded as a valid quantitative concept.[3]

Even as a metaphor, the hourglass may be misleading. Variability at the phylotypic stage may be small but decisive, already containing the properties that indicate or even induce larger differences in subsequent developmental pathways, and eventually leading to highly differentiated adult morphologies of different taxa. Based on similar arguments, Richardson et al. (1997) even questioned the existence of a phylotypic stage at all.

Fitness Landscapes

The eminent American population geneticist Sewall Wright (1932) introduced the notion of *fitness landscapes* or *adaptive landscapes* in an address at the Sixth International Congress of Genetics. Wright's landscapes were based on a genotype space, or more specifically, on a space of allele combinations, where a fitness or adaptiveness value was assigned to each element of this space. In population genetics and game theory, fitness landscapes are also used on a space of allele frequencies or phenotype frequencies, so that a point on this landscape represents an entire population (Hofbauer and Sigmund 1998). In an attempt to synthesize genetical and paleontological approaches, Simpson (1944) applied adaptive landscapes to a space of phenotypic traits, which became the basis of the statistical approaches in evolutionary quantitative genetics (Arnold et al. 2001). The basic statement associated with fitness landscapes or adaptive landscapes is that evolution proceeds toward the "peaks" in this landscape by natural selection. Wright specifically introduced fitness landscapes to illustrate the problem of local fitness optima (peaks), which are separated by the areas of low fitness (valleys), so that evolution may be trapped at a local maximum that is lower than an adjacent peak. Wright concluded that genetic drift of small subpopulations has to be crucial in escaping from suboptimal fitness peaks—his shifting-balance theory, which has been subject to heavy debates (see, e.g., Johnson 2008 for review).

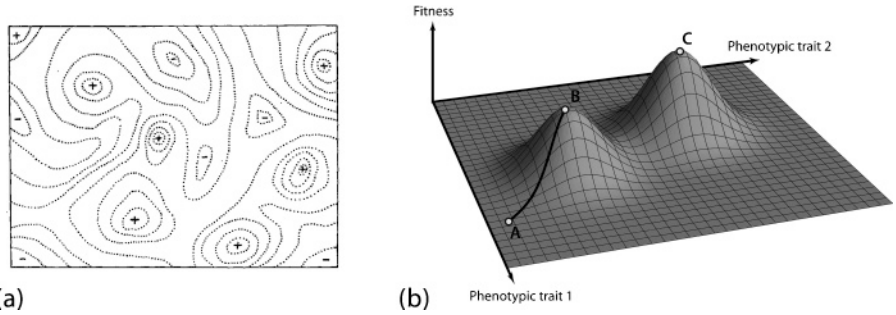

Figure 10.
(a) Wright's 1932 diagram of a fitness landscape with several peaks and valleys. His original figure caption was, "Diagrammatic representation of the field of gene combinations in two dimensions instead of many thousands. Dotted lines represent contours with respect to adaptiveness." (b) A three-dimensional representation of a simple adaptive landscape based on a two-dimensional phenotypic space. A population with a mean phenotype at the point A (and approximately uncorrelated traits) tends to evolve toward the nearest peak B. However, fitness at the point C is higher than at B, but every path from B to C leads through a valley of low fitness (peak-shift problem).

Formally, a fitness landscape (or adaptive landscape) consists of

(1) a set X of configurations (genotypes or phenotypes),
(2) a notion of nearness, neighborhood, distance, or accessibility, and
(3) a function $f : X \to R$ that assigns a real number—the fitness or adaptiveness value—to each element.

The set X together with the geometric or algebraic structure on X forms the configuration space (Stadler 2002). A fitness landscape on the genotype space can be related to the fitness landscape on the phenotype space via the so-called genotype–phenotype map, assigning each genotype to a phenotype or to a set of phenotypes (Alberch 1991; Wagner and Altenberg 1996).

Fitness landscapes have remained highly influential until today and may be the most widespread metaphor in evolutionary biology and genetics. They are used both in mathematical modeling (Gavrilets 2004), and even more so in the graphical description of various arguments in evolutionary biology (e.g., McGhee 2007). But there are a number of problems and confusions associated with fitness landscapes, which we do not want to reiterate here (see, e.g., Gavrilets 2004; Kaplan 2008; Pigliucci 2008). The point we would like to emphasize is that the usefulness and the meaningfulness of statements about fitness landscapes partly depend on the geometrical and topological properties of the underlying phenotype or genotype space.

When using fitness landscapes, it is often assumed (such as in McGhee 2007) that the space on which a fitness function is defined is essentially Euclidean, i.e., all notions associated with distance and direction hold. As should be clear from what we have argued for so far, this assumption will most often be unwarranted. Already in the most benign cases of morphospaces, the space will only be locally Euclidean as in Kendall's shape space. But fitness landscapes are often applied in a macroevolutionary context to a large phenotypic diversity, so that the underlying space would have a relatively weak geometric structure, which may not warrant quantitative notions of distance or direction at all. Wright (1988: 116) wrote that the space which he envisioned for his fitness landscape consists of genotypes that "are packed, side by side in a two-dimensional space in such a way that each is surrounded by genotypes that differ by only one gene replacement." Apparently, this relationship (which he adequately depicted by the Hamming graphs as shown in Figure 7) cannot be mapped onto a two-dimensional surface, let alone a Euclidean surface. As we have seen, the space of nucleotide sequences is a metric space without a vector structure, and more general genotype configurations are described best by a pretopological space.

That is, in most realistic applications of fitness landscapes, familiar properties like distance, direction, or length are not warranted to characterize the geometry of the landscape or of evolutionary trajectories on the fitness surface. However, certain concepts may still hold even for weak topological structures. As we have remarked above, in affine morphospaces the concept of a point lying in between two points is fundamental and invariant under affine transformations. For example, the two fitness peaks B and C in Figure 10(b) can be identified as peaks that are separated by a valley even under affine transformation of the morphospace, such as rescaling or rotation of the axes. While most concepts in quantitative genetics are based on a state space equipped with a vector structure or on a space of gene frequencies, the basic concepts related to fitness landscapes can be formulated even for pretopological spaces (Stadler 2002). Defining a peak as a point where all neighbors have lower fitness, and *mutatis mutandis* for valleys, is based only on the qualitative relationship of neighborhood. Accordingly, the peak shift problem can be phrased as the problem of evolving from genotype A to genotype B, when all possible paths between A and B consist of at least one genotype with a fitness lower than that of both A and B. Other properties of fitness landscapes and genotype–phenotype maps can also be

formulated when employing a pretopological structure only (see, e.g., the discussion of RNA shape space above).

Summary

Morphospaces are mathematical spaces describing and relating the phenotypic configuration or the morphology of biological organisms. The distribution of phenotypes in such a space may allow inferences about the pattern of phenotypic relationships, which may not have been evident from a pairwise comparison of the phenotypes themselves. Our intuitive concept of a space is usually based on Euclidean vector spaces and gives rise to a language that includes statements about distance, direction, angle, and length. However, most realistic morphospaces are not Euclidean but have weaker geometrical and algebraic structures. Ignorance about the actual properties of a morphospace may lead to an inappropriate use of this space—as a metaphor and as a computational tool. We showed that the biological notions that can be inferred from these spaces differ profoundly across the range of possible topologies. In particular, we find that moving from local to global comparisons among specimens—from microevolutionary to macroevolutionary spaces—is accompanied by a significant weakening of meaningful geometrical structure.

Landmark-based morphometrics gives rise to a Riemannian shape space, which locally is Euclidean. Most classic morphospaces are affine spaces, which have a vector structure but no notion of a distance. Other spaces, such as sequence space, may have a metric but no vector structure. Finally, the most general sort of morphospace, like the skeleton space, does not support a quantitative measure of similarity at all, but may allow a qualitative notion of neighborhood or nearness (topological or pretopological spaces). While landmark-based morphometrics can only describe a relatively limited range of phenotypic variability, a pretopological space can be based on very general classification schemes that apply to a very large range of diversity.

These geometrical properties of morphospaces may even limit the meaningfulness of common metaphors, such as the developmental hourglass, which typically covers a large range of ontogenetic and phylogenetic variability. In contrast, the core ideas related to fitness landscapes can also be formulated in terms of neighborhood or accessibility and hence apply to weak morphospaces as well.

Notes

1. McGhee (1999) emphasized the distinction between "empirical morphospaces," denoting the statistical constructs typically used in morphometrics, and "theoretical morphospaces," such as Raup's space of coiled shells. This distinction seems to originate mainly from a confusion between ordinations and the morphospace itself. There are no mathematical differences between empirical and theoretical morphospaces; basically, the applicability of both spaces depends on the underlying topology. Differences between the classical approaches in morphometrics and in theoretical morphology are more in terms of how measurements are selected and the style of post hoc interpretation of results (cf. Huttegger and Mitteroecker in preparation).

2. More formally, a pretopology on X can be defined by sets of neighborhoods $N(x)$ for each $x \in X$. $N(x)$ is a neighborhood system if it meets the following criteria:

(1) $x \in N$ for all $N \in N(x)$.

(2) If $N_1, N_2 \in N(x)$, then $N_1 \cap N_2 \in N(x)$.

(3) If $N_1 \in N(x)$ and $N_1 \subset N_2$, then $N_2 \in N(x)$.

If N is the mapping that assigns the neighborhood system $N(x)$ to each x, then (X, N) is called the pretopological space. The neighborhood relationships can be represented by a directed graph, consisting of vertices from each element to its neighbors. The structure of a pretopological space is more local than the structure of a topological space. One can get a topological space from a pretopological space by imposing more stringent requirements on the neighborhood systems, e.g., by adding the following axiom:

(4) For all $N \in N(x)$ there is an $N \in N(x)$ such that $N \in N(y)$ for all $y \in N$. For more details see, e.g., Stadler et al. 2001.

3. In an empirical study on the phylotypic stage, Bininda-Emonds et al. (2003) circumvented the problem of homologous measurements by comparing variability in the timing of homologous developmental events across different ontogenetic stages, but ignored spatial variation among these developmental processes or their products. This interesting approach hence covers only a single dimension of the morphospace.

References

Alberch P (1991) From genes to phenotype: Dynamical systems and evolvability. Genetica 85: 5–11.

Amari S (1985) Differential-Geometrical Methods in Statistics. Berlin: Springer.

Arnold SJ, Pfrender ME, Jones A (2001) The adaptive landscape as a conceptual bridge between micro and macroevolution. Genetica 112–113: 9–32.

Bininda-Emonds ORP, Jeffery EJ, Richardson MK (2003) Inverting the hourglass: Quantitative evidence against the phylotypic stage in vertebrate development. Proceedings of the Royal Society London B 270: 341–346.

Blackith RE, Reyment RA (1971) Multivariate Morphometrics. London: Academic Press.

Bookstein FL (1989) Principal warps: Thin plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11: 567–585.

Bookstein F (1991) Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge: Cambridge University Press.

Bookstein F (1998) A hundred years of morphometrics. Acta Zoologica Academiae Scientarium Hungaricae 44: 7–59.

Dryden IL, Mardia KV (1998) Statistical Shape Analysis. New York: Wiley.

Duboule D (1994) Temporal colinearity and the phylotypic progression: A basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. Development (Suppl): 135–142.

Fontana W, Schuster P (1998a) Continuity in evolution: On the nature of transitions. Science 280: 1451–1455.

Fontana W, Schuster P (1998b) Shaping space: The possible and the attainable in RNA genotype–phenotype mapping. Journal of Theoretical Biology 194: 491–515.

Foote M (1993) Contributions of individual taxa to overall morphological disparity. Paleobiology 19: 403–419.

Galis F, Metz JAJ (2001) Testing the vulnerability of the phylotypic stage: On modularity and evolutionary conservation. Journal of Experimental Zoology 291: 195–204.

Galton F (1888) Co-relations and their measurement, chiefly from anthropometric data. Proceedings of the Royal Society 45: 135–145.

Galton F (1907) Classification of portraits. Nature 76: 617–618.

Gavrilets S (2004) Fitness Landscapes and the Origin of Species. Princeton, NJ: Princeton University Press.

Gould SJ (1977) Ontogeny and Phylogeny. Cambridge, MA: Harvard University Press.

Hofbauer J, Sigmund K (1998) Evolutionary Games and Population Dynamics. Cambridge: Cambridge University Press.

Huttegger S, Mitteroecker P (in preparation) The geometry of phenotype spaces: Invariance and meaningfulness.

Johnson N (2008) Sewall Wright and the development of shifting balance theory. Nature Education 1(1).

Johnson RA, Wichern DW (1998) Applied Multivariate Statistical Analysis. Upper Saddle River, NJ: Prentice Hall.

Kaplan J (2008) The end of the adaptive landscape metaphor? Biology and Philosophy 23: 625–638.

Kendall D (1981) The statistics of shape. In: Interpreting Multivariate Data (Barnett V, ed), 75–80. New York: Wiley.

Kendall D (1984) Shape manifolds: Procrustean metrics and complex projective spaces. Bulletin of the London Mathematical Society 16: 81–121.

Klingenberg CP, McIntyre GS (1998) Geometric morphometrics of developmental instability: Analyzing patterns of fluctuating asymmetry with Procrustes methods. Evolution 52: 1363–1375.

Manfreda E, Mitteroecker P, Bookstein FL, Schaefer K (2006) Functional morphology of the first cervical vertebra in humans and non-human primates. Anatomical Record, Part B: The New Anatomist 289: 184–194.

Marcus LF (1990) Traditional morphometrics. In: Proceedings of the Michigan Morphometrics Workshop (Rohlf FJ, Bookstein FL, eds), 77–122. Ann Arbor, MI: University of Michigan Museums.

Marcus LF, Hingst-Zaher E, Zaher H (2000) Application of landmark morphometrics to skulls representing the orders of living mammals. Hystrix, Italian Journal of Mammology 11: 27–47.

Mardia KV, Bookstein F, Moreton I (2000) Statistical assessement of bilateral symmetry of shapes. Biometrika 87: 285–300.

Mardia KV, Kent JT, Bibby JM (1979) Multivariate Analysis. London: Academic Press.

McGhee GR (1999) Theoretical Morphology: The Concept and Its Applications. New York: Columbia University Press.

McGhee GR (2007) The Geometry of Evolution: Adaptive Landscapes and Theoretical Morphospaces. Cambridge: Cambridge University Press.

Milne-Edwards H (1844) Considérations sur quelques principes relatifs à la classification naturelle des animaux. Annales des Sciences Naturelles (Zoologie) (Série 3) 1: 65–99.

Mitteroecker P, Bookstein FL (2007) The conceptual and statistical relationship between modularity and morphological integration. Systematic Biology 56: 818–836.

Mitteroecker P, Bookstein FL (2008) The evolutionary role of modularity and integration in the hominoid cranium. Evolution 62: 943–958.

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

Mitteroecker P, Gunz P (2009) Advances in geometric morphometrics. Evolutionary Biology 36: 235–247.

Mitteroecker P, Gunz P, Bernhard M, Schaefer K, Bookstein F (2004) Comparison of cranial ontogenetic trajectories among great apes and humans. Journal of Human Evolution 46: 679–697.

Niklas KJ, Kerchner V (1984) Mechanical and photosynthetic constraints on the evolution of plant shape. Paleobiology 10: 79–101.

Pearson K, Morant GM (1934) The Wilkinson head of Oliver Cromwell and its relationship to busts, masks and painted portraits. Biometrika 26: 1–116.

Pigliucci M (2008) Sewall Wright's adaptive landscapes: 1932 vs. 1988. Biology and Philosophy 23: 591–603.

Raff R (1996) The Shape of Life: Genes, Development, and the Evolution of Animal Form. Chicago: University of Chicago Press.

Rasskin-Gutman D, Buscalioni AD (1996) Affine transformation as a model of virtual form change for generating morphospaces. In: Advances in Morphometrics (Marcus LF, Corti M, Loy A, Slice D, Naylor G, eds), 169–178. New York: Plenum Press.

Raup DM (1966) Geometric analysis of shell coiling: General problems. Journal of Paleontology 40: 1178–1190.

Raup DM, Michelson A (1965) Theoretical morphology of the coiled shell. Science 147: 1294–1295.

Reyment RA (1991) Multidimensional Paleobiology. New York: Pergamon Press.

Richardson MK, Hanken J, Gooneratne ML, Pieau C, Raynaud A, Selwood L, Wright GM (1997) There is no highly conserved embryonic stage in the vertebrates: Implications for current theories of evolution and development. Anatomy and Embryology 196: 91–106.

Rohlf FJ (1999) Shape statistics: Procrustes superimpositions and tangent spaces. Journal of Classification 16: 197–223.

Rohlf FJ, Marcus LF (1993) A revolution in morphometrics. Trends in Ecology and Evolution 8: 129–132.

Rohlf FJ, Slice DE (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. Systematic Zoology 39: 40–59.

Rohlf FJ, Sokal RR (1965) Coefficients of correlation and distance in numerical taxonomy. University of Kansas Science Bulletin 45: 3–27.

Schindel DE (1990) Unoccupied morphospace and the coiled geometry of gastropods: Architectural constraints or geometric covariation? In: Causes of Evolution (Ross RA, Allmon WD, eds), 270–304. Chicago: University of Chicago Press.

Simpson GG (1944) Tempo and Mode in Evolution. New York: Columbia University Press.

Slice DE (2001) Landmark coordinates aligned by procrustes analysis do not lie in Kendall's shape space. Systematic Biology 50: 141–149.

Slice DE (2005) Modern Morphometrics in Physical Anthropology. Dordrecht, the Netherlands: Kluwer.

Small C (1996) The Statistical Theory of Shape. New York: Springer.

Sneath P, Sokal R (1973) Numerical Taxonomy. San Francisco, CA: Freeman.

Sokal RR (1961) Distance as a measure of taxonomic similarity. Systematic Zoology 10: 70–79.

Stadler PF (2002) Fitness landscapes. In: Biological Evolution and Statistical Physics (Lässig M, Valleriani A, eds), 187–207. Berlin: Springer.

Stadler BMR, Stadler PF (2004) The topology of evolutionary biology. In: Modeling in Molecular Biology (Ciobanu G, Rozenberg G, eds), 267–286. Berlin: Springer.

Stadler BMR, Stadler PF, Shpak M, Wagner GP (2002) Recombination spaces, metrics, and pretopologies. Zeitschrift für Physikalische Chemie 216: 217–234.

Stadler BMR, Stadler PF, Wagner G, Fontana W (2001) The topology of the possible: Formal spaces underlying patterns of evolutionary change. Journal of Theoretical Biology 213: 241–274.

Suppes P, Krantz DH, Luce RD, Tversky A (1989) Foundations of Measurement, Vol. II: Geometrical, Threshold, and Probabilistic Representations. New York: Academic Press.

Thomas RDK, Reif W-E (1993) The skeleton space: A finite set of organic designs. Evolution 47: 341–360.

Thomas RDK, Shearman RM, Stewart GW (2000) Evolutionary exploitation of design options by the first animals with hard skeletons. Science 288: 1239–1242.

Thompson DAW (1917) On Growth and Form. Cambridge: Cambridge University Press.

von Baer KE (1828) Entwicklungsgeschichte der Thiere: Beobachtung und Reflexion. Königsberg, Germany: Bornträger.

Wagner GP, Altenberg L (1996) Complex adaptations and the evolution of evolvability. Evolution 50: 967–976.

Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the Sixth International Congress of Genetics 1: 356–366.

Wright S (1988) Surfaces of selective value revisited. American Naturalist 131: 115–123.