

PalaeoMath 101

Size & Shape Coordinates

In the last column I tried to show how the same information we had captures previously using linear distances between landmarks could be captured, summarized, and used for ordination studies based on the coordinate positions of the landmarks themselves. Actually, because these coordinate positions are linked to a common reference—the origin of the coordinate system—any localized feature represented by a landmark is automatically located in both absolute and relative senses to all other features described by landmarks. In other words, keeping your data in the landmark mode of representation allows you to access all contrasts between all landmark locations simultaneously.

One could, of course, ask ‘Why not simply measure all distances between all landmarks and use that as the basis for your analysis?’ Indeed, a morphometric approach exists—Euclidean Distance Matrix Analysis (EDMA, see Lele and Richtsmeier 1991) that employs precisely this strategy. There’s been quite bit of controversy about EDMA in the morphometric literature and this isn’t the place to review those issues (perhaps I’ll, do that in a future column). But one classic and practical concern has to do with the geometries of completely specified measurement networks. Figure 1 shows a completely specified distance network for a landmark set defined for the *Calymene* specimen from the trilobite dataset.

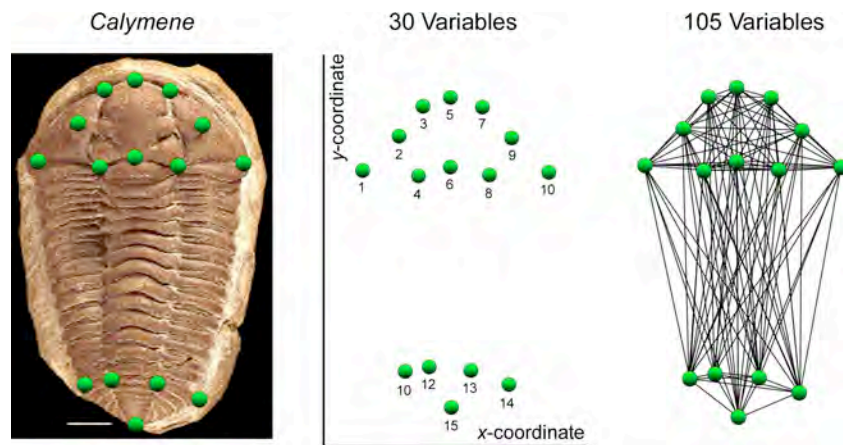


Figure 1. Alternative approaches to the quantification of shape using landmarks. Right: Digital image of a *Calymene* specimen with the locations of 15 landmarks superposed (scale bar = 7.87 mm). Centre: representation of its form using positions of landmark points in a Cartesian coordinate system. This representation specifies the relative locations of all points exactly and requires 30 variables. Right: representation of the same form using all linear distances between all landmarks. This representation requires 105 variables and contains much measurement redundancy.

Obviously the coordinate-point representation is much more compact than the inter-landmark distance representation. The former requires only 30 variables for a 15-landmark set (the x and y coordinate values) whereas the latter requires 105 variables to capture the same information. But aside from this, the coordinate-point representation is better in that much of the extra information specified by the complete distance network is redundant (e.g., the distance from the tip of the pygidium to the most anterior glabellar landmark is much the same regardless of whether it’s being measured to the right or left sides). Also, unlike the coordinate-point data, the scalar values representing inter-landmark distances carry no geometric information about relative landmark placement. Knowing the distance from landmark 1 to landmark 2 is 2.33 mm places landmark 2 somewhere along a circle centred at landmark 1 with a radius of 2.33 mm, but does not indicate where landmark 2 is located on that circle. However, if landmark 1 is located at coordinate position (7.62 mm, 8.16 mm) and landmark 2 at position (9.86 mm, 7.53 mm) the placement of these two points, and the morphological features they represent, has been established precisely.

What we now want to do is develop some means of comparing sets of landmark points with one other that gives us maximum control over the factors responsible for form variation.

There are four such factors: position, orientation, scale, and shape. In the last essay I showed you an easy way of gaining control over the positional and orientational aspects of different landmark sets. In 1986 Fred Bookstein introduced a simple modification to these equations that allowed sets of landmarks to be brought into common alignment in terms of position, orientation, and scale. It's a little appreciated fact these days, but that 1986 article started the geometric morphometrics revolution.

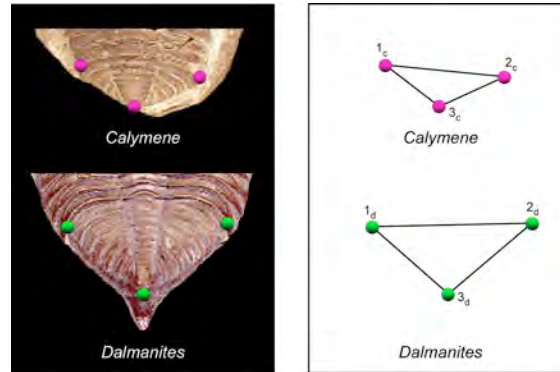


Figure 2. Landmarks used to define triangles that summarize the gross form of *Calymene* and *Dalmanites* pygidia.

To illustrate his method consider the trilobite pygidium (Fig. 2). The pygidium is a roughly triangular structure whose gross shape, in most cases, can be estimated by specifying three landmarks, two at the lateral maxima on either side of the axial lobe and a third at the distal terminus. This suits our illustrative purposes nicely as a triangle is the simplest geometric figure to have a complete form; to have position, orientation, size, and a shape. Points have positions. Lines have positions, orientations, and sizes (lengths). But only triangles and more complex polygons have all four descriptive form features.

Table 1. Cartesian coordinates of triangles shown in Figure 2.

| Vertex | x-coordinate | y-coordinate |
|--------|--------------|--------------|
| 1_c | 1.94 | 7.77 |
| 2_c | 4.98 | 7.47 |
| 3_c | 3.30 | 6.73 |
| 1_d | 1.63 | 3.65 |
| 2_d | 5.68 | 3.77 |
| 3_d | 3.53 | 1.97 |

Table 1 gives the coordinate positions of the two sets of triangle vertices shown in Figure 2. The tools we developed in the last column are sufficient to match these triangles along a user-selected axis or baseline. Bookstein (1986) developed the following equations that not only accomplish this operation in a more compact form, but that also adjust the sizes of the triangles via rigid scaling of a baseline to a unit value.

$$\eta_1 = \frac{(2_x - 1_x)(3_x - 1_x) + (2_y - 1_y)(3_y - 1_y)}{(2_x - 1_x)^2 + (2_y - 1_y)^2} \quad (15.1)$$

$$\eta_2 = \frac{(2_x - 1_x)(3_y - 1_y) - (2_y - 1_y)(3_x - 1_x)}{(2_x - 1_x)^2 + (2_y - 1_y)^2}$$

In this equation the denominator of the ratios contains the landmarks that define the baseline. For our triangles the most reasonable baseline choice would be the chord joining landmarks 1 and 2. Accordingly, we can use the equations in 15.1 to analyse the data in Table 1 without having to adjust the notation. Results of these calculations (see the *Palaeo-math-101-2* spreadsheet) are shown in Figure 3.

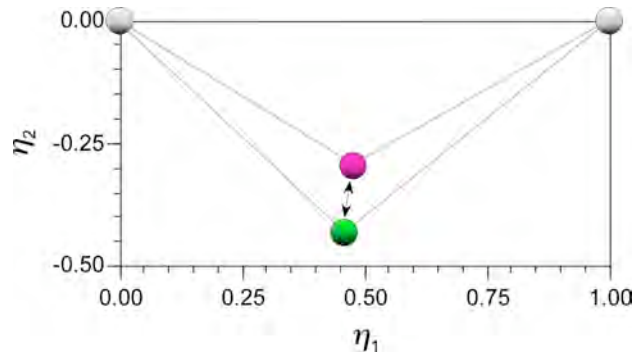


Figure 3. Results of correcting the triangular pygidial shapes shown in Figure 2 for position, orientation and scaling using the baseline (Bookstein) shape coordinate method. Note all shape differences are subsumed in the position of the free (non-baseline) landmark. Color codes for genera as in Figure 2.

The comparison illustrated in Figure 3 accords well with our intuition based on a traditional qualitative comparison of the pygidial triangles in Figure 2. The *Calymene* pygidial shape is shallower than that of the *Dalamanites*, at least for these specimens. But note also that we have now removed all variation between the triangles due to position, orientation, and size. What we are left with is a summary of variation due solely to shape differences.

Naturally, variation also exists in the positions of the baseline landmarks (1 and 2). But because these landmarks serve as the basis of the standardization of position, orientation, and scale, all the shape-difference information in this comparison has been focused on the positional difference in the single non-baseline landmark, landmark 3. This might strike some as arbitrary (what if you're interested in knowing about patterns of variation at landmarks 1 and 2?), but it does have the advantage of greatly simplifying an otherwise complex pattern of variation at three locations into a contrast between two points in space. For this simple system the vector between the *Calymene* and *Dalamanites* positions for landmark 3 in the shape space quantifies how much the forms differ, the overall directions of the difference, and suggests a simple procedure through which one shape can be transformed into the other.

The η values on each axis in Figure 3 represent new variables that have been produced for sets of landmarks once the effects of position, orientation, and scale have been removed. Since shape is what these variables express, they are termed 'shape variables'. Perhaps the best way to think of them are transformations of the original coordinate values in which we've emphasized one aspect of the information present in those original values (shape differences) by removing the effects of the other three.

Because we've used a baseline between landmarks to correct for position, orientation and scaling, and because Bookstein (1986) developed this approach to shape-coordinate calculation, what we've calculated in the *Palaeo-math-101-2* spreadsheet to this point and, drawn in Figure 3, are the Bookstein shape coordinates. If we had more than three landmarks in our system we'd still select a baseline and then use the equations in 15.1 to calculate the Bookstein shape coordinates for all the non-baseline landmarks. Similarly, if we had more than a two pygidia in our sample we'd be able to plot—and so make comparisons between—a larger number of pygidia in the Bookstein shape-coordinate space. Figure 4 shows these shape coordinates for all the images in our trilobite dataset that include pygidia.

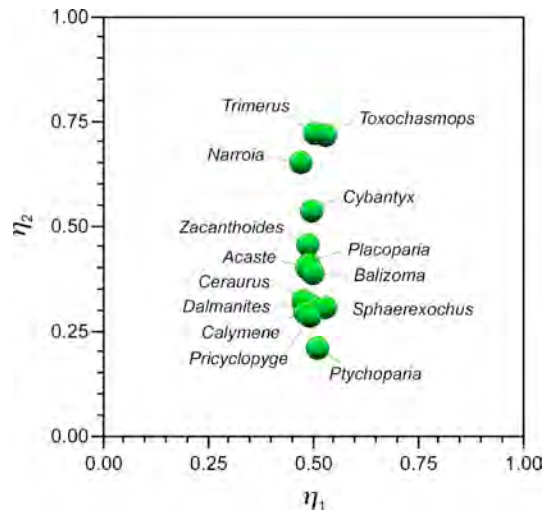


Figure 4. Ordination of the pygidial shape variation in 14 trilobite genera using Bookstein shape coordinates. See text for discussion.

In this ordination the free coordinates form a trace up the center of the plot because, on the whole, trilobite pygidia are bilaterally symmetrical. *Ptychoparia* exhibits the most flattened pygidial shape in this sample, *Trimerus*, the deepest. Although the distribution in the shape space appears quasi-continuous, in places there is a suggestion that some marked gaps in the shape distribution may also be present. For example, a gap seems to be present between shallow pygidium of *Ptychoparia* and all other genera, between a set of genera with deep pygidial shapes (*Trimerus-Toxochasmops-Narroia*) and all other genera, and between the intermediate pygidial shape of *Cybantyx* and all other genera. Two more diverse groups of genera with moderately deep pygidia are also evident on this plot. If the gaps between these putative shape groups remained after additional sampling they could be used to more objectively and reproducibly assign these genera to pygidial shape categories, say for a phylogenetic analysis (see MacLeod 2002 for further discussion of this approach to character-state recognition).

To this point we've ignored size variation in the context of shape coordinates. In previous columns dealing with traditional multivariate analysis we've seen that size and shape are often interwoven in complex ways. The mathematical definition of shape is 'the factor that remains after position, orientation, and scaling factors have been removed.' But that definition begs the question "what is variation due to scaling?"

As was noted in the previous column on allometry and PCA (Newsletter 59), the best conceptual definition of size change is an increase or decrease in the magnitude of linear distances between features (= landmarks) that occurs at the same rate in all regions of the form. Shape change, then, is a localized increase or decrease in the magnitude of linear distances between features. But regardless of how it's defined conceptually, there are at present, and will likely remain, a multiplicity of operational ways 'size' can be measured or represented (e.g., volumes, weights, distances, combinations of distances). In terms of Bookstein shape coordinates (equations 15.1), the relevant scaling factor is provided by the absolute length of the baseline. Curiously, despite the clear implications of his shape-coordinate method the specification of size, Bookstein (1986) proposed a radically different morphometric size index: centroid size (*S*).

Since 1986 centroid size has been defined in several different ways. Bookstein originally described it as "the sum of all squared distances between landmarks in pairs" (p. 190). This value was deemed statistically equivalent to "the sum of distances from each landmark to their joint centroid, each distance weighted by its own sample mean." (p. 190). Later in that same article Bookstein introduced the concept of taking the square root of *S* in order to linearize the index and place it in the same units as the original coordinate values. Later authors (e.g., Zelditch *et al.* 2004) have tended to define 'S' as the square root of the sum of squared distances of all landmarks from their joint centroid, despite the inevitable confusion this causes with the older literature. To avoid that pitfall I'll use the symbol *S'* for the non-weighted

sum of squared distances from the joint centroid and ‘root centroid size’ (RCS) to describe the square root of this convenient morphometric size index. The RCS corresponds to the following equation.

$$RCS = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 + (y_i - \bar{y})^2} \quad (15.2)$$

In this equation n corresponds to the number of landmarks.

The *Palaeo-math-101-2* spreadsheet calculates all these size indices for the trilobite pygidial dataset shown in Figure 4 to demonstrate they are all very highly correlated with one another. This gives empirical support to the assertion (made explicitly in Bookstein 1986, but not discussed at all in Zelditch *et al.* 2006) that the concept behind each is the same. All three size indices proposed for use in geometric morphometrics represent size as the sum of distances between landmark points, thus operationalizing the network shown in Figure 1 as an appropriate procedure for estimating size (but not shape).

Both Bookstein (1986) and Zelditch *et al.* (2006) claim these size measures are uncorrelated with shape. This is correct, but perhaps in more subtle manner than it first appears. Zelditch *et al.* (2006) in particular confuses matters for many readers by couching their discussion of centroid size in terms of isometric shape change. The fact is the centroid size concept—and the centroid size computation—is entirely agnostic when it comes to the question of isometric or allometric shape change. There is no way to tell from the outset whether landmarks used as the basis for the centroid size computation experience little, a moderate amount, or a great deal of allometric change over a sample. Consequently, centroid size is not an inherently isometric size index (and so uncorrelated with shape change for that reason). Rather, what Bookstein (1986) means when he speaks of centroid size being uncorrelated with shape is simply that any shape described by the same number of landmarks may be compared in terms of its size using the centroid size index. In this sense the ‘shape’ of the form being quantified is, in a sense, irrelevant from the point- of-view of centroid size.

Centroid size is clearly a better size index than any index constructed from an arbitrary subset of landmarks collected on a form. As shown in the *Palaeo-math-101-2* spreadsheet, the *RCS* index is conceptually synonymous with (but of course not computationally equivalent to) coordinate-based or distance-based size indices that employ all landmarks to obtain estimate. This is what distinguishes the *RCS* from the size estimates we employed in the previous columns on regression and multivariate analyses. In those cases we are always dealing with an arbitrarily selected subset of all possible distances between landmark points. the *RCS* differs from these in that all information from all regions of the form is employed in the size estimate. But does this mean centroid size always corresponds to our intuitive notion of what size represents?

Because centroid size is a obtained via summation, it represents a theoretical range of values that is not only unbounded, but guaranteed to increase if the number of landmarks used to estimate it increases.¹ This leads to some awkwardness and plainly counterintuitive results. Take, for example, the three identical triangles shown in Figure 5.

¹ I thank Jonathan Krieger for pointing this fact out to me originally and hope he will publish a more complete review of centroid size than I have had space to do here.

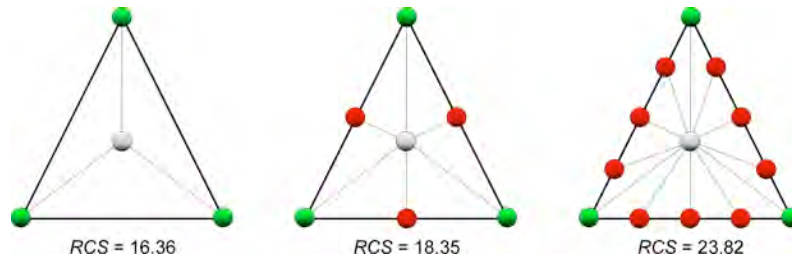


Figure 5. Unbounded nature of the RCS index. The two triangles have exactly the same dimensions. However, when the RCS index is used to estimate their size this value is tied to the number of landmarks used to represent the form. Green = landmarks, White = centroid. Red = constructed landmarks.

In one landmark has been placed at the three vertices. In another, additional, constructed landmarks have been placed at the midpoints of the sides. And in the last these half-side chords have been further subdivided into equal-length segments. If the RCS (or S , or S') is calculated for these three landmark sets the size values will differ, substantially. Note neither the total lengths of the sides have changed, nor the area of the triangle. Moreover, the magnitude of the differences between size estimates will depend entirely on how many landmarks are used to represent the form and on the placement of the landmarks relative to the centroid, even for forms that have exactly the same dimensions. Now, provided the landmark points chosen are good landmarks (= unambiguously locatable positions on the forms that are strictly comparable to one another in a biological sense), and provided you are comfortable with the idea of using these, and only these positions, to estimate body size, there is no problem. But if your landmarks are less precisely determined (= represent approximations of the positions of questionably comparable features), or if the idea of using only a few landmarks to determine renders the estimate of body size more a caricature than an observation, they may well be a problem. In these cases, the independence of centroid size from shape is not all that clear cut. To illustrate this using the opposite of Figure 5, consider the comparisons shown in Figure 6.

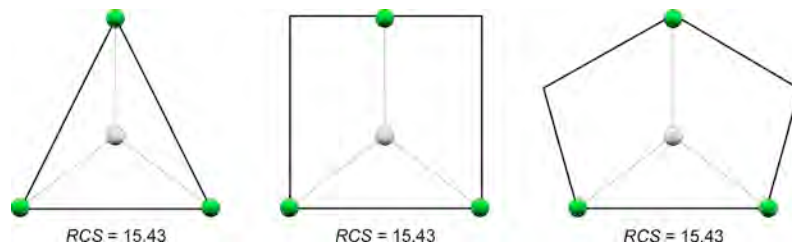


Figure 6. Forms that would be represented as having the same size as measured by a three-landmark centroid size estimate.

These idiosyncrasies of the centroid size index should be kept in mind when designing landmark sets that will be used to estimate size and shape in morphometric studies and when comparing RCS values for different landmark sets. Suffice it to say, there is no 'perfect' size index and the decision as to which of the many size indices is most appropriate for a particular study will, in most cases, depend on the details of the forms being investigated and the purposes of the study.

Turning now to a consideration of the relation between size and shape for the trilobite pygidial data we can test the allometry model by performing a multiple linear regression of the two shape variables (η_1 , η_2) on RCS (see the *Palaeo-math-101* column in *Newsletter 58* for a discussion of multiple linear regression). The 3D scatter plot for this regression is shown in Figure 7 and the regression ANOVA in Table 2.

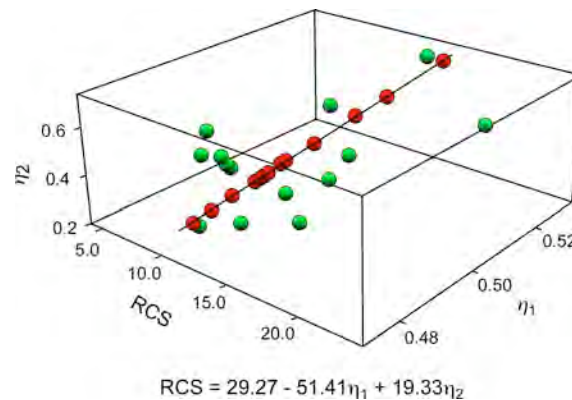


Figure 7. Multiple linear regression scatterplot of the two trilobite pygidial shape coordinates (η_1 , η_2) on root centroid size (RCS). Green = observed values. Red = predicted values.

Table 2. ANOVA results for multiple linear regression of shape coordinates on size.

| Source | DoF | SSQ | Mean Squares | F |
|------------|-----|---------|--------------|-------|
| Total | 13 | 345.676 | | |
| Regression | 2 | 141.269 | 70.634 | 3.801 |
| Error | 11 | 204.407 | 18.582 | |

Looking up the critical value of the F statistic we find the regression is just slightly non-significant at the traditional 95 percent confidence level ($\alpha = 0.056$), but close enough to be interesting. Inspection of further statistics for this regression indicates that η_1 does not exhibit a significant partial regression slope ($t = -0.747$), but η_2 does ($t = 2.702$). The fact that there is a distinction between the two variables is obvious from Figure 4, but still it's nice to see the significance of η_2 confirmed.

Lastly we can use Bookstein shape coordinates to obtain an picture of relations between taxa in a form space (size + shape) or in a size-free shape space. For this analysis let's go back to an analysis of the trilobite cranidia, using the landmarks for that structure shown in Figure 1. To construct the form matrix we would simply select a baseline (e.g., anterior and posterior glabellar mid-line landmarks), calculate the shape coordinates for all non-baseline landmarks, decide whether we wanted to include information about cranial left-right asymmetry (if not we should either use only landmarks from the right or left sizes, or possibly reflect one side onto the other and then average the corresponding landmarks), and submit the resulting matrix with a size variable (= form space) or without (= shape space) to a covariance-based PCA (see *Palaeo-math-101* column in *Newsletter* 59 for a discussion of PCA). Results of the first two shape axes for an analysis that averaged left and right landmarks to correct for single-side asymmetry is shown in Figure 8 and Table 3.

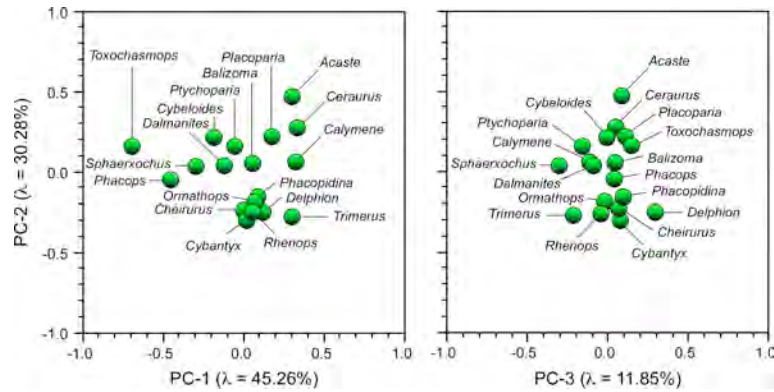


Figure 8. Results from a principal component analysis of non-baseline Bookstein shape coordinates for 18 trilobite crania (see Fig. 1 for an illustration of the landmarks used).

Table 3. Principal component loadings for cranial shape-coordinate variables.

| Landmarks | PC-1 | PC-2 | PC-3 |
|-----------|--------|--------|--------|
| 1x | 0.841 | 0.042 | 0.494 |
| 1y | -0.176 | 0.797 | 0.040 |
| 2x | 0.240 | 0.105 | -0.203 |
| 2y | -0.129 | 0.459 | 0.388 |
| 3x | 0.252 | 0.282 | -0.486 |
| 3y | -0.297 | -0.083 | 0.561 |
| 4x | 0.183 | 0.223 | -0.107 |
| 4y | -0.043 | -0.061 | 0.029 |

It's instructive to compare Figure 8 with Figure 3B from the previous column (Newsletter 68). In that result we had corrected an analogous set of trilobite landmark data for position and orientation, but not for scale. Obviously the inclusion of size matters a great deal in terms of overall partitioning of the observed shape variance. But more to the point, we have now developed a tool that can partition size and shape much more cleanly in terms of the conceptual distinctions between the two, and much more elegantly in terms of the mathematics. Best of all, it makes a real difference when we do this as patterns not evident in the previous analysis have been revealed here.

In particular, note the tight cluster of taxa with scores close to the lower limit on PC-2 axis, consisting of *Phacopidina*, *Delphion*, *Trimerus*, *Rhenops*, *Cybantyx*, *Cheirurus*, and *Ormathops*. While these specimens have very different sizes—*Trimerus* is the largest specimen in the sample, *Ormathops* the smallest—our shape coordinate results suggest this group shares an underlying shape similarity that we have not seen previously. What is this similarity?

Inspection of the loading table (Table 3) shows that landmark 1 contributes the most to total shape variance. This marks the position of the posterior peripheral terminus of the free cheek. The importance of this characteristic can be graphically assessed by plotting the shape coordinates and labelling them to show (1) the scatter of points for each landmark location (2) the identification of specimens as belonging to the putative group of taxa listed above on the basis of a qualitative inspection of the PCA optimized shape space (Fig. 9).

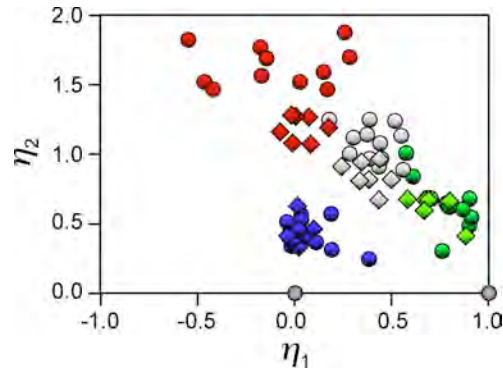


Figure 9. Plot of averaged crandial shape coordinates. Red = landmark 1; white = landmark 2, green = landmark 3, blue = landmark 4, gray = baseline landmarks (see Fig. 1 for landmark locations referenced to morphology). Diamond symbols mark points included in the *Phacopidina*, *Delphion*, *Trimerus*, *Rhenops*, *Cybantyx*, *Cheirurus*, and *Ormathops* putative subgroup. See text for discussion.

Note that landmark 1 exhibits the greatest variance of all the non-baseline landmarks and that it shows a marked separation between those specimens with fixed cheeks whose distal, lateral, posterior termini are located relatively close to the baseline, and those whose distal margins are located further away from the baseline. Landmark 1, along with landmark 2 where the same pattern is developed but occupies a smaller range of variation, appear to be the two most important sources of shape variability over the second shape coordinate (η_2). With respect to η_1 , landmarks 1 and 3 exhibit the greatest range of shape variation with landmark 1 variation predominating. Taken together it is clear that the pattern of loadings in Table 3 reflects these aspects of variation in the shape coordinate data in a simple, straight-forward, and highly informative manner.

Shape coordinates represent a fundamental part of what distinguishes geometric morphometrics from previous approaches. Bookstein shape coordinates were the first type of shape coordinates to be formulated and much of the early theoretical work in geometric morphometrics was inspired by experiments performed using them. These days the term ‘shape coordinate’ has become more-or-less synonymous with a different approach to shape coordinate computation, one that we will discuss in the next column in this series. Nevertheless Bookstein-style shape coordinates continue to be employed in several different contexts, in particular studies that employ morphometric approaches in the analysis of ontogenetic series (e.g., Webster et al. 2001; Kim et al. 2002).

Relative to the ‘other’ sort of shape coordinates, the original Bookstein formulation is mathematically very simple to compute and highly useful in a wide range of situations. Complications do arise when using Bookstein shape coordinates, most notably with respect to the fact that shape variation in the baseline coordinates is transferred to the non-baseline shape coordinates, sometimes in complex ways. A conceptual distinction also exists between this approach to shape specification and use of the centroid size index for size specification. This distinction needs to be kept in mind when using Bookstein shape coordinates and centroid size in the same study, as do the more counterintuitive aspects of the centroid size index in general. However, as I hope I’ve shown, Bookstein shape coordinates and centroid size are good places to begin an exploration of what geometric morphometrics is all about and why it marks such a radical departure from the previous distance-based morphometric approaches.

Norman MacLeod
Palaeontology Department, The Natural History Museum
N.MacLeod@nhm.ac.uk

REFERENCES

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

- Kim, K., H. D. Sheets, R. A. Haney, and J. T. Mitchell.** 2002. Morphometric analysis of ontogeny and allometry of the Middle Ordovician trilobite *Triarthrus becki*. *Paleobiology*, **28**(364–377).
- Lele, S. and J. T. Richtsmeier.** 1991. Euclidean Distant Matrix Analysis: a coordinate free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology*, **86**, 415–427.
- MacLeod, N.** 2002. *Phylogenetic signals in morphometric data*. In N. MacLeod and P. L. Forey, eds. *Morphology, shape and phylogeny*. Taylor & Francis, London. 100–138 pp.
- MacLeod, N.** 2005. Regression 4: Going Multivariate. *The Palaeontological Association Newsletter*, **58**, 44–53.
- MacLeod, N.** 2005. Principal components analysis (eigenanalysis & regression 5). *Palaeontological Association Newsletter*, **59**, 42–54.
- Webster, M., H. D. Sheets, and N. C. Hughes.** 2001. *Allometric patterning in trilobite ontogeny: testing for heterochrony in Nephrolemlus*. In M. L. Zelditch, ed. *Beyond heterochrony*. John Wiley & Sons, New York. 105–142 pp.
- Zelditch, M. L., D. L. Swiderski, H. D. Sheets, and W. L. Fink.** 2004. *Geometric morphometrics for biologists: a primer*. Elsevier/Academic Press, Amsterdam. 443 pp.

Don't forget the *Palaeo-math 101-2* web page, now at a new home at:
http://www.palass.org/modules.php?name=palaeo_math&page=1