# A DESCRIPTION OF THE GENERATING CURVE OF BIVALVES WITH STRAIGHT HINGES

by MARGARET JENNIFER ROGERS

ABSTRACT. A method of describing the whole of the generating curve of a lamellibranch is sought. When lengths and angles are used to describe an outline, much of the outline remains undefined. A curve can be fitted to an outline and the coefficients in the particular approximation employed then define the outline. Reasons are given for fitting a Tchebychev polynominal, rather than a spline, or Fourier series containing both sine and cosine terms. Polar coordinates r and θ are calculated for each point on a digitized outline. $\cos \theta$, rather than θ is chosen as the independent variable when the Tchebychev coefficients are calculated. It is found that about 100 irregular spaced data points are required to produce stable coefficients, and that adequate numerical accuracy is obtained when the outline is described by the first six coefficients. These six coefficients can also be used as shape discriminators. As the value of c_0 is a measure of size, it can be used to standardize the other coefficients. The standardized coefficients can be used to compare the shape of the generating curve of shells of different sizes.

DESCRIPTIONS of the morphology of organisms should, ideally, be objective, reproducible, and enable different forms to be distinguished. In this article the description of the lateral outline, or generating curve (Raup 1966) of the Carboniferous non-marine bivalve genera *Naiadites* and *Curvirimula* is discussed. The general form of the generating curve is elliptical. A point (the origin of growth, or tip of the umbo) and a line (the hinge-line, which is straight, or only slightly curved in the genera studied) define the orientation of these shells to which further descriptions can be related. A coarse separation of shapes can be made qualitatively by simple terms such as outline approximately triangular (text-fig. 1a), rectangular (text-fig. 1b), or semicircular (text-fig. 1c, d). The distinction between text-fig. 1c and text-fig. 1d is more difficult to describe qualitatively. Measurements of various lengths and angles have proved useful in describing and distinguishing shell forms
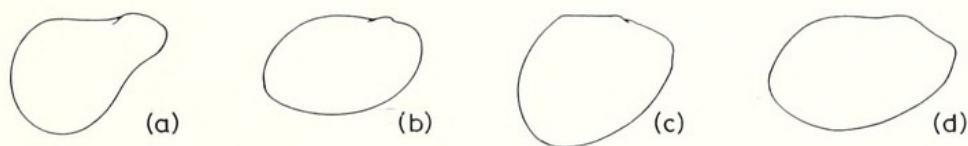
TEXT-FIG. 1. Outlines of four shells which can be described qualitatively as approximately (*a*) triangular, (*b*) rectangular, (*c* and *d*) semicircular.

(Davies and Trueman 1927; Deleers and Pastiels 1947; Trueman and Weir 1955; Eager 1973; Hajkr, Lukasova, Ruzicka, and Rehor 1974), but only a few points on the shell outline are defined. As Bookstein (1978) notes, techniques which use 'landmarks', inter-landmark distances, and angles ignore much of the shape of the outline. Between the landmarks the outline is undefined and cannot be reproduced.

Papin and Khoroshev (1974) measured the radius of curvature (text-fig. 2, rc_i) of successive parts of the outline, indicating that sections of the curve should be described, rather than interpoint distances. The value of the radius of curvature at P_i, when calculated from the position of three successive points P_{i-1}, P_i, P_{i+1} (text-fig. 2) on the curve, is highly sensitive to the actual position of the points P_{i-1}, P_{i+1}. Accordingly, it is virtually impossible to obtain a reliable estimate of this
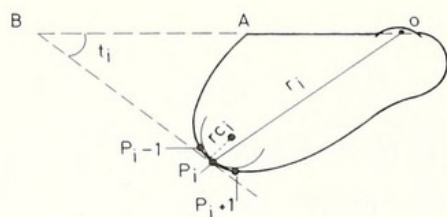
TEXT-FIG. 2. Measurements on a shell outline. The hinge OA is produced to B. P_{i-1}, P_i, P_{i+1} are three consecutive points on the outline. rc_i is the radius of curvature, t_i the tangent angle, and r_i the distance from O of point P_i.

quantity. For this reason the radius of curvature, and a related measurement, the tangent angle t_i (text-fig. 2) are not suitable as shape descriptors. Error in the measurement of the length r_i (text-fig. 2), from the umbo to the point P_i is due to the accuracy of the measuring device used and is not related to the shape of the curve at P_i. It is essential to choose a property of the curve which can be measured accurately as a basis of a reproducible description. The length r_i satisfies this condition.

With the advent of more sophisticated devices such as a digitizing table, the (x, y) coordinates (text-fig. 6) of many points on an outline can be recorded with ease. A large number of unequally spaced data points will yield an accurate reproduction. If the recorded points are used to redraw the outline (text-fig. 3), a few equally spaced data points may not reproduce the outline accurately. Although a large number of data points produce an accurate reproduction, for the purposes of comparison it is necessary to look for a more economical way in which to represent the outline. A variety of curve-fitting techniques exist which allow a numerical approximation to the fossil outline to be derived. A set of numbers which represent the coefficients in the particular approximation employed can then define the outline. For reasons which are discussed below, the curve-fitting method chosen here is the Tchebychev polynomial. The purpose of this article is to determine the reliability of the Tchebychev coefficients as shape descriptors and discriminators in the practical context of bivalves with a straight hinge.
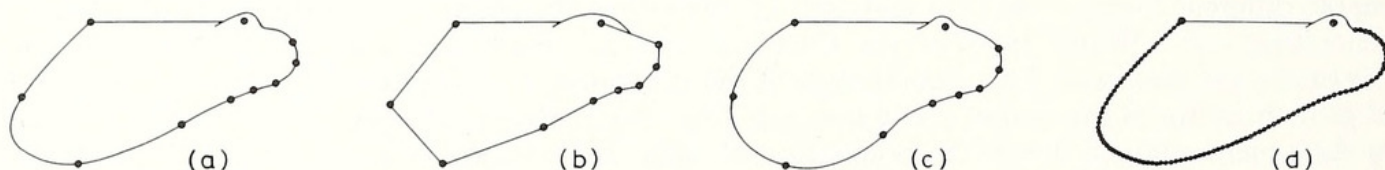
TEXT-FIG. 3. Reproduction of shell (a) from data points taken at 20° intervals and joined by (b) curved and (c) straight lines. (d) is reproduced from many unequally spaced data points.

CURVE FITTING TECHNIQUES

All descriptive, curve fitting techniques seek a solution to the equation

$$y = f(x) \text{ over a specified range of } x$$

where x is the independent variable and y the dependent variable. A length, r_i (text-figs. 2 and 6), called the radial length below, can be measured with reasonable accuracy and is a suitable dependent variable. The angle θ_i (text-fig. 6) at the origin between the radial length and the reference line AOB is an appropriate independent variable since, once this angle is known, the value of the dependent variable is precisely defined.

With these coordinates, the equation representing a circle is $r = a \cos \theta$, but the outline of a bivalve shell is, in general, too complex to be represented by such a simple equation. Accordingly, techniques such as splines and Fourier series have been used. A spline curve may be fitted to a series of data points provided there are no marked inflections. If such inflections occur, the data points are chopped up into segments at the turning-points. These turning-points are called knots. Suitable positions for knots (x) are shown on two shell outlines in text-fig. 4. Each part of the shells would be accurately described by the coefficients of the appropriate spline function, but these coefficients cannot be used to compare the shapes of the shells (which may belong to the same species) because each function relates to a different range of the independent variable.

TEXT-FIG. 4. Fitting spline curves to a shell outline. **x** mark the position of knots.
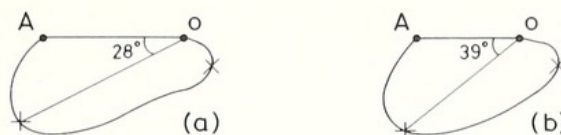
Fourier series coefficients have been used to define the outlines of ostracodes (Kaesler and Waters 1972), Bryozoa (Anstey and Delmet 1973), blastoids (Waters 1977), and sand grains (Ehrlich and Weinberg 1970). As yet palaeontologists cannot readily understand what outline is being described by a particular set of coefficients. This led Scott (1980), in his study of Foraminifera, to calculate radii from the origin to the outline at $10°$ intervals from the Fourier series representation of an outline. The thirty-six measurements allow the outline to be visualized, but the description is not economical and the outline is undefined between the radii. A Fourier series in which both sine and cosine terms are used is periodic, the value of the independent variable ranging from 0 to $360°$, and is well suited for describing the closed curves found in the above examples, provided the origin and orientation are defined. In the case of a sand grain (text-fig. 5a) a choice of origin and orientation can be made on the basis of its geometry. The line of orientation could be its longest axis (AOB) and the origin its mid-point. In radially symmetric organisms, the centre is a point of morphologic significance and the orientation can be defined in morphologic terms. As in the case of the sand grain, the origin and orientation of a bivalve can be defined geometrically (text-fig. 5b), but it is not clear that such definitions would be homologous, or have any morphologic relevance. Biological reference points within the curve, such as muscle insertions, are rarely preserved in fossil material and are thus unsuitable. In those bivalves having a straight hinge, an origin and orientation can be chosen which are morphologically homologous and are usually visible in fossil material. Since they lie on, rather than within, the outline, a periodic function may not fit the outline most economically.
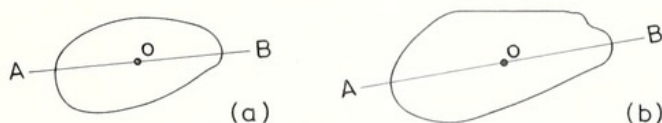
TEXT-FIG. 5. Geometric method of defining the origin and orientation of an outline. AOB = longest axis, O being the centre. (a) sand grain, (b) bivalve.

The Tchebychev series provides a function in which the independent variable lies in the range -1 to $+1$ and the NAG Library routine EO2ADF enables this function to be fitted to an arbitrary set of data points. For these reasons it was decided to use the coefficients of the polynomials of a Tchebychev series to describe the generating curve of several bivalves. Experiments were designed to discover how many data points are needed to ensure that the calculated coefficients are the same wherever the data points are taken, and by whom. Further, how many coefficients are needed to give a good representation, whether θ or a transform of θ as the independent variable gives better results; how sensitive the coefficients are to a change of range in the initial data points, and how effective they are as shape discriminators. The results of these experiments are given below.

The basic objective of this study is to create attributes (shape descriptors) which can be used to study variation within populations and as aids to classification.

METHODS OF TAKING DATA POINTS AND CALCULATING THE TCHEBYCHEV COEFFICIENTS

A digitizing table consists of a plane surface over which a pointer may be moved. In the device used for this study, the pointer consists of the intersection of a pair of cross wires. When a button is pressed the x, y coordinates of the pointer are recorded to an accuracy of 0.025 mm. The cursor must lie on the table when the coordinates are recorded. Thus the outline to be digitized must be drawn on a piece of paper, be photographed, or be a projected image. Material embedded in matrix cannot be digitized directly.

Initially, outlines of shells, scaled to be about 4 cm square, were drawn using a camera lucida. The origin of the polar coordinates system, 0, was taken as the projection of a perpendicular from the tip of the umbo on to the hinge (text-fig. 6b) and this was the first point digitized. When the specimen consisted of an external representation of the shell, this origin is concealed by the umbonal swelling and its position must be inferred from growth lines. The hinge AOB (text-fig. 6a) is the line of orientation and the posterior end of the hinge (A) was the second point digitized. Subsequent points, i.e. P_i (text-fig. 6a) were digitized from A to B.
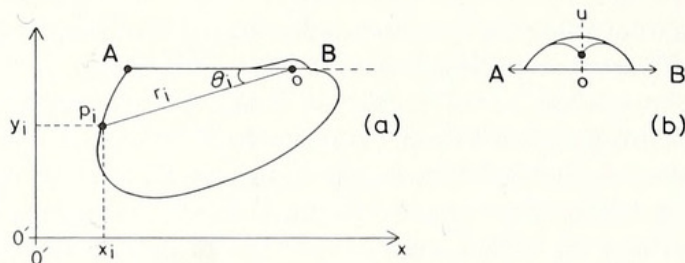
TEXT-FIG. 6. Diagram showing the orientation of a shell for digitizing. AOB = hinge line, u = tip of umbo, O = origin of the polar coordinates, O' = origin of x, y coordinates. x_i, y_i are the rectangular coordinates, and r_i, θ_i the polar coordinates of the point P_i.

The (x_i, y_i) coordinates can be used to calculate the polar coordinates (r_i, θ_i) of the point P_i. The shape is then defined by the polar equation:

$$r = r(\theta)$$

where θ is the independent variable and r is the single valued function of θ. An approximation to this function may be made by a Tchebychev series of the form

$$r = \tfrac{1}{2}c_0 T_0(v) + \sum_{m=1}^{\infty} c_m T_m(v)$$

where v, the independent variable, is θ or a transform of θ, $T_m(v)$ is the Tchebychev polynomial of degree m, and the c_m are the coefficients whose values define the particular function $r(\theta)$ in each case. A full account of the Tchebychev polynomials can be found, for example, in Fröberg (1965). The general polynomial is defined as $T_m(v) = \cos(m \arccos v)$ and the first few polynomials are

$T_0(v) = 1$
$T_1(v) = v$
$T_2(v) = 2v^2 - 1$
$T_3(v) = 4v^3 - 3v$

When $v = \cos\theta$, $T_m v = \cos(m\theta)$ and the expansion is effectively a Fourier cosine series, the form of which is

$$r = \tfrac{1}{2}b_0 + \sum_{n=1}^{\infty} b_n \cos(n\theta)$$

where b_n are the coefficients of the Fourier cosine series. In this case it is a simple matter to calculate the area of a specimen, provided θ ranges from 0 to π radians (0–180°).

$$\text{The area} = \frac{\pi}{4}\left(\frac{c_0^2}{2} + \sum_{m=1}^{\infty} c_m^2\right)$$

The coefficients were calculated using the NAG library routine EO2ADF which computes the least-squares approximation to an arbitrary set of data points. The introductory remarks to this routine states that more points should be recorded where the outline changes markedly, and also at the ends of the range. The independent variable is scaled so that its value ranges from -1 to $+1$ by the linear transform
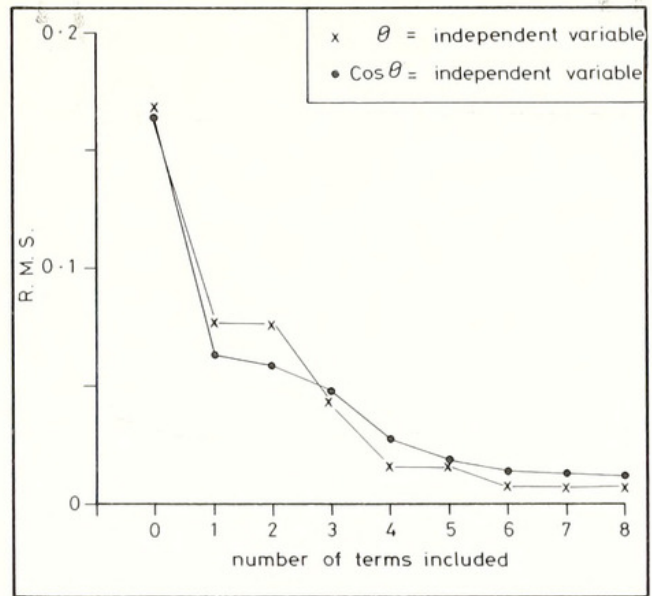
$V = (2v - v_{max} - v_{min})/(v_{max} - v_{min})$ where V = scaled independent variable
v_{max} = maximum value of v
v_{min} = minimum value of v

Thus, if θ ranges from 0 to 180° and $v = \cos\theta$ then $V = \cos\theta$.

TEXT-FIG. 7. Root mean square residual (R.M.S.) plotted against number of terms included.

The root mean square residual (R.M.S.) is a measure of the departure of the fitted curve from the original data points. Initially the residuals decrease rapidly as successive terms are used (text-fig. 7); thereafter they decrease more slowly, and indeed may increase slightly as unwanted fluctuations are produced. The number of terms in the series required to give adequate accuracy is the number after which the residuals decrease only slowly. In the case shown in text-fig. 7, six or seven terms give an adequate fit.

NUMBER OF COEFFICIENTS REQUIRED AND CHOICE OF THE INDEPENDENT VARIABLE

The polar coordinates r and θ of a digitized outline were used to calculate the coefficients of the polynomial of the Tchebychev series from degree zero to degree eight. The plots of the R.M.S. (text-fig. 7) for the two cases when $v = \theta$ and $v = \cos \theta$ are typical and show that c_2 has little effect on the accuracy of the fit of the polynomial. Thereafter the accuracy of the fit improves continuously as terms are added when $v = \cos \theta$. The R.M.S. decreases irregularly when $v = \theta$, and it is difficult to identify where significant flattening of the curve occurs. The values of the coefficients when $v = \cos \theta$ are given in Table 1. It can be seen that when a new coefficient (c_i) is added, the value of the preceding one, c_{i-1} changes considerably but the value of c_{i-2} is only slightly altered. The addition of a new coefficient can cause a marked change in the value of all the preceding coefficients when $v = \theta$.

TABLE 1. Value of the coefficients when polynomials of different degree are evaluated $v = \cos \theta$ is the independent variable

Degree, m of polynomial	Terms included								
	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8
0	0·798								
1	0·605	−0·251							
2	0·609	−0·240	0·032						
3	0·599	−0·239	0·054	0·050					
4	0·606	−0·235	0·052	0·026	−0·063				
5	0·606	−0·237	0·049	0·030	−0·048	0·032			
6	0·607	−0·237	0·048	0·031	−0·048	0·026	−0·016		
7	0·607	−0·237	0·049	0·030	−0·049	0·026	−0·013	0·007	
8	0·605	−0·237	0·050	0·030	−0·048	0·028	−0·013	0·003	−0·012
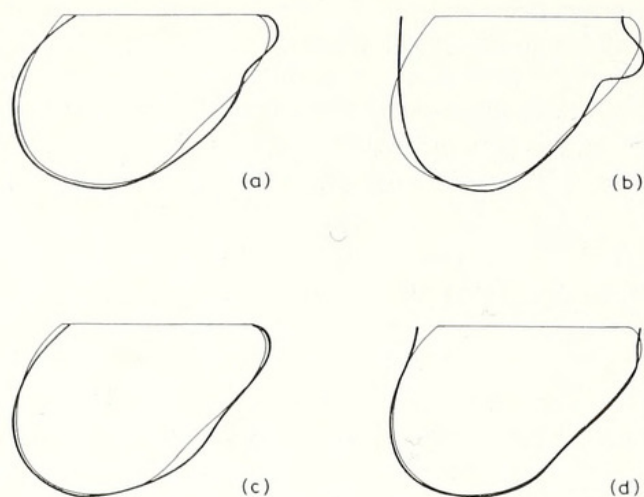
TEXT-FIG. 8. Original outline of a shell overlain by the outline reproduced using 5 terms (a, b) and 8 terms (c, d). The independent variable is θ (a, c) and $\cos \theta$ (b, d).

The outline was recreated from the coefficients of the polynominal, the polar coordinates of the fitted points being calculated using the NAG Library routine EO2AEF. When $v = \theta$ the outline is somewhat more circular than the original (text-fig. 8a, c) whereas the ends of the outline are somewhat more inaccurate when $v = \cos \theta$ (text-fig. 8b, d). This is an unavoidable consequence of using $\cos \theta$; the calculated curve must meet the hinge at an angle of 90° when $\theta = 0$ or 180°. $\cos \theta$ is preferred as the independent variable because an accurate portrayal of the ends of the outline is considered of less importance than the general shape which is recreated more accurately when $v = \cos \theta$. For this reason and because of the behaviour of the R.M.S. and the stability properties of the coefficients noted above, it was decided to adopt $\cos \theta$ as the independent variable, rather than θ.

Inspection of the residuals suggests that the five coefficients, c_0–c_4, are sufficient to give a good approximation (text-fig. 8a, b). However, it was found that the outlines recreated using the eight coefficients, c_0–c_7, look much more like *Naiadites* (text-fig. 8c, d). The three coefficients, c_5–c_7, although small in value and accounting for little of the residual error, improve the detailed representation of the outline.

NUMBER OF DATA POINTS

About 100 data points were recorded on the generating curve of one shell. A random set of thirty of these 100 points were used to calculate the coefficients of the Tchebychev polynomial to degree 8. This was done ten times for each of 30, 40, 50, 60, and 80 data points. The variability of c_0 is shown in text-fig. 9, the value of the coefficient being plotted against the number of data points used. It is seen that when thirty data points are used the value of the coefficient varies from 0·5 to 1·34 and is 0·94 when the full data set is used. About 100 data points yield a stable value of 0·94 for c_0.

The same outline drawn and digitized by two further people gave values for c_0 ranging from 0·89 to 0·94 when about 100 data points were recorded. An error of about 5% was considered satisfactory, particularly considering that either the material or the technique was unfamiliar to the participants.
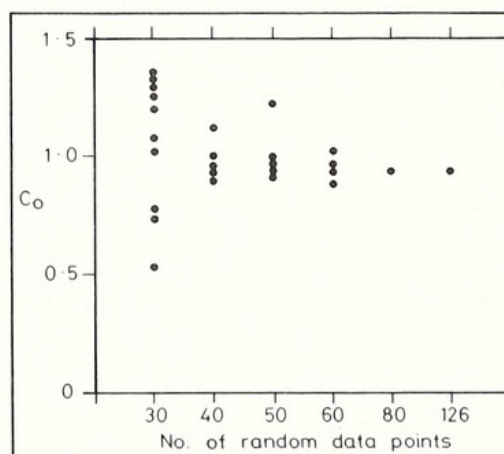
TEXT-FIG. 9. Value of c_0 plotted against number of random data points selected from a set of about 100 digitized on an outline. In each case coefficients were calculated to degree 8. When forty or more data points are used, the ten values of c_0 cannot be recorded separately at this scale.

THE EFFICIENCY OF THE COEFFICIENTS AS DISCRIMINATORS OF SHAPE

Two rather different shells (text-fig. 10) were selected, and drawn and digitized by different operators. The coefficients were calculated to degree eight for each digitized outline. When the value of the coefficients is plotted (text-fig. 10c), it is seen that certain coefficients (c_0 and c_3) have distinctly different values for the two shells, whereas other coefficients (c_1) are approximately the same for both shells. These two shells are of similar size, and the zeroth coefficient is a measure of size; it is the radius of the semicircle which fits the outline giving the smallest root mean square residual. Using this criterion as a measure of size, c_0 can be used to standardize the remaining coefficients. Plots of the standardized coefficients (c_1/c_0, c_3/c_0) (text-fig. 10c) show that the first coefficient cannot be used to distinguish the two shells, but c_3 becomes more effective as a discriminator. This result is expected, because c_1 is a crude measure of asymmetry which is refined by successive odd-numbered coefficients. Negative values of c_1 indicate that the shells are asymmetric; that is, the posterior lobe is larger than the anterior lobe. These results indicate that shapes of *Naiadites* shells of any size may be discriminated using at most seven coefficients.
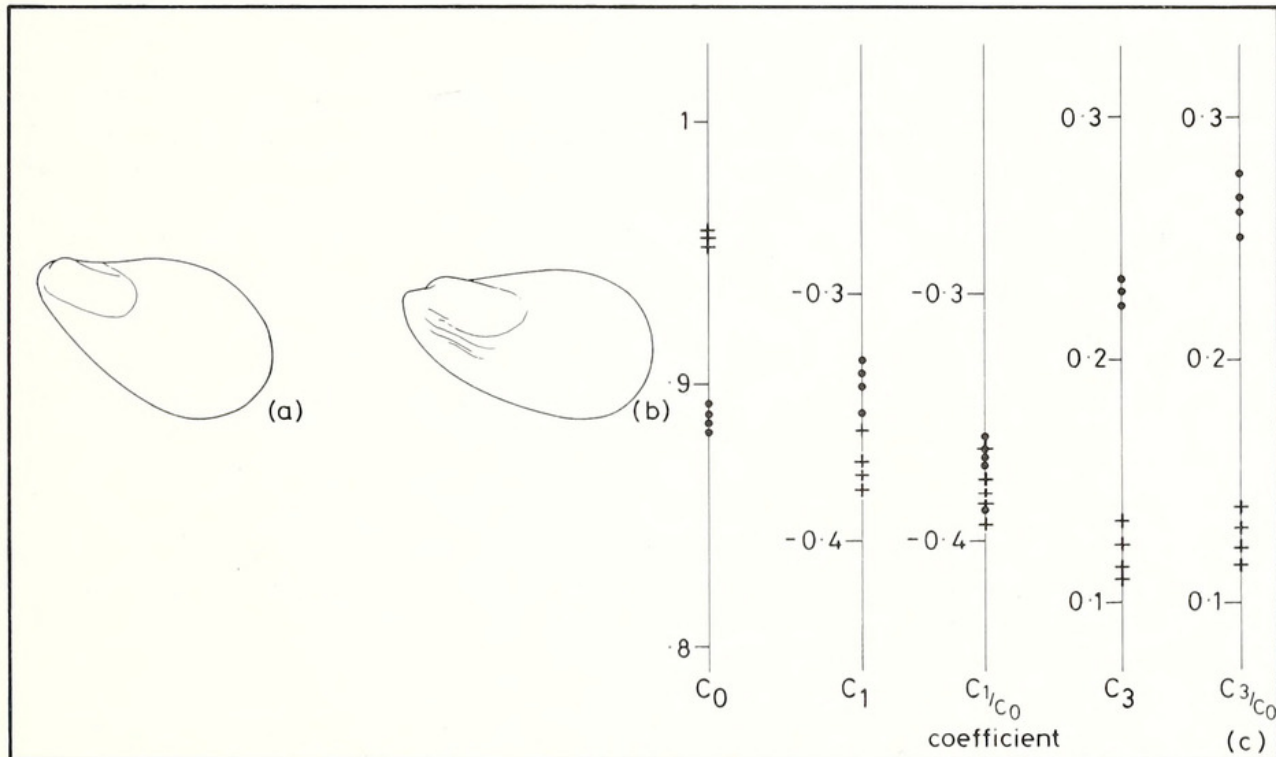
TEXT-FIG. 10. (a) and (b) two shells digitized and drawn by three different people. (c) value of the coefficients c_0–c_3 and the ratios c_1/c_0 and c_3/c_0 ● = shell a. + = shell b.

EFFECTS OF POOR PRESERVATION

As noted earlier, the origin of growth is often obscured in fossil material. However, provided the umbo is not very large, the tests made above suggest that different operators will choose approximately the same origin. A more serious problem is the fact that the anterior lobe is often distorted, and thus its outline cannot be digitized with confidence. Although in *Naiadites* the anterior lobe is small, it was found that very different interpretations of the anterior outline gave rise to markedly different coefficients. Sometimes the distortion brings the anterior lobe below the level of the hinge, thus it is not possible to measure the radial length when $\theta = 180°$. It was found that provided $v = \cos\theta$ and θ ranged from 0 to about $170°$ the coefficients obtained were similar to

those obtained when θ ranged from 0 to 180°. In reasonably well-preserved specimens of *Naiadites* it is usually possible to use a range of θ of 0–170°.

In *Curvirimula* it was found that the hinge rarely extended beyond the umbo, and often the maximum value of θ was 120°. As noted earlier, it is impossible to compare the coefficients of two curves if the ranges to which they fit are very different.

Further, not only should the range of θ be the same, but the morphologic structures described should be the same if the coefficients are to be used as discriminators. Thus, if the range of θ chosen is 0 to 120°, the coefficients calculated for one shell (text-fig. 11a) describe the whole of the generating curve, whereas those for a second shell (text-fig. 11b) fail to describe the anterior lobe, and the two sets of coefficients cannot be used to compare the generating curve of the two shells. In the genera studied the umbo and the hinge are the only structures on the outline which are homologous. However, the anterior lobe contains the anterior adductor muscles, part of the foot and its associated musculature, and the size of the anterior lobe is a good indicator of the mode of life of the organism. A suitable way of comparing these shells may be to describe the posterior lobe in terms of the coefficients which define its outline, and to describe the anterior lobe in terms of area, a parameter which can easily be calculated from the digitized points. Such a description would still be economical, and a comparison of shells described in this way would be justified on theoretical grounds. This method has not yet been tested.

TEXT-FIG. 11. Two specimens of *Curvirimula* having very different maximum value of θ.

SUMMARY AND CONCLUSION

The shape of the generating curve of those shells (including *Naiadites*) which have a straight hinge extending beyond the umbo may be described economically using the coefficients of a polynomial of the Tchebychev series of the eighth degree. Such descriptors are very stable provided that about 100 data points are used and that $\cos \theta$ is used as the independent variable. C_0 is a measure of size, and tests indicate that if it is used to standardize the coefficients of higher degree, the first five scaled coefficients can be used as shape discriminators.

Describing shells in this way has the further advantage that, because the whole of the generating curve can be reproduced, any other features of the shell such as the longest radial length can be calculated from these nine coefficients. If θ ranges from 0 to 180°, other features such as the area can be calculated directly from the coefficients.

When the hinge is not straight, or does not extend beyond the umbo, the shape can still be described, but it is difficult to specify the range of the independent variable in such a way that different species can be compared. In the case of *Curvirimula* the hinge rarely extends beyond the umbo. For this genus it is suggested that the coefficients of a very restricted portion of the generating curve, together with a further parameter, the area of the anterior lobe, which has functional significance, may be used in order to compare shells. For shells which lack a straight hinge it is possible that a periodic function may be more appropriate as a shape discriminator. The umbo could be retained for the purpose of orientation, but it may be necessary to define a geometric origin.

REFERENCES

ANSTEY, R. L. and DELMET, D. A. 1973. Fourier analysis of zooecial shapes in fossil tubular bryozoans. *Bull. geol. Soc. Amer.* **84**, 1753–1764.

BOOKSTEIN, F. L. 1978. The measurement of biological shapes and shape change. *Lecture notes in biomathematics*, **24**, 1–191.

DAVIES, J. H. and TRUEMAN, A. E. 1927. A revision of the non-marine lamellibranchs of the Coal Measures. *Quart. Jl geol. Soc.* **83**, 210–259.

DELEERS, C. and PASTIELS, A. 1947. Étude biométrique des *Anthraconauta* du Houiller de la Belgique. *Publ. Assoc. strat. Houillères*, **2**, 1–99.

EAGAR, R. M. C., 1973. Variation in shape of shell in relation to palaeoecological station in some non-marine Bivalvia of the Coal Measures of south east Kentucky and of Britain. *C.R. 7th Cong. Int. Strat. Géol. Carb., Krefeld*, 1971, **2**, 387–416.

ERLICH, R. and WEINBERG, B. 1970. An exact method for characterization of grain shape. *J. sed. Pet.* **40**, 205–212.

FROBERG, C. E. 1965. *Introduction to numerical analysis*. Reading, Mass., Addison-Wesley publ. Co., x + 340 pp.

HAJKR, O., LUKASOVA, A., RUZICKA, B., and REHOR, F. 1974. Die gattung *Citothyris* (Bivalvia) aus dem Karbon und ihr statistich-geometrisches Modell. *Freiberg, Fortschungsh.* **C306**, 1–159.

KAESLER, R. L. and WATERS, J. A. 1972. Fourier analysis of the ostracode margin. *Bull. geol. Soc. Amer.* **83**, 1169–1178.

PAPIN, YU. S. and KHOROSHEV, N. G. 1974. Numerical expression of the shape of shells outlined by a convex or convexo-concave smooth line. *Zh Paleont.* **8**, 121–124. [In Russian.]

SCOTT, G. H. 1980. The value of outline processing in the biometry and systematics of fossils. *Palaeontology*, **23**, 757–768.

RAUP, D. M., 1966. Geometric analysis of shell coiling: general problems. *J. Paleont.* **40**, 1178–1190.

TRUEMAN, A. E. and WEIR, J. 1955. The Carboniferous non-marine Lamellibranchia. *Palaeontogr. Soc. Monograph*, **8**, 207–242.

WATERS, J. A., 1977. Quantification of shape by use of Fourier analyses: the Mississippian Blastoid genus *Pentremites*. *Paleobiology*, **3**, 288–299.

NAG LIBRARY 1978. Mark 6. Numerical Algorithms Group Ltd., Oxford, U.K.

M. J. ROGERS
Department of Geology
Queen's Building
University Walk
Bristol BS8 1TR