Biological Journal of the Linnean Society (2000), *70:* 161–187. With 8 figures

doi:10.1006/bijl.1999.0395, available online at http://www.idealibrary.com on IDEAL®

Three into two doesn't go: two-dimensional models of bird eggs, snail shells and plant roots

J. M. C. HUTCHINSON

School of Biological Sciences, University of Bristol, Woodland Road, Bristol, BS8 1UG

Received 26 August 1998; accepted for publication 7 July 1999

Three published two-dimensional analyses of three-dimensional structures are reinvestigated. (1) In their 1997 paper, Barta & Székely sought shapes of bird eggs that maximized the size of eggs that packed under a circular brood patch. Inappropriately they measured egg size by cross-sectional area; maximizing volume implies different optima. Also their genetic algorithm mislocated their optima, probably because of too much mutation. (2) In this journal in 1985, Heath considered a section of an idealized snail shell; altering the overlap of adjacent whorls alters the ratio of shell material to volume enclosed. Heath located an overlap that minimized perimeter/area of the cross-section, which is different from minimizing surface-area/volume. I derive the surface area of a logarithmic helicospiral; some formulae used previously are slightly incorrect. Heath's alteration of overlap changed shell volume and aperture area; a more meaningful reanalysis keeps these characters constant. (3) In 1987, Fitter used a two-dimensional model of plant roots to show that topology, growth rate, and nutrient diffusion rate affect exploitation efficiency (area of nutrient depletion zone/volume of root). Recalculation and reanalysis show that only part of the effect of topology depends on overlap of depletion zones. I compare Fitter's model to coplanar root systems with three-dimensional depletion zones and then to fully three-dimensional branching. Finally I discuss further examples in which two-dimensional models could mislead.

ADDITIONAL KEY WORDS:—branch – egg shape – Gastropoda – mollusc – root architecture – spatial model – spiral – theoretical morphology – oology – optimality model.

CONTENTS

E-mail: John.Hutchinson@bristol.ac.uk

0024–4066/00/050161+27 $35.00/0

INTRODUCTION

Considering that we live in a three-dimensional world, I am surprised by how much easier it is to think in two dimensions than in three. This conceptual simplification is no doubt one reason why theoreticians often try to reduce a three-dimensional problem to two dimensions, other reasons being that it is easier to sketch diagrams than build solid models, and that computations run faster. Those structures which exhibit rotational symmetry about an axis particularly invite this approach, since considering just a section down the axis then involves no loss of information about the complete structure.

This paper's aim is to highlight that, whilst such a dimensional reduction can often be useful as a tool for thought, it need not be a valid short cut when the analysis aims to be quantitative. To demonstrate the reasons why, I consider in detail three published cases where the two-dimensional analyses were quantitatively inappropriate for conclusions about the three-dimensional structures. The first two cases concern the shape of bird eggs designed to pack optimally and the shape of snail shells that minimize the requirements for shell material; both these analyses are based on axial sections. A third case concerns a model of plant roots where the inappropriate two-dimensional analysis was not of a section but of a three-dimensional branching structure that had been considered squashed flat. The aim here was to relate the pattern of branching to efficiency in absorbing nutrients from the soil. In each case I recalculate the results when allowing for the three-dimensional structure and compare them with the published results. The three cases are analysed and discussed in three separate sections, and in the General Discussion I mention some other examples to emphasize that the mismatch of results from two-dimensional and three-dimensional models is a widespread phenomenon.

Pappus' theorem

The bird-egg and snail-shell examples require the use of Pappus' theorem, which allows calculation of volume and surface area of a solid of revolution from its cross-section. Imagine a small square that will be rotated about a coplanar axis, thus tracing out a ring-shaped solid of revolution (Fig. 1). Pappus' theorem states:

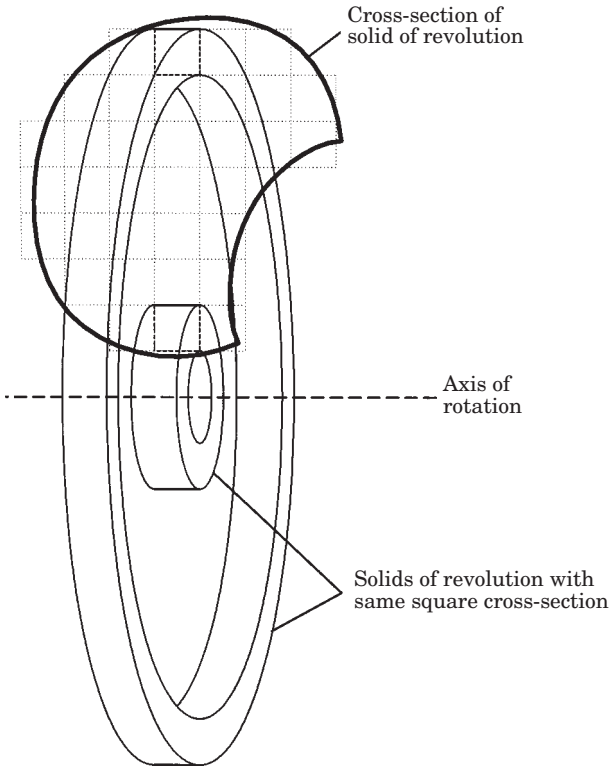

Figure 1. Imagine a solid of revolution generated by rotating the crescent-shaped cross-section about the axis. Its volume can be approximated by dividing the cross-section into squares and generating solids of revolution from each. Squares of the same size but a different distance from the axis of revolution generate different volumes.

$$\begin{aligned}\text{Volume of the ring} &= \text{area of the square} \times \text{distance traversed by centre of square} \\ &= \text{area of the square} \times \text{distance of centre of square from axis} \times 2\pi.\end{aligned}$$

A cross-section of more irregular outline can be approximated as made of a number of small squares each of the same size (Fig. 1). The important point is that the volume generated by different squares is not the same. Squares nearer the axis contribute less to the total volume of the solid of revolution than those farther away. Thus the cross-sectional area of a solid of revolution is not enough information to calculate its volume. We also need to know the distance of the centre of area to the axis, which will depend partly on the cross-section's shape and orientation. For instance, rotation of an ellipse about its minor axis generates a larger volume than rotation about its major axis.

To calculate the volume generated by more irregular shapes than a square, again consider the shape divided into small squares, and sum the volumes of the ring-shaped volumes generated by each.

$$\text{Total volume} \approx \Sigma(\text{area of square} \times \text{distance to axis of square's centre} \times 2\pi).$$

It is also true that

$$\text{Total volume} = \text{total area} \times \text{distance to axis of centre of area of whole shape} \times 2\pi.$$

From comparison of these two equations it can be seen that the centre of area of a figure is defined as a weighted average of the centres of all constituent parts of the figure, with the weighting equal to the area of each part.

There is another version of Pappus' theorem that enables calculation of the surface area of a solid of revolution from a cross-section coplanar with the axis. If we approximate the outline of the cross-section by a series of short straight lines:

$$\text{Surface area} \approx \Sigma(\text{length of straight line} \times \text{distance to axis of straight line's centre} \times 2\pi).$$

Also, similarly to above,

$$\text{Surface area} = \text{total perimeter} \times \text{distance to axis of centre of perimeter} \times 2\pi,$$

where the centre of the perimeter is the weighted average of the centre of component parts of the outline, each weighted by the part's length.

BIRD-EGG SHAPE

Barta & Székely (1997) recently published a stimulating paper on the optimal shape of bird eggs. What they optimized was the efficiency with which a clutch of eggs (all of the same shape and size) packed into a circular area. The argument was that the higher this efficiency, the larger the eggs that could pack under the mother's brood patch. Barta & Székely showed that the optimal egg shape depended on the number of eggs in the clutch.

An assumption of their study was that the eggs sat so that their long axes were all coplanar, thus simplifying the problem from three to two dimensions. This simplification is unrealistic in that: (i) in the nests of many species the inside forms a cup; (ii) even for eggs lying on a flat surface, an orientation parallel to that plane is not exactly the way many egg shapes settle and nor will it optimize packing (Andersson, 1978); (iii) eggs can potentially overlie each other at their pointed ends or even form overlying layers. Barta & Székely were aware of these limitations of their model, but they had to assume coplanar axes in order for their computations to run in a practicable time (Barta, pers. comm.). I discuss here a different limitation of their two-dimensional perspective which can be avoided at no computational cost.

Barta & Székely measured efficiency of packing by looking down on the nest and calculating the area of the eggs divided by the area of the circle that circumscribed the clutch. This is a measure of the size of the eggs than can fit within a circular brood patch of unit area. But it is strange to measure egg size by cross-sectional area. The more appropriate measure of egg size is volume, as this is directly proportional to the resources available to the embryo. Because of Pappus' theorem, maximizing cross-sectional area will not normally maximize volume.

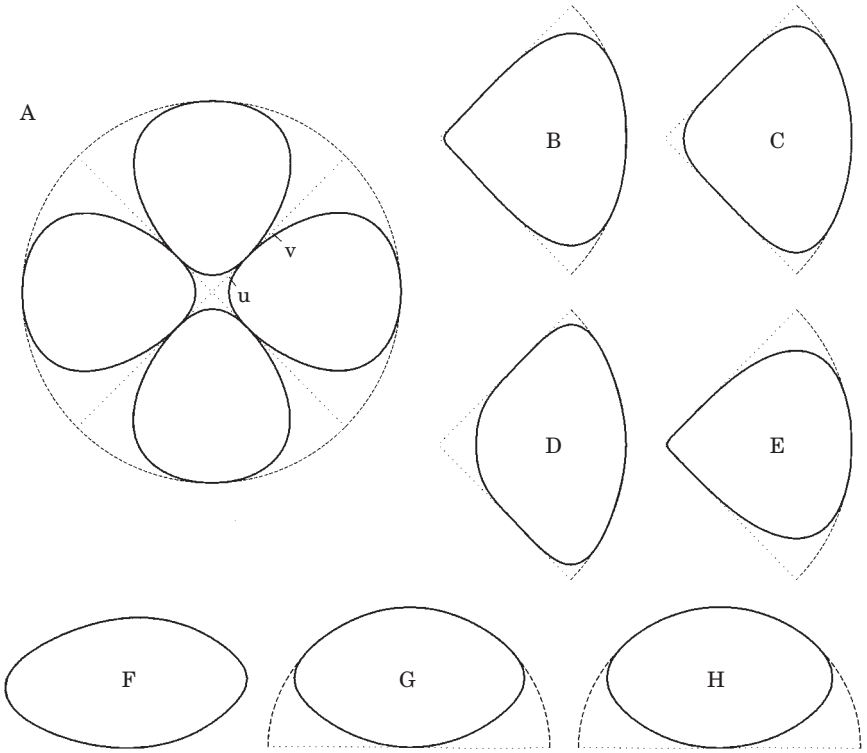

Figure 2. (A)–(E) concern the case of four eggs packing symmetrically into a circular nest with their axes coplanar. (A) The shape illustrated by Barta & Székely (1997), calculated from the median parameter values of their genetic algorithms (GAs). (B) The shape that their GA should have found—maximizing cross-sectional area, subject to certain limits on parameter values, and with the periphery described by only 22 points. (C) Area maximized, parameters unconstrained. (D) Volume maximized, parameters unconstrained. (E) Volume maximized, length along the axis constrained to be not shorter than breadth. (F)–(H) concern the two-egg case. (F) Output of the Barta & Székely GA that maximized area; the two eggs packed 5° off antiparallel. (G) Area maximized, parameters unconstrained. (H) Volume maximized, parameters unconstrained. Parameter values, given as a vector $[c_0, c_1, c_2, c_3]$, were: (A) [0.865, 0.272, 0.014, 0.005]; (B) [1, 0.659, −0.091, 0.020]; (C) [1.189, 0.725, −0.110, −0.288]; (D) [1.469, 0.700, −0.184, −0.610]; (E) [0.885, 0.508, −0.037, 0.131]; (F) [0.544, 0.056, −0.100, −0.084]; (G) [0.613, 0, −0.139, 0]; (H) [0.625, 0, −0.103, 0].

The four-egg case

Consider a clutch of four pointed eggs that pack together symmetrically in the way observed in many wader nests (Fig. 2A). Suppose that the shape illustrated does maximize the cross-sectional area of eggs that fit into a unit circle (the shape is the one that Barta & Székely calculated for the four-egg case). If eggs were instead to maximize volume, it would be better to swell the outline of the egg around region v in the figure, even at the expense of losing slightly more cross-sectional area at u. This is because region v is farther from the axis of rotation than u and so the same cross-sectional area contributes more to volume.

Let us now see how important this effect is quantitatively. Barta & Székely allowed eggs to have the range of shapes whose *x*-*y* co-ordinates are specified by the following equation,

$$y=\sqrt{1-x^2}(c_0+c_1x+c_2x^2+c_3x^3),$$

where they kept c_0, c_1, c_2 and c_3 within the arbitrary limits $0\leq(c_0,\ c_1)\leq 1$ and $-0.1\leq(c_2,c_3)\leq 0.1$. Another constraint was that no region of the outline could be concave. Barta & Székely had extra parameters that determined the orientation of each egg relative to its neighbours, and to solve this multivariate optimization problem they used a genetic algorithm (Goldberg, 1989). For simplicity, I instead constrained the eggs to pack symmetrically with their narrower ends towards the centre, which meant that I could search for the global optimum simply with a systematic search. As a test of my method I initially attempted to replicate Barta & Székely's results, and so searched for the shape that minimized cross-sectional area.

It transpired that the shapes found by Barta & Székely's genetic algorithm were far from my optimum both in shape and efficiency. Barta & Székely's maximum efficiency (cross-sectional area of four eggs in a circle of unit area) was 0.77, whereas for exactly the same problem I found a more wedge-shaped egg with an efficiency of 0.91 (Fig. 2B). Moreover this shape is up against their arbitrary constraint that $c_0\leq 1$. The shape maximizing area is rather different if the parameters are allowed to take any value (Fig. 2C).

I compared this with the shape maximizing the volume of eggs that would fit in the unit circle (Fig. 2D). Not surprisingly the latter is also wedge shaped, but it is noticeably different in the way predicted above—the fit is less good at the narrower end near the axis. Still, the maximum-area eggs achieved 96.5% of the volume of the maximum-volume eggs.

The optimal eggs that I found are wider than long, which would probably pose a bird insuperable problems to manufacture. Accordingly I also found the optima when width was constrained to be no greater than length (Fig. 2E). With this extra constraint the shape maximizing area is almost identical with the shape maximizing volume.

The two-egg case

I also analysed optimal shape when there are only two identical eggs in a circular nest. Figure 2F shows Barta & Székely's solution. For this case I could prove analytically that one globally optimal arrangement must have the axes of the eggs lying precisely antiparallel to each other and the outlines touching in the centre of the nest. By systematic searching of parameter space I then discovered that optimizations of both area and volume resulted in exactly symmetrical eggs (i.e. $c_1=c_3=0$). The shape maximizing area (Fig. 2G) is only a little different from that maximizing volume (Fig. 2H), and the volume of the former is 99.5% of the maximum volume.

Discussion

Eggs that maximize volume will tend to stick up farther above the substrate than those that maximize area, and this may mean that a brood patch may not be able to envelop them so well. The limited ability of the brood patch to stretch and wrap over such lumpy objects as eggs means that the eggs really should be constrained

to fit not into a circle but into some other partly deformable three-dimensional space that is restricted in surface area. However, in reality the shape of the nest and of the brood patch can evolve to optimize packing as much as can egg shape.

Obviously the problem that Barta & Székely bravely tackled is but a caricature of the optimization that has occurred through natural selection, but nevertheless their approach does highlight that packing efficiency will select for particular egg shapes, and that the optimum depends on clutch size. Unfortunately without repeating their very time-demanding genetic algorithm, I cannot say for other clutch sizes whether the optimal shapes are roughly alike when volume rather than cross-sectional area is maximized. But it seems likely that a shape that maximizes cross-sectional area will also perform well in maximizing volume, so their conceptual error need not make much quantitative difference.

A digression on the performance of Barta & Székely's genetic algorithm

The genetic algorithm (GA) is a computational optimization technique that mirrors evolutionary processes of natural selection (Goldberg, 1989). The parameter values are typically encoded as binary digits along a 'chromosome', with the values able to differ between individuals in a population. The best individuals in a population are selected to breed the next generation, and both mutation and recombination continually provide variation.

I was surprised that Barta & Székely's GA did not get very close to finding the optimum shape in the four-egg case, only achieving 84% of the maximum efficiency. Perhaps more surprising is that in the two-egg case the GA, despite getting quite close, nevertheless failed to climb to the top of the adaptive landscape (e.g. the orientation of Barta & Székely's best solution was 5% off the optimal antiparallel arrangement—Barta, pers. comm.). Similarly their one-egg optimization did not quite achieve the clearly optimum solution of a spherical egg.

I suspect that the reason was that Barta & Székely maintained a high mutation rate (5% for each parameter). Their surviving configurations had thus had to cope well with maladaptive changes in shape or orientation parameters. A set of parameters that is robust in this sense need not be close to the parameters that maximize efficiency when there is no disruption. This is a recognized problem with GAs that use a conventional binary code. It can be ameliorated by various techniques that reduce the magnitude of the effect of one-bit point mutations, such as Gray coding (Caruana & Schaffer, 1988), successive narrowing of the range of values onto which a binary code maps (Schraudolph & Belew, 1992), or explicitly modelling mutation and recombination in ways independent of the binary coding (Bäck, 1996). Barta & Székely did in fact use the latter strategy but they made mutation equally likely to lead to any point in parameter space, rather than more likely to have a minor effect (Barta, pers. comm.).

One could argue that it is more realistic to include the possibility of occasional misorientated or misshapen eggs in a model, as this is what happens in real nests. In that case the results are likely to depend very much on the error structure specified. Barta & Székely's model did not well represent the real sources of stochasticity, both because large changes in parameter value were made as likely as small, and because the individual shape parameters were not representative of real biological processes. Also one major source of biological stochasticity, variation in

clutch size, was not included. Measuring what stochasticity is realistic would be the necessary first step before constructing a more sophisticated new generation of models.

SNAIL SHELLS

Snail shells consist of a tube that steadily expands in cross-section, while coiling around an axis. There is a long tradition of modelling snail shells as logarithmic helicospirals to which many species indeed approximate quite closely (McGhee, 1998). The essential property of this geometrical figure is that all dimensions of the shell increase by the same proportion with each revolution of growth, so that shape remains constant. A section down the coiling axis reveals the full structure of the shell (Fig. 3A). The shell in Figure 3A has a circular generating curve, and by altering just three other parameters a wide variety of shapes can be produced (Raup, 1966). The three parameters represent the rate of expansion each revolution, the relative distance of the generating curve from the coiling axis, and the translation rate of the generating curve down the coiling axis.

Heath (1985) investigated whether terrestrial snail shells were designed to minimize the amount of shell material required to enclose a particular volume of flesh. In particular he considered the overlap of the generating curve between successive whorls. By incorporating the preceding whorl as one wall of the aperture, the snail avoids having to lay down so much new shell material. However, if the overlap is high, increasing overlap does not achieve much extra reduction in the shell material required, whereas the preceding whorl bulges far into the new whorl and thus considerably reduces the volume enclosed (Fig. 3B–E). There is therefore some optimum overlap.

Heath intended to find the overlap that minimized the amount of shell material required to enclose a particular volume of flesh. Instead he considered the cross-section of the tube sectioned down the coiling axis (Fig. 3A) and minimized the perimeter of new shell divided by the area enclosed. Because of Pappus' theorem, minimizing perimeter/area (P/A) is not the same as minimizing surface-area/volume (SA/V_i). The reason is that for the crescent-shaped cross-section the perimeter and the area have centres different distances from the coiling axis.

Heath found that real shells showed more overlap than the optimum that minimized P/A. I first examine whether the agreement is better if we instead calculate the optimum that minimizes SA/V_i. Then I will explain that the absolute volume of the shell and the taper of its interior also affect these ratios, and that optimum overlap differs when these quantities are held constant.

Calculating the volume and surface area of a helicospiral

The equations in the Introduction are for a solid of revolution where there is no expansion and no translation along the coiling axis. However, Pappus' theorem may be extended to calculate the volume of helicospirals (e.g. Raup & Graus, 1972). For a short segment of the helicospiral,

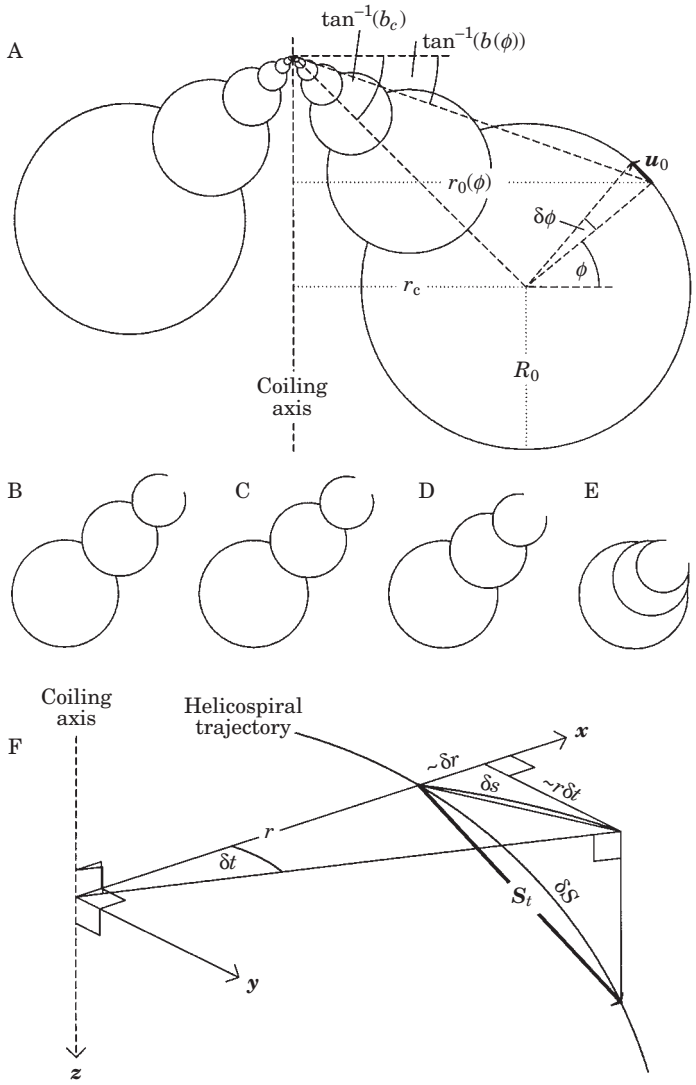

Figure 3. (A) An isometric snail shell sectioned down its coiling axis. (B)–(E) show the sections of three consecutive whorls which grow by a factor of 1/0.7 each revolution and where $b_c = 1$. Overlap (measured by F) increases across the page: (B) $F = 0.125$, maximizing SA/V_i or V_s/V_i if aperture diameter and per-revolution expansion rate are kept constant; (C) $F = 0.139$, the optimum overlap predicted by Heath from maximizing P/A; (D) $F = 0.276$, maximizing SA/V_i if volume and aperture area are kept constant; (E) $F = 0.624$, maximizing SA/V_i if volume and per-revolution expansion rate are kept constant. (F) Co-ordinate system and variables applied to a small step δt of a helicospiral trajectory.

$$\delta\text{volume} \approx \text{area of cross-section} \times \delta\text{distance moved by centre of area of cross-section.}$$

The cross-section considered here is coplanar with the coiling axis, not perpendicular to the growth direction of the tube. For a logarithmic spiral

$$\text{area of cross-section} = a_0 e^{2mt},$$

where a_0 is this area at some arbitrary reference point on the shell and t is the number of revolutions from that reference point. t is measured in radians and m measures the proportionate increase in the shell's linear dimensions with each radian of growth.

When calculating the distance moved by the centre of area we should consider length δs (Fig. 3F), which is unaffected by translation down the coiling axis. This is because a shear transformation leaves volume unchanged (think of pushing over a stacked pack of cards). If we had instead measured area of a section cut perpendicular to the direction of growth, we would have had to have incorporated translation rate. It is a standard result for all two-dimensional curves that

$$\delta s^2 \approx (r\delta t)^2 + (\delta r)^2.$$

Here we want r to be r_a the distance of the centre of area to the coiling axis, which at $t = 0$ I call r_{a0}. For a logarithmic spiral $r_a = r_{a0} e^{mt}$, so the two-dimensional movement of the centre of area around the axis is

$$\mathrm{d}s_a = r_{a0}(1 + m^2)^{1/2} e^{mt} \mathrm{d}t.$$

Thus

$$\mathrm{d}(\text{volume}) = \text{area of cross-section} \times \mathrm{d}s_a = a_0 e^{2mt} \times r_{a0}(1 + m^2)^{1/2} e^{mt} \mathrm{d}t.$$

Integrating this from $t = -\infty$ to 0, yields a total volume for the shell of

$$\frac{1}{3} a_0 r_{a0}(1/m^2 + 1)^{1/2}.$$

For cross-sections of irregular shape $a_0 r_{a0}$ must be found by integrating over component parts of the cross-section.

Calculating surface area is less straightforward, essentially because a shear transformation does affect surface area. Pappus' theorem works only for surfaces of revolution in which there is no expansion. Additional text in the appendix explains the reason and shows how surface area can be calculated for a logarithmic helicospiral. Raup & Graus (1972) and Stone (1997) overlooked this limitation of Pappus' theorem, although quantitatively their error is insignificant for realistic parameter values. In the appendix I also correct some earlier work of my own (Hutchinson, 1990a; see also Rice, 1998), which suffered from a related fallacy.

An alternative approach is not to calculate surface area but to calculate the volume of shell material under the assumption that the shell is proportionately thicker in larger whorls. The volume of shell material can be then calculated by an expression similar to that for volume of the shell contents. For thick shells the equivalent of $a_0 r_{a0}$ must be calculated directly by integrating component areas of shell thickness. For thin shells whose generating curve is coplanar with the coiling axis and uniform in thickness, $a_0 \approx$ (shell thickness at $t = 0$) $\times$ (the relevant length of perimeter of cross-section), and $r_{a0} \approx$ the centre of this length of perimeter. Thickness

here is measured in the plane of the cross-section, so will not be uniform measured perpendicular to the shell surface, but the discrepancy is small for realistic parameter values.

Results of incorporating Pappus' theorem

For illustrative purposes I analysed the case of a generating curve with a circular outline, with linear dimensions increasing each revolution by a factor of between 1.11 and 3.33, and with a translation rate between $b_c = 0$ and $b_c = 1$ (if the centre of the generating curve is r_c from the coiling axis, its distance from the apex measured along the coiling axis is $b_c r_c$: see Fig. 3A). For the initial analysis I follow Heath in comparing shells with the same radius of aperture.

Figure 4A shows how the optimal overlap depends on whether we minimize either perimeter/area ($=P/A$), or (surface area)/(interior volume) ($=SA/V_i$), or (volume of shell material)/(interior volume) if the shell is thin and increasing from early to late whorls ($=V_s/V_i$). The translation rate affects only SA/V_i and makes rather little difference. For all three cases the expansion rate increases our measure of optimum overlap, but the ordering of the three optima remains the same (Fig. 4B). Heath's P/A optimum consistently occurs at larger overlaps than the two optima based on proper three-dimensional analyses, although the difference is slight (cf. Fig. 3C with 3B). Since Heath observed that real shells had greater overlaps than his predictions, incorporating Pappus' theorem seems to make it even less plausible that overlap optimizes the use of shell material.

Further analysis 1: comparison at constant volume

There are two factors other than overlap that affect SA/V_i, but that have not yet been properly taken into account. One is that because SA/V_i is not dimensionless it will decrease with absolute size. In the above analysis, when overlap is increased the volume of the shell decreases. It would seem more meaningful to compare SA/V_i at a constant shell volume by allowing the shell to grow further. This is equivalent to making the measure dimensionless by multiplying SA/V_i by volume$^{1/3}$. Since shell thickness is proportional to volume$^{1/3}$ in an isometric shell, this adjusted value is proportional to the amount of shell material required to make a shell of unit volume a particular thickness all over. Adjusting V_s/V_i in a similar way is also meaningful because V_s has so far included an arbitrary constant reflecting shell thickness at the aperture, which we can also now replace by a constant $\times$ volume$^{1/3}$. Henceforth we will be comparing shells of the same volume that have the same thickness at the aperture.

When these adjusted values are plotted against overlap, it turns out that for rapidly expanding shells the more overlap the better (Fig. 4C,D). The reason is that as overlap increases the outer whorl takes up most of the shell diameter. Rather than being a thin disk, the shape is then much more rounded and thus requires less shell to enclose a particular volume. For shells that expand more slowly there is an optimal overlap, but this is very high (Fig. 3E).

Heath did also attempt to compare shells that enclosed a constant volume. Unfortunately his two-dimensional perspective led him to believe erroneously that keeping aperture area constant would achieve this. He kept aperture area constant

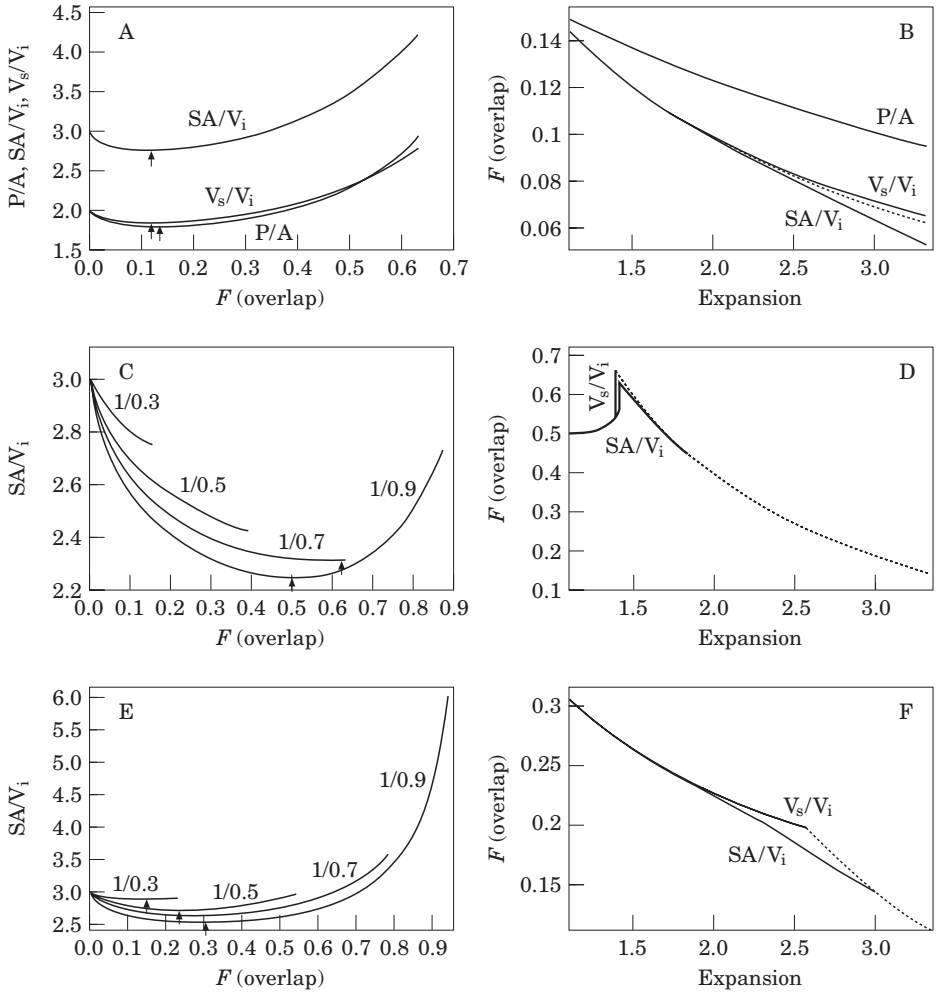

Figure 4. Optimal whorl overlap in a snail shell growing as a logarithmic helicospiral. Each row of the diagram corresponds to a different set of constraints: (A)–(B) aperture diameter and per-revolution expansion rate both constant; (C)–(D) volume and per-revolution expansion rate constant; (E)–(F) volume and aperture area constant. Graphs in the left column show the effect of altering overlap on various measures of the ratio between the amount of shell material used and the volume enclosed. An arrow indicates where the overlap maximizes these measures. These optima depend on the expansion rate (the four curves in (C) and (E) correspond to four different values; the expansion rate in (A) is 1/0.7, a typical value for a terrestrial snail). Graphs in the right column show the effect of expansion rate on the optimal overlap.

F is the proportion of the diameter of the larger whorl that is overlapped by the smaller. All values for SA/V_i and P/A assume that when $F=0$ the aperture has unit radius. Values for V_s/V_i should be multiplied by (shell thickness/$V_i^{1/3}$) measured at the aperture; this factor is assumed to be independent of F. Expansion rate is the proportionate increase in linear dimensions each revolution. In (E) the curves are labelled with the expansion rate when $F=0$; expansion rate will increase along each curve as overlap increases; this increased value is used in (F). I take $b_c=1$ (as in Fig. 3A), except for the dashed line in (B) corresponding to the SA/V_i curve for the planispiral case $b_c=0$ (the V_s/V_i optima are unaffected by b_c). Since $b_c\neq 0$, the maximum overlap is constrained by my not allowing the whorl cross-sections to cross the coiling axis; in (D) and (F) the parts of the curves generated by this constraint are dotted.

by adjusting expansion rate and ironically there are other reasons for doing this that are valid. I explain them in the following section.

Further analysis 2: comparison at constant aperture area and constant volume

When changing the overlap in the above analyses, I have maintained a constant increase in dimensions each revolution. However, by thinking in three dimensions, one realizes that when overlap is high, and thus when the aperture is relatively close to the coiling axis, the length of the one revolution of tube connecting consecutive cross-sections is relatively shorter than when there is less overlap. Logarithmic helicospirals are coiled cones, and with increasing overlap, the more rapidly tapering is this cone. In shells of like volume, a more rapid taper means a larger aperture. The relative size of the aperture is important for snails because it influences the access for predators, the rate of dehydration, and the size of the foot, and thus the strength of adhesion. Therefore there is a strong argument for keeping the relative size of the aperture (and thus the interior taper of the coiled cone) constant when we alter overlap, and adjusting the per-revolution rate of expansion accordingly. The per-revolution rate of expansion seems likely to have less direct importance to a snail. This is what I have done in Figure 4E & F, again comparing shells of the same total volume.

An optimum whorl overlap now exists for all but the fastest expansion rates. The optima are very similar whether the shell is of uniform thickness (SA/V_i) or whether inner whorls have proportionately thinner shells (V_s/V_i). The translation rate has very little effect. Shells with relatively larger apertures have larger optimal overlaps (as measured by *F*, the proportion of the diameter of the larger whorl occupied by the smaller).

If we compare shells with the same per-revolution expansion rate, shells that have optimal overlaps under these new criteria have larger overlaps than Heath found optimal (under his criteria of minimizing P/A with either per-revolution expansion rate or aperture area held constant). So my three-dimensional analysis could explain why Heath observed larger overlaps in real snails than his optimality model predicted. The overlaps predicted by the last version of the model certainly look the most realistic (Fig. 3D).

Discussion

I shall not test more carefully whether my analysis accurately predicts overlap in real shells, because a model realistic enough to give quantitative predictions would be much more complicated. Real shells do not have circular apertures and shell thickness is not uniform around the aperture. Also when a new whorl is laid down the penultimate whorl to which it attaches is covered in a reasonably thick layer of new shell material. A further complication is that snails typically redistribute shell material as they grow, thinning the shell walls that have become internalized, and thickening the walls of the early whorls that remain external (Kohn, Myers & Meenakshi, 1979; Hutchinson, 1990b). These processes do affect the overlap that maximizes the average shell thickness attainable with a set volume of shell material. Thus a much more detailed model is required based on measurements of relative

thicknesses of different parts of the shell, how much shell material is redistributed and the cost of redistributing it.

In such a model overlap would not be a particularly meaningful parameter to alter because snails seem unlikely to alter overlap whilst keeping every other growth parameter constant. More direct ways in which snails change their shell morphology are through altering the shape, expansion and site of attachment of the aperture (Hutchinson, 1989). It is more plausible that each of these parameters could vary independently of the others, and so it would be more meaningful to ask how altering these parameters affects the shell thickness attainable with a set volume of shell material.

Although more complicated, such analyses should certainly make use of the three-dimensional approach developed in this paper. So I now summarize the ways in which a three-dimensional perspective changed the analysis. We realized that: (i) minimizing P/A would not necessarily minimize SA/V_i; (ii) we should compare shells of the same volume; and (iii) increasing overlap increased the rate of taper of the shell interior so that we were still not really comparing like with like. On the other hand, Heath's two-dimensional perspective gave him the original insight that there might be an optimum overlap.

PLANT ROOTS

There have been many models in the botanical and palaeontological literature examining the consequences of different branching patterns in plants. Most branching patterns and most models of them are fully three-dimensional (a justified exception being models of ramets spreading horizontally over the ground). Other models concern conduction rather than exploitation of space and so can be topological rather than geometrical. However, a paper by Fitter (1987) is one example of where a two-dimensional model again misrepresents some aspects of a three-dimensional situation. This misrepresentation is conceptually similar to those identified in the earlier examples, even though a network of branches does not form a solid of revolution and may even be asymmetrical.

Fitter was concerned with the effect of a root system's topology on the exploitation of dissolved solutes in a uniform soil. Herring-bone arrangements in which all side roots branch directly from a single axial root have a high topological index (TI), compared with the low TI of topologies in which every offshoot can itself branch (in Fig. 5 compare E with B). Fitter constructed a two-dimensional model in which branching occurred stochastically but was constrained by specified parameters. For the results that he reported branching angle was kept at 75°, but it was chance whether this was to the left or right (except that an upwards component in the direction of growth was disallowed). As the root grew it drained nutrients from a zone around it, this zone growing with time both because the root reached new regions and because nutrients continued to diffuse towards the older parts of the root (Fig. 5B–G). Once a zone was depleted through the proximity of one branch, it could not be exploited again, so branches that grew too close to others would lower the efficiency. Exploitation efficiency was measured as the area of the depletion zone divided by the volume of the root.

Fitter varied the probabilities of branching and showed a positive relationship

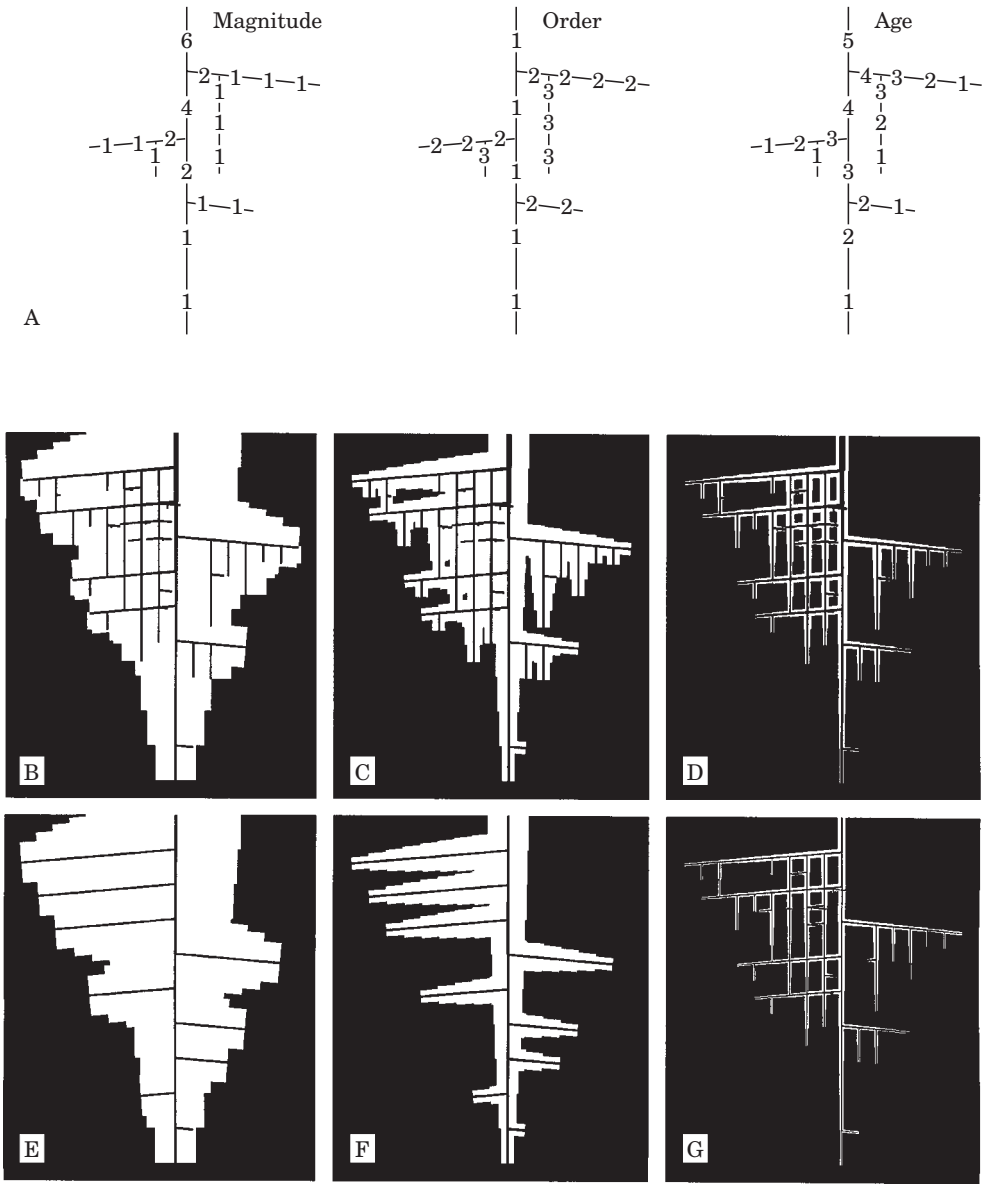

Figure 5. (A) Magnitudes, orders and ages for each link of the same root network (see page 179 for definitions). (B)–(G) show depletion zones (white) with the root thickness drawn to scale inside them. The network in (B)–(D) and (G) has a low TI of 0.52; that in (E) and (F) has the maximum TI of 1.0. In (B)–(F) axial root growth = 10 mm/day; in (G) growth = 50 mm/day but the representation is at 1/5 the scale, and hence the roots appear thinner. Diffusion rate is set by the parameter D where $D = 10^{-6}\text{cm}^2\text{s}^{-1}$ in (B) and (E); $D = 10^{-7}\text{cm}^2\text{s}^{-1}$ in (C), (F) and (G); $D = 10^{-8}\text{cm}^2\text{s}^{-1}$ in (D).

between TI and exploitation efficiency. The slope of this relationship was lower if the diffusion rate was low, which he explained by the depletion zones around each branch then being too small to overlap often (Fig. 5D). Similarly if the root grew fast relative to its branching rate, roots were too far apart for the depletion zones

to be likely to overlap (Fig. 5G). However, if the root grew slowly and diffusion rate was fast, Fitter found no effect of topology above a TI of about 0.7.

Except for this last result, these patterns are not particularly surprising, and seem likely to hold qualitatively whether a model is two- or three-dimensional. However, Fitter's model used realistic values of diffusion rates and growth rates, so there is a temptation to draw quantitative conclusions, and these may be misleading because his model is two-dimensional. It is obvious that there is more space in three dimensions and thus less likelihood of depletion zones overlapping, but I wish to establish here how important this is quantitatively. Therefore I have reconstructed Fitter's model but then extended it to three dimensions.

Fitter and others have since constructed fully three-dimensional models of nutrient depletion around root systems (Fitter *et al.*, 1991; Lynch & Nielsen, 1996), so one might argue that the earlier two-dimensional model is now not worth detailed criticism. But the details of these models are different and their output is reported in a different way, so they do not tell us to what extent the two-dimensional model was misleading. Therefore the temptation to revert to a two-dimensional simplification may remain, especially when more difficult tasks are tackled (such as the optimization of branching rules, or the incorporation of competition with neighbouring root systems). Also the acceptance of the need for three-dimensional models by root modellers does not necessarily mean that modellers of other structures realize the dangers (models of branching in above-ground parts of plants and in bryozoans were fully three-dimensional before Fitter published his two-dimensional model—e.g. Niklas & Kerchner, 1984; Cheetham & Hayek, 1983). By identifying the several different reasons why the two-dimensional simplification gives different results, I aim to make it apparent that the same differences can apply in other sorts of models. A further reason to compare two- and three-dimensional models of root systems in particular is that roots are often studied growing in rhizotrons, which forces the root system to become almost two-dimensional.

Results from the two-dimensional model

Instructions in Fitter (1987) allow reconstruction of his program. Slight ambiguities were resolved through personal communication (the first node lies at the ground surface and cannot branch; $t=1$ for the most recently grown segment of root). Unfortunately the values of exploitation efficiency in Fitter (1987) appear to be erroneous by a factor of 60% or more, so I give a corrected version of his results in Figure 6A & B. The broad patterns that Fitter observed do hold, but the surprising result that TIs above a value of 0.7 make no difference when growth rate is slow and diffusion rate large turns out to be erroneous.

It also transpires that a major reason why exploitation efficiency increases with TI has nothing to do with the decreasing overlap of depletion zones as TI increases (although the latter is indeed another reason). To demonstrate this the analysis can be repeated with the root broken up into isolated segments so that there is no overlap at all. The exploitation efficiency for these disarticulated roots still increases strongly with TI. To distinguish this effect from the reduction in exploitation efficiency due to overlap, I define 'topological efficiency' as the exploitation efficiency of the disarticulated roots, and 'overlap efficiency' as the ratio of the exploitation efficiency of the intact root system to the exploitation efficiency of the disarticulated

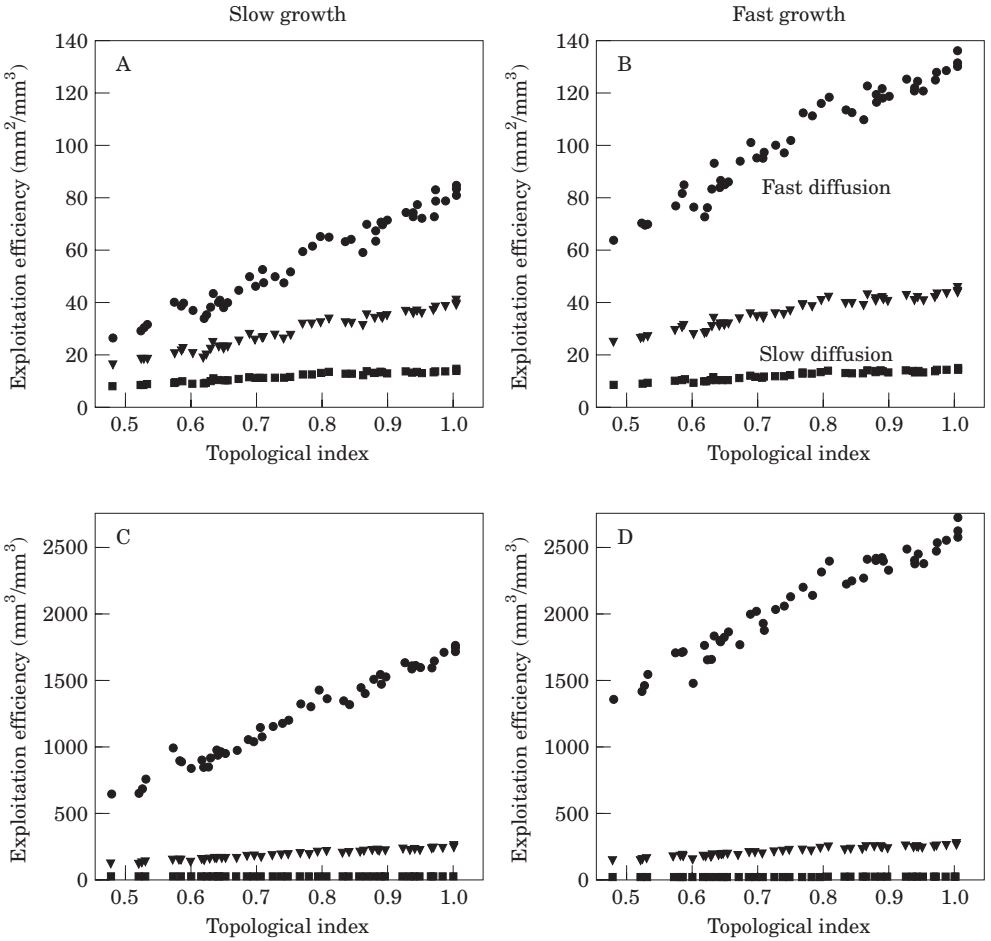

Figure 6. The relationship between exploitation efficiency and TI is plotted for two different values of root growth rate (A & C, slow; B & D, fast) and three different values of solute diffusion rate. (A) and (B) are corrected versions of Figure 9 of Fitter (1987), in which exploitation efficiency is measured as (area of depletion zone)/(root volume) and the network is two-dimensional. In (C) and (D) exploitation efficiency is measured as the ratio of volumes and the network is fully three-dimensional. In this and succeeding diagrams the same collection of networks is used for each combination of parameters. Diffusion rate is set by the parameter D; $\bullet \Rightarrow D = 10^{-6}\text{cm}^2\text{s}^{-1}$, $\blacktriangledown \Rightarrow D = 10^{-7}\text{cm}^2\text{s}^{-1}$, $\blacksquare \Rightarrow D = 10^{-8}\text{cm}^2\text{s}^{-1}$. Axial root growth = 10 mm/day in (A) and (C), 50 mm/day in (B) and (D).

root system. Overlap efficiency is high if the depletion zones around adjacent roots overlap little. From these definitions it follows that

$$\text{exploitation efficiency} = \text{topological efficiency} \times \text{overlap efficiency}.$$

Figure 7 shows log[exploitation efficiency] and log[topological efficiency] for a range of root systems and for three combinations of growth and diffusion rates. For each value of TI the difference between these two values is log[overlap efficiency]. Topological efficiency is unaffected by the growth rate of the root, and the effect of a change in diffusion rate is simply to translate values of log[topological efficiency]

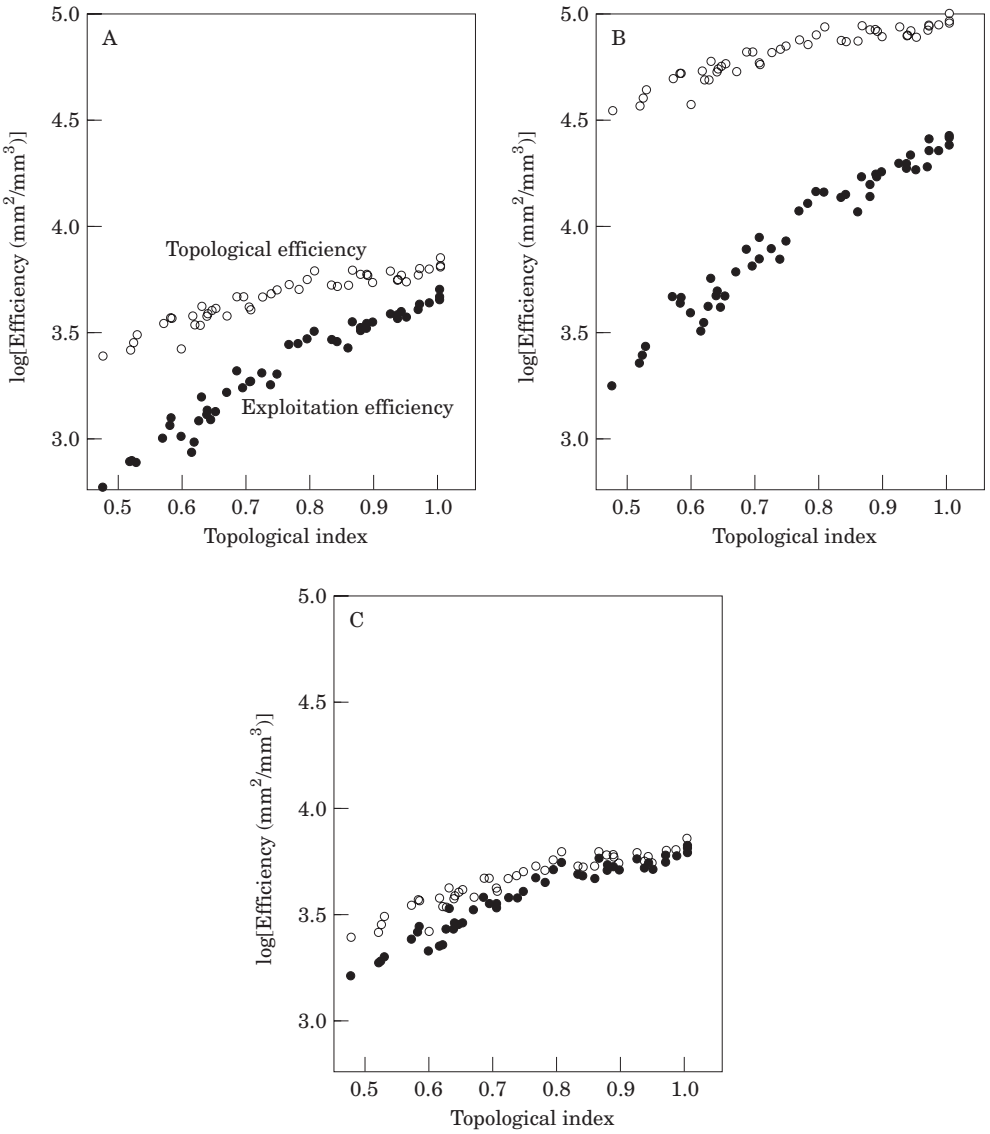

Figure 7. For three combinations of root growth rate and solute diffusion rate, the relationship of log[exploitation efficiency] (●) and log[topological efficiency] (○) with TI. (A) slow growth, medium diffusion; (B) slow growth, fast diffusion; (C) fast growth, medium diffusion. For each network the difference between these values is log[overlap efficiency]. Exploitation efficiency is measured as (area of depletion zone)/(root volume) and the network is two-dimensional. (A) axial root growth = 10 mm/day, diffusion rate with $D = 10^{-7}\text{cm}^2\text{s}^{-1}$; (B) growth = 10 mm/day, $D = 10^{-6}\text{cm}^2\text{s}^{-1}$; (C) growth = 50 mm/day, $D = 10^{-7}\text{cm}^2\text{s}^{-1}$.

by a constant amount. Figures 7 and 8 show that both topological efficiency and overlap efficiency increase with TI.

The reason that topological efficiency increases is that Fitter's model made width of the depletion zone dependent on the age of that length of root, the length of a root dependent on its 'order' and the thickness dependent on its

'magnitude'. Order is 1 for the axial root, 2 for its branches, 3 for the branches off these branches, etc. (Fig. 5A). If we consider the root divided at each branch point into separate lengths, magnitude is a technical term defined as follows: lengths terminating in a meristem have magnitude 1 and all other lengths have a magnitude equal to the sum of their two daughter branches (Fig. 5A). As TI changes the distribution of orders and magnitudes changes (cf. Fig. 5B & E). The many extra peripheral roots in a low TI network have similar root thicknesses and widths of depletion zones as the roots in a high TI network, and in any case have relatively little effect on overall efficiency because they are short. But the magnitudes (and thus thicknesses) of the axial root and its immediate branches are higher when TI is low, whilst the depletion zone around them remains the same. This inevitably lowers the topological efficiency and thus also exploitation efficiency. This effect depends only on the topology of the root system, not at all on how much depletion zones overlap.

Like topological efficiency, overlap efficiency also increases with increasing TI, so both factors contribute to the increase in exploitation efficiency. However, when growth rates are high and diffusion rates are low, overlap efficiency starts to plateau at medium values of TI (Fig. 8C). Further increases in TI increase the probability of overlap by only a little.

Three-dimensional depletion zones around two-dimensional networks

As a first step to analyse the effect of converting the two-dimensional model to three dimensions, I consider identical two-dimensional networks, but allow the depletion zones to be three-dimensional cylinders around each length of root, rather than their rectangular cross-sections. Exploitation efficiency is now defined as a ratio of volumes, which is dimensionally more logical. Fitter *et al.* (1991) also made this change in a later three-dimensional model.

Because of this change the topological efficiency also changes. Nevertheless the relationship of topological efficiency to TI remains very similar, and its relationship to growth rate is unchanged. But the use of the volume of a cylindrical depletion zone rather than its cross-sectional area will square the effect of diffusion rate on topological efficiency (area $\propto$ diffusion rate^1; volume $\propto$ diffusion rate^2). This dominates the change in exploitation efficiency and is responsible for the most obvious differences between the top and lower halves of Figure 6 (the lower half applies to the fully three-dimensional network discussed presently).

Technically more interesting is the change in overlap efficiency when depletion zones are made three-dimensional. Consider two parallel cylinders that we move towards each other. Up to now what has mattered is the area of overlap of their silhouettes (or cross-sections down the axis) as a proportion of their summed cross-sections. This always overestimates the volume of overlap as a proportion of the summed volume. The overestimate is higher the more different the radii of the cylinders. It is this factor that is responsible for the overlap efficiency increasing when we compare volumes of depletion zones rather than areas (Fig. 8). This increase in overlap efficiency is by a factor of up to 19% when depletion zones are wide and TI low.

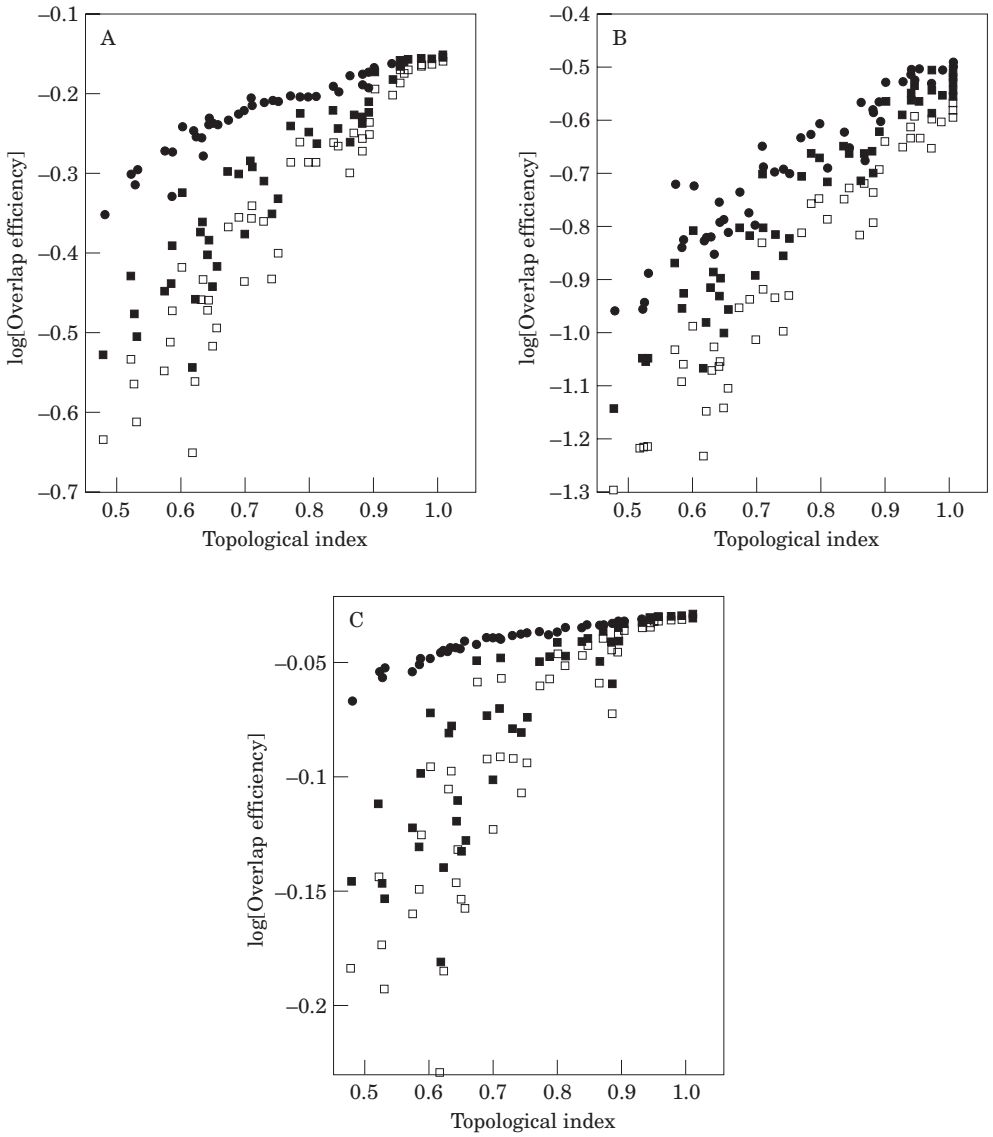

Figure 8. For three combinations of root growth rate and solute diffusion rate, the relationship of overlap efficiency with TI. (A)–(C) as in Figure 7. (□) overlap efficiency based on areas for a two-dimensional network. (■) overlap efficiency based on volume for a two-dimensional network. (●) overlap efficiency based on volume for a three-dimensional network. (A) axial root growth = 10 mm/day, diffusion rate with $D = 10^{-7}\text{cm}^2\text{s}^{-1}$; (B) growth = 10 mm/day, $D = 10^{-6}\text{cm}^2\text{s}^{-1}$; (C) growth = 50 mm/day, $D = 10^{-7}\text{cm}^2\text{s}^{-1}$. Note that the vertical scale differs between graphs.

Fully three-dimensional networks

I now make the networks fully three-dimensional by allowing the radial angle of branching to take any value as long as the branch does not then grow upwards. This model is not unreasonable since the equivalent of phyllotaxy (a regular sequence

of branching directions) has not been demonstrated in roots (Jourdan & Rey, 1997). To facilitate the comparison of the two- and three-dimensional models, I keep the collection of topologies identical.

Exploitation efficiency still increases with TI (Fig. 6C,D). Even more than in the two-dimensional models, the main reason for its increase with TI is the increase in topological efficiency with TI, and this relationship is quantitatively unaffected by making the network three-dimensional. Overlap efficiency is higher in three-dimensional than in two-dimensional networks, and this increase of overlap efficiency is highest at low TI (Fig. 8). This means that the relationship of overlap efficiency with TI is flatter, and consequently so is the relationship of exploitation efficiency with TI.

One reason why the overlap efficiency is usually higher in the fully three-dimensional network seems obvious. The three-dimensional networks extend into a larger volume of soil (a full cylinder or cone, rather than a thin slice of one), and thus the likelihood of depletion zones overlapping is less. I tried modelling this by cutting up the root networks into straight lengths and calculating the expected overlap efficiency if these disarticulated roots were placed at random at two densities. These densities corresponded to the three-dimensional root system in a cylinder of soil just enclosing it, and to the two-dimensional root system in the thin rectangular prism of soil just enclosing it. The difference in overlap efficiency between the two disarticulated root systems was comparable with the difference in overlap efficiency between the intact three-dimensional and two-dimensional networks. The agreement was very good with wide depletion zones; with narrow depletion zones and networks of low TI the random model underestimated the increase in overlap efficiency when the branching structure became three-dimensional.

Such disagreements are to do with how the intact root systems are organized, in ways that can either increase or decrease the extent of overlap compared with random arrangements. In Fitter's model some aspects of the organization matter more in the two-dimensional model than in the three-dimensional version. Branching angle is always 75°, so that in two dimensions most roots of order 3 are growing directly on a collision course with roots of order 2 (they form a sort of rhomboid grid: Fig. 5). In three dimensions the extra degree of freedom in growth direction makes such collisions much less likely. In the random model, in contrast, when TI is low and depletion zones are narrow these zones are quite unlikely to meet even at the higher root concentrations corresponding to two dimensions; thus the random model then underestimates the effect of increasing the dimensions from two to three.

This effect only occurs with networks of low TI, because when TI is high there are no roots of order 3. Because of the constant branching angle and regular spacing of branches, roots of order 2 lie parallel and overlap occurs little more in two dimensions than in three (e.g. Fig. 5F). Therefore with networks of high TI the random model tends to overestimate the effect of increasing the dimensions.

Discussion

To summarize, the reanalysis of Fitter's two-dimensional model using a fully three-dimensional perspective generated the same broad conclusion: exploitation

efficiency increased with increasing TI. Indeed this result seems intuitively obvious enough that to argue the case qualitatively would perhaps not require a model. However, the quantitative relationships that Fitter reported were in large part due to differences in topological efficiency, which resulted solely from the different laws of growth of depletion zones and root volumes. It was this that caused much of the consistency between the three and two-dimensional models. The overlap of depletion zones (as measured by overlap efficiency) changed much less with TI in three dimensions than in two (Fig. 8). Increases in TI above a medium value of 0.75 had little effect on overlap efficiency if diffusion rate was low or root growth fast.

There were three distinct components to why the overlap efficiency was different in three dimensions than in two: (i) with two overlapping cylinders the volume of overlap is proportionately less than the cross-sectional area of the overlap; (ii) a three-dimensional network occupies more volume of soil, so the greater 'dilution' of roots will reduce the probability of overlap; and (iii) the extra degree of freedom of movement in three dimensions disrupts the grid-like arrangements generated in two dimensions.

The lessons from theoretical morphology are also relevant to experimental design. Because of the difficulty of visualizing root systems non-destructively, it is usual to perform experiments where the roots grow in rhizotrons, sandwiched between two sheets of glass so that they are constrained to lie almost in a plane (e.g. Berntson & Woodward, 1992). Either plants will have a fixed root topology adapted to perform well in three dimensions, in which case this topology need not perform well in two dimensions. Or plants will alter their branching decisions according to local nutrient conditions, in which case the topologies that develop in a nearly two-dimensional space should not correspond to those that develop in an unconstrained space. Hendrick & Pregitzer (1996) review several potential reasons why root systems grown in rhizotrons commonly differ from those growing unconstrained, but this explanation is not mentioned.

GENERAL DISCUSSION

Necessarily this paper has so far consisted of three disjunct analyses. I now note some common themes.

Although it is clear that it can be logically incorrect to use a two-dimensional model, it is not so straightforward to say whether or not it is a good approximation. In all the above cases the two-dimensional models clearly demonstrated that the original intuition had some logical consistency. Moreover the quantitative results were often not that dissimilar. However, in other cases two- and three-dimensional models can give results that are very different, even qualitatively different.

One example arises in a paper by Cannings & Cruz Orive (1975). They were modelling whether a male that leaves a group attracts more or fewer females. The females' rules of movement, specified in the model, mean that females coming from some directions will encounter the solitary male rather than the group. The size of the angle for which this is true determines whether it pays males to leave. In two dimensions a 90° angle corresponds to 1/4 of incoming females reaching the solitary male. But in three dimensions an angle of 90° corresponds to only just over 1/7 of incoming females. This can make the difference whether a group is stable.

Another example is a more abstract example from probability theory, but I mention it because it emphasizes a further difference about three dimensions that is relevant to many problems: there is not only more space but there is an extra degree of freedom of movement. Imagine an animal on a two-dimensional square grid. It moves one grid point at a time, and which of the four the directions it chooses is random. One can prove that, given infinite time, it returns to the starting point with probability 1: i.e. it is certain to do so. On a three-dimensional cubic lattice this probability is not 1 but about 0.34 (Hughes, 1995). Durrett & Levin (1994) discuss a different problem, also set on a square grid, in which the state of each grid point has some probability of flipping to the state of a randomly chosen neighbour. Such models have been applied to the spatial spread of genes in a population and of genotypes in a vegetatively spreading plant. On a three-dimensional cubic lattice a patchwork of different genotypes will persist, whereas in one and two dimensions a monoculture will emerge.

Biomechanics provides further examples. For instance consider dome structures, such as limpet shells or the human skull. One might naively suppose that the mechanical properties of a dome could be gauged from those of an arch corresponding to the dome's cross-section. In fact there are important differences (Salvadori & Heller, 1963). For instance, when a circular arch is loaded tensile stresses arise, whereas in symmetric loading of a shallow spherical dome the stresses are all compressive, meaning that structures weak in tension can safely be used. The inadequacy of a two-dimensional analysis is also seen by comparing the properties of a cylindrical barrel vault and a dome, both of which have the same cross-section. The buckling load, for instance, is proportional to the thickness-to-radius ratio in a barrel shell, rather than to the square of this ratio as in a dome.

Conversely, there are some surprising examples, of interest to biologists, of results in three dimensions holding exactly in two dimensions. Thus the mean square displacement after n of steps of a random walk is the same in three dimensions as in two (Hughes, 1995). For perfect fractals the fractal dimension can be estimated equally well from a two-dimensional projection as from the full three-dimensional structure (Nielsen, Lynch & Weiss, 1997). And for a collection of randomly placed roots, whatever their pattern of orientation, the mean distance from a random point to the n'th nearest root is the same in three dimensions as in a plane section with the same density of roots (Marriott, 1972). However, all these results are special cases. Thus the random-walk result does not hold if the mean displacement is considered. No biological structure is a true fractal, and roots are never randomly placed. So these analytic results should be taken only as an indication that two dimensions may be a reasonable approximation of three.

The general lesson that I draw is that two-dimensional models may be valuable as an informal non-quantitative aid to thinking about three-dimensional problems. But quantitative results from two-dimensional models often do not apply to three dimensions, and one should beware even of qualitative results. In fact making a model three-dimensional is often not that much more difficult; the problem has been not realizing that it can make a difference.

In addition I would make two subsidiary points that come out from these reanalyses. The first is that one should treat ratios that are not dimensionless with care, if not with suspicion (cf. Stephens, 1994). If such a ratio is to be maximized it is normally important that comparisons are made at a standard size. This is then equivalent to maximizing a dimensionally balanced ratio. Thus in the first example

I was maximizing the volume of eggs contained within a constant area, which is equivalent to maximizing volume/area$^{1.5}$ when area is not fixed.

The second point is that two of these studies presented numerical computer-generated results that were incorrect even within the terms of the two-dimensional models used by the authors. In geometrical models, as in other quantitative models, it is always worth calculating simple cases by hand as a check.

ACKNOWLEDGEMENTS

Thanks to Zoltan Barta and Alastair Fitter for supplying me with unpublished details of their work, and to John Currey, Alasdair Houston, Heike Reise, David Heath and two other referees for their thoughtful comments on the manuscript. John Currey also supervised my SERC-funded doctorate at the University of York when I first reanalysed Heath's work. The Schools of Mathematics and of Biological Sciences at the University of Bristol generously continued to provide facilities after my contracts there had ceased.

REFERENCES

Andersson M. 1978. Optimal egg shape in waders. *Ornis Fennica* **55:** 105–109.

Bäck T. 1996. Evolution strategies: an alternative evolutionary algorithm. *Lecture Notes in Computer Science* **1063:** 3–20.

Barta Z, Székely T. 1997. The optimal shape of avian eggs. *Functional Ecology* **11:** 656–662.

Berntson GM, Woodward FI. 1992. The root system architecture and development of *Senecio vulgaris* in elevated CO_2 and drought. *Functional Ecology* **6:** 324–333.

Cannings C, Cruz Orive LM. 1975. On the adjustment of the sex ratio and the gregarious behaviour of animal populations. *Journal of Theoretical Biology* **55:** 115–136.

Caruana RA, Schaffer JD. 1988. Representation and hidden bias: Gray vs. binary coding for genetic algorithms. In: Laird J, ed. *Proceedings of the 5th international conference on machine learning*. San Mato, CA: Morgan Kaufmann, 153–161.

Cheetham AH, Hayek LC. 1983. Geometric consequences of branching growth in adeoniform Bryozoa. *Paleobiology* **9:** 240–260.

Durrett R, Levin SA. 1994. Stochastic spatial models: a user's guide to ecological applications. *Philosophical Transactions of the Royal Society Series B* **343:** 329–350.

Fitter AH. 1987. An architectural approach to the comparative ecology of plant root systems. *New Phytologist* **106** (Suppl.)**:** 61–77.

Fitter AH, Stickland TR, Harvey ML, Wilson GW. 1991. Architectural analysis of plant root systems. 1. Architectural correlates of exploitation efficiency. *New Phytologist* **118:** 375–382.

Goldberg DE. 1989. *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison Wesley.

Heath DJ. 1985. Whorl overlap and the economical construction of the gastropod shell. *Biological Journal of the Linnean Society* **24:** 165–174.

Hendrick RL, Pregitzer KS. 1996. Applications of minirhizotrons to understand root function in forests and other natural ecosystems. *Plant and Soil* **185:** 293–304.

Hughes BD. 1995. *Random walks and random environments. Volume 1: random walks*. Oxford: Oxford University Press.

Hutchinson JMC. 1989. Control of gastropod shell shape; the role of the preceding whorl. *Journal of Theoretical Biology* **140:** 431–444.

Hutchinson JMC. 1990a. Control of gastropod shell form via apertural growth rates. *Journal of Morphology* **206:** 259–264.

Hutchinson JMC. 1990b. Design in the shell shape of a terrestrial snail, *Trichia hispida*. Unpublished D.Phil. thesis, University of York, U.K.

Jourdan C, Rey H. 1997. Modelling and simulation of the architecture and development of the oil-palm (*Elaeis guineensis* Jacq.) root system. 2. Estimation of root parameters using the RACINES postprocessor. *Plant and Soil* **190:** 235–246.

Kohn AJ, Myers ER, Meenakshi VR. 1979. Interior remodeling of the shell by a gastropod mollusc. *Proceedings of the National Academy of Sciences of the U.S.A.* **76:** 3406–3410.

Lynch J, Nielsen KL. 1996. Simulation of root system architecture. In: Waisel Y, Eshel A, Kafkafi U, eds. *Plant roots: the hidden half*, 2nd edition. New York: Marcel Dekker, 247–257.

Marriott FHC. 1972. The distance distribution of a random point from randomly spaced lines. *Biometrics* **28:** 874–875.

McGhee GR Jr. 1998. *Theoretical morphology*. New York: Columbia University Press.

Moseley H. 1842. On conchyliometry. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* **21:** 300–305.

Nielsen KL, Lynch JPL, Weiss HN. 1997. Fractal geometry of bean root systems: correlations between spatial and fractal dimension. *American Journal of Botany* **84:** 26–33.

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

Raup DM. 1966. Geometric analysis of shell coiling: general problems. *Journal of Paleontology* **40:** 1178–1190.

Raup DM, Graus RR. 1972. General equations for volume and surface area of a logarithmically coiled shell. *Mathematical Geology* **4:** 307–316.

Rice SH. 1998. The bio-geometry of mollusc shells. *Paleobiology* **24:** 133–149.

Salvadori M, Heller R. 1963. *Structure in architecture*. Englewood Cliffs, NJ: Prentice-Hall.

Schraudolph NN, Belew RK. 1992. Dynamic parameter encoding for genetic algorithms. *Machine Learning* **9:** 9–21.

Stephens DW. 1994. The scale of nature: fitted parameters and dimensional correctness. *Behavioral and Brain Sciences* **17:** 150–152.

Stone JR. 1997. Mathematical determination of coiled shell volumes and surface areas. *Lethaia* **30:** 213–219.

APPENDIX: CALCULATION OF THE SURFACE AREA OF A HELICOSPIRAL

Figure 3F defines a three-dimensional Cartesian co-ordinate system (x, y, z), where the coiling axis is the z axis. In cylindrical co-ordinates, r measures the distance of a point from the coiling axis, and t measures the angle about the coiling axis.

Our logarithmic helicospiral is generated by rotating, expanding, and translating a circular generating curve. Prior to these operations $t=0$, the centre of the generating curve is at (r_c, 0, $b_c r_c$), and its radius is R_0. Consider a very short arc of this curve, from ϕ to $\phi+\delta\phi$ around the curve (Fig. 3A). Its location is ($r_c+R_0\cos\phi$, 0, $b_c r_c - R_0\sin\phi$), which can be rewritten in the form ($r_0(\phi)$, 0, $b(\phi)r_0(\phi)$). Because $\delta\phi$ is very short, this arc is well approximated by its chord, of length and direction given by the vector

$$\mathbf{u}_0 = \begin{pmatrix} -R_0\sin\phi\delta\phi \\ 0 \\ -R_0\cos\phi\delta\phi \end{pmatrix}.$$

The suffix 0 in $\mathbf{u}_0$ signifies that $\mathbf{u}_0$ is a special case of $\mathbf{u}_t$, when $t=0$.

Now consider rotating this straight line about the coiling axis through a very short angle δt, whilst also expanding and translating it. The change of position will be

$$\mathbf{S}_t \approx \begin{pmatrix} \delta r \\ r\delta t \\ \delta z \end{pmatrix}.$$

For a point at angle ϕ around the generating curve, following a logarithmic helicospiral, $r = r_0(\phi)e^{mt}$ and $z = b(\phi)r_0(\phi)e^{mt}$. Thus

$$\mathbf{S}_t \approx \begin{pmatrix} r_0(\phi)me^{mt}\delta t \\ r_0(\phi)e^{mt}\delta t \\ b(\phi)r_0(\phi)me^{mt}\delta t \end{pmatrix} = \begin{pmatrix} m \\ 1 \\ b(\phi)m \end{pmatrix} r_0(\phi)e^{mt}\delta t.$$

To a close approximation the rotation by δt has translated a straight-line section of the outline along a straight-line path given by the vector $\mathbf{S}_t$. The surface area generated is thus approximately the area of the parallelogram with edges formed by the vectors $\mathbf{u}_0$ and $\mathbf{S}_t$. When we are considering a simple surface of revolution (rather than a spiral or helix),

$$\mathbf{S}_t \approx \begin{pmatrix} 0 \\ x\delta t \\ 0 \end{pmatrix},$$

so that $\mathbf{S}_t$ and $\mathbf{u}_0$ are then the edges of a rectangle and we can simply calculate the area as the product of the lengths of $\mathbf{S}_t$ and $\mathbf{u}_0$. In that case Pappus' theorem holds. It also holds for a spiral or helical shell in which the aperture is configured so that $\mathbf{u}_0$ is always perpendicular to the direction of growth $\mathbf{S}_t$; with a helical shell such apertures are not coplanar with the coiling axis and with a spiral shell they are not planar, so Pappus' theorem should not be applied to a cross-section. In other cases the direction of $\mathbf{u}_0$ in the x–z plane affects the angle that $\mathbf{u}_0$ makes with $\mathbf{S}_t$, and thus the area of the parallelogram. This area is given by

$$|\mathbf{u}_0 \times \mathbf{S}_t| = \sqrt{1 + m^2(b(\phi)\sin\phi - \cos\phi)^2}R_0r_0(\phi)e^{mt}\delta\phi\delta t.$$

This formula holds for $t \neq 0$, except that the length of $\mathbf{u}_t$ will increase from $R_0\delta\phi$ to $R_0e^{mt}\delta\phi$, so that the above expression must be multiplied by e^{mt}. The surface area generated by rotating, expanding and translating the vector $\mathbf{u}_0$ through an angle t is this modified expression integrated with respect to t. Thus the surface area of this helical strip of shell from $t = -\infty$ up to the aperture ($t = 0$) is

$$\left[0.5\sqrt{1/m^2 + (b(\phi)\sin\phi - \cos\phi)^2}R_0r_0(\phi)e^{2mt}\delta\phi\right]_{-\infty}^{0} = 0.5\sqrt{1/m^2 + (b(\phi)\sin\phi - \cos\phi)^2}R_0r_0(\phi)\delta\phi.$$

This is a special case of the formula derived by Moseley (1842) for logarithmic helicospirals with generating curves of arbitrary shape. We then have to integrate around the aperture with respect to ϕ to calculate the entire surface area of the shell. Unfortunately this must be done numerically.

As explained above, with a spiral or helix we cannot use Pappus' theorem to calculate surface area as the product of the periphery of the cross-section coplanar with the coiling axis and the arc length traversed by the centre of the periphery. That approach would lead to the expression

$$0.5\sqrt{1/m^2 + 1 + b(\phi)^2}R_0r_0(\phi)\delta\phi,$$

which overestimates or equals the correct formula given above. However, the discrepancy makes little practical difference for realistic parameter values. An extreme case would be where the shell tripled in diameter each revolution, $b_c = 1$, and overlap is high ($F = 1/6$): the discrepancy is then only 1.5%. Raup & Graus (1972) advocated this erroneous approach, and their formula contains an additional error in that the $b(\phi)^2$ term is divided by m^2. Stone's (1997) approach of converting the cross-section to a plane curve perpendicular to the helical trajectory works exaxtly only if the helix is not an expanding spiral.

Similarly in an earlier paper (Hutchinson, 1990a) I was in error in equating shell secretion rates at different sites around the aperture with the arc length traversed by points at those sites (Rice, 1998, follows me in this). Consider a narrow helical strip of shell generated by one short length of the

aperture. If the strip of shell is considered fixed, but the aperture becomes more inclined to the direction of growth, the length of aperture generating this strip increases, so that each mantle cell has to secrete less shell material. Different parts of a planar aperture make different angles with the direction of growth, and I had overlooked that this affects the rate at which different parts of a planar aperture need to secrete shell.

However, the main point of that paper still holds; relative secretion rates around an aperture would have to be controlled impossibly accurately for this to be the mechanism that regulates the apical angle of a shell ($\propto b_c$). As in Hutchinson (1990a), let us consider parameters typical of a terrestrial helicid ($m = 0.068$, and $R_0/r_c = 0.415$) and compare growth rates around a circular aperture coplanar with the coiling axis. It is obvious that in the planispiral case ($b_c = 0$) the sites of maximum and minimum shell secretion are at $\phi = 0°$ and $\phi = 180°$. When $b_c = 1$ the maximum and minimum are at $\phi = -0.65°$ and $\phi = 180.66°$, only very slightly off a diameter; at $\phi = 0°$ secretion rate is only 0.1% below that at the maximum.