Supplementary material for: Hierarchical mechanical patterns in morphogenesis: from mollusc shells to plants, fungi and animals

D.E. Moulton, A. Goriely, & R. Chirat

Contents

1 Shell secretion in molluscs

The external shell of molluscs is composed of an outer organic layer, the periostracum, and underlying calcified layers. During growth, the mantle moves forward slightly beyond the calcified shell edge while secreting the periostracum, which isolates the supersaturated extrapallial fluid from which the calcified shell is precipitated [49]. The periostracum is secreted in the periostracal groove, between the outer and inner lobe when the mantle edge is bi-lobed as in gastropods (e.g. [21]), or between the outer and middle lobes when the mantle edge is three-lobed (i.e., outer, middle and inner lobes) as in bivalves (e.g. [5]). During its secretion, the periostracum closely adheres to the secreting epithelium of a mantle lobe [8, 11]. The growing sheet of periostracum is extruded between the mantle lobes, while thickening and stiffening through sclerotization. It subsequently reaches its external position where it serves as template for carbonate crystal nucleation and shell formation and becomes fixed on the outer shell surface [19, 38].

Mismatch in length. The structure of the generative zone basically corresponds to a configuration in which a thin and stiff layer (the periostracum), bound to a thicker and softer foundation (the lobe of the mantle edge), is continuously secreted during the accretionary growth. In physical systems,

this configuration is well-known to generate wavy patterns when a mismatch in lenght/surface between both layers generates compressive stresses, triggering a wrinkling instability minimizing the energy of the system [31, 27]. Wrinkling instabilities are ubiquitous in nature, and generate multiscale wavy patterns not only in a wide range of physical systems, but are also involved in the development of traits as diverse as fingerprints, pathological mucosal folding in airways, or lung, brain and gastro-intestinal morphogenesis [23, 33, 40, 6, 44, 2, 17]. In most of the bilayered physical systems studied so far, wrinkling instabilities emerge either when the upper layer swells (when immersed in a solvent) and expands relative to the substrate, or when the upper layer is bound to a pre-stretched substrate so that compressive stresses are generated on relaxation of the system. In both cases, the upper layer displays an excess of surface relative to the substrate, which mimics differential growth of the two tissues in biological systems. Both situations have been used as a physical simulacrum for studying biological pattern formation, for example, brain morphogenesis in which higher growth rate of the cortical layer relative to sublayers generates compressive stress and folding of the cortex [10, 20]. In the case of the folding of the generative zone of molluscs, an analogous situation is unlikely as it would imply that the periostracum would grow at a higher rate than the mantle that secretes it. In physical systems, the mismatch in surface length between layers triggering a mechanical instability may be also generated by the generic behaviour of shrinkage triggered by polymerization and cross-linking (chemical bonds linking polymers to each other). In inducing different polymerization rates and cross-linking between the surface and the bulk of a film, a mismatch of shrinkage generates compressive stresses triggering a wrinkling instability [37]. As seen above, the periostracum starts to be secreted in the periostracal groove, and then it moves forward attached to the mantle lobe, while thickening and stiffening through sclerotization [19, 38]. Sclerotization is a complex biochemical process of cross-linking fibrous proteins, causing close packing of the polymers, dehydration and stiffening. Lateral shrinkage of the stiffening and thickening periostracum during the cross-linking process of sclerotization can clearly induce continuous compressive stresses on the underlying softer growing mantle epithelium, and trigger a mechanical instability folding the generative zone.

Memory. In molluscs, a part of the shell edge may be broken as a result of trauma. In that case, the mantle retracts behind the broken edge and secrets new shell material replacing the missing part. Yet, the spiral ridges pattern reappears totally unchanged after shell trauma, which shows that the setting up of this pattern involves a memory/irreversibility condition, i.e. that once a ridge secreting zone is formed in the mantle epithelium, it does not disappear during development. The fact that the mantle epithelium is irreversibility patterned explains why in Muricidae, the spiral ridges pattern remains unaffected by the repeated extension and retraction of the mantle edge that characterizes the episodic shell growth of this family. The spiral ridges are secreted by zones of the mantle edge that remain folded once the body of the animal is extracted from its shell [21, 47]. Also, they are secreted by thickened zones of the secreting epithelium, with columnar cells taller and more tightly packed than in the adjacent inter-ridges spaces, increasing the amount of shell material secreted locally and the thickness of outer prismatic layer [47]. In summary, the mantle epithelium is compressed and irreversibly remodelled at the level of the folded zones of the mantle edge secreting the ridges, which provides a memory to the system.

1.1 A note on the mechanical hypothesis

The above descriptions comprise known facts about mollusc shell growth and the potential for mechanical interactions. The model developed in this paper is based on an idealisation of these

physical interactions, following the hypothesis that forces induced by a length mismatch between different components of the shell secreting system can account for the morphogenesis of spiral ridges and spines. In this way, the present work conceptually follows similar ideas as [13], which specifically studied the formation of a single spine. That model was based on the well-known fact that spines emerge episodically during phases of excess in marginal growth rate, that is when the marginal growth of the shell increases relative to the spiral growth. In [13], the mechanical outcome of this excess of marginal length of the mantle was investigated, by imposing a length mismatch with the shell secreted in the previous increment. By varying growth and stiffness parameters in the morphomechanical model, a wide variety of realistic spine forms were generated.

The model for spines presented in the current paper follows a similar conceptual approach as that developed in [13], though extended from analysis of a single spine to the more complex hierarchical ridges and multi-spine forms to which the ridges amplify upon a burst of growth. While both the model we develop here and the work [13] are based on known biological and physical processes, we note than an important step remains missing: an experimental validation. Unfortunately, though mollusc shells constitute a model system in theoretical biology, experimental approaches on these organisms remains a huge challenge for three main reasons: the slowness of the shell secretion process (with growth rates of only a few tenths of a mm per day), the micrometer dimensions of the generative zone, and especially the fact that when a mollusc is disturbed by an observer, it retracts into its shell and stops secreting. Added to this is the fact that most species with spines are marine molluscs that are difficult to maintain in aquariums under optimal growth conditions, the challenges posed to any experimental approach are vast. Still, despite the difficulties in a full experimental validation, a theoretical approach based on mechanics nevertheless has a predictive power, through comparison of model output with observed shell patterns, and strong potential to elucidate key interactions and mechanisms.

2 Modelling approach

2.1 Conceptual idea

We consider a planar system consisting of two elastic layers, termed Layer 1 and Layer 2, each growing, i.e. increasing in length, such that Layer 1 grows at a faster rate. Attachment of the layers means that Layer 1 is constricted to the same domain is Layer 2. For inextensible material, an initial length difference will thus immediately create a wrinkling instability in Layer 1 in order to accommodate the excess length. Our objective is to first classify the instability and then to continue growing both layers well beyond the initial buckling. With continued growth of the layers, expansion of the domain will create space between existing wrinkles; our aim is to characterize the occurrence of secondary, tertiary, etc. buckling events that form a hierarchical pattern. In order to isolate the key minimal ingredients for the generation of such an evolving pattern, we make several simplifying assumptions, that will be justified a posteriori:

A1 We seek a buckled shape within the class of Gaussian functions.

A2 We assume small deformations, enabling for a valid linearization of energy terms.

Assumption A1 is motivated by the prevalence of examples in nature with patterns consisting of shapes with a simple 'bump-like' geometry, all oriented in the same direction (see examples described in main text). To approximate such patterns, the key first step to our approach is to

characterize the pattern as a series of 'Gaussian bumps'. That is, we suppose that the shape of the beam is described by the curve $(S, y(S))$, where y consists of a sum of terms of the form

$$a_i(t) \exp\left(\frac{(S-S_i)^2}{\sigma_i^2}\right).$$

As we show below, the advantage of this approach is that by constricting the shape, and in conjunction with assumption A2, we obtain a simple analytical expression for the mechanical energy, which enables to easily and explicitly determine both the amplitudes a_i as functions of time as well as bifurcation points, i.e. time points at which a new set of Gaussian bumps appears, for the evolving pattern. While the framework we develop provides a novel approach to hierarchical patterns in a growing two-layer system, the basic idea of constricting the class of deformation within energy methods dates back to the classic work of Timoshenko [41].

2.2 Growth

We model the two layer system as a growing planar elastic beam, Layer 1, elastically attached to a foundation, Layer 2. Let the centerline of Layer 1 be given by

$$\mathbf{r}(S) = S\mathbf{e}_x + y(S,t)\mathbf{e}_y, \quad -L(t) < S < L(t). \tag{1}$$

The correspondence between S and the horizontal coordinate follows A2 and reflects the notion that our primary interest is in patterns for which the deformation from the flat state is not very large, i.e. for which the patterned state may be well-approximated by a graph $y(x)$; though we note that when considering fractal patterns (Sec 5) we must revisit this and take a more complex form. The domain $[-L(t), L(t)]$ defines the length of Layer 2, $2L(t)$, which is related to the initial length $L(0) = L_0$ by the growth factor $\hat{\gamma}$, i.e. $L(t) = \hat{\gamma} L_0$ and such that $\hat{\gamma}$ satisfies a given growth law. To make this system dimensionless, we scale time on the growth time scale of Layer 2, and we scale the domain lengths by the initial length L_0 so that in dimensionless form $\partial_t \hat{\gamma} = 1$. Therefore, $\hat{\gamma} = 1 + t$ and thus

$$L(t) = 1 + t. \tag{2}$$

Layer 1 is also growing, such that its dimensionless length is $2l(t)$ which relates to the initial length l_0 via $l(t) = \gamma l_0$. The growth stretch γ is also assumed to be spatially uniform and increasing at a constant rate $g > 1$, i.e.

$$\partial_t \gamma = g,$$

which gives

$$l(t) = (1 + gt) l_0. \tag{3}$$

The condition $g > 1$ means that Layer 1 is growing at a faster rate than Layer 2, which is a key ingredient for the evolving pattern that we study. We denote 2δ as the evolving length differential between the layers:

$$\delta(t) = l(t) - L(t) = \delta_0(1 + (g-1)t), \tag{4}$$

where $\delta_0 = l_0 - 1$ is the length differential, if any, at $t = 0$. We primarily consider in this work an inextensible rod, for which the arc length of the curve $\mathbf{r}$ must equal the length $2l(t)$, that is at all times we have the condition

$$\int_{-L(t)}^{L(t)} \sqrt{1 + \epsilon \left(\partial_S y(S,t)\right)^2}\, dS = 2l(t). \tag{5}$$

Here $\epsilon := A^2/L_0^2$ relates a characteristic length A of the pattern height, i.e. a length scale for dimensional y, and the domain length scale L_0. For the patterns we consider, Assumption A2 implies that $\epsilon \ll 1$, which enables to approximate the length constraint as

$$\int_{-L(t)}^{L(t)} 1 + \frac{1}{2}\epsilon \left(\partial_S y(S,t)\right)^2 \, dS = 2l(t). \tag{6}$$

2.3 Energy

For given mechanical parameters, initial lengths L_0, l_0, and growth parameter g, the curve y is determined by minimizing the mechanical energy, which we assume consists of bending energy and a foundation energy that characterizes the resistance of Layer 2 to deformation of Layer 1. Note that implicit in this approach is the assumption that growth is occurring on a much longer time scale than the elastic time scale, so that at all times the system is assumed to be in quasi-static mechanical equilibrium.

The bending energy is one half the product of the bending stiffness E_b and the square of the curvature of the centerline $\mathbf{r}(S)$; under assumption A2 the linearized bending energy is proportional to the second spatial derivative of y, $\partial_{SS} y(S,t)$. We define $\frac{K}{2} f(y)$ as a potential energy for the foundation, i.e. for the interaction of the layers, where K has units of energy per length, f is a dimensionless function satisfying $f(0) = 0$, and $f(y) > 0$ for $y > 0$, i.e. the interaction energy is minimized in the flat state[1]. Specific functional forms of f are considered below. Combining these components, we define the energy

$$\mathcal{E}[y(S,t)] := \int_{-L(t)}^{L(t)} \frac{\epsilon E_b}{2L_0} \left(\frac{\partial^2 y(S,t)}{\partial S^2}\right)^2 + \frac{KL_0}{2} f(y(S,t)) \, dS, \tag{7}$$

which should be minimized subject to the length requirement (6).

2.4 Gaussian reduction

Following assumption A1, we seek $y(S,t)$ as a series of Gaussian bumps:

$$y(S,t) = \sum_{i=1}^{N} a_i \exp\left(\frac{-(S-S_i)^2}{\sigma_i^2}\right). \tag{8}$$

Here a_i defines the amplitude, S_i the location, and σ_i the width of the ith Gaussian bump. Each of these may in principle vary with time, though as we show below the variation of amplitude is the most relevant component for energy minimization and pattern evolution. Despite an assumed functional form, for any more than a single Gaussian, the energy (7) remains very cumbersome to compute analytically (and energy minimization may not be feasible analytically at all depending on the form of $f(y)$). Therefore, we make two additional assumptions to enable progress:

A3 The Gaussian bumps do not significantly overlap.

A4 Each Gaussian bump "fits" inside the domain.

[1]The potential energy f should also satisfy $f > 0$ for $y < 0$, though our analysis explicitly restricts to unidirectional patterns $y > 0$.

These assumptions will be defined more explicitly below. Conceptually, the benefit of A3 is that each energy component may be integrated separately for each individual Gaussian, i.e. we remove the coupling in the integration as well as any dependence in the energy on the specific locations of the Gaussians encoded in the S_i. Assumption A4 further implies that we may approximate each Gaussian integral by an integral from $-\infty$ to ∞, since the 'tails' will have negligible contribution. Following these simplifying assumptions, for the bending energy we may write

$$\int_{-L(t)}^{L(t)} \left(\frac{\partial^2 y(S,t)}{\partial S^2}\right)^2 \, dS \approx \sum_{i=1}^{N} a_i^2 \int_{-\infty}^{\infty} \left[\frac{\partial^2}{\partial S^2}\left(\exp\left(\frac{-S^2}{\sigma_i^2}\right)\right)\right]^2 \, dS. \tag{9}$$

Expanding the derivative, the integrals on the right may be computed explicitly, giving

$$\sum_{i=1}^{N} \frac{3\sqrt{\pi} a_i^2}{\sqrt{2}\sigma_i^3}. \tag{10}$$

Turning to the length constraint (6), the combination of assumptions A1, A3, and A4 enables the simplification

$$\int_{L(t)}^{L(t)} \left(\partial_S y(S,t)\right)^2 \, dS \approx \sum_{i=1}^{N} a_i^2 \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial S}\left(\exp\left(\frac{-S^2}{\sigma_i^2}\right)\right)\right]^2 \, dS. \tag{11}$$

Again the right hand side may be evaluated explicitly, and recalling the definition (4), the length constraint reduces to the algebraic condition

$$\sum_{i=1}^{N} \frac{\sqrt{\pi}\epsilon a_i^2}{2\sqrt{2}\sigma_i} = 2\delta(t). \tag{12}$$

If the interaction (foundation) energy is described by a polynomial form for f, then a similar reduction can be made. For instance, a linear foundation has quadratic energy $f(y) = \frac{K}{2}y^2$, where the constant K describes the foundation stiffness, while a nonlinear foundation may be described by $f(y) = \frac{K}{2}y^m$ for integer $m > 2$ (In dimensional form, we could write $f(y) = (y/A)^m$, so that K has units of energy for all values of m. Here, we have assumed that y has already been scaled by A.) A generic f could be represented by a polynomial spline with multiple terms, though we restrict our analysis to single term polynomials. For this, an explicit form can be obtained for arbitrary integer $m \geq 2$ as follows:

$$\int_{-L(t)}^{L(t)} y^m \, dS \approx \sum_{i=1}^{N} a_i^m \int_{-\infty}^{\infty} \left(\exp\left(\frac{-S^2}{\sigma_i^2}\right)\right)^m \, dS = \sum_{i=1}^{N} \frac{\sqrt{\pi} a_i^m \sigma_i}{\sqrt{m}}. \tag{13}$$

2.5 Hierarchical construction

A hierarchical pattern is defined by a sequence of bifurcations each leading to increasing complexity in the pattern. Here we suppose that each bifurcation event is characterized by the appearance of a new level of Gaussian bumps all with the same characteristics. We define Level 1, the initial bifurcation, by the form

$$y(S,t) = \sum_{i=1}^{N_1} a_1 \exp\left(\frac{-(S-S_i)^2}{\sigma_1^2}\right). \tag{14}$$

Notice that each of the N_1 bumps in Level 1 all have the same amplitude a_1 and width σ_1, varying only in their location, which is defined by the S_i. For a Level 2 pattern, we add to the form above a second family of Gaussians:

$$y(S,t) = \sum_{i=1}^{N_1} a_1 \exp\left(\frac{-(S-S_{i_1})^2}{\sigma_1^2}\right) + \sum_{i=1}^{N_2} a_2 \exp\left(\frac{-(S-S_{i_2})^2}{\sigma_2^2}\right). \tag{15}$$

All N_2 of the Gaussians in the second family have amplitude a_2 and width σ_2, and with more than one family we distinguish the locations with subscripts S_{i_1} and S_{i_2}. Bifurcations to higher levels follow the same pattern.

As outlined in the main text, the numbers of Gaussians in each family follows a simple rule: each new level appears in between each bump of the already present pattern. Thus, if the initial pattern has $N_1 = 2$, which we numerate as a 1.1 pattern, then at Level 2 the pattern reads 1.2.1, and at Level 3 we have 1.3.2.3.1. According to this structure, $N_2 = N_1 - 1$, $N_3 = 2(N_1 - 1)$, and we may extrapolate to $N_k = 2^{k-2}(N_1 - 1)$. Alternatively, depending on how the edges are treated, the Level 2 pattern could read 2.1.2.1.2, with Level 3 having the form 3.2.3.1.3.2.3.1.3.2, which yields the general rule $N_k = 2^{k-2}(N_1 + 1)$. We restrict our analysis to these two formulations in this paper, though we note that the framework is easily extended to more generic scenarios, e.g. if the domain is expanding non-uniformly so that higher level patterns are not uniformly spread between lower levels.

2.6 Energy minimization

Following the Gaussian reduction and hierarchical construction, the energy at Level k is expressed as an algebraic function of k amplitudes $\{a_1, a_2, \ldots, a_k\}$, and k widths $\{\sigma_1, \sigma_2, \ldots, \sigma_k\}$. Despite the reduction from a function space to a finite algebraic space, minimizing the energy and locating the bifurcation points at which the system transitions to higher levels remains quite complex. To simplify further, and to more precisely define the assumptions of the Gaussians fitting in the domain without significant overlap, A3 and A4, we consider more explicitly the Level 1 pattern.

2.7 Level 1

The energy at Level 1, for a foundation potential $f(y) = \frac{K}{2}y^m$, is given by

$$\mathcal{E}_1(a_1, \sigma_1) = N_1 \left(\frac{3\sqrt{\pi}\epsilon E_b}{2\sqrt{2}L_0} \frac{a_1^2}{\sigma_1^3} + \frac{KL_0\sqrt{\pi}}{2\sqrt{m}} a_1^m \sigma_1 \right). \tag{16}$$

To simplify the notation, we divide the energy by the coefficient of the first term, giving the dimensionless form

$$E_1(a_1, \sigma_1) = N_1 \left(\frac{a_1^2}{\sigma_1^3} + \mu a_1^m \sigma_1 \right), \tag{17}$$

where

$$\mu := \frac{KL_0^2\sqrt{2}}{3E_b\epsilon\sqrt{m}}$$

is a dimensionless parameter characterizing the ratio of foundation to bending stiffness. This should be minimised subject to the length constraint

$$N_1 \frac{\sqrt{\pi}\epsilon}{2\sqrt{2}} \frac{a_1^2}{\sigma_1} = 2\delta(t). \tag{18}$$

Again, for notational simplicity, we may absorb the coefficient $\sqrt{\pi}\epsilon/(2\sqrt{2})$ into the definition of $\delta(t)$ by appropriately redefining δ_0 and g. Thus, in what follows we represent the length constraint as

$$N_1 \frac{a_1^2}{\sigma_1} = 2\delta(t) = 2(\delta_0 + gt). \tag{19}$$

Solving (19) for a_1 and inserting into (17) gives an equation for σ_1, which has a single minimum located at

$$\sigma_1 = 16^{\frac{1}{m+6}} \left(\mu(m+2)\delta^{\frac{m}{2}-1} N_1^{1-\frac{m}{2}}\right)^{-\frac{2}{m+6}}. \tag{20}$$

In the case $m = 2$, this reads $\sigma_1 = \mu^{-1/4}$. For $m = 4$, a case we analyze extensively, the formula reads

$$\sigma_1 = \left(\frac{2N_1}{3\delta\mu}\right)^{1/5}.$$

Note that time only enters in $\delta = \delta(t)$. Thus, in the case $m = 2$, the energy minimizing value of σ_1 is fixed for all time. In the case $m = 3$, $\sigma_1 \sim \delta^{-1/9}$, while for $m = 4$, $\sigma_1 \sim \delta^{-1/5}$. Since δ is a linear function of t, for moderate values of m, σ is a strongly sublinear function of t. That is, there is a slow decrease in σ_1 with time, implying a slight decrease in Gaussian width as the pattern evolves, though the width remains close to constant. This apparent near constancy of width of each bump is also exhibited in many biological hierarchical patterns, though note in all. Indeed, spiral ridges in mollusc shells increase both in amplitude and width during development. The increase in width is likely a consequence of the expanding domain. These ridges correspond to irreversibly remodeled and thickened zones of the secreting epithelium; these thickened zones likely grow along with the rest of the mantle epithelium, so it's not surprising that they increase in width. In principle, an alternative formulation of our model could involve setting a rule for the increase in the width of each bump based on domain growth, though clearly extra space between bumps must still be created, i.e. it is not the case that the initial wrinkling pattern simply expands outward with domain growth, or else bifurcation to higher levels could not occur. We anticipate that under such a rule, the qualitative features of the model would not change, only the quantitative features, such as bifurcation points. Lacking a clear logic for a rule for growth-induced width expansion, here we stick with the simplest formulation of constant width. Thus, in order to aid further analytical progress, we add an additional assumption:

A5 The width of the Gaussian bumps at each level do not change in time.

That is, once we have determined the value of σ_i for a given level, it is kept constant, and thus the energy is only a function of the amplitude a_i.

2.7.1 Mode selection

It remains to determine the number of Gaussians, N_1, which we may term the mode of the pattern. To do this, we first characterize the width of a single Gaussian bump. Consider a Gaussian centered at the origin,

$$y(S) = a \exp\left(\frac{-S^2}{\sigma^2}\right).$$

For small parameter $0 < \varepsilon \ll 1$, $y(S) = \varepsilon$ may be solved for S, yielding

$$S_\varepsilon = \sigma\sqrt{\ln\left(\frac{a}{\varepsilon}\right)}.$$

Since $\sqrt{\ln a}$ is a slowly increasing function of a, and a is assumed to not grow too large, and σ is effectively constant, the distance from the Gaussian center to the point at which the height is S_ϵ is close to constant. From this, we thus define the half-width of a Gaussian bump $h = h(\sigma)$ as

$$h(\sigma) = \alpha\sigma,$$

where $\alpha \approx 2$ is taken as a fixed parameter that comes from a choice of a/ε; e.g. $a = 1$, $\varepsilon = 0.1$ gives $\alpha = 1.517$, while $a = 1$, $\varepsilon = 0.01$ gives $\alpha = 2.146$.

Having defined a width of each Gaussian, the number of Gaussians in Level 1 may now be determined from the requirement that the pattern fits within the domain. At $t = 0$, there is an excess of length $2\delta_0$ between the layers. Since the beam is inextensible, an initial length differential implies that the pattern must appear immediately, at $t = 0$, when the domain length is $2L(0) = 2$. At this point, N_1 Gaussian bumps of width $\alpha\sigma_1$ fit within the domain if

$$2N_1\alpha\sigma_1 \leq 2. \tag{21}$$

Inserting the energy minimizing value of σ_1 back into the energy, and also replacing a_1 by the length constraint, we may express the energy at $t = 0$ in terms of N_1 (and the parameters δ_0 and μ). We find for $m > 2$ that the energy is a monotonically decreasing function of N_1, e.g. in the case $m = 4$ we have

$$E_1 = 5\left(\frac{\delta_0^7\mu^2}{18N_1^2}\right)^{1/5} \tag{22}$$

The observed mode is the one with the smallest energy that still satisfies the constraint of fitting in the domain. That is, the mode N_1 is determined as the largest value that satisfies the condition

$$N_1\sigma_1(N_1) \leq 1 \tag{23}$$

where $\sigma_1(N_1)$ comes from (20) with $\delta = \delta(0) = \delta_0$.

The argument above holds for $m > 2$. In the case $m = 2$, the energy minimizing value of σ_1, $\sigma_1 = \mu^{-1/4}$, is independent of N_1, as is the energy at $t = 0$. This is because for $m = 2$, both terms in the energy are proportional to $N_1a_1^2$, and the length constraint fixes $N_1a_1^2$ in terms of δ. To be consistent, in this case we still fix the mode as the largest value that satisfies the constraint $\alpha N_1\sigma_1 \leq 1$. While taking a smaller value of N_1 would technically have the same energy, it would leave more of the domain flat and would require a larger initial amplitude $a_1(0)$ at the locations of the Gaussians, which is unrealistic in terms of physical patterns. In any case, as we show below, the case $m = 2$ is irrelevant for hierarchical pattern formation.

2.8 Level 1 amplitude

To summarise the above subsections, the pattern at Level 1 is specified by the following steps:

1. The mode, or number of Gaussians N_1, is given by the smallest integer N_1 satisfying (23).
2. The width σ is given by (20), with $\delta = \delta_0$.
3. The amplitude $a_1 = a_1(t)$ is determined from (19).

Note that since $\delta = \delta_0 + gt$ and σ_1 and N_1 are constant, the amplitude curve has parabolic form.

2.9 Level 2 energy and bifurcation

To determine if and when the pattern will bifurcate to Level 2, we first define the Level 2 energy. In dimensionless form, this reads

$$E_2(a_1, a_2; \sigma_1, \sigma_2) = \frac{N_1 a_1^2}{\sigma_1^3} + \frac{N_2 a_2^2}{\sigma_2^3} + \mu \left(N_1 a_1^m \sigma_1 + N_2 a_2^m \sigma_2 \right), \tag{24}$$

Here the Level 2 pattern includes the same Gaussian terms as in Level 1, plus an additional N_2 Gaussians of amplitude a_2 and width σ_2. The first two terms account for the bending energy, while the latter two account for the foundation energy. As noted earlier, for a uniformly expanding domain, N_2 is determined by N_1, either satisfying $N_2 = N_1 + 1$ or $N_2 = N_2 - 1$. The energy (24) is to be minimized subject to the length constraint

$$N_1 \frac{a_1^2}{\sigma_1} + N_2 \frac{a_2^2}{\sigma_2} = 2\delta(t). \tag{25}$$

Following the analysis at Level 1, we express the energy E_2 as a function only of the amplitudes a_1 and a_2, and depending *parametrically* on the widths σ_1 and σ_2. By assumption A_5, the widths are taken to be constant. Thus, σ_1 is already determined at Level 1. To determine σ_2, we utilize assumptions A_3 and A_4. For N_1 Gaussians of width $2\alpha\sigma_1$ and N_2 Gaussians of width $2\alpha\sigma_2$ to fit within the domain, it must hold that

$$2N_1\alpha\sigma_1 + 2N_2\alpha\sigma_2 \leq 2L(t) = 2(1+t). \tag{26}$$

Defining t_2 as the bifurcation time at which the Level 2 pattern appears (t_2 to be determined), we can replace the inequality in (26) with equality and thus define σ_2 via

$$\sigma_2 := \frac{1 + t_2 - N_1\alpha\sigma_1}{N_2\alpha}. \tag{27}$$

To determine t_2 and the energy minimizing amplitudes a_1, a_2, in principle one can solve the length constraint (25) for a_1, insert into (24) and then minimize E_2. For general m this cannot be done analytically. Thus, to proceed we consider the specific cases $m = 2$ and $m = 4$.

2.9.1 m=2

In the case $m = 2$, upon solving the length constraint for a_1 and inserting in the energy, E_2 has the form

$$E_2 = \frac{\delta\sigma_1^3 \left(\mu\sigma_1^4 + 1\right) + a_2^2 N_2 \left(\sigma_1^2 - \sigma_2^2\right)\left(1 - \mu\sigma_1^2\sigma_2^2\right)}{\sigma_1^2\sigma_2^3}, \tag{28}$$

which is minimized at $a_2 = 0$. We are led to conclude that within this framework, a bifurcation to Level 2 does not occur for $m = 2$. That is:

A linear foundation does not generate a hierarchical pattern.

2.9.2 m=4

In the case $m = 4$, the same steps lead to an energy

$$E_2 = \frac{\mu\sigma_1^3 \left(a_2^2 N_2 - \delta\sigma_2\right)^2}{N_1\sigma_2^2} + \frac{\delta - \frac{a_2^2 N_2}{\sigma_2}}{\sigma_1^2} + \frac{a_2^2 N_2 \left(a_2^2\mu\sigma_2^4 + 1\right)}{\sigma_2^3}. \tag{29}$$

This form yields an explicit and non-trivial minimum, given by

$$a_2 = \frac{\sqrt{2\delta\mu\sigma_2^2\sigma_1^5 - N_1\sigma_1^2 + N_1\sigma_2^2}}{\sqrt{2}\sqrt{\mu N_2\sigma_2\sigma_1^5 + \mu N_1\sigma_2^4\sigma_1^2}}. \tag{30}$$

(Similarly, an explicit and non-trivial solution may be found in the case $m = 3$.)

Given a_2, we also have an explicit expression for a_1:

$$a_1 = \sqrt{\frac{\sigma_1}{N_1}\left(2\delta(t) - \frac{N_2 a_2^2}{\sigma_2}\right)}. \tag{31}$$

It remains to determine t_2, the bifurcation time, which lives implicitly in the formulas above via the form for σ_2 given in (27). The width σ_2 is also to be determined, but this is defined in terms of t_2 via (27). At $t = t_2$, $a_2 = 0$ by definition, therefore we may find t_2 by setting $a_2 = 0$ in (30), with σ_2 replaced by (27), and with $\delta = \delta_0 + gt_2$. The resulting equation may be solved explicitly for t_2, though with a formula that is too cumbersome to reproduce here.

By construction, the energy $E_1(a_1)$ matches exactly $E_2(a_1, a_2)$ at the bifurcation point $t = t_2$, since $a_2 = 0$ at that point. It is readily verified, for instance by direct plotting of both E_1 and E_2 against time past the bifurcation point, or by the shape of E_2 as a function of a_2 in (30), that $E_2 < E_1$ for $t > t_2$; in other words, the appearance of the Level 2 curve a_2 does indeed represent a physical bifurcation to a lower energy state.

2.10 Level 3 and beyond

The procedure described above is naturally extended to Level 3 and beyond. Indeed, in the case $m = 4$, an explicit analytical expression may be obtained for the amplitudes, widths, and bifurcation points.

At Level k, the requirement for the pattern to fit is given by

$$\sum_{i=1}^{k} 2N_i\alpha\sigma_i \leq 2L(t) = 2(1 + t), \tag{32}$$

where we recall that for a uniformly expanding domain, $N_k = 2^{k-2}(N_1 + 1)$ or $N_k = 2^{k-2}(N_1 - 1)$. Since $\{\sigma_1, \sigma_2, \ldots \sigma_{k-1}\}$ are assumed fixed from the previous levels, this gives a formula for σ_k in terms of the bifurcation time t_k:

$$\sigma_k := \frac{1 + t_k - \sum_{i=1}^{k-1} N_i\alpha\sigma_i}{N_k\alpha}. \tag{33}$$

The length constraint at Level k is

$$\sum_{i=1}^{k} N_k\frac{a_k^2}{\sigma_k} = 2\delta(t), \tag{34}$$

which may be solved for a_1 and then inserted in the energy

$$E_k(a_1, a_2, \ldots, a_k; \sigma_1, \sigma_2, \ldots, \sigma_k) = \sum_{i=1}^{k}\left(\frac{N_i a_i^2}{\sigma_i^3} + \mu N_i a_i^m\sigma_i\right). \tag{35}$$

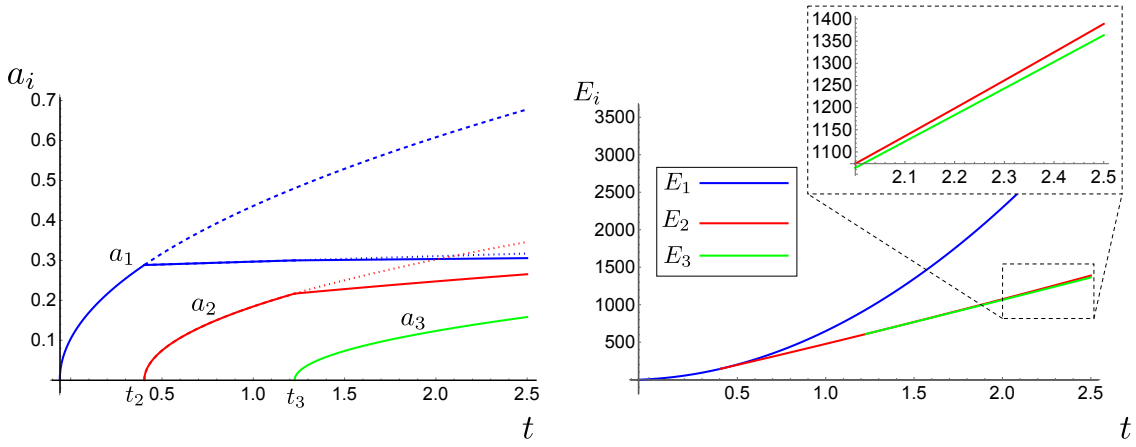

Figure 1: Left: amplitude curves including bifurcations up to Level 3. Right: corresponding energies, showing that bifurcation to higher levels represents a path with lower energy.

The amplitudes $\{a_2, a_3, \ldots, a_k\}$ come from solving simultaneously the equations

$$\frac{\partial E_k}{\partial a_i} = 0, \quad i = 2, 3, \ldots, k$$

and the bifurcation time t_k is determined by taking the explicit expression for a_k, setting $a_k = 0$, setting $\delta = \delta_0 + gt_k$, replacing σ_k by (33), and solving the resulting expression for t_k.

Using a computer algebra software, explicit expressions are easily produced, at least in the cases $m = 3$ and $m = 4$. Using Mathematica, we have produced explicit expressions up to Level 5; in principle the algebra enables to continue to higher levels, though predictably the algebraic expressions become increasingly complex.

2.11 Energy comparison

To demonstrate that the bifurcation to higher hierarchical levels does indeed represent a physical bifurcation, i.e. a transition to a path with lower energy, we compare in Fig 1 the total energy for the first three levels. At left we plot the amplitude curves, $a_1(t)$ (blue), $a_2(t)$ (red), and $a_3(t)$ (green), where the dashed blue curve shows the amplitude a_1 following the Level 1 path, i.e. ignoring the bifurcation to Level 2, and similarly the dotted red and blue curves show the amplitudes a_1 and a_2 following Level 2 past the Level 3 bifurcation. At right are the corresponding energies $E_i(t)$, $i = 1, 2, 3$. The Level 2 energy only appears past the bifurcation time t_2, at which point it is clearly lower than the Level 1 energy. Similarly, the Level 3 energy is lower than the Level 2 energy once it appears at bifurcation time t_3. The difference between E_2 and E_3 is smaller, but it is clear from the inset that E_3 is indeed lower.

While this demonstrative example of course does not constitute formal proof, we note that the energy ordering $E_{i+1} < E_i$ for $t > t_{i+1}$ has held in all cases that we have examined.

2.12 The physical pattern

At a given level k, the energy minimizing approach above specifies the amplitudes $\{a_1(t), a_2(t), \ldots, a_k(t)\}$ as functions of time and the widths $\{\sigma_1, \sigma_2, \ldots, \sigma_k\}$. In order to display the pattern, it remains to

specify the physical locations of the Gaussian bumps. That is, at Level k the pattern is described by $(x, y) = (S, y(S, t))$ for $S \in [-L(t), L(t)]$ where

$$y(S,t) = \sum_{i=1}^{k} \sum_{j=1}^{N_i} a_i(t) \exp\left(\frac{-(S - S_{i_j})^2}{\sigma_i^2}\right). \tag{36}$$

Here the outer sum is over each of the levels 1 to k, while the inner sum is over the Gaussians within a level. The value S_{i_j} describes the location of the jth bump of the ith level. For a uniformly expanding domain, as noted, the next level appears in between each bump of the previous level. Consider first the case where new bumps appear also on the outside edges. That is, if the initial pattern is 1.1, this is continued until bifurcation point t_2, at which point the pattern is 2.1.2.1.2. This holds until the next bifurcation point t_3, at which time the pattern transitions to 3.2.3.1.3.2.3.1.3.2.3. Subsequently, at $t = t_4$, we have 4.3.4.2.4.3.4.1.4.3.4.2.4.3.4.1.4.3.4.2.4.3.4, and so on.

At the bifurcation points t_k, the next level pattern exactly fits within the domain by construction. But in between bifurcation points, for $t_k < t < t_{k+1}$, the pattern remains at Level k, but the domain has expanded, meaning that there is extra space in the domain, but not yet enough for the next bifurcation. In line with a uniformly expanding domain, to place the centers of the Gaussians, we use the principle that the extra space is evenly distributed between each Gaussian bump. For $t_k < t < t_{k+1}$, there are

$$G_k := \sum_{i=1}^{k} N_i = N_1 + \sum_{i=2}^{k} 2^{i-2}(N_1 + 1) = 2^{k-1}(N_1 + 1)$$

Gaussians, which have a total width of

$$2W_k := \sum_{i=1}^{k} 2N_i \alpha \sigma_i = 2\alpha \left(N_1 \sigma_i + \sum_{i=2}^{k} 2^{i-2}(N_1 + 1)\sigma_i \right),$$

recalling that $\alpha\sigma_i$ is the half-width of a Gaussian. We define $\hat{\delta}$ as the half-width of extra space between two bumps. For instance, at Level 2, the pattern would be situated as $\hat{\delta}.2.(2\hat{\delta}).1.(2\hat{\delta}).2.(2\hat{\delta}).1.(2\hat{\delta}).2.\hat{\delta}$. In this construction, it follows that

$$2\hat{\delta}G_k = 2L(t) - 2W_k,$$

which explicitly defines $\hat{\delta}$ via the formulas for G_k and W_k above. The Gaussian centers can then be defined by starting at the left end, $S = -L(t)$, and translating by the appropriate number of widths and half-widths for each level. To this end, it is useful to first define the width of the *repeating unit* for each level. The repeating unit is the sub-pattern that repeats N_1 times. For instance, in the case $k = 4$, the repeating unit is the pattern 4.3.4.2.4.3.4.1, or including the half-widths, $\hat{\delta}.4.(2\hat{\delta}).3.(2\hat{\delta}).4.(2\hat{\delta}).2.(2\hat{\delta}).4.(2\hat{\delta}).3.(2\hat{\delta}).4.(2\hat{\delta}).1.\hat{\delta}$. This repeating unit has total width

$$R_4 := 16\hat{\delta} + 2\alpha(\sigma_1 + \sigma_2 + 2\sigma_3 + 4\sigma_4).$$

Within this structure, the centers can be defined by moving some number of repeating units plus the appropriate distance within the repeating unit. Continuing with the case $k = 4$, the centers of the 1's are located at

$$S_{1_j} = -L(t) + (j-1)R_4 + 15\hat{\delta} + 2\alpha(\sigma_2 + 2\sigma_3 + 4\sigma_4) + \alpha\sigma_1, \quad j = 1, \ldots N_1,$$

while the centers of the 2's are located at

$$S_{2_j} = -L(t) + (j-1)R_4 + 7\hat{\delta} + 2\alpha(\sigma_3 + 2\sigma_4) + \alpha\sigma_2, \;\; j = 1, \dots N_1 + 1.$$

We must distinguish between the first and second appearing 3's, which we denote with subscripts a and b, and between the first, second, third and fourth 4's, for which we use subscripts a through d. For the 3's, the centers are at

$$\begin{aligned} S_{3_{a_j}} &= -L(t) + (j-1)R_4 + 3\hat{\delta} + 2\alpha(\sigma_4) + \alpha\sigma_3, \;\; j = 1, \dots N_1 + 1, \\ S_{3_{b_j}} &= -L(t) + (j-1)R_4 + 11\hat{\delta} + 2\alpha(3\sigma_4 + \sigma_3 + \sigma_2) + \alpha\sigma_3, \;\; j = 1, \dots N_1 + 1 \end{aligned}$$

while the centers of the 4's are located at

$$\begin{aligned} S_{4_{a_j}} &= -L(t) + (j-1)R_4 + \hat{\delta} + \alpha\sigma_4, \;\; j = 1, \dots N_1 + 1, \\ S_{4_{b_j}} &= -L(t) + (j-1)R_4 + 5\hat{\delta} + 2\alpha(\sigma_4 + \sigma_3) + \alpha\sigma_4, \;\; j = 1, \dots N_1 + 1 \\ S_{4_{c_j}} &= -L(t) + (j-1)R_4 + 9\hat{\delta} + 2\alpha(2\sigma_4 + \sigma_3 + \sigma_2) + \alpha\sigma_4, \;\; j = 1, \dots N_1 + 1, \\ S_{4_{d_j}} &= -L(t) + (j-1)R_4 + 13\hat{\delta} + 2\alpha(2\sigma_4 + 2\sigma_3 + \sigma_2) + \alpha\sigma_4, \;\; j = 1, \dots N_1 + 1 \end{aligned}$$

The above construction may be extended indefinitely as needed, and is also readily adapted to expression of a pattern in which new bumps do not appear on the edges, e.g. 1.1 to 1.2.1 to 1.3.2.3.2.1.

3 Validating approach

To validate our Gaussian bumps framework, we conduct two separate analyses of comparison to a more complex formulation.

Validation 1 We first compare amplitude curves for the initial buckling against the weakly nonlinear analysis of a growing elastic rod on a foundation provided in reference [1]. To do so, we must adapt the model slightly: in [1], the rod is extensible, thus some length differential can be accommodated by a compression of the longer layer, and buckling does not occur until this compression reaches a critical value. To extend our framework to an extensible regime for comparison to this work, we replace the length constraint by a stretching/compression energy. In the continuous rods approach, one defines the stretch

$$\alpha := \frac{\partial s}{\partial S}, \tag{37}$$

where S is the arc length parameter of the rod in the grown (stress-free) state, and s is the arc length parameter in the current state [32]. The stretching energy is then given by

$$\frac{E_s}{2} \int_{-L}^{L} (\alpha - 1)^2 \; dS \tag{38}$$

where E_s is the stretching stiffness. In the above, the stretch α is a local property. Such a quantity is not well-defined in our Gaussian approach; however we can still define an analogous stretching energy as a global property. To do this, we return to the length constraint, (6)

$$\int_{-L(t)}^{L(t)} 1 + \frac{1}{2}\epsilon \left(\partial_S y(S,t)\right)^2 \; dS = 2l(t).$$

The left hand side is the linearization of the arc length of Layer 1, the right hand side is the imposed length of Layer 1, and the equality reflects the length constraint. To replace the constraint by an energy, we define the stretching energy

$$\frac{E_s}{2}\left(\int_{-L(t)}^{L(t)} 1+\frac{1}{2}\epsilon\left(\partial_S y(S,t)\right)^2\ dS-2l(t)\right)^2.$$

In the limit $E_s \to \infty$, clearly the extensible rod limits to the inextensible case we have thus far considered.

In the framework of [1], time is not an explicit variable; rather the pattern evolves due to increase in the growth parameter γ, which is defined as

$$\gamma := \frac{\partial S}{\partial S_0}.$$

Here S_0 is the arc length parameter of the growing rod in its initial state, and S is the arc length parameter in the grown (unstressed) state. Since γ is constant in the calculation we consider, and the domain length is fixed, γ is equivalent to l/L in our notation.

In Fig 2A, we display the amplitude curves for a Level 1 pattern with $N_1 = 1$ (left) and $N_1 = 3$ (right), as a function of γ. The inset plots show the buckled shape at the indicated points. The solid curves are computed from the weakly-nonlinear analysis (WNA) approach, and the dashed curves are computed from our Gaussian approach. In this context, $\gamma = 1$ represents no growth, so that the rod (Layer 1) length exactly matches the foundation (Layer 2) length, and the rod is stress-free in the trivial flat state. As γ increases, the rod remains flat, but with compression energy building, until bifurcation occurs at a critical value γ^*, indicated by the point where the amplitude curve meets the axis. The WNA curve is computed from the full system of differential equations for a growing elastic rod, including geometric non-linearity. To summarize the calculation: first a linear stability analysis is conducted with eigenvalue γ to determine γ^*, then the unknown amplitude is computed in the neighbourhood of γ^* by perturbing as $\gamma = \gamma^* + \epsilon\gamma_1$ for $0 < \epsilon \ll 1$ and solving the full system up to $O(\epsilon^3)$ to obtain a solvability condition on the amplitude. It is worth stressing that the calculations in the WNA approach, outlined in full detail in [1], are very involved and lengthy, while calculation of the amplitude curve from the Gaussian approach involves a couple of lines of basic algebra.

Despite the conceptual difference between a local and global stretching energy, and a huge difference between the two approaches in the form and difficulty of the calculations, the Gaussian approach agrees with the full non-linear model remarkably well. Note that there are no fitting parameters in this comparison – all parameters in the WNA approach have an exact analogue that was passed to the Gaussian approach (full calculations available in the accompanying Mathematica notebook ComparisonWithWNA.nb; see data sharing description in main text for link). Though the complexity of the WNA approach does not include domain growth and thus does not include calculation of the bifurcation to Level 2, this comparison lends strong validity to the Gaussian reduction.

Validation 2 As a second form of validation, we relax Assumption A3 that each Gaussian is separated from the others and consider whether the amplitude curves and bifurcation to the next level are changed. That is, we maintain the assumption that the shape consists of a sum of Gaussians (as justified above), but removing the assumption that the Gaussians do not significantly overlap.

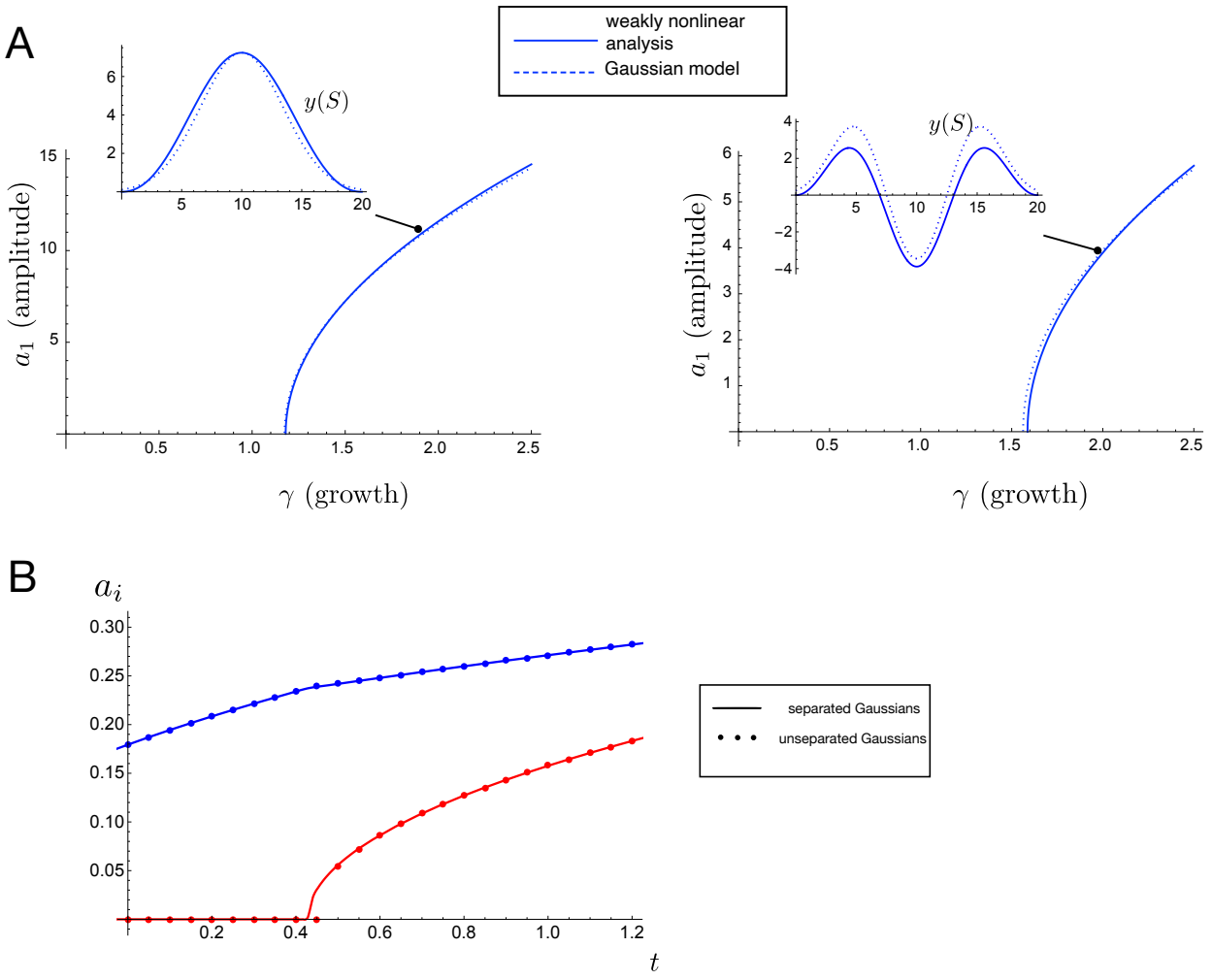

Figure 2: Validation of the Gaussian bumps approach. A. Comparison of initial amplitude curves for Gaussian model and weakly nonlinear analysis of growing elastic rod [1], both for $N_1 = 1$ (left) and $N_1 = 3$ (right). Inset plots show the buckled shape at the indicated points. B. Amplitude curves for Levels 1 and 2 as computed via separated Gaussians and Assumption A3 (solid curves) and numerically computed energy minimum without assuming separated Gaussians (dots). The Level 1 curve has amplitude a_1 in blue, and Level 2 has amplitude a_2 in red. Parameter values provided in Sec 9.

Recall that this assumption was used to replace integrals of sums of Gaussians in the energy by a sum of separated integrals with limits extended to $-\infty$ to ∞. In relaxing this assumption, we must integrate the length constraint and energy components numerically. Specifically, we consider the bifurcation from Level 1 to Level 2 such that $N_1 = 2$, $N_2 = 1$. Here, the pattern at Level 2 has the form

$$y_2(S) = a_1 \left(\exp\left(\frac{-(S - S_1)^2}{\sigma_1^2} \right) + \exp\left(\frac{-(S + S_1)^2}{\sigma_1^2} \right) \right) + a_2 \exp\left(\frac{-S^2}{\sigma_2^2} \right). \tag{39}$$

(This same form applies at Level 1, if $a_2 = 0$.) The length constraint reads

$$\int_{-L}^{L} y_2'(S)^2 \, dS = \delta(t). \tag{40}$$

Note that since the Gaussians are coupled in the integrals, the locations of the Gaussians cannot be ignored. We suppose here that the Gaussians are centered according to the rules outlined in Sec 2K, through which we define $S_1(t) = L(t) - \alpha\sigma$, and by symmetry the Level 2 Gaussian remains centered at $S = 0$. Our approach is as follows:

1. We fix the Gaussian widths σ_1 and σ_2 equal to the values obtained in the 'separated Gaussians' approach.
2. We solve the length constraint (40) at Level 2 for a_1, giving an explicit solution in terms of $\delta(t)$, σ_1 and σ_2, and the unknown $a_2(t)$.
3. We thus express the energy as an integral involving the unknown a_2 and otherwise known quantities.
4. We integrate the energy numerically over a range of a_2 values, and determine the value of a_2 for which the energy is minimized.

Within this framework, we expect that if we are far enough past the bifurcation time t_2 as predicted by the separated Gaussians, the two approaches should give the same energy minimizing values of a_1 and a_2, as there will be sufficient extra space in the domain such that the energy and length computed from both approaches will be nearly identical. The question then is whether the amplitudes agree for t close to t_2 and whether the predicted bifurcation point also agrees. For instance, if the unseparated Gaussians model predicts a bifurcation to Level 2 at an time that is appreciably earlier ($t < t_2$), when there is not yet sufficient space for a separated new Gaussian to appear, then Assumption A3 may be considered to be invalidated.

An investigation of this question in full detail or generality is beyond the scope of this work. Indeed, in some instances, such as the analysis of period doubling presented in [9], or hierarchical wrinkling analysed in [51], a bifurcation to a higher level of complexity does involve an overlapping of new modes with previous modes. However, in each of the examples appearing in main text Fig 1, a clear separation of new levels is observed. To demonstrate that this is mechanically consistent, we present an example in Fig 2B. Here we plot the amplitude curves a_1 and a_2 as computed for the separated and unseparated Gaussian models. The curves are in near perfect agreement, and the bifurcation point for the unseparated Gaussians is very close to t_2 (in fact, slightly later, though this will depend on the specific value of α that is used in determining the width of each Gaussian and thus the bifurcation condition). We leave it as a question for future work to determine exactly

which conditions lead a mechanical system to bifurcate in an overlapping versus non-overlapping manner.

4 Evolving foundation

In the framework outlined so far, the foundation/substrate energy at each Gaussian is proportional to a_i^m, which derives from the dimensional energy potential $\frac{K}{2}\left(\frac{y}{A}\right)^m$. This represents a *static* resistance to deformation of the growing beam, that is the energy is always at a minimum state when $y = 0$. A more realistic approach in many systems is to include a local remodeling of Layer 2 in response to deformation of Layer 1. To include such a remodelling in the model, we may replace the interaction potential energy (in dimensional form) with

$$\frac{K}{2}\left(\frac{y(S,t)-\hat{y}(S,t)}{A}\right)^m.$$

Here $\hat{y}(S,t)$ describes locally the substrate shape, and gives the shape at which the foundation energy is minimized. A natural choice is to evolve the substrate shape towards the shape of the growing beam, which in the simplest linear setting is described by the evolution law

$$\frac{\partial \hat{y}(S,t)}{\partial t} = \eta\left(y(S,t)-\hat{y}(S,t)\right) \tag{41}$$

where η characterizes the substrate relaxation or remodeling rate.

Within the Gaussian bumps framework, the function $\hat{y}$ will have the same shape as y, as expressed in (36). That is, $\hat{y}$ is a sum of separated Gaussians, and with the same widths σ_i as y, differing only in the amplitudes. Denoting the amplitudes of the substrate by $\hat{a}_i$, the substrate at Level k will have the form

$$\hat{y}(S,t) = \sum_{i=1}^{k}\sum_{j=1}^{N_i}\hat{a}_i(t)\exp\left(\frac{-(S-S_{i_j})^2}{\sigma_i^2}\right). \tag{42}$$

Then, the dimensionless foundation energy terms $\mu a_i^m \sigma_i$ are replaced by $\mu(a_i-\hat{a}_i)^m\sigma_i$, and the $\hat{a}_i(t)$ evolve according to

$$\hat{a}_i'(t) = \eta\left(a_i(t)-\hat{a}_i(t)\right). \tag{43}$$

This evolution equation is combined with an 'initial' condition that the substrate begins flat. Here we must be careful that the amplitudes a_i only appear at bifurcation time t_i, thus the correct condition is

$$\hat{a}_i(t_i) = 0, \tag{44}$$

which includes the implicit statement that $\hat{a}_i(t) \equiv 0$ for $t < t_i$.

The addition of the $\hat{a}_i$ to the energy unfortunately disrupts the analytical solvability of the system: even though each $\hat{a}_i$ is a fixed constant at each time step, we find that Mathematica is unable to solve analytically for the energy minimizing values $\{a_i\}$. Thus, when including an evolving foundation, we resort to a semi-analytical, semi-numerical approach, which we outline below.

4.1 Simulating energy minimization with evolving foundation

Since the foundation is assumed to begin in a flat state, the initial bifurcation condition and mode selection remain as before, solved with $\hat{a}_1 = 0$. Once the mode N_1 and width σ_1 have been selected, the amplitude curve $a_1(t)$ at Level 1 is determined explicitly by the length constraint, and is given by

$$a_1 = \sqrt{\frac{\delta(t)\sigma_1}{N_1}}. \tag{45}$$

We may insert this form directly into the evolution equation $\hat{a}_1'(t) = \eta\,(a_1(t) - \hat{a}_1(t))$, which may be solved exactly, though with a cumbersome formula.

At Level 2, the energy has the form

$$E_2(a_1, a_2; \hat{a}_1, \hat{a_2}, \sigma_1, \sigma_2) = \frac{N_1 a_1^2}{\sigma_1^3} + \frac{N_2 a_2^2}{\sigma_2^3} + \mu\left(N_1(a_1 - \hat{a}_1)^m \sigma_1 + N_2(a_2 - \hat{a}_2)^m \sigma_2\right), \tag{46}$$

while the length constraint reads

$$N_1 \frac{a_1^2}{\sigma_1} + N_2 \frac{a_2^2}{\sigma_2} = 2\delta(t). \tag{47}$$

To obtain the bifurcation time t_2 at which the system transitions to Level 2, we solve (47) for a_1, and insert into (46). We also insert the form for $\hat{a}_1(t)$, and set $t = t_2$ and $\hat{a_2} = 0$. The resulting energy E_2 is a function only of a_2, but involving the unknown t_2. The bifurcation time t_2 corresponds to the point where the energy minimizing branch of E_2 has $a_2 = 0$. That is, we seek t_2 such that

$$\lim_{a_2 \to 0} \left(\frac{1}{a_2} \frac{\partial E_2}{\partial a_2} \right) := g(t_2) = 0.$$

Here, the division by a_2 eliminates the trivial solution branch.

Once t_2 is determined, we compute the amplitude curves incrementally via the following steps:

1. Increment time by Δt, so that $t = t_2 + i\Delta t$ and numerically find the value of a_{2_i} by solving

$$\frac{\partial E_2}{\partial a_2} = 0,$$

   where a_1 is eliminated from E_2 via the length constraint and with $\hat{a}_1 = \hat{a}_{1_{i-1}}$, $\hat{a}_2 = \hat{a}_{2_{i-1}}$.

2. Compute a_{1_i} via the length constraint and with $a_2 = a_{2_i}$ as determined from Step 1.

3. Increment $\hat{a}_1$ and $\hat{a}_2$ by a forward Euler step:

$$\hat{a}_{1_i} = \hat{a}_{1_{i-1}} + \eta\Delta t(a_{1_i} - \hat{a}_{1_i}), \;\; \hat{a}_{2_i} = \hat{a}_{2_{i-1}} + \eta\Delta t(a_{2_i} - \hat{a}_{2_i})$$

4. Repeat.

The above procedure defines discrete values for each of a_1, a_2, $\hat{a}_1$, and $\hat{a}_2$, from which we generate interpolating functions defined for $t_2 < t < t_{\text{end}}$, where t_{end} is the end point of the simulation time.

At Level 3, we follow a similar procedure. First, to find the bifurcation time t_3, the length constraint is solved for a_1, which is inserted into the energy E_3. We also insert $\hat{a}_3 = 0$, $t = t_3$, and the interpolating functions for a_2, $\hat{a}_1$, and $\hat{a}_2$ computed above. We seek the time t_3 at which

$$\lim_{a_3 \to 0} \frac{1}{a_3} \frac{\partial E_3}{\partial a_3} := g(t_3) = 0,$$

where the division by a_3 eliminates the trivial solution branch and the root is found numerically. Having determined t_3, we compute the amplitude curves as follows

1. Increment time by Δt, so that $t = t_3 + i\Delta t$ and numerically find the values of a_{2_i}, a_{3_i} by numerically minimizing
$$E_3(a_2, a_3)$$
where a_1 is eliminated from E_3 via the length constraint and with $\hat{a}_1 = \hat{a}_{1_{i-1}}$, $\hat{a}_2 = \hat{a}_{2_{i-1}}$, $\hat{a}_3 = \hat{a}_{3_{i-1}}$.

2. Compute a_{1_i} via the length constraint and with $a_2 = a_{2_i}$, $a_3 = a_{3_i}$ as determined from Step 1.

3. Increment $\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$ by the same forward Euler step defined above.

4. Repeat.

As before, from the discrete points we generate interpolating functions, now defined on $t_3 < t < t_{\text{end}}$. This basic procedure is naturally extended to higher levels.

5 Fractal pattern

In adapting our framework to the generation of a fractal structure[2], the starting point, based on our aim of describing fractal-like spines in molluscs, is to suppose that the domain length remains fixed, i.e. the length of Layer 2, $2L$, does not increase with time, while Layer 1 continues to increase in length, i.e. $l = l(t)$[3]. Thus, whereas in the linear pattern considered up to now the appearance of new levels occurs in the free space created by the expanding domain, for a fractal pattern new levels can only appear emerging from the sides of the pattern that is already present. The two scenarios we will consider are: (i) the domain length $L = L(0) = 1$ at all times, or (ii) a pre-pattern first forms following the framework outlined above for $t < t_{\text{end}}$, at which point the domain stops expanding, i.e. $L = L(t_{\text{end}})$ for all $t > t_{\text{end}}$ while Layer 1 continues to increase in length, that is $l(t)$ is still monotonically increasing for all t. Scenario (i) is simpler, as the initial buckling can only produce a Level 1 pattern, after which continued growth can only generate increasing complexity in a fractal structure. In scenario (ii), a wider variety of patterns may emerge, as the pattern at t_{end} may already have a hierarchical structure with multiple levels, any of which may support the addition of fractal levels.

As we will show below, the energy minimizing framework we have outlined above works in the same basic way for a fractal pattern, only needing changes in how the widths σ_i are defined and in the bifurcation condition defining when the next level of pattern 'fits'.

5.1 Basic structure and length constraint

Consider the schematic shown in Fig 3A, which shows the emergence of a Level 2 fractal, with amplitude b_1 and width ξ_1, appearing on top of a Gaussian with amplitude a_1 and width σ_1.

[2]We use the word fractal loosely here to describe a hierarchical structure that has a fractal-like structure, though this will not strictly speaking create a fractal in the usual mathematical sense.

[3]This assumption is not strictly necessary for the development that follows, but is evidently valid in mollusc spines, and is useful in simplifying the calculations since it restricts how space for new levels can appear.

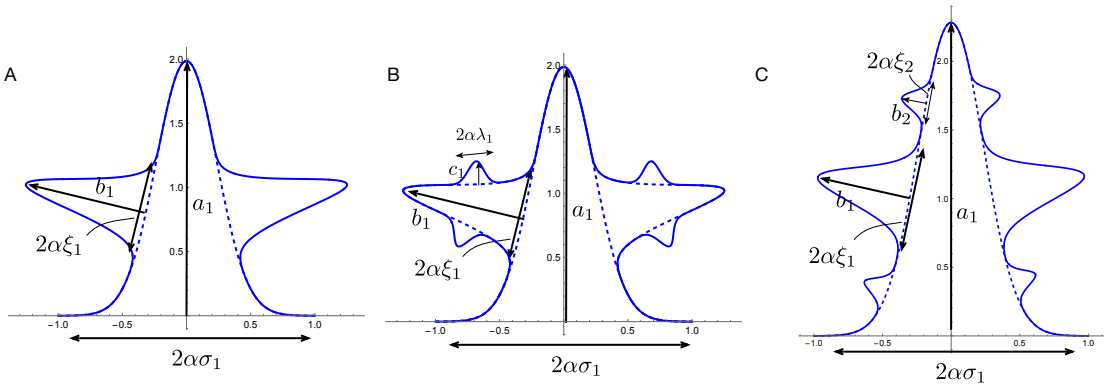

Figure 3: Fractal formation within Gaussian approach. In A, Level 2 spines with amplitude b_1 and width ξ_1 emerge from the side of a Level 1 spine with amplitude a_1 and width σ_1. In B, Level 3 spines with amplitude c_1 and width λ_1 emerge from the sides of the Level 2 spines. In C, an alternative bifurcation to higher level is shown, in which additional Level 2 spines, with amplitude b_2 and width ξ_2, emerge from the side of the Level 1 spine.

Clearly, the geometry is no longer described by

$$\mathbf{r}(S,t) = S\mathbf{e}_x + y(S,t)\mathbf{e}_y,$$

as in the linear pattern with separated Gaussians. Rather, the fractal shape has the more generic form

$$\mathbf{r}(S,t) = x(S,t)\mathbf{e}_x + y(S,t)\mathbf{e}_y,$$

where the exact form of x and y describe the geometry of the fractal shape and shall be derived explicitly below; however, at present we shall demonstrate that the Gaussian approach still enables the fractal shape to be computed via separated Gaussians. To this end, consider the length constraint for this schematic, which formally reads

$$\int_{-L}^{L} \epsilon\sqrt{\left(\partial_S x(S,t)\right)^2 + \left(\partial_S y(S,t)\right)^2}\, dS = 2l(t). \tag{48}$$

This integral is intractable to compute exactly analytically, due to the complex form that x and y will have (to be elucidated in Section 7). To make progress, we revisit and adapt assumptions A3 and A4 for fractal patterns:

FA3 Gaussian bumps for fractal levels can only appear on the edge of a present Gaussian.

FA4 Each fractal Gaussian bumps "fits" on the edge of a present Gaussian.

Under these assumptions, we may approximate the length of the fractal curve shown in Fig 3 by integrating the Gaussians separately, and accounting for the length that is lost from the flat edge of the a_1 Gaussian by the presence of the b_1 Gaussian. Defining

$$y_1 = a_1(t)\exp\left(\frac{-S^2}{\sigma_1^2}\right), \quad y_2 = b_1(t)\exp\left(\frac{-S^2}{\xi_1^2}\right),$$

we may approximate the fractal length as

$$\underbrace{\int_{-\alpha\sigma_1}^{\alpha\sigma_1} 1+\epsilon\frac{1}{2}\left(\partial_S y_1(S,t)\right)^2\ dS}_{\text{length of } a_1 \text{ Gaussian}} - \underbrace{2(2\alpha\xi_1)}_{\text{extra length}} + 2\underbrace{\int_{-\alpha\xi_1}^{\alpha\xi} 1+\epsilon\frac{1}{2}\left(\partial_S y_2(S,t)\right)^2\ dS}_{\text{length of } b_1 \text{ Gaussian}}. \tag{49}$$

Here we have subtracted 2 times $2\alpha\xi_1$ from the a_1 length, which is approximately the width of the b_1 Gaussian that is removed. Assumption FA4 implies that the b_1 Gaussian entirely fits within the edge of the a_1 Gaussian. Similarly, by the original A4 assumption, the a_1 Gaussian fully fits within the domain. Therefore, we may replace the limits of the integrals of $\partial_S y_i(S,t)^2$ with $-\infty$ to ∞ and apply the integral formula presented in (12). The length thus reduces to

$$2\alpha\sigma_1 + \frac{3\sqrt{\pi}\epsilon a_1^2}{2\sqrt{2}\sigma_1} + 2\left(\frac{3\sqrt{\pi}\epsilon b_1^2}{2\sqrt{2}\xi_1}\right). \tag{50}$$

For this simple mode 1 pattern, the domain length $2L = 2\alpha\sigma_1$ by construction. Therefore, setting the length equal to $2l(t)$, and scaling out the factor of $\frac{3\sqrt{\pi}\epsilon}{2\sqrt{2}}$ as before, the length constraint is reduced to

$$\frac{a_1^2}{\sigma_1} + 2\frac{b_1^2}{\xi_1} = 2\delta(t), \tag{51}$$

which has the same structure as in the linear pattern. Indeed, the argument above easily generalizes, as the number of Gaussians in a fractal level is 2 times the number of Gaussians on which they appear. Suppose for instance that the original linear pattern consists of N_1 Gaussians of amplitude a_1 and width σ_1 and N_2 Gaussians of amplitude a_2 and width σ_2, and that a fractal pattern appears with Gaussians of amplitude b_1 and width ξ_1 appearing on the edges of the a_1 Gaussians, and Gaussians of amplitude b_2 and width ξ_2 appearing on the edges of the a_2 Gaussians. The length constraint for this pattern would have the form

$$\frac{N_1 a_1^2}{\sigma_1} + \frac{2N_1 b_1^2}{\xi_1} + \frac{N_2 a_2^2}{\sigma_2} + \frac{2N_2 b_2^2}{\xi_2} = 2\delta(t). \tag{52}$$

5.2 Fractal energy

In terms of the energy, we maintain that under assumptions FA3 and FA4, the mechanical energy of a fractal pattern may be well-approximated by the exact same form as we have used in the linear pattern. Considering first the bending energy, and returning to the schematic Fig 3A, the integral of the square of the curvature may be approximated as

$$\int_{-\infty}^{\infty}\left(\frac{\partial^2 y_1}{\partial S^2}\right)^2\ dS + 2\int_{-\infty}^{\infty}\left(\frac{\partial^2 y_2}{\partial S^2}\right)^2\ dS \tag{53}$$

The second term accounts for the curvature of the b_1 Gaussian, and the integration limits follow from FA4. The first term similarly accounts for the curvature of the a_1 Gaussian, but also uses assumption FA3 in order to neglect the "missing portion" of the a_1 Gaussian, since the b_1 Gaussian appears on the flat edge of the a_1 Gaussian where the curvature is very minimal. Note that in defining the shape fully, the functions y_1 and y_2 do not share the same coordinate system; rather y_2 represents the distance from the edge of the first Gaussian in the normal direction. In terms of energy minimization, this distinction is irrelevant due to the modular nature of our approach.

However, it is of key importance in constructing the energy minimizing shape (details in Sec 77.4 below).

We may treat the interaction energy in a similar way. In particular, the interaction energy for the original a_1 Gaussian is the same as the linear case treated before, and we suppose that deformation of the b_1 Gaussian *from the* a_1 *edge* is resisted. Here, it may be natural to include a different foundation constant for the fractal edge and the original linear pattern. Defining $\hat{K}$ for the latter, the foundation energy in this schematic would take the form

$$\int_{-\infty}^{\infty} \frac{K}{2} y_1^m \, dS + 2 \int_{-\infty}^{\infty} \frac{\hat{K}}{2} y_2^m \, dS \tag{54}$$

Applying the same integral formulas as in Sec 22.4, we obtain an energy with the same structure as in the linear case. For instance, returning to the case of N_1 Gaussians of parameters (a_1, σ_1), on top of which a fractal Gaussian of parameters (b_1, ξ_1) appears, and N_2 Gaussians of parameters (a_2, σ_2), on top of which a fractal Gaussian of parameters (b_2, ξ_2) appears, the dimensionless energy would take the form

$$E_2(a_1, a_2, b_1, b_2; \sigma_1, \sigma_2, \xi_1, \xi_2) = \frac{N_1 a_1^2}{\sigma_1^3} + \frac{N_2 a_2^2}{\sigma_2^3} + \frac{2N_1 b_1^2}{\xi_1^3} + \frac{2N_2 b_2^2}{\xi_2^3} \tag{55}$$

$$+ \mu \left(N_1 a_1^m \sigma_1 + N_2 a_2^m \sigma_2 \right) + \hat{\mu} \left(2N_1 b_1 \xi_1 + 2N_2 b_2^m \xi_2 \right), \tag{56}$$

Extension to an evolving foundation may be accommodated by defining $\hat{a}_i$ and $\hat{b}_i$ in the same way as outlined in Sec 4.

5.3 Bifurcation condition

It remains only to define the bifurcation condition for a new level of fractal Gaussians to appear. Bifurcation occurs only when sufficient space appears on the edge of the already present pattern. Considering again the schematic in Fig 3, the a_1 Gaussian must have sufficient amplitude for the new level to appear on its edge. This may be modeled by the condition

$$2\alpha\xi_1 \leq \beta a_1(t) \tag{57}$$

where $\beta < 1$ is a constant that fixes how much height the Gaussian must attain before there is space for the fractal pattern to appear. Defining t_2 as the bifurcation time for the b_1 Gaussian, we thus define ξ_1 via

$$\xi_1 = \frac{\beta a_1(t_2)}{2\alpha}, \tag{58}$$

where $a_1(t_2)$ is defined by the amplitude curve prior to the appearance of the b_1 fractal. To determine t_2, we solve the length constraint (51) for a_1, and insert into the energy

$$E_2 = \frac{N_1 a_1^2}{\sigma_1^3} + \frac{2N_1 b_1^2}{\xi_1^3} + \mu N_1 a_1^m \sigma_1 + \hat{\mu} 2N_1 b_1 \xi_1, \tag{59}$$

while also replacing ξ_1 by (58). The bifurcation time t_2 corresponds to the point where the energy minimizing branch of E_2 has $b_1 = 0$. Thus, we seek t_2 such that

$$\lim_{b_1 \to 0} \left(\frac{1}{b_2} \frac{\partial E_2}{\partial b_1} \right) := g(t_2) = 0,$$

where the division by b_2 eliminates the trivial solution branch. In some cases, the solution may be found analytically, but if the linear pattern is more complex, or in the presence of an evolving foundation, the root of $g(t_2) = 0$ must be performed numerically.

5.4 Higher level fractal pattern

The ideas outlined above may naturally be extended to higher level fractal patterns. For instance, if a Gaussian with amplitude c_1 and width λ_1 appears on the edges of the b_1 Gaussian, this would constitute a Level 3 Gaussian, with energy

$$E_3 = \frac{N_1 a_1^2}{\sigma_1^3} + \frac{2N_1 b_1^2}{\xi_1^3} + \frac{4N_1 c_1^2}{\lambda_1^3} + \mu N_1 a_1^m \sigma_1 + \hat{\mu} 2N_1 b_1 \xi_1 + \hat{\hat{\mu}} 4N_1 c_1 \lambda_1, \tag{60}$$

Note that when extending from a linear pattern with separated Gaussians to a fractal regime, the number of potential forms increases, such that even classifying the level of the pattern requires care. For instance, return to the case of the a_1 and a_2 initial Gaussians with b_1 and b_2 fractal Gaussians. Here, the initial pattern represents a Level 2 Linear pattern, on top of which the b_1 and b_2 Gaussians appear, also Level 2. It is most natural to classify this as a Level 2 fractal, in that the additional Gaussians only appear on the edges of initial Gaussians, and consisting of Level 2 patterns at each Gaussian level. Note that the ordering b_1, b_2 implies that the bifurcation to the b_1 Gaussian occurred prior to the b_2 Gaussian. But note also that in principle the b_2 Gaussian could appear in three different places: (i) on the edge of the a_2 Gaussian, (ii) on the edge of the a_1 Gaussian, or (iii) on the edge of the b_1 Gaussian[4]. Case (iii) corresponds to a Level 3 Fractal, for which notationally it is more consistent to use c_1 and λ_1 than b_2 and ξ_2. In any case, this leaves the question: *what sets the location of the next bifurcation?* The short answer, following our framework, is that the next level will appear wherever sufficient space first exists. In case (i), for the b_2 Gaussian, with width $2\alpha\xi_2$, to appear on the edge of the a_2 Gaussian, would require sufficient amplitude in a_2, which can be approximated by the requirement:

$$2\alpha\xi_2 \leq \beta a_2(t). \tag{61}$$

In case (ii), for the b_2 Gaussian to appear on the edge of the a_1 Gaussian, which already has the b_1 Gaussian present would require sufficient amplitude in a_1, expressed as

$$2\alpha\xi_2 \leq \beta(a_1(t) - 2\alpha\xi_1). \tag{62}$$

And in case (iii), for the b_2 Gaussian (called in this case c_1 with width λ_1) to appear on the edge of the b_1 Gaussian would require

$$2\alpha\lambda_1 \leq \beta b_1(t). \tag{63}$$

For each of these requirements, setting t equal to the bifurcation time t_4 and replacing inequality by equality and finding the energy minimizer leads to a calculation for t_4. The actual location will correspond to that which gives the least value of t_4; thus in practice t_4 would need to be computed and compared for each possibility. While these calculations are straightforward, extending a fractal pattern beyond the initial levels quickly gets very involved.

[4]If the domain L is allowed to expand, an additional option exists in which the next level appears in the linear space created by domain expansion.

6 Pattern amplification

In this section we provide further details on main text Fig 5A, in which a pre-pattern of ridges is dilated in both a linear and fractal-like manner. In this analysis, we consider both a low and high density ridges pattern, by creating a Level 3 pattern with the form 3.2.3.1.3.2.3 on a domain of width $2L$, where $L = 3$ for the low density and $L = 1$ for high density ridges. The domain width $2L$ is held fixed, while the Layer 1 length $l(t)$ increases. The linear spines have Gaussian width equal to the width of the ridges in the pre-pattern; the spine amplitude is then computed as the energy minimizers, with evolving foundation, following the approach of Section 4. For the fractal-like spines, we suppose that the ridges pattern is dilated to a Level 3 fractal, taking the form shown in Fig 3B, i.e. with the Level 2 spine bifurcating from the edge of the Level 1 spine, and the Level 3 spine bifurcating from the edge of Level 2. For this construction, the width of the Level 1 spine is not the width of the Level 1 ridge, but rather defined by the domain width, that is $\alpha\sigma_1 = L$. From there, σ_2 and σ_3 are defined by the requirement of fitting on the side of the lower level spine, and the amplitude curves are found by minimizing the energy as described above. Details on plotting the resulting shapes follow the same procedure as described in Sec 7D below.

7 Shell simulations

In this section we outline the procedure for constructing 3D shell simulations appearing in main text Fig 5B and C (while also providing the procedure for constructing a solution profile in the fractal-like case). These simulations consist of four steps: (i) construction of the geometry of the shell as a smooth surface, (ii) construction of the ridges pattern, (iii) amplification of the ridges pattern to either linear or fractal-like manner, (iv) creating a continuous surface with ridges dilated to spines, and (v) wrapping the ridges and spine surface onto the geometry of the shell. We discuss each aspect in turn below.

7.1 Base geometry

The base geometry of the shell is constructed following an extrinsic approach, in which an aperture shape is wrapped around a central helicospiral curve. We begin with the following parameterization for a helicospiral

$$\mathbf{x}(\tau) := \begin{pmatrix} x_1(\tau) \\ x_2(\tau) \\ x_3(\tau) \end{pmatrix} = \begin{pmatrix} r_0 \exp(g_w \tau) \cos\tau \\ r_0 \exp(g_w \tau) \sin\tau \\ h_0 \exp(g_w \tau) \end{pmatrix}. \tag{64}$$

From this, the tangent-normal-binormal frame, denoted $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$ may be computed following standard formulas.

We next define a planar curve with parameterization $(\rho_1(s), \rho_2(s))$, which defines the shape of the aperture. To construct the smooth shell surface, in the simplest form, at each point along the central curve, i.e. for each value of τ, the planar curve is dilated by factor $\Lambda(\tau) = \exp(c\tau)$ and wrapped around the central curve in the **N**-**B** plane. Here, we incorporate extra degrees of freedom by allowing for rotations about the $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$ axes. In particular, we define the rotated frame

$\{\boldsymbol{\beta}, \boldsymbol{\Gamma}, \boldsymbol{\Theta}\}$:

$$\boldsymbol{\beta} = \cos\theta_1 \mathbf{B} + \sin\theta_1 \mathbf{T}, \tag{65}$$

$$\boldsymbol{\Gamma} = \cos\theta_2 \mathbf{N} + \sin\theta_2(\cos\theta_1 \mathbf{T} - \sin\theta_1 \mathbf{B}, \tag{66}$$

$$\boldsymbol{\Theta} = \cos\theta_2 \cos\theta_1 \mathbf{T} - \cos\theta_2 \sin\theta_1 \mathbf{B} - \sin\theta_2 \mathbf{N}. \tag{67}$$

The smooth shell surface is then parameterized as

$$\mathbf{R}(s,\tau) = \mathbf{x}(\tau) + \Lambda(\tau)\left(\rho_1(s)(\cos\theta_3 \boldsymbol{\Gamma}(\tau) + \sin\theta_3 \boldsymbol{\beta}(\tau)) + \rho_2(s)(\cos\theta_3 \boldsymbol{\beta}(\tau) - \sin\theta_3 \boldsymbol{\Gamma}(\tau))\right) \tag{68}$$

Specific functions and parameter choices may be found in the Mathematica notebook ShellSim.nb accompanying these simulations (see data sharing description in main text for link).

7.2 Ridges pattern

The ridges pattern follows the model formulation outlined in Sec 2. For the shell with linear spines, the parameters used (see Sec 9) generate a pattern with $N_1 = 9$ ridges of Level 1, and the ridges attain Level 3. For the fractal-like shell, to generate a denser ridges pattern, we use a smaller initial domain length, and an increased value of μ; this generates a pattern with $N_1 = 5$, and attains Level 5 hierarchy. Here, for ease of calculation, we do not incorporate an evolving foundation, as this enables to reach Level 5 hierarchy analytically without requiring numerical energy minimization. However, the parameters used are such that the hierarchical structure is maintained even without the memory component of the model.

7.3 Amplification of the ridges pattern

The amplification of the ridges pattern to linear or fractal-like spines follows the approach outlined in Section 6, with the domain length fixed and equal to the length $L(t)$ at the end of the ridges stage, and with a significant increase in the rate of increase of $l(t)$ (specific parameters given in Sec 9). For the linear spines, the Gaussian widths are given by those determined at the ridges stage, and the amplitudes are determined as energy minimizers following Section 4. For the fractal-like spines, we implement a pattern similar to that shown in Fig 3C. The *Chicoreus ramosus* shell shown in main text Fig 1 displays a spine pattern in which the Level 1 and Level 2 ridges each form the center of spines, with the Level 3 ridges dilating on the side of the Level 1 spine, with the Level 4 ridges dilating only slightly on the side of both Level 1 and Level 2 spines, and with Level 5 ridges appearing intercalated throughout but with almost no dilation. We simulate the emergence of this pattern in the following simplified way:

1. Since both the Level 1 and Level 2 ridges dilate to Level 1 spines, here for simplicity we combine Level 1 and Level 2 within the same initial energy and length constraint, such that the amplitude and width of the Level 2 spines are smaller but with fixed proportion to Level 1 [5]. Letting a_1 and σ_1 be the amplitude and width of the Level 1 spine, respectively, the

[5] This simplification is made for two reasons. One is so that the energy minimization can be taken up to Level 4 with memory included but without numerical issue. The second reason is that since we are fixing the domain length, and both Level 1 and Level 2 spines occupy the same fixed domain in this idealization, we do not have a clear means of signaling when the bifurcation to Level 2 should occur in the fractal pattern. As our objective is to demonstrate consistency in principle, a full resolution of these complex morphologies is left for a future study.

Level 2 spine is assigned amplitude κa_1 and width $\xi\sigma_1$, where we have used $\kappa = \xi = 0.5$ in our simulation. Thus, the length constraint is given by

$$N_1 \frac{a_1^2}{\mu_1} + (N_1 - 1)\frac{(\kappa a_1)^2}{\xi\sigma_1} = 2\delta(t),$$

while the energy is given by

$$E_1 = N_1 \frac{a_1^2}{\sigma_1^3} + (N_1 - 1)\frac{(\kappa a_1)^2}{(\xi\sigma_1)^3} + \hat{\mu}\left(N_1(a_1 - \hat{a}_1)^4\sigma_1 + (N_1 - 1)\kappa(a_1 - \hat{a}_1)^4\xi\sigma_1\right)$$

To determine the amplitude and width, we solve the length constraint for a_1, in terms of σ_1, insert the solution for a_1 into the energy E_1 while also setting $\hat{a}_1 = 0$, and then compute analytically the value of σ_1 for which $\partial E_1/\partial\sigma_1 = 0$. Remodeling is then included by updating $\hat{a}_1$ via

$$\hat{a}_1'(t) = \hat{\eta}(a_1 - \hat{a}_1).$$

2. The widths and amplitudes for Levels 3 and 4 are computed following the procedure outlined in Sec 5. In particular, since the Level 3 pattern dilates on the edge of the Level 1 spine, the width σ_3 is defined by the amplitude a_1 via

$$2\alpha\sigma_3 = \beta a_1,$$

where we have used $\beta = 0.5$ in our simulation. The Level 4 ridge dilates on the edge of both the Level 1 and Level 2 spines. For simplicity, we define the bifurcation to Level 4 spine by a single condition

$$2\alpha\sigma_4 = \beta(a_1 - 2\alpha\sigma_3),$$

which is the requirement to fit on the edge of the Level 1 spine, where the Level 3 spine already appears. With these definitions for width in place, the computation of amplitudes as energy minimizers proceeds as we have outlined, with memory included via updating $\hat{b}_3$ and $\hat{b}_4$. Bifurcation to Level 5 spines is not included, reflecting the very small degree of dilation of Level 5 ridges on real shell specimens, though in principle such bifurcation could be included in the framework.

7.4 Creating ridges and spines surface

Once the amplitudes and widths of both ridges and spines have been determined, it remains to connect the patterns into a single continuous surface. In the case of the linear spines, this is straightforward. The ridges surface is generated by defining the locations of the Gaussian centers following the procedure outlined in Section 2.12. The spines then simply require an increase in amplitude of the ridges, with both the domain length and the location of the centers remaining constant.

In the fractal-like case, however, it is far more complicated to generate a surface. In particular, we require a consistent means of defining the locations of Gaussians that emerge from the side of existing Gaussians, which requires careful consideration of arc lengths as the fractal-like structure develops. Our general approach consists in the following steps:

1. Create the Level 1 spines pattern.

2. Determine the total arc length and arc length parameter of the Level 1 spines pattern, denoted $\hat{l}(t)$ and $\hat{S}(t)$, respectively, and determine the normal vector for the Level 1 spines pattern, denoted $\hat{\boldsymbol{\nu}}$.

3. Superimpose the Level 2 spines on the side of the Level 1 spines, with Gaussians appearing in the direction $\hat{\boldsymbol{\nu}}$, and with locations defined in terms of $\hat{S}$.

4. Determine the total arc length and arc length parameter of the combined Level 1 and Level 2 spines, denoted $\hat{\hat{l}}(t)$ and $\hat{\hat{S}}(t)$, respectively, and determine the normal vector for the combined Level 1 and Level 2 pattern, denoted $\hat{\hat{\boldsymbol{\nu}}}$.

5. Superimpose the ridges pattern on the combined Level 1 and Level 2 spines pattern, with Gaussians appearing in the direction $\hat{\hat{\boldsymbol{\nu}}}$, and with locations defined by $\hat{\hat{S}}$.

More info on Step 2 The pattern at Step 1 has the form of separated Gaussians appearing in the y-direction and spaced along the x-axis, i.e. with parameterization $\mathbf{r}(S,t) = (S, y(S,t))$ where $y(S,t)$ has the general form

$$y(S,t) = \sum_{i,j} a_i(t) \exp\left(-\frac{(S - S_{ij})^2}{\sigma_i^2}\right) := \sum_{i,j} y_{ij}. \tag{69}$$

Suppose that the domain of S is $[-L, L]$, with L defined by the domain length at the end of the ridges formation. The arc length $\hat{S}$ is related to S via the stretch

$$\hat{\lambda} := \frac{\partial \hat{S}}{\partial S} = \sqrt{1 + \left(\frac{\partial y}{\partial S}\right)^2}.$$

We may use the separation of the Gaussians to approximate

$$\hat{\lambda} \approx \sqrt{1 + \sum_{i,j} \left(\frac{\partial y_{i,j}}{\partial S}\right)^2}.$$

Thus, to determine $\hat{S}$, we integrate numerically the system

$$\frac{\partial \hat{S}}{\partial S} = \hat{\lambda}(S,t), \quad \hat{S}(-L,t) = 0.$$

From this, the total arc length is given by $\hat{l}(t) = \hat{S}(L,t)$.

The unit normal vector to the curve $(S, y(S,t))$ is straightforwardly defined as

$$\hat{\boldsymbol{\nu}} = \frac{\left(-\sum_{i,j} \frac{\partial y_{i,j}}{\partial S}, 1\right)}{\sqrt{1 + \sum_{i,j} \left(\frac{\partial y_{i,j}}{\partial S}\right)^2}}.$$

More info on Step 4 The pattern at Step 3 has the form

$$\hat{\mathbf{r}}(S,t) = S\mathbf{e}_x + \sum_{i,j} y_{ij}\mathbf{e}_y + \hat{\boldsymbol{\nu}}(S,t)\sum_{k,l}\hat{y}_{k,l}(\hat{S}(S,t)). \tag{70}$$

Here the first two terms express the Level 1 pattern in terms of Cartesian unit vectors $\mathbf{e}_x$ and $\mathbf{e}_y$. The Level 2 spines are defined in the third term by the functions $\hat{y}_{k,l}$, which form a set of separated Gaussians defined on the $\hat{S}$ domain, mapped onto the S domain via the arc length conversion and extending in the normal direction $\hat{\boldsymbol{\nu}}$. To impose the ridges pattern (or equivalently to impose a Level 3 spine) on top, we require the normal vector to the curve $\hat{\mathbf{r}}$. To avoid the exceedingly tedious direct calculation, we first note that the arc length $\hat{\hat{S}}$ relates to S via

$$\hat{\hat{\lambda}} := \frac{\partial\hat{\hat{S}}}{\partial S},$$

where the stretch $\hat{\hat{\lambda}}$ satisfies

$$\frac{\partial\hat{\mathbf{r}}}{\partial S} = \hat{\hat{\lambda}}\hat{\hat{\boldsymbol{\tau}}}. \tag{71}$$

Here $\hat{\hat{\boldsymbol{\tau}}}$ is the unit tangent to $\hat{\mathbf{r}}$. To compute $\hat{\hat{\boldsymbol{\tau}}}$ and subsequently $\hat{\hat{\boldsymbol{\nu}}}$, we first note that the tangent $\hat{\boldsymbol{\tau}}$ and normal $\hat{\boldsymbol{\nu}}$ to the curve at Level 1 satisfy the Frenet equations

$$\frac{\partial\hat{\boldsymbol{\tau}}}{\partial S} = \hat{\lambda}\hat{\kappa}\hat{\boldsymbol{\nu}} \tag{72}$$

$$\frac{\partial\hat{\boldsymbol{\nu}}}{\partial S} = -\hat{\lambda}\hat{\kappa}\hat{\boldsymbol{\tau}} \tag{73}$$

where $\hat{\lambda}$ is the geometric stretch as defined above and $\hat{\kappa}$ is the curvature of the Level 1 pattern, which is given by

$$\hat{\kappa} = \frac{1}{\hat{\lambda}^3}\sum_{i,j}\frac{\partial^2 y_{i,j}}{\partial S^2}.$$

Therefore, we may write

$$\frac{\partial\hat{\mathbf{r}}}{\partial S} = \hat{\lambda}\hat{\boldsymbol{\tau}} + \sum_{k,l}\frac{\partial\hat{y}_{k,l}}{\partial S}\hat{\boldsymbol{\nu}} - \hat{\lambda}\hat{\kappa}\sum_{k,l}\hat{y}_{k,l}\hat{\boldsymbol{\tau}}. \tag{74}$$

Combining (71) and (74), we may express the tangent at Level 2 in terms of the tangent and normal at Level 1:

$$\hat{\hat{\boldsymbol{\tau}}} = \frac{a\hat{\boldsymbol{\tau}} + b\hat{\boldsymbol{\nu}}}{\hat{\hat{\lambda}}} \tag{75}$$

where the stretch $\hat{\hat{\lambda}} = \sqrt{a^2+b^2}$ is given in terms of the easily computed quantities

$$a := \hat{\lambda}\left(1 - \hat{\kappa}\sum_{k,l}\hat{y}_{k,l}\right), \quad b := \sum_{k,l}\frac{\partial\hat{y}_{k,l}}{\partial S}.$$

From here, the normal vector $\hat{\hat{\boldsymbol{\nu}}}$ is given by

$$\hat{\hat{\boldsymbol{\nu}}} = \frac{-b\hat{\boldsymbol{\tau}} + a\hat{\boldsymbol{\nu}}}{\hat{\hat{\lambda}}} \tag{76}$$

where the choice of sign corresponds to an outward pointing normal.

The arc length parameter at Level 2 can be determined by integrating (numerically) the system

$$\frac{\partial \hat{\tilde{S}}}{\partial S} = \hat{\tilde{\lambda}}(S,t), \quad \hat{\tilde{S}}(-L,t) = 0.$$

More info on Step 5 The full pattern, including Level 3 fractal, can now be expressed in the form

$$\hat{\tilde{\mathbf{r}}}(S,t) = S\mathbf{e}_x + \sum_{i,j} y_{ij}\mathbf{e}_y + \hat{\boldsymbol{\nu}}(S,t)\sum_{k,l}\hat{y}_{k,l}(\hat{\tilde{S}}(S,t)) + \hat{\tilde{\boldsymbol{\nu}}}(S,t)\sum_{m,n}\hat{\tilde{y}}_{k,l}(\hat{\tilde{S}}(S,t)). \tag{77}$$

Here, the Level 3 pattern, appearing in the final term, is in principle expressed as a sum of separated Gaussians $\hat{\tilde{y}}_{k,l}$ defined in terms of the Level 2 arc length parameter $\hat{\tilde{S}}$, with spacing determined by the total arc length $\hat{\tilde{l}}(t) = \hat{\tilde{S}}(L,t)$. However, we found that the complexity of the structure at Level 2 led to numerical inaccuracies in integrating for $\hat{\tilde{S}}$, which generated a surface that was very computationally expensive to render and also often with an unsmooth appearance. Therefore, in the shell simulation presented in the main text, we instead expressed the locations of the Level 3 Gaussians in terms of the Level 1 arc length $\hat{S}$, which led to a significantly faster computation, conceptually changes very little, and visually produced a better surface. For specific details, the reader is referred to the Mathematica notebook accompanying these simulations (see data sharing description in main text for link).

7.5 Wrapping pattern onto shell

The construction above defines the ridges/spines as an evolving curve in the x-y plane; equating the time variable t with the z-axis thus defines a surface. It remains to wrap this surface onto the coordinates of the shell.

To this end, we first construct an orthonormal frame $\{\mathbf{d_1}, \mathbf{d_2}, \mathbf{d}_3\}$ attached to the smooth shell surface and oriented such that $\mathbf{d}_3$ is tangent to the shell aperture, $\mathbf{d}_1$ is normal to the shell aperture, and $\mathbf{d}_2$ points in the direction of shell growth. Considering the parameterization $\mathbf{R}(s,\tau)$ in (68), the aperture lives in the $\boldsymbol{\beta}$-$\boldsymbol{\Gamma}$ plane, and thus $\mathbf{d}_2 = \boldsymbol{\Theta}$. Then, $\mathbf{d}_3$ is computed by normalizing $\partial\mathbf{R}/\partial s$, and $\mathbf{d}_1$ completes the orthonormal frame. This gives

$$\mathbf{d}_1 = \frac{\rho_2'(s)(\cos\theta_3\boldsymbol{\Gamma} + \sin\theta_3\boldsymbol{\beta}) - \rho_1'(s)(\cos\theta_3\boldsymbol{\beta} - \sin\theta_3\boldsymbol{\Gamma})}{\sqrt{\rho_1'(s)^2 + \rho_2'(s)^2}} \tag{78}$$

$$\mathbf{d}_2 = \boldsymbol{\Theta} \tag{79}$$

$$\mathbf{d}_3 = \frac{\rho_1'(s)(\cos\theta_3\boldsymbol{\Gamma} + \sin\theta_3\boldsymbol{\beta}) + \rho_2'(s)(\cos\theta_3\boldsymbol{\beta} - \sin\theta_3\boldsymbol{\Gamma})}{\sqrt{\rho_1'(s)^2 + \rho_2'(s)^2}}. \tag{80}$$

The parameter s increases moving along the shell aperture. This parameter should form a one-to-one map with the parameter S in the ridges/spine construction. To do this, we define $\hat{s}$ as the arc length parameter moving along the aperture in the smooth shell. This parameter satisfies

$$\hat{s}'(s) = \sqrt{\rho_1'(s)^2 + \rho_2'(s)^2}, \quad \hat{s}(s_0) = 0, \tag{81}$$

where s_0 is the lower bound for the range of s. If s_1 is the upper bound for the range of s, then the total arc length of the aperture is $\hat{s}_1 = \hat{s}(s_1)$. The domain of the ridges/spine pattern is $S \in [-L, L]$,

from which we suppose that the $\hat{s}$ domain is linearly mapped to the S domain via

$$S(\hat{s}) = -L + \frac{2L}{\hat{s}_1}\hat{s}.$$

We also require a one-to-one map between the 'time' variable τ in the smooth shell parameterization and the 'time' variable t in the ridges/spine construction. This is complicated by the fact that a shell will generate spines at discrete and usually periodic points during development. Aiming for simplicity, here we consider the formation of a single row of spines, which then may be recreated at different periodic points in the shell parameterization. Supposing that the ridges/spines are generated as t varies from t_0 to t_1, and that the pattern occurs over a range of $\Delta\tau$ in the τ variable, starting from $\tau = \tau_0$, we define the linear map

$$t(\tau) = t_0 + \frac{t_1 - t_0}{\Delta\tau}(\tau - \tau_0).$$

We may now construct a single row of spines by mapping the parametric curve $\hat{\hat{\mathbf{r}}}(S,t)$ into the variables (s,τ) of the smooth shell by equating the $\mathbf{d}_3$ direction with $\mathbf{e}_x$ and the $\mathbf{d}_1$ direction with $\mathbf{e}_y$. Specifically, the surface $\hat{\hat{\mathbf{r}}}(S,t)$, describing a row of spines dilated from a ridges pattern, may be imposed onto the shell surface $\mathbf{R}(s,\tau)$ given by Equation (68) as follows:

$$\mathbf{R}_{\text{full}}(s,\tau) = \mathbf{R}(s,\tau) + \Lambda(\tau)\left[\left(\hat{\hat{\mathbf{r}}}\cdot\mathbf{e}_x - S(\hat{s}(s))\right)\mathbf{d}_3(s,\tau) - \hat{\hat{\mathbf{r}}}\cdot\mathbf{e}_y\mathbf{d}_1(s,\tau)\right], \quad s_0 \le s \le s_1,\ \tau_0 \le \tau \le \tau_0 + \Delta\tau. \tag{82}$$

In this expression, note that $\hat{\hat{\mathbf{r}}}$ means $\hat{\hat{\mathbf{r}}}(S(\hat{s}(s), t(\tau))$. The term

$$\hat{\hat{\mathbf{r}}}\cdot\mathbf{e}_x - S$$

is needed in the tangent $\mathbf{d}_3$ direction to wrap the fractal shape with non-monotonic change in the x-coordinate $x(S,t) := \hat{\hat{\mathbf{r}}}\cdot\mathbf{e}_x$ onto the aperture; effectively, the component $S\mathbf{d}_3$ is already included in $\mathbf{R}(s,\tau)$, so it is only deviations of x from S that need to be mapped into the tangential direction. For the linear spines pattern, on the other hand, $x \equiv S$ and thus the ridges/spines can be added to the smooth shell by moving only in the $\mathbf{d}_1$ direction by $y(S,t) := \hat{\hat{\mathbf{r}}}\cdot\mathbf{e}_y$ (in fact, it is the minus $\mathbf{d}_1$ direction, since $\mathbf{d}_1$ points into the shell).

7.5.1 Extra details

Two other small details were included in the shell simulations in the main text, in order to achieve a closer comparison with real shells. Namely, in the fractal-like shell simulation, (i) the amplitude of the Level 1 spines was artificially made to decrease with increasing aperture arc length (i.e. as a function of s), and (ii) the fractal-like spines were artificially made to curl back in the direction of shell growth (the $\mathbf{d}_2$ direction in the description above). The change in amplitude likely reflects a non-uniform shell secretion along the arc length, while the curling back has never been previously quantified or explained, but may also be a consequence of secretion rate: since points at the tip of the fractal have a larger radius from the shell aperture, maintaining a planar profile would require an increased secretion at the tip; thus a constant secretion rate would imply a curling back. These features do not impact the shell in a significant way, but do improve the visual comparison. While they could in principle be modeled explicitly, we have opted to add them artificially in this work.

8 Different geometries

In this section we outline model adaptations for simulating the other geometries appearing in main text Fig 6.

8.1 Circular geometry

Main text Fig 6A and D simulate the formation of a hierarchical pattern on a circular geometry. This requires only a modification in the presentation of the pattern. That is, energy minimization proceeds in the same way as in the linear case of Sec 2, where we identify the domain length $L(t)$ with the radius of the expanding circle. Denoting the latter by $r(t)$, we have

$$r(t) = \Delta\theta L(t), \tag{83}$$

where $\Delta\theta$ is the angular sector of the circular geometry (e.g. $\Delta\theta = 2\pi$ for the full circle modeled in simulating the lamellae under the mushroom cap in main text Fig 6A). Similarly, the transverse parameter S in the original linear formulation is translated to angle θ in the circular formulation via $S = r(t)(\theta - \theta_0)$, where θ_0 is the angle corresponding to the edge of the domain.

Under these simple translations, the full pattern is then created as a surface in cylindrical polar coordinates. For instance, a linear pattern of Gaussians

$$y(S,t) = \sum_{i=1}^{N_1} a_1(t) \exp\left(-\frac{(S - S_i(t))^2}{\sigma_1^2}\right) \tag{84}$$

would be expressed in circular geometry as the surface

$$\mathbf{R}(S,t) = (r(t)\cos\theta(S,t), r(t)\sin\theta(S,t), y(S,t)) \tag{85}$$

where $\theta(S,t) = \theta_0 + \frac{S}{r(t)}$.

8.2 Cylindrical geometry

In simulating the columnals of the stalk of a fossil crinoid in main text Fig 6C, we take a simplified approach, applying the same energy minimization as in Sec 2, where the length $L(t)$ is identified with the length of the stalk, or more generally the axial domain in cylindrical geometry. This amounts to treating the pattern as an evolving 1D pattern that is then wrapped onto a cylinder as a surface of revolution.

A more detailed calculation could in principle be undertaken in which it is not an excess of length between two layers (idealized as lines), but as an area excess between two surfaces: Layer 2 would be a cylindrical core with radius R and length $L = L(t)$, surrounded by Layer 1, treated as a surface of revolution made by rotating the curve $R + y(S,t)$, where y is a sum of Gaussians, around the central axis. In this view the bending energy would need to be replaced by an integral over the surface of revolution of the square of the mean curvature, though with suitable simplifications based on assumption of small amplitude. While the main structure would work similarly, e.g. bending energy of a given Gaussian would be proportional to a^2, the σ dependence would be impacted by the change in geometry. Such an analysis is beyond the scope of this work, and we thus leave such calculations for future investigations focused on this or other similar geometries.

9 Parameters

In this section, we provide all parameters used in constructing all output appearing in this paper. We note that while mathematically it is natural to scale the initial domain length to $L_0 = 1$, in practice we found that in exploring the parameter space, it was preferable to leave the initial length as a free variable. For instance, increasing the buckling mode requires a significant increase in μ, but can be accomplished with a more moderate value of μ if L_0 is increased (noting from Sec 2 that μ scales as L_0^2). With L_0 different from 1, $L = L_0(1 + t)$ and the model works in the exact same way otherwise. The values reported below and in the accompanying Mathematica notebooks have L_0 different from 1.

9.1 Main text Fig 3

Main text Fig 3 A shows amplitude curves for interaction energy with both $m = 2$ and $m = 4$. For both cases, we have used parameter values $\mu = 100$, $\delta_0 = 0.3$, $g = 5$, $L_0 = 4.5$, and $\alpha = 2.5$, solving the system up to $t = 0.75$. It is also worth noting that since the point of the figure is to compare the different values of m, under otherwise equivalent parameters, we have had to change the energy form slightly: namely, recall from (17) that we absorbed a coefficient coming from the integral of the interaction energy into the constant μ. However, a different coefficient is absorbed in the cases of $m = 2$ and $m = 4$, which would effectively change the value of μ in moving between the two cases. To keep a consistent comparison, we have thus kept an additional constant in front of the interaction energy that is different for $m = 2$ than $m = 4$. In particular, noting that

$$\int_{-\infty}^{\infty} \left(a \exp\left(-\frac{S^2}{\sigma^2} \right) \right)^2 \, dS = a^2 \sigma \sqrt{\frac{\pi}{2}}, \tag{86}$$

while

$$\int_{-\infty}^{\infty} \left(a \exp\left(-\frac{S^2}{\sigma^2} \right) \right)^4 \, dS = a^2 \sigma \frac{\sqrt{\pi}}{2}, \tag{87}$$

for $m = 2$ the interaction energy had the form

$$\mu a^2 \sigma \sqrt{\frac{\pi}{2}}$$

, while for $m = 4$ we used

$$\mu a^2 \sigma \frac{\sqrt{\pi}}{2}.$$

Main text Fig 3B showed amplitude curves for an evolving foundation with different values of η, as displayed on the fig. Other parameter values for this figure were: $\mu = 100$, $\delta_0 = 0.108$, $g = 5$, $L_0 = 4.5$, and $\alpha = 2.5$, with amplitudes solved up to $t = 0.8$.

9.2 Main text Fig 4

Main text Fig 4A shows 4 different patterns, produced with different values of μ and g, as indicated on the figure. Other parameter values for these plots were $\delta_0 = 0.07$, $L_0 = 4.5$, $\alpha = 2.5$, and foundation relaxation parameter $\eta = 0.1$, Amplitudes were computed up to $t = 2.5$ in each case.

9.3 Main text Fig 5

Main text Fig 5A shows both a linear and fractal-like energy minimizing amplification in both a lower density (i) and higher density (ii) case. we have used parameter values $\mu = 80$, $\delta_0 = 0.01$, $g = 5$, $L_0 = 1$, $\eta = 0.1$, and $\alpha = 3$, solving the system up to $t = 2$ in (i) and $t = 8$ in (ii). The pre-pattern has form 3.2.3.1.3.2.3, with $(\sigma_1, \sigma_2, \sigma_3) = (0.28, 0.2, 0.075)$ in (i) and $(\sigma_1, \sigma_2, \sigma_3) = (0.09, 0.6, 0.021)$ in (ii).

The shell simulations in main text Fig 5B require parameters for (i) the formation of the initial ridges pattern, (ii) the amplification to spines, (iii) parameterization of the shell, and (iv) wrapping the ridges/spines pattern onto the shell.

(i) Ridges pattern For the linear shell, we have used parameter values $\mu = 4 \times 10^4$, $\delta_0 = 0.1$, $g = 5$, $L_0 = 7$, $\eta = 0.1$, and $\alpha = 2.9$, solving the system up to $t = 1.2$.

For the fractal-like shell, we have used parameter values $\mu = 1 \times 10^6$, $\delta_0 = 0.4$, $g = .7$, $L_0 = 1$, and $\alpha = 2$, solving the system up to $t = 8$ (this simulation did not include foundation relaxation, as noted in Sec 7C, which is equivalent to $\eta = 0$).

(ii) Amplification to spines As noted in the main text, amplification of the pattern to spines conceptually changes the interpretation of the two layers: the mantle and periostracum are taken together to comprise Layer 1, while the evolving shell edge plays the role of Layer 2. Consistent with this, Layer 1 should have an increased width and thus increased bending stiffness. Since μ is a dimensionless parameter with bending stiffness in the denominator, it follows that μ should take a smaller, and potentially significantly smaller, value, compared to the ridges formation. Also, as spines form under a burst of growth, the excess rate g should be significantly increased. As foundation relaxation models a completely different process in spine formation (calcification of shell material), η was also allowed to change. However, the domain length in spine formation was set by the end of the ridges computation, as was the excess length (δ_0) and Gaussian widths (in the linear case, as discussed in Sec 7 above). Accordingly, we modified the parameter values as follows.

For the linear shell, we used $\mu = 400$, $g = 1000$, $\eta = 0.5$. The pattern was computed up to $t = 2.5$ (where we reset time to $t = 0$ at the start of spine formation.

For the fractal-like shell, we used $\mu = 10$, $\alpha = 2$, $g = 500$, $\eta = 2$. The pattern was computed up to $t = 0.26$.

(iii) Smooth shell parameterization For the shell with linear spines, the parameters described in Sec 7A were taken to be: $r_0 = 0.65$, $g_w = 0.093$, $h_0 = 1.84$, $c = 0.10$ $\theta_1 = 0.26$, $\theta_2 = -0.11$, $\theta_3 = 1.61$. The functions ρ_1 and ρ_2 had form

$$\rho_1(s) = \frac{1}{2}\left(\tanh\left(\frac{s-a}{b}\right)+1\right)(dx_0 s - a_0\cos(a)) - \frac{1}{2}a_0\cos(s)\left(1-\tanh\left(\frac{s-a}{b}\right)\right) + x_0,$$

$$\rho_2(s) = \frac{1}{2}\left(\tanh\left(\frac{s-a}{b}\right)+1\right)(a_0\sin(a) + dy_0 s) + \frac{1}{2}a_0\sin(s)\left(1-\tanh\left(\frac{s-a}{b}\right)\right) + y_0,$$

with $a = 3.27, a_0 = 0.32, b = 0.1, x_0 = -4.94, dx_0 = 0.34, y_0 = -0.36, dy_0 = -0.03$, plotted over the range $s \in (0.25, 4.41)$.

The fractal-like shell had the same parameterization and the same functional form for ρ_1 and ρ_2, varying only in parameter choices, which were: $r_0 = 0.70$, $g_w = 0.096$, $h_0 = 1.95$, $c = 0.11$

$\theta_1 = 0.26$, $\theta_2 = -0.11$, $\theta_3 = 1.49$, $a = 3.2$, $a_0 = 0.6$, $b = 0.45$, $x_0 = -0.39$, $dx_0 = .34$, $y_0 = -0.30$, $dy_0 = -0.03$, and range $s \in (0, 3.32)$.

(iv) Ridges/spines on shell The only parametric choices needed in connecting the ridges/spines pattern and the smooth shell were in the map between t (ridges/spines time) and τ (the shell parameter that increases with development), as well as a scale parameter. Following the notation in Sec 7E above, we need to set t_0, t_1, and δ_t.

For the linear shell, we set $t_0 = -30$, $t_1 = 3.7$, and $\delta_t = 2\pi/3$. The ridges/spines were also scaled by a factor of 0.075.

For the fractal-like shell, we set $t_0 = 5$, $t_1 = 8.26$, and $\delta_t = 1.88$. The ridges/spines were scaled by a factor of 0.1.

9.4 Main text Fig 6

In the simulations appearing in main text Fig 6, we have used the following parameters:

A. $\mu = 1.546 \times 10^7, \delta_0 = 0.07, g = 0.72, L_0 = 2.76, \alpha = 2.5, \eta = 0.01$, solved up to $t = 4$.
B. $\mu = 2.7 \times 10^4, \delta_0 = 0.28, g = 0.20, L_0 = \pi, \alpha = 2.5, \eta = 0.75$, solved up to $t = 1$.
C. $\mu = 1.94 \times 10^3, \delta_0 = 0.75, g = 3.12, L_0 = 2.67, \alpha = 2.5, \eta = 0.05$, solved up to $t = 1$.
D. $\mu = 1.5 \times 10^6, \delta_0 = 0.07, g = 1.56, L_0 = \pi, \alpha = 2.5, \eta = 0.01$, solved up to $t = 1.3$.

9.5 Fig S2

SM Fig S2A computed amplitude curves compared against the weakly non-linear (WNL) analysis described in [1]. For these plots, we do not scale out integration constants, as doing so would complicate comparison of variables with the WNL calculation. The bending and stretching stiffness are both set to 1, and the interaction energy (which has $m = 2$ in correspondence with the linear foundation in [1]) has stiffness coefficient 2.39×10^{-3} for the plot on the left (producing mode $N_1 = 1$), and 3.98×10^-2 for the plot on the right ($N_1 = 3$). In both cases, the domain length is fixed at $L = 20$.

SM Fig S2B consists of computing amplitudes up to Level 2, with no foundation relaxation. Parameter values are $\mu = 5 \times 10^4$, $\delta_0 = 0.4$, $g = 0.7$, $L_0 = 1$, and $\alpha = 2.3$.

10 On the possibility of mechanical patterning in other organisms

Here we outline the process by which hierarchical patterns may be understood to form in the organisms appearing in main text Fig 1.

10.1 Agaricomycetes

In the fungal kingdom, Agaricomycetes is a highly diversified class of mushroom-forming fungi (the term "mushroom" classically refers to the spore-bearing fruiting body of fungi). Many species have lamellae under the cap that display a hierarchy of size reflecting their sequential development (main text Fig 1F). Primary lamellae form first, radiating all over the inner surface of the cap, while intercalary lamellae (with at least 6 levels) emerge sequentially in the widening spaces between preceding lamellae as the cap grows [26]. Patterns generated with the formalism of local-activation with long-range inhibition have been considered reminiscent of the fertile part of fruiting bodies

such as lamellae, but the lack of true tissues in fungi suggest that a Turing-like process based on diffusible molecules in an isotropic domain is unlikely, contrarily to a mechanical hypothesis [24]. Moreover, lamellae are not a 2D molecular pattern, but 3D structures that emerge sequentially as folds in the expanding concave annular zone under the cap, and display an outer layer made of inflated cells [14], two characteristic suggesting that a mismatch induced by growth differential could induce compressive stresses involved in their development.

10.2 Micrasterias

in the unicellular green alga *Micrasterias*, asexual reproduction occurs via mitosis, a process during which two symmetric flat semi-cells give rise to two daughter cells. As a semi-cell grows and expands, its margin becomes notched by a series of folds pointing inward and displaying a hierarchy of size (5 levels) recording their sequential development (main text Fig 1C). Although the development of *Micrasterias* has been investigated for decades, the formation of the complex patterned cell wall and the precise localization of fold indentations remain unexplained [29]. Interestingly, the cell wall is thicker and shows a denser network of cellulose microfibrils at the high curvature points of the indentations which increases its stiffness and so decrease its extensibility locally [16]. Also, overexpression of a gene closely related to the expansin genes that modulates the mechanical properties of the cell walls in plants, results in cell shape malformations [43]. A mechanical instability of the expanding cell margin, coupled with an irreversibility condition resulting from the thickening and stiffening of the wall at the level on the high curvature indentations, is a plausible working hypothesis for future research on *Micrasterias* morphogenesis.

10.3 Fungiidae

Among Cnidaria, Fungiidae is a family of solitary corals whose name refers to the fungus–like appearance of these animals also known as "mushroom corals" (main text Fig 1G). The coral skeleton (corallum) grows through accretion, with concentric growth lines from the center to the outside. Vertical calcareous septa radiate from the mouth area on the center of the upper surface and show the hierarchical pattern (at least up to 5 levels), their size reflecting their sequential development during growth and expansion of the corallum. The same hierarchical pattern of septa may be seen in other Scleractinia (reef-building corals), including colonial species. These septa are secreted at the level of folds in an outer epithelium. They alternate with mesenteries, that are radial inward-growing folds of tissue increasing the surface area of the gastrovascular cavity (coelenteron) [50]. It is well-known that calcareous septa develop following the pattern of the radial mesenteric folds, but to our knowledge, their morphogenesis remains unexplained. In vertebrates, a wide variety of folding patterns of the gastrointestinal tract emerge as a result of mechanical instabilities of spatially constrained, growing inner tissue [6]. Interestingly, the pharynx and gastrodermis (the tissue forming the mesenteric folds) are the area of highest cell proliferation in the developing scleractinian polyp [25]. Whether a constrained growth generates mechanical stresses involved in the development of the folding pattern of both mesenteries and septa could be an interesting possibility to explore.

10.4 Portuguese man of war

Still among Cnidaria, the "Portuguese man of war" *Physalia physalis* (main text Fig 1H) is a siphonophore, a group of colonial hydrozoans formed by genetically identical, but morphologically

different and functionally specialized zooids (predation, nutrition, reproduction). Zooids are attached below a gas-filled float called pneumatophore, also used as a sail by the colony sometimes found stranded on beaches. The pneumatophore is a multi-layered structure, consisting of an outer codon, an inner gas-sac called pneumatosaccus, and a gas gland that produces carbon monoxide filling the float. Muscle contractions of the outer codon increases pressure within the pneumatosaccus that expands into, and erects the sail-like crest of the pneumatophore [34]. The crest of the pneumatophore displays a series of folds connected by inner septa showing a hierarchy (up to 4 levels) [42]. A comparison of individuals at different stages of development shows that this hierarchy reflects the sequential development of the septa in the growing pneumatophore, though the mechanisms involved remain unknown. In physical systems, a hierarchy of compressive folds may develop perpendicular to stretch at the edge of inflated air bag [12], and it suggests that the inflation of the growing pneumatophore during its development could generate sufficient mechanical stresses involved in the setting up of the hierarchical pattern of folds and septa.

10.5 Lobster claw denticles

Hierarchical patterns are also seen in the row of denticles of one of the claws of the lobster *Homarus* (main text Fig 1I). Initially, lobster claws are identical, but during the juvenile stages, differential reflex and greater activity causes the development of powerful muscle fibers and a stouter crusher claw (equiprobably on the left or right), while the remaining cutter claw keep a slender shape [18]. In older lobsters, only the cutter claw displays the size hierarchy of denticles (at least up to 4 levels), while there is no detectable pattern in the greater and more bulbous denticles of the crusher claw. A similar hierarchical pattern of denticles may be seen in the cutter claw of langoustines *Nephrops norvegicus*. In crabs, successive developmental stages show that these denticles emerge sequentially over the growing claws [30] and correspond to folds separated by localized furrows in the cuticle [45], a morphology of gyri and sulci reminiscent of those generated by mechanical instabilities in physical and biological systems, such as the brain (see [20]). Like other arthropods, Crustacea grow by episodic molting (ecdysis). The new uncalcified cuticle is secreted underneath the old hardened one, so that it grows constrained in a volume less than that of the old and too tight cuticle it is to replace and outpace. This explains why the secreting epithelium and the newly secreted cuticle at the pre-molt stage are highly folded [4]. When the older hardened cuticle is shed, the new soft and flexible cuticle is inflated by increased hydrostatic pressure created by rapid water uptake, and is subsequently calcified. Whether these repeated folding episodes of the epithelium-cuticle complex at the pre-molt stage are associated with an irreversible tissue remodeling and the formation of the denticular folds of the claws is an interesting possibility for further studies to explore.

10.6 Crinoidea

Crinoidea belong to the echinoderm phylum, like starfish or sea urchins. Most crinoids have a stalk fixed to the substrate, and supporting the crown, consisting of the cup-shaped theca (with the mains body parts, viscera, mouth...) from which arms radiate. The mesodermal endoskeleton is composed of articulated series of calcareous pieces. The stalk is composed of sometimes hundreds of pieces called columnals differentiated into nodals and internodals. New nodals develop in the proximal region, just beneath the crown. They are thin initially, and increase in height as they grow and migrate in a more distal position along the stalk. At the same time, secondary, tertiary, quaternary ... internodals are successively intercalated [46]. Thus, columnals display a hierarchy

of size (at least up to 5 levels) reflecting their sequential development and intercalation dynamics over an expanding, growing region of the stalk (main text Fig 1J). Crinoids are recognized for their striking regenerative capability, and stalk regeneration occurs also by proximal nodals development and subsequent intercalation of internodals [35]. But current stalked crinoids live in the relatively deep sea, their development remains rarely studied [3] and to our knowledge, the mechanisms underlying this ordered spatial pattern of columnals are unknown.

10.7 Sawsharks

A distinctive feature of sawsharks is their elongate cartilaginous saw-like rostrum, an extension of the skull that carries on each side, a row of tooth-like spikes that are actually modified skin denticles (main text Fig 1K). The largest denticles grow first, followed by the medium-sized ones as the rostrum grows and space becomes available, and finally the smaller ones [48] generating a hierarchy of size (up to 3 levels) reflecting their sequential development. How this hierarchical pattern is generated remains unexplained. As oral teeth, and other dermal appendages such as scales, spines, hair or feathers, dermal denticles emerge from placodes, folds formed by the thickening of the epithelium and condensation of the underlying mesenchymal cells. Reaction-diffusion processes have been supposed to control the spatial distribution of placodes of both sharks dermal denticles and chick feathers [15]. Under this hypothesis, molecular dynamics first sets up a pattern that serves as a blueprint governing the local behavior of cells, triggering in turn morphological changes. However, the morphogenetic processes that actually shape three-dimensional forms are overlooked in most studies based on the chemical pre-pattern hypothesis. A recent study shows instead that, at least in the case of the developing chick feather buds, mechanical forces do not "read out" a prefiguring molecular pattern, but actively set up and modulate patterning. Indeed, the spontaneous formation of aggregates of mesenchymal cells (the dermis), focally compresses and bends the overlying epithelial layer (the epidermis), activating in turn a mechanotransductive pathway that induces the gene expression initiating formation of feather primordium [39]. An *ex vivo* tissue reconstitution shows that mesenchymal cells spontaneously self-organize into a periodic multicellular aggregates, providing evidence that contractility-dependent mechanical interplay between cells and the extracellular matrix is sufficient to spontaneously generate long-range morphological pattern [36]. Whether similar mechanisms underlie patterning of shark dermal denticles remains unknown, but it is worth noting that our theoretical framework is consistent with this mechanical patterning process in the case of rostral denticles in sawsharks.

10.8 Cynodontidae

A similar mechanochemical process could also explain the dentition patterning in the "vampire fishes" Cynodontidae, a family of freshwater species from South America. Representatives of these family (main text Fig 1L) display a hierarchy of teeth size (up to 4 levels). Although the process of tooth replacement in these fishes has been partly described [7], the dentition has not been studied and whether the self-similar pattern reflects the sequential development of dental placodes over the expanding, growing jaw remains to be verified.

10.9 Pelagornis teeth

Lastly, the hierarchy of size (up to 5 different levels) may be seen in the "teeth" of the fossil genus *Pelagornis* (main text Fig 1M), the largest known flying bird (up to 6,4 m wingspan) [22]. These

conic spikes are actually not true teeth, but bony excrescences from the beak itself. It has been assumed that the development of these pseudoteeth involved epithelial/mesenchymal molecular interactions delayed to the late stage of development, once the beak was fully grown. Assuming therefore that they emerge on a domain of fixed size, the sequential development of these pseudoteeth and their size hierarchy have been interpreted as resulting from a dynamic in which the radius of inhibition of a diffusing molecule gradually decreases around primary, secondary, tertiary and quaternary pseudoteeth [28]. This hypothesis however, lacks both a theoretical support and experimental clues showing that a similar molecular dynamic is involved in other living species. Our theoretical framework and comparison with other organisms lead us to privilege a mechanochemical hypothesis and the sequential development of these pseudoteeth over the expanding, growing beak, a hypothesis that could be partly tested by discovery of very young specimens of this fossil bird.

References

[1] Axel A Almet, Helen M Byrne, Philip K Maini, and Derek E Moulton. Post-buckling behaviour of a growing elastic rod. *Journal of mathematical biology*, 78:777–814, 2019.

[2] Davide Ambrosi, Martine Ben Amar, Christian J Cyron, Antonio DeSimone, Alain Goriely, Jay D Humphrey, and Ellen Kuhl. Growth and remodelling of living tissues: perspectives, challenges and opportunities. *Journal of the Royal Society Interface*, 16(157):20190233, 2019.

[3] Shonan Amemiya, Akihito Omori, Toko Tsurugaya, Taku Hibino, Masaaki Yamaguchi, Ritsu Kuraishi, Masato Kiyomoto, and Takuya Minokawa. Early stalked stages in ontogeny of the living isocrinid sea lily m etacrinus rotundus. *Acta Zoologica*, 97(1):102–116, 2016.

[4] AL Arsenault, JD Castell, and FP Ottensmeyer. The dynamics of exoskeletal—epidermal structure during molt in juvenile lobster by electron microscopy and electron spectroscopic imaging. *Tissue and Cell*, 16(1):93–106, 1984.

[5] Jorge A Audino, José Eduardo AR Marian, Andreas Wanninger, and Sônia GBC Lopes. Mantle margin morphogenesis in nodipecten nodosus (mollusca: Bivalvia): new insights into the development and the roles of bivalve pallial folds. *BMC developmental biology*, 15:1–22, 2015.

[6] Valentina Balbi, Ellen Kuhl, and Pasquale Ciarletta. Morphoelastic control of gastro-intestinal organogenesis: Theoretical predictions and numerical insights. *Journal of the Mechanics and Physics of Solids*, 78:493–510, 2015.

[7] Barry Berkovitz and Peter Shellis. *The teeth of non-mammalian vertebrates*. Academic Press, 2016.

[8] Ulrich Bielefeld, Werner Peters, and Wilhelm Becker. Ultrastructure and cytochemistry of periostracum and mantle edge of biomphalaria glabrata (gastropoda, basommatophora). *Acta Zoologica*, 74(3):181–193, 1993.

[9] F. Brau, H. Vandeparre, A. Sabbah, C. Poulard, A. Boudaoud, and P. Damman. Multiple-length-scale elastic instability mimics parametric resonance of nonlinear oscillators. *Nature Physics*, 7(1):56–60, 2011.

[10] S. Budday, C. Raybaud, and E. Kuhl. A mechanical model predicts morphological abnormalities in the developing human brain. *Scientific reports*, 4:5644, 2014.

[11] Antonio G Checa, Carmen Salas, Elizabeth M Harper, and Juan de Dios Bueno-Pérez. Early stage biomineralization in the periostracum of the 'living fossil'bivalve neotrigonia. *PLoS One*, 9(2):e90033, 2014.

[12] Zhen Chen, Hsiao-Yu Chen, Danny M Kaufman, Mélina Skouras, and Etienne Vouga. Fine wrinkling on coarsely meshed thin shells. *ACM Transactions on Graphics (TOG)*, 40(5):1–32, 2021.

[13] Régis Chirat, Derek E Moulton, and Alain Goriely. Mechanical basis of morphogenesis and convergent evolution of spiny seashells. *Proceedings of the National Academy of Sciences*, 110(15):6015–6020, 2013.

[14] Siu Wai Chiu and David Moore. Development of the basidiome of volvariella bombycina. *Mycological Research*, 94(3):327–337, 1990.

[15] Rory L Cooper, Alexandre P Thiery, Alexander G Fletcher, Daniel J Delbarre, Liam J Rasch, and Gareth J Fraser. An ancient turing-like patterning mechanism regulates skin denticle development in sharks. *Science advances*, 4(11):eaau5484, 2018.

[16] Martin Felhofer, Konrad Mayr, Ursula Lütz-Meindl, and Notburga Gierlinger. Raman imaging of micrasterias: new insights into shape formation. *Protoplasma*, 258(6):1323–1334, 2021.

[17] A. Goriely. *The Mathematics and Mechanics of Biological Growth*. Springer Verlag, New York, 2017.

[18] CK Govind and Joanne Pearce. Mechanoreceptors and minimal reflex activity determining claw laterality in developing lobsters. *Journal of experimental biology*, 171(1):149–162, 1992.

[19] J Herbert Waite, ASM Saleuddin, and SO Andersen. Periostracin—a soluble precursor of sclerotized periostracum in mytilus edulis l. *Journal of comparative physiology*, 130:301–307, 1979.

[20] Maria Holland, Silvia Budday, Alain Goriely, and Ellen Kuhl. Symmetry breaking in wrinkling patterns: Gyri are universally thicker than sulci. *Physical review letters*, 121(22):228002, 2018.

[21] Daniel J Jackson, Carmel McDougall, Kathryn Green, Fiona Simpson, Gert Wörheide, and Bernard M Degnan. A rapidly evolving secretome builds and patterns a sea shell. *BMC biology*, 4(1):1–10, 2006.

[22] Daniel T Ksepka. Flight performance of the largest volant bird. *Proceedings of the National Academy of Sciences*, 111(29):10624–10629, 2014.

[23] Michael Kücken and Alan C Newell. Fingerprint formation. *Journal of theoretical biology*, 235(1):71–83, 2005.

[24] Francisco Kuhar, Leticia Terzzoli, Eduardo Nouhra, Gerardo Robledo, and Moritz Mercker. Pattern formation features might explain homoplasy: fertile surfaces in higher fungi as an example. *Theory in Biosciences*, 141(1):1–11, 2022.

[25] Agathe Lecointe, Isabelle Domart-Coulon, Alain Paris, and Anders Meibom. Cell proliferation and migration during early development of a symbiotic scleractinian coral. *Proceedings of the Royal Society B: Biological Sciences*, 283(1831):20160206, 2016.

[26] Michael Levine. The origin and development of lamellae in agaricus campestris and in certain species of coprinus. *American Journal of Botany*, pages 509–533, 1922.

[27] Bo Li, Yan-Ping Cao, Xi-Qiao Feng, and Huajian Gao. Mechanics of morphological instabilities and surface wrinkling in soft materials: a review. *Soft Matter*, 8(21):5728–5745, 2012.

[28] Antoine Louchart, Vivian de Buffrénil, Estelle Bourdon, Maitena Dumont, Laurent Viriot, and Jean-Yves Sire. Bony pseudoteeth of extinct pelagic birds (aves, odontopterygiformes) formed through a response of bone cells to tooth-specific epithelial signals under unique conditions. *Scientific reports*, 8(1):12952, 2018.

[29] Ursula Lütz-Meindl. Micrasterias as a model system in plant cell biology. *Frontiers in Plant Science*, 7:999, 2016.

[30] Nobufumi Masunari, Kazuki Sekiné, Bong Jung Kang, Yoshitake Takada, Masatsugu Hatakeyama, and Masayuki Saigusa. Ontogeny of cheliped laterality and mechanisms of reversal of handedness in the durophagous gazami crab, portunus trituberculatus. *The Biological Bulletin*, 238(1):25–40, 2020.

[31] Myoung-Woon Moon, Sang Hoon Lee, Jeong-Yun Sun, Kyu Hwan Oh, Ashkan Vaziri, and John W Hutchinson. Wrinkled hard skins on polymers created by focused ion beam. *Proceedings of the National Academy of Sciences*, 104(4):1130–1133, 2007.

[32] DE Moulton, TH Lessinnes, and A Goriely. Morphoelastic rods. part i: A single growing elastic rod. *Journal of the Mechanics and Physics of Solids*, 61(2):398–427, 2013.

[33] Derek E Moulton and Alain Goriely. Possible role of differential growth in airway wall remodeling in asthma. *Journal of applied physiology*, 110(4):1003–1012, 2011.

[34] Catriona Munro, Zer Vue, Richard R Behringer, and Casey W Dunn. Morphology and development of the portuguese man of war, physalia physalis. *Scientific reports*, 9(1):15522, 2019.

[35] Hiroaki Nakano, Taku Hibino, Yuko Hara, Tatsuo Oji, and Shonan Amemiya. Regrowth of the stalk of the sea lily, metacrinus rotundus (echinodermata: Crinoidea). *Journal of Experimental Zoology Part A: Comparative Experimental Biology*, 301(6):464–471, 2004.

[36] Karl H Palmquist, Sydney F Tiemann, Farrah L Ezzeddine, Sichen Yang, Charlotte R Pfeifer, Anna Erzberger, Alan R Rodrigues, and Amy E Shyer. Reciprocal cell-ecm dynamics generate supracellular fluidity underlying spontaneous follicle patterning. *Cell*, 185(11):1960–1973, 2022.

[37] Juan Rodriguez-Hernandez. Wrinkled interfaces: Taking advantage of surface instabilities to pattern polymer surfaces. *Progress in Polymer Science*, 42:1–41, 2015.

[38] ASM Saleuddin and HENRI P PETIT. The mode of formation and the structure of the periostracum. In *The mollusca*, pages 199–234. Elsevier, 1983.

[39] Amy E Shyer, Alan R Rodrigues, Grant G Schroeder, Elena Kassianidou, Sanjay Kumar, and Richard M Harland. Emergent cellular self-organization and mechanosensation initiate follicle pattern in the avian skin. *Science*, 357(6353):811–815, 2017.

[40] Amy E Shyer, Tuomas Tallinen, Nandan L Nerurkar, Zhiyan Wei, Eun Seok Gil, David L Kaplan, Clifford J Tabin, and L Mahadevan. Villification: how the gut gets its villi. *Science*, 342(6155):212–218, 2013.

[41] Stephen P Timoshenko and James M Gere. *Theory of elastic stability*. Courier Corporation, 2012.

[42] AK Totton. Studies on physalia physalis (l.). part 1. *Natural history and morphology. Discovery Reports*, 30:301–368, 1960.

[43] Katrijn Vannerum, Marie JJ Huysman, Riet De Rycke, Marnik Vuylsteke, Frederik Leliaert, Jacob Pollier, Ursula Lütz-Meindl, Jeroen Gillard, Lieven De Veylder, Alain Goossens, et al. Transcriptional analysis of cell growth and morphogenesis in the unicellular green alga micrasterias (streptophyta), with emphasis on the role of expansin. *BMC Plant Biology*, 11:1–17, 2011.

[44] Victor D Varner and Celeste M Nelson. Computational models of airway branching morphogenesis. In *Seminars in cell & developmental biology*, volume 67, pages 170–176. Elsevier, 2017.

[45] David A Waugh, Rodney M Feldmann, Amanda M Schroeder, and Matt HE Mutel. Differential cuticle architecture and its preservation in fossil and extant callinectes and scylla claws. *Journal of Crustacean Biology*, 26(3):271–282, 2006.

[46] GD Webster. Crinoid pluricolumnal noditaxis patterns. *Journal of Paleontology*, pages 1283–1288, 1974.

[47] Nicole B Webster and A Richard Palmer. Connecting pattern to process: Growth of spiral shell sculpture in the gastropod nucella ostrina (muricidae: Ocenebrinae). *Evolution & Development*, 20(5):160–171, 2018.

[48] Monique Welten, Moya Meredith Smith, Charlie Underwood, and Zerina Johanson. Evolutionary origins and development of saw-teeth on the sawfish and sawshark rostrum (elasmobranchii; chondrichthyes). *Royal Society Open Science*, 2(9):150189, 2015.

[49] KARL M WILBUR and ASM Saleuddin. Shell formation. In *The mollusca*, pages 235–287. Elsevier, 1983.

[50] Cheryl M Woodley, Craig A Downs, Andrew W Bruckner, James W Porter, and Sylvia B Galloway. Diseases of coral. 2016.

[51] Yeguang Xue, Won-Kyu Lee, Jianghong Yuan, Teri W Odom, and Yonggang Huang. Mechanics modeling of hierarchical wrinkle structures from the sequential release of prestrain. *Langmuir*, 34(51):15749–15753, 2018.