Geometry and Elasticity of Shells

Daniel Duffy, dlduffy@umich.edu August 27, 2025

What is a shell? It's any object that's thin enough that it can be modelled as a 2D surface. We'll consider shells made of elastic solid materials.

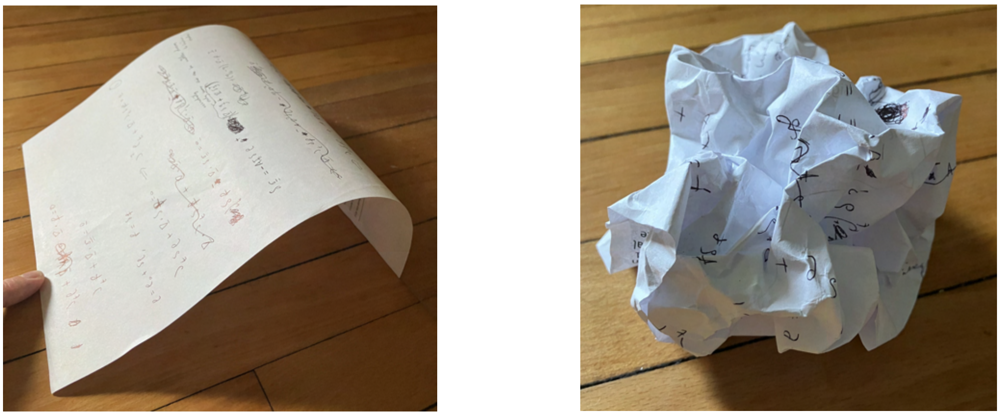

Figure 1: Left: A beautiful shell. Right: Another shell, beautiful in its own special way.

Why study them? First, they can *bend* easily because they're thin, and thereby change *shape* in dramatic or interesting ways. Thus they have fascinating geometry-driven mechanics, so shell theory is a deeply mathematical subject, but one that often involves making or looking at pretty pictures, which is a nice combination. Second, they are used all over the place, to achieve mechanical goals with a minimum of weight and space. Roofs and walls, musical instruments (especially the parts that radiate sound efficiently), hulls of rockets/boats/cars, origami/paper, fabrics/textiles, pressure vessels, ping-pong balls, leaves/flowers, . . . etc! Leaves and other similar examples in biology are particularly interesting, because they are 'active': they change shape of their own accord as they grow. That's very cool, but we'll stick to non-active shells for now.

Since shells are all about shape, we'll need to begin with some geometry. Sadly we'll only cover the absolute bare minimum required for our examples (or possibly a little less).[1] We'll use the Einstein summation convention. Latin indices will run over $\{1, 2, 3\}$, Greek indices will run over $\{1, 2\}$, and $\partial_\alpha \equiv \partial/\partial x^\alpha$. It's useful to define a fixed *reference state*/configuration for the 2D shell; usually this is just the state before any deformation has occurred. We inscribe it with some

[1]If you want to read more about differential geometry, Hobson *et al*'s *General Relativity: An Introduction for Physicists* is a good place to start. Marsden and Hughes's *Mathematical Foundations of Elasticity* is a good place to continue.

coordinates x^1, x^2. These coordinates are best thought of as labels for 'material points': each tiny physical piece of the shell is assigned a unique label, which is just its coordinate pair (x^1, x^2) (imagine this being physically painted onto each infinitesimal element).

Any realised/deformed state of the shell, can be described by specifying the location in 3D of each material point. We will be considering only Euclidean 3D space, so we can use Cartesian 3D coordinates and interpret them as components of a position vector $\boldsymbol{R} = (X, Y, Z)$. We thus describe any realised shell with a vector function of the reference coordinates: $\boldsymbol{R}(x^1, x^2)$. As the shell deforms, the reference state does not change; it's just the function $\boldsymbol{R}(x^1, x^2)$ that changes. See Fig. 2.

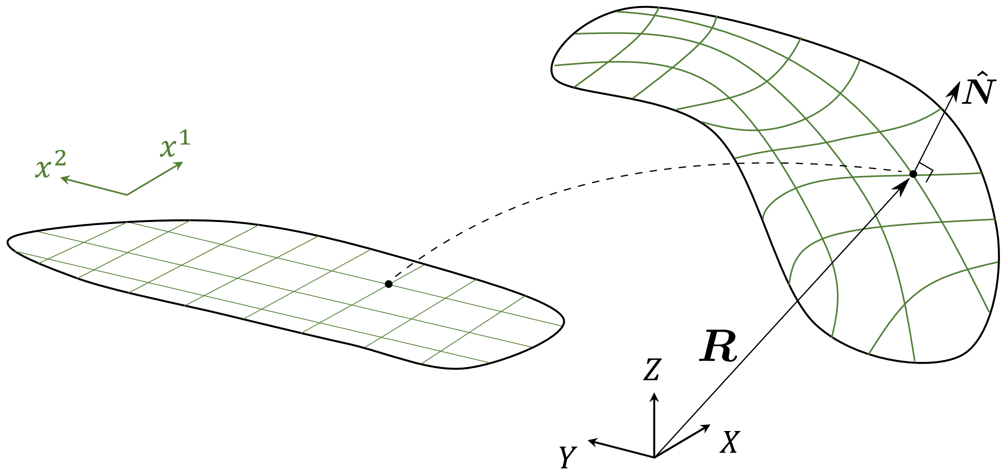

Figure 2: Left: A reference configuration of a shell, flat in this case, with coordinates x^α inscribed. Right: A realised shell described by $\boldsymbol{R}(x^\alpha)$.

What might $\boldsymbol{R}(x^\alpha)$ look like? Well, for example, if θ ranges from 0 to π and ϕ ranges from 0 to 2π, then

$$\boldsymbol{R}(\theta, \phi) = \begin{pmatrix} L \sin\theta \cos\phi \\ L \sin\theta \sin\phi \\ L \cos\theta \end{pmatrix} \tag{1}$$

describes a full sphere of radius L. If z ranges from 0 to H and ϕ ranges from 0 to 2π then

$$\boldsymbol{R}(\phi, z) = \begin{pmatrix} L \cos\phi \\ L \sin\phi \\ 2z \end{pmatrix} \tag{2}$$

describes a cylinder of height $2H$ and radius L. Similarly,

$$\boldsymbol{R}(x,y) = \begin{pmatrix} 0 \\ y \\ x \end{pmatrix} \tag{3}$$

describes some portion of the $Y - Z$ plane, depending on the ranges of x and y. (I promise a reference state is a helpful concept, even though I didn't specify one in any of those examples. For one thing, the reference state is what fixes the ranges of the coordinates x^α.)

Now, the tiny vector in the realised shell between material point x^α and material point $x^\alpha + \mathrm{d}x^\alpha$ is $\mathrm{d}\boldsymbol{R} = \mathrm{d}x^\alpha\, \partial_\alpha \boldsymbol{R}$. Thus the realised squared distance between two such neighbouring material points

$$|\mathrm{d}\boldsymbol{R}|^2 = \mathrm{d}\boldsymbol{R} \cdot \mathrm{d}\boldsymbol{R} = \mathrm{d}x^\alpha\, (\partial_\alpha \boldsymbol{R}) \cdot (\partial_\beta \boldsymbol{R})\, \mathrm{d}x^\beta \equiv \mathrm{d}x^\alpha\, a_{\alpha\beta}\, \mathrm{d}x^\beta\,, \tag{4}$$

where we've defined the downstairs components $a_{\alpha\beta}$ of the 'metric' tensor field a. At each material point, $a_{\alpha\beta}$ are the components of a symmetric 2×2 matrix. For our earlier sphere example, where $x^1 = \theta$ and $x^2 = \phi$,

$$a_{\alpha\beta} = \begin{pmatrix} L^2 & 0 \\ 0 & L^2 \sin^2\theta \end{pmatrix}. \tag{5}$$

Check you can calculate this, and find the metrics for the cylinder and plane examples too (exercise)! Since a measures distances between points, it is invariant under rigid-body motions of the realised shell, but is affected by stretches/compressions. Moreover, a is purely 'intrinsic': it measures only in-surface distances, i.e. only those that could be measured by a short-sighted bug crawling on the surface. W define a's upstairs components $a^{\alpha\beta}$ via a matrix inverse: $a^{\alpha\beta} a_{\beta\gamma} \equiv \delta^\alpha_\gamma$, where as usual δ^α_γ is the Kronecker delta, i.e. the components of the identity matrix. I sometimes use the symbol a^{-1} for the matrix with elements $a^{\alpha\beta}$.

At each point on the realised surface there is a unit vector $\hat{\boldsymbol{N}}$ normal (perpendicular) to the surface and a 'tangent plane' that's tangential to the surface. The two vectors $\partial_\alpha \boldsymbol{R}$ form a basis for the tangent plane at each point, and the normal vector is given by

$$\hat{\boldsymbol{N}} \equiv \frac{\partial_1 \boldsymbol{R} \times \partial_2 \boldsymbol{R}}{|\partial_1 \boldsymbol{R} \times \partial_2 \boldsymbol{R}|}. \tag{6}$$

We can now define another key tensor field κ: the 'second fundamental form' or 'curvature tensor'.[2] Its components

$$\kappa_{\alpha\beta} \equiv -\hat{\boldsymbol{N}} \cdot \partial_\beta \partial_\alpha \boldsymbol{R}. \tag{7}$$

Like a, κ is manifestly invariant under rigid-body motions of the realised surface. Also, κ is symmetric due to the symmetry of $\boldsymbol{R}$'s mixed second derivatives. However, unlike a, κ depends on the 'extrinsic' geometry: bending a flat sheet into a cylinder will change κ while leaving a unchanged.

What does κ have to do with curvature? Well, suppose $\boldsymbol{R}(x^\alpha)$ is a linear function. This means the realised shell is a flat plane (as in one of our earlier examples). Then the second derivatives of $\boldsymbol{R}$ are zero, so $\kappa = 0$. Thus we've established that $\kappa = 0$ for flat shells!

Going further, let's examine a surface in the vicinity of some material point p, whose realised position is $\boldsymbol{R}|_p$. Taylor-expanding about p, we get

$$\boldsymbol{R} = \boldsymbol{R}|_p + \mathrm{d}x^\alpha \left(\partial_\alpha \boldsymbol{R}\right)|_p + \frac{1}{2}\mathrm{d}x^\alpha \left(\partial_\beta \partial_\alpha \boldsymbol{R}\right)|_p \, \mathrm{d}x^\beta + \text{higher order terms.} \tag{8}$$

Now, how far is each point $\boldsymbol{R}$ from p's tangent plane? In other words, what's the local 'height function' describing the surface, with height measured from p's tangent plane? Well, we just write down the displacement vector from p to $\boldsymbol{R}$, which is $\boldsymbol{R} - \boldsymbol{R}|_p$, and then we compute the component normal to the plane, i.e. the component along $\hat{\boldsymbol{N}}|_p$. Thus the height function is

$$\begin{aligned} h &= \hat{\boldsymbol{N}}|_p \cdot (\boldsymbol{R} - \boldsymbol{R}|_p) && (9) \\ &= -\frac{1}{2}\mathrm{d}x^\alpha \mathrm{d}x^\beta \left(\kappa_{\alpha\beta}\right)|_p + \text{higher order terms,} && (10) \end{aligned}$$

where we used the fact that $\hat{\boldsymbol{N}}|_p$ is perpendicular to $(\partial_\alpha \boldsymbol{R})|_p$, and substituted in our definition of κ. The above formula captures the essence of κ: Locally around any p, every surface is described by a quadratic function giving the perpendicular displacement from p's tangent plane, and κ encodes the coefficients of that quadratic.

Here's another consequence of this local-quadratic picture: each straight line lying in p's tangent plane and passing through p becomes a parabola locally, when projected along $\hat{\boldsymbol{N}}|_p$ onto the surface (i.e. cast the line's shadow onto the surface). There's a circle that locally lines up with that parabola, called the 'osculating'

[2] Warning: much of the terminology commonly draped around κ is an outdated mess. The symbol b is often used instead.

circle; see Fig. 3. A quick sketch and a Taylor expansion reveals that, at a point

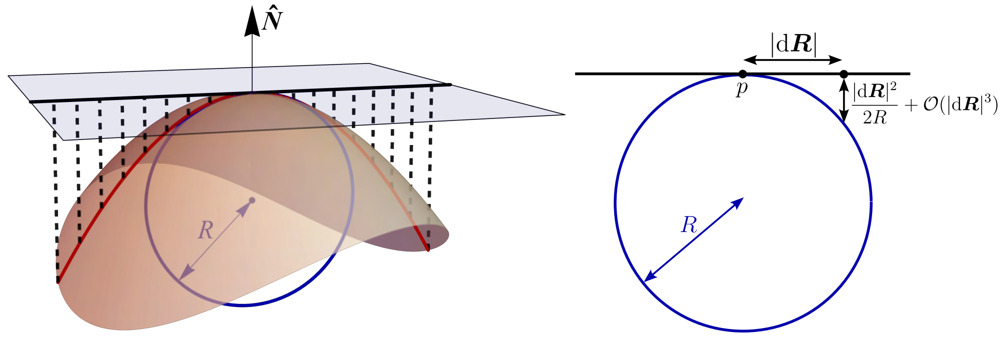

Figure 3: A paraboloid, which every sufficiently smooth surface looks like locally. Any (black) straight line in the tangent plane, projected along $\hat{\boldsymbol{N}}$ onto the paraboloid, yields a (red) parabola. The parabola's curvature $1/R$ is the inverse radius of its (blue) osculating circle.

in the tangent plane a distance $|\mathrm{d}\boldsymbol{R}|$ from p, the perpendicular distance to the circle is $|\mathrm{d}\boldsymbol{R}|^2/2R$ to leading order, where R is the circle's radius (please excuse the poor choice of symbol). We equate this distance to that given by (10) so that the circle and the parabola match to leading order, finding

$$\frac{1}{R} = \frac{\mathrm{d}x^\alpha \, \mathrm{d}x^\beta \, (\kappa_{\alpha\beta})|_p}{|\mathrm{d}\boldsymbol{R}|^2} = \frac{\mathrm{d}x^\alpha \, \mathrm{d}x^\beta \, (\kappa_{\alpha\beta})|_p}{\mathrm{d}x^\rho \, \mathrm{d}x^\sigma \, a_{\rho\sigma}}, \tag{11}$$

where the second equality is just subbing in (4). Now, the above equation yields a different $1/R$ for each $\mathrm{d}x^\alpha$. In each case $\mathrm{d}x^\alpha$ corresponds to some direction (the black line), and $1/R$ is called the 'normal' curvature in that direction. If you plug in every possible $\mathrm{d}x^\alpha$, you'll find that $1/R$ takes a maximum and a minimum value at each point; we call these the 'principal curvatures', κ_1 and κ_2 respectively. Now, just defining $v^\alpha \equiv \mathrm{d}x^\alpha$ for a moment to avoid notational insanity, you can differentiate (11) with respect to v^α to find when these principal curvatures occur; in a few lines you can (exercise!) show that they occur when v^α is an eigenvector of the matrix $a^{\gamma\beta}\kappa_{\beta\alpha}$, with the corresponding eigenvalue being $1/R$. Thus the principal curvatures are just the eigenvalues of that matrix.[3]

Now, for any matrix, the sum of the eigenvalues is the trace, and the product of the eigenvalues is the determinant. Thus the 'mean' or Germain curvature (after Sophie Germain)

$$H \equiv \frac{\kappa_1 + \kappa_2}{2} = \frac{1}{2}\mathrm{tr}[a^{-1}\kappa]\,, \tag{12}$$

[3]This matrix really represents the same tensor as κ; it's just the 'mixed components'. However it's often called the 'shape operator', annoyingly.

and the Gauss curvature

$$K \equiv \kappa_1 \kappa_2 = \det[a^{-1}\kappa], \tag{13}$$

where within square brackets a^{-1} and κ refer to matrices of $a^{\alpha\beta}$ and $\kappa_{\alpha\beta}$ respectively.

Let's get a little intuition for H and K: First, note that the sign of H is essentially arbitrary, because you could always reverse the (arbitrary) direction of $\hat{\boldsymbol{N}}$ everywhere, flipping the sign of κ, and thus of H. In contrast, the sign of K is real and important: it determines whether the surface is locally saddle-like, cylinder-like, or sphere-like; see Fig. 4. If $K < 0$ that means the principal curvatures must

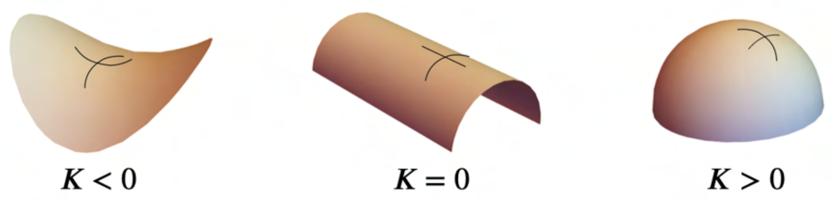

Figure 4: Surfaces with different signs of Gauss curvature K. Arcs of the principal-curvature osculating circles are shown in black.

have opposite signs. What can that mean, since no osculating circle can have a negative radius? Well, we use the word 'radius' slightly loosely, defining it to be negative if $\hat{\boldsymbol{N}}$ points towards the circle's centre, and positive otherwise. Thus $K < 0$ corresponds to the two principal osculating circles curving in opposite senses to each other, like a saddle. If either principle curvature is zero (meaning the corresponding osculating circle is just a straight line), then $K = 0$ necessarily. Finally, note that $H = 0$ is only possible if $K < 0$ (or if the shell's completely flat).

It's straightforward to see that H is an extrinsic quantity: Imagine bending a flat sheet of paper into half a cylinder. The flat sheet had $H = 0$, while the half-cylinder has one non-zero principal curvature (Fig. 4 center), so has $H \neq 0$. Thus H has detected the deformation even though the in-material distances between material points didn't change, i.e. the metric a didn't change; it can't have, because to a good approximation paper can't be stretched (it'll tear first). Such constant-a deformations are called 'isometric', and two surfaces are said to be 'isometries' of each other if they're related by such a deformation.

(Exercise: Given a flat sheet $\boldsymbol{R}(x, y) = (x, y, 0)$, write down an $\boldsymbol{R}(x, y)$ that would correspond to the sheet being isometrically wrapped around a cylinder of radius L. Compute $a_{\alpha\beta}$ to show that it's unchanged by the deformation.)

In contrast, K is intrinsic; to change K you have to stretch or compress the material tangentially, changing the in-material distances between material points. In fact, K can be computed directly just from the metric, without writing down κ at all![4] This result is called Gauss's Theorema Egregium. It's not at all obvious, and in some ways it's pretty counter-intuitive. However, it should hopefully be fairly familiar from everyday life that you can't deform any one of the surfaces in Fig. 4 into any of the others without stretching. This fact has frustrated mapmakers and gift wrappers for centuries: a flat sheet of paper is essentially inextensible, and has $K = 0$, so cannot be wrapped onto the $K > 0$ surface of a sphere. As a result, highly distorting 'projections' must be used to draw flat maps of the Earth's curved surface.

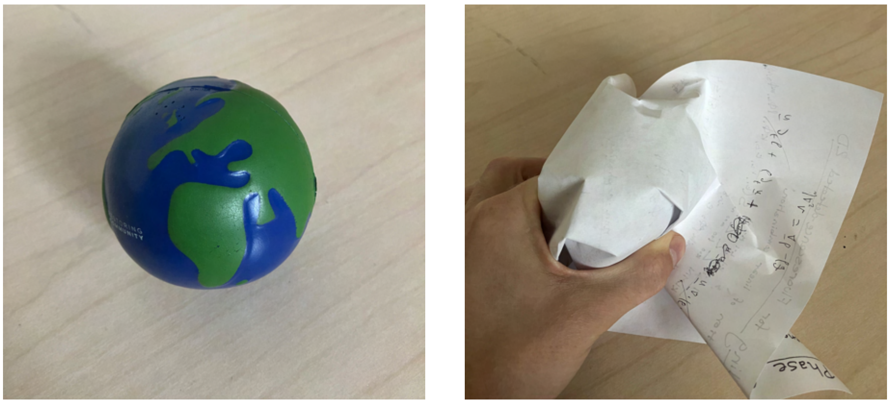

Figure 5: Left: A globe ($K > 0$) that we'd like to map onto a flat sheet ($K = 0$). Maybe to achieve this we can just wrap the sheet around it without stretching? Right: Wrapping not going according to plan. Sad!

Here are some more non-obvious facts:

1. The tensors a and κ are not completely independent; there are three scalar 'Gauss–Codazzi' equations that they must satisfy if they are to correspond to a real surface. These constraints just encode the requirement that the (third) mixed partial derivatives of $\boldsymbol{R}(x^\alpha)$ are symmetric; if they are satisfied, a and κ are termed 'compatible'. One of the three equations ends up equating the

[4]The formula is a nonlinear mess involving second derivatives of $a_{\alpha\beta}$. It's called the Brioschi formula.

$K = \det[a^{-1}\kappa]$ with an expression involving only derivatives of the metric a, and thus encapsulates (and proves) Gauss's Theorema Egregium.

2. If I give you an a field and a κ field, you can in principle work backwards to work out the shape of the corresponding surface. In other words, together a and κ fully determine a surface up to rigid-body transformations (roughly speaking at least; there can be topological complications). This is sometimes called the fundamental theorem of surface theory, or the Bonnet theorem.

Now, deformations of an elastic shell cost energy. This energy E is given as an integral over the reference-state surface: $E = \int W(\boldsymbol{R}(x^\alpha))\mathrm{d}A$, where W is some 'energy density'. Due to the Bonnet theorem I just mentioned, a and κ actually fully describe the shape of a surface, so we can replace $W(\boldsymbol{R}(x^\alpha))$ with some $W(a, \kappa)$. The standard W is

$$W = \frac{Y}{8(1-\nu^2)} Q[\bar{a}^{-1}(a - \bar{a})] + \frac{D}{2} Q[\bar{a}^{-1}(\kappa - \bar{\kappa})]\,, \tag{14}$$

where the quadratic operator $Q(\tau) \equiv \nu \mathrm{tr}(\tau)^2 + (1-\nu)\mathrm{tr}(\tau \cdot \tau)$. Here ν is Poisson ratio, while $Y \equiv \mathtt{E}t$ and $D \equiv \mathtt{E}t^3/(12(1-\nu^2))$ are respectively the stretching modulus and bending modulus ('flexural rigidity') for Young's modulus $\mathtt{E}$. The parameter t is the thickness of the true 3D shell, which we otherwise pretend is just a 2D surface. The tensor fields $\bar{a}$ and $\bar{\kappa}$ are the 'preferred' or 'target' metric and curvature tensor respectively. Why those names? Because the energy is zero if $a = \bar{a}$ and $\kappa = \bar{\kappa}$, so they are the values of a and κ that the shell 'wants' to have; thus they are usually just the a and κ of the reference state.

Rewriting W very schematically, it looks like

$$W \sim t[a - \bar{a}]^2 + t^3[\kappa - \bar{\kappa}]^2\,. \tag{15}$$

so deviations $a - \bar{a}$ are only picked up by the first term in the energy, and deviations $\kappa - \bar{\kappa}$ are only *directly* picked up by the second term (though remember that a does secretly know all about Gauss curvature due to the Theorema Egregium). The first term is called the stretch term, because it measures deviations in in-material distances from their preferred values, which is what stretch/compression is. The second term is called the bend term because it is affected by bending, i.e. curvature changes. *Crucially* the stretch term $\propto t$, while the bend term $\propto t^3$.

Since in shells, t is small (in some appropriate sense), bend is much cheaper than stretch; this fact governs almost all of shell mechanics. It ensures that a is almost always close to $\bar{a}$ in a shell. Let's now define the strain tensor $\varepsilon \equiv (a - \bar{a})/2$ and the bend tensor $\beta \equiv \kappa - \bar{\kappa}$, so W becomes

$$W = \frac{Y}{2(1-\nu^2)} Q[\bar{a}^{-1}\varepsilon] + \frac{D}{2} Q[\bar{a}^{-1}\beta] \,. \tag{16}$$

We can now view the Theorema Egregium more optimistically than mapmakers do: We can engineer *strong* shells, if we arrange their shape such that a certain deformation will require changing K; then it will require changing a, and thus paying expensive stretch energy, which the shell will resist strongly! We all do this in our daily lives. Holding a slice of pizza, for example, we tend to pinch it into a partial cylinder, imbuing it with 'curvature-induced rigidity': The undeformed slice has $K \approx 0$ since pizzas are (usually!) baked while planar, and pinching is easy because it doesn't change K, so costs only pure bend energy. For a pinched slice to then droop excessively under gravity, it would have to stretch to form a $K < 0$ saddle-like shape, whereas an un-pinched slice can easily droop via pure bend into a $K = 0$ partial cylinder; see Fig. 6.

Figure 6: Left: A slice drooping via cheap pure bend. Right: Rigidified by curvature, this slice cannot droop without expensive stretch, so it droops much less.

How do we use (16) to predict a shell's shape in a statics problem? Well, as usual in physics, we minimize energy to find equilibrium states. (More generally one could also write down a kinetic energy, and an action, and find equations of motion ... and so on.) Our energy will be $E = \int W(\boldsymbol{R}(x^\alpha))\mathrm{d}A$, plus some extra term that corresponds to a load (force).

In simulations we can minimise our energy in its full glory. For theory work, the game is to throw away lots of unimportant terms in the energy to get something tractable. Which terms to chuck is problem dependent, not obvious, and sometimes tedious to work out. Different possibilities correspond to the huge zoo of different 'shell theories' with names like Föppl–von Kármán theory, Donnell-Mushtari-Vlasov theory, shallow-shell theory, moderate-rotation theory, membrane theory, . . . we won't worry too much about the names or the term chucking process; we'll just get on with solving problems.

1 Examples

1.1 A cantilever plate

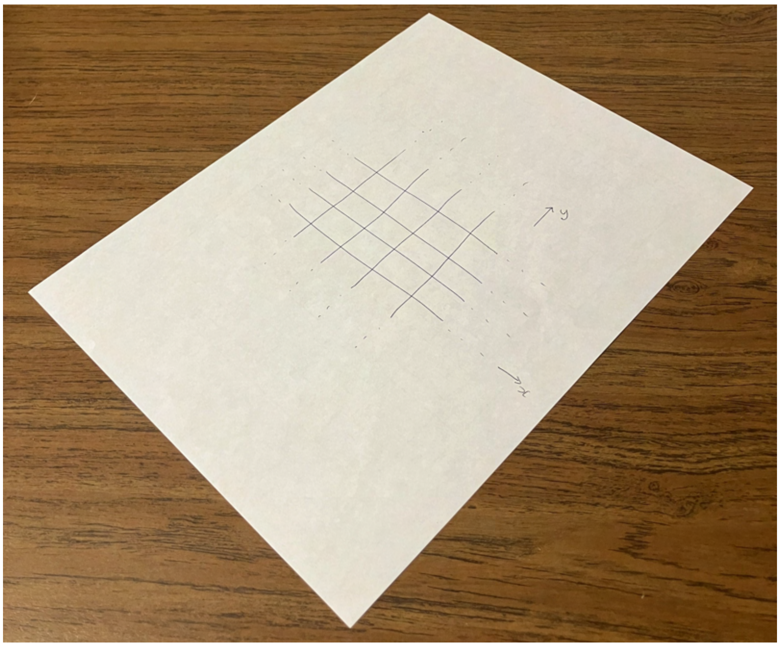

Figure 7: What a plate looks like. A 'paper plate' is this thing; ask anyone.

We're going to start with plates, meaning shells whose zero-energy state is planar. Thus $\bar{\kappa}_{\alpha\beta} = 0$. The reference state will be a flat rectangle, inscribed with Cartesian coordinates $x^1 = x$, $x^2 = y$. Thus $\bar{a}_{\alpha\beta}$ is just the identity matrix. We'll describe the deformation in terms of small displacement fields $\boldsymbol{u}(x, y) = (u(x, y),\ v(x, y))$ and $w(x, y)$ respectively in the tangential and normal directions to the reference state, so

$$\boldsymbol{R}(x,y) = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \tag{17}$$

where that first term is just the $\boldsymbol{R}$ of the reference state.

We'll do 'linear elasticity', meaning we assume that perturbations from the reference state are small enough that we can take ε and β to be linear in displacements and their derivatives, with higher-order terms negligible.

The correct linear expressions for a plate are

$$\varepsilon_{\alpha\gamma} = \frac{1}{2}\left(\nabla_\alpha u_\gamma + \nabla_\gamma u_\alpha\right), \tag{18}$$

$$\beta_{\alpha\gamma} = -\nabla_\alpha \nabla_\gamma w. \tag{19}$$

The first line should look familiar from standard 3D elasticity. The second should remind you of (7), our definition of κ. Indeed, we must have $\hat{\boldsymbol{N}} = (0,0,1)+$ terms first order in displacements and their derivatives, and thus the β_{ij} above comes quickly from plugging (17) into (7) and working to leading order (exercise). Now we can plug ε and β into (16) to get

$$W = \frac{Y}{2(1-\nu^2)}\, Q[\left(\nabla_\alpha u_\gamma + \nabla_\gamma u_\alpha\right)/2] + \frac{D}{2}\, Q[\nabla_\alpha \nabla_\gamma w]\,. \tag{20}$$

Starting to look tractable!

For our first example, we're going to look at a 'cantilever'; see Fig. 8, The end of the plate at $x = L$ will be unconstrained, while the $x = 0$ end will be 'clamped', meaning it can neither move nor rotate. Finally, I'll assume a uniform load force-per-unit-length is applied along the free end, in the normal direction.

Our setup allows us to make a huge simplification: translational invariance in the y direction (i.e. nothing depends on y).[5] Then, by mirror symmetry, $v = 0$. Thus, representing derivatives with respect to x by primes, the energy density (20) becomes

$$W = \frac{Y}{2(1-\nu^2)}\,(u')^2 + \frac{D}{2}\,(w'')^2. \tag{21}$$

[5] Even in this problem translational invariance is in truth broken near the shell's boundaries unless $\nu = 0$, but we'll ignore that.

Figure 8: A plate cantilever (white) is 'built in' / clamped at its left-hand edge, and loaded at its right-hand edge by a load (red).

To get an energy, we integrate W over area, but since W doesn't depend on y, we can just do the y-integral trivially, leaving a single integral over x. That integral is the plate's elastic energy, but we need to add a term to account for the load force. Let's make the standard assumption that the load doesn't change as deformation occurs (usually a very good approximation); for example the force could be the weight of a wire glued along the free edge. In that case the gravitational potential energy would be $Fw(L)$ for a total force F, and something like this almost always the form used for any applied force in continuum mechanics; it's minus the work done by the force as the deformation occurs. Thus the total energy of the system is

$$E = Fw(L) + S\int_0^L \left(\frac{Y}{2(1-\nu^2)}(u')^2 + \frac{D}{2}(w'')^2\right)\mathrm{d}x, \tag{22}$$

where S is the plate's length in the y direction. We just need to minimise this over all fields u and w, subject to the constraints implied by the 'clamping' BCs at $x = 0$, which are $u(0) = w(0) = w'(0) = 0$. The way I think about the $w' = 0$ BC is that in reality the plate extends to $x < 0$, it's just 'glued down' flat there, so has $w' = 0$ there, so we need the $w' = 0$ BC to avoid a discontinuity in w' at $x = 0$, which would yield a discontinuous normal vector, and hence infinite bend energy.

We're ready to minimise E, which we do via standard calculus-of-variations handle turning: We perturb about some already-deformed state by making the substitutions $u \to u + \delta u$ and $w \to w + \delta w$, and require that $\delta E = 0$ to leading order for all δu and δw. When that holds, the deformed state is a stationary point of the energy, and we trust that it's a minimum. We only consider perturbations that satisfy $\delta u(0) = \delta w(0) = \delta w'(0) = 0$, because any others are forbidden by our boundary conditions; the system can't explore them so neither should we. To leading order,

$$\begin{aligned}\delta E &= F\delta w(L) + S\int_0^L \Big(\frac{Y}{1-\nu^2}\, u'\delta u' + D\, w''\delta w''\Big)\mathrm{d}x \\ &= F\delta w(L) + S\Big[\frac{Y}{1-\nu^2}\, u'\delta u + D\, w''\,\delta w'\Big]_0^L - S\int_0^L \Big(\frac{Y}{1-\nu^2}\, u''\,\delta u + Dw'''\,\delta w'\Big)\mathrm{d}x \\ &= F\delta w(L) + S\Big[\frac{Y}{1-\nu^2}\, u'\delta u + D\, w''\,\delta w' - D\, w'''\,\delta w\Big]_0^L - S\int_0^L \Big(\frac{Y}{1-\nu^2}\, u''\,\delta u - D\, w''''\,\delta w\Big)\mathrm{d}x\end{aligned} \tag{23}$$

where we've integrated by parts to get the second line, then done it again to get the third line. The boundary terms between square brackets automatically evaluate to zero at $x = 0$ due to our BCs there, so

$$\delta E/S = \Big(F\delta w/S + \frac{Y}{1-\nu^2}\, u'\delta u + D\, w''\,\delta w' - D\, w'''\,\delta w\Big)\Big|_{x=L} - \int_0^L \Big(\frac{Y}{1-\nu^2}\, u''\,\delta u - D\, w''''\,\delta w\Big)\mathrm{d}x.$$

We require that the above $\delta E = 0$ *for all* BC-satisfying δu and δw.

Let's first consider perturbations that have $\delta u(L) = \delta w(L) = \delta w'(L) = 0$, for which the boundary terms are zero, leaving just the integral. Then, since δu and δw in the integrand are arbitrary and independent, $\delta E = 0$ requires that their coefficients are zero everywhere, yielding the Euler-Lagrange equations

$$u'' = 0, \tag{24}$$

$$w'''' = 0. \tag{25}$$

Now that we know the integral vanishes all by itself, let's consider more general perturbations, having arbitrary boundary values $\delta u(L)$, $\delta w(L)$, and $\delta w'(L)$. Since these three quantities are independent from each other, $\delta E = 0$ requires that their coefficients are separately zero. Thus we read off

$$u'(L) = 0, \tag{26}$$

$$F/S - Dw'''(L) = 0, \tag{27}$$

$$w''(L) = 0. \tag{28}$$

These are the BCs for our Euler-Lagrange equations at $x = L$. BCs that emerge naturally in this way from a variational principle are inventively called 'natural' BCs.

It's straightforward (exercise!) to solve our Euler-Lagrange equations and impose our BCs to find

$$u = 0, \tag{29}$$

$$w = \frac{F\left(x^3 - 3Lx^2\right)}{6DS}. \tag{30}$$

The resulting shape is plotted in Fig. 9.

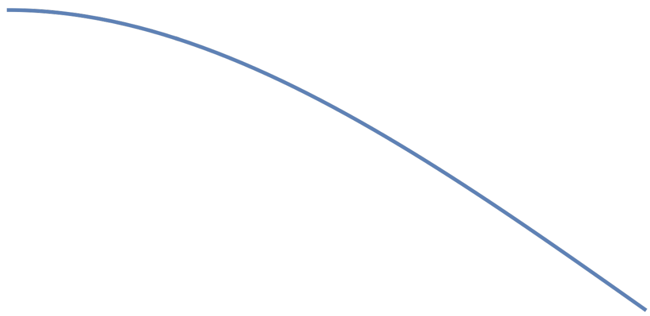

Figure 9: The normal displacement w (and thus the physical shape), of a cantilever plate rectangular plate subjected to load at the free edge. The vertical scale was chosen arbitrarily.

Ok, but what are the key takeaways? First, the bend energy contained second derivatives, which led to a fourth-order equation; almost all shell problems have this feature. Second, to leading order we only get normal displacement when we load in the normal direction — unsurprising but worth pointing out. Third, $w \propto 1/t^3$ because of the bending modulus D in the denominator, so even a small force generates a very large displacement! That t^3 is inherited from the bend energy, which is indeed cheap, as discussed earlier. The geometric point is that in this problem the load force can deform the plate *isometrically* i.e. without any stretch; the plate can only resist with weak bending forces, so it has to bend a lot to support the load!

Exercise: Remove the load force at the end of the cantilever. Instead, consider a cantilever with a significant weight, add an appropriate term to the energy, and minimise to find the displacements that occur as the cantilever deforms under its own weight.

1.2 A simply supported plate

Let's now solve a more complicated problem, without translational invariance. We'll again load a rectangular plate in the normal direction, but this time we'll apply some force $f(x, y)\mathrm{d}A$ to each area element. For the sake of variety, let's apply 'simply supported' boundary conditions on all four edges. This means that displacements all along the boundary are constrained to be zero ($u_\alpha = w = 0$), but rotations are free (the edges are 'hinged', if you like). You might think this latter freedom means ∇w is completely unconstrained on the boundary, but that can't be quite right, because if $w = 0$ along the boundary, then the along-the-boundary component of ∇w must also be zero; however the other component of ∇w is indeed unconstrained.

Picking up from (20), the total energy is

$$E = \int \left(\frac{Y}{2(1-\nu^2)} Q[(\nabla_\alpha u_\gamma + \nabla_\gamma u_\alpha)/2] + \frac{D}{2} Q[\nabla_\alpha \nabla_\gamma w] + f w \right) \mathrm{d}A, \tag{31}$$

and our task is to minimize E over all possible $\boldsymbol{u}$ and w that satisfy $u_\alpha = w = 0$ on the boundary. As a first step, note that differentiating $Q[\,]$ with respect to its argument yields

$$\frac{\partial Q[\tau]}{\partial \tau_{\alpha\gamma}} = 2\nu \,\mathrm{tr}(\tau)\delta_{\alpha\gamma} + 2(1-\nu)\tau_{\alpha\gamma} \equiv 2L_{\alpha\gamma}[\tau]. \tag{32}$$

Thus if we define the symmetric tensors

$$N = \frac{Y}{1-\nu^2} L[\varepsilon], \tag{33}$$

$$M = DL[\beta], \tag{34}$$

our energy variation becomes

$$\delta E = \int \Big(N_{\alpha\gamma} \nabla_\alpha \delta u_\gamma - M_{\alpha\gamma} \nabla_\alpha \nabla_\gamma \delta w + f \delta w \Big) \mathrm{d}A. \tag{35}$$

As before, we integrate the δu_γ term by parts once, and the δw term twice, where 'by parts' means we apply the following identity (which follows quickly from the 2D divergence theorem):

$$\int \boldsymbol{p} \cdot \nabla q \,\mathrm{d}A = \oint q\, \boldsymbol{p} \cdot \hat{\boldsymbol{n}} \,\mathrm{d}s - \int q \,\nabla \cdot \boldsymbol{p} \,\mathrm{d}A, \tag{36}$$

where s is arc length along the boundary, and $\hat{\boldsymbol{n}}$ is the unit vector that's normal to the boundary curve but tangential to the shell, and points outward away from the shell. We get (exercise)

$$\delta E = \oint \Big(N_{\alpha\gamma}\delta u_\gamma - M_{\alpha\gamma}\nabla_\gamma \delta w + \delta w \nabla_\gamma M_{\alpha\gamma}\Big)\hat{n}_\alpha \,\mathrm{d}s + \int \Big(-\delta u_\gamma \,\nabla_\alpha N_{\alpha\gamma} - \delta w\,\nabla_\alpha \nabla_\gamma M_{\alpha\gamma} + f\delta w\Big)\mathrm{d}A.$$

Now, looking at the boundary term, we see that $N_{\alpha\gamma}\hat{n}_\alpha\,\mathrm{d}s$ has the interpretation of a force on the boundary element $\mathrm{d}s$, because it dots with a displacement to give an energy change. For this reason, $N_{\alpha\gamma}$ is a kind of stress tensor; it's called the 'membrane stress' because it relates to stretch but not bend. For an analogous reason, $M_{\alpha\gamma}$ should probably be called the 'torque tensor' or 'moment tensor', but sometimes it's called the 'stress-couple' tensor.

Anyway, making exactly the same kind of argument as in the previous section, we can read off the Euler-Lagrange equations from the area integral:

$$\nabla_\alpha N_{\alpha\gamma} = 0, \tag{37}$$

$$\nabla_\alpha \nabla_\gamma M_{\alpha\gamma} - f = 0. \tag{38}$$

Then our $\delta u_\alpha = \delta w = 0$ BCs ensure that all boundary terms are zero except that involving the component of $\nabla \delta w$ perpendicular to the boundary; that component is unconstrained, so its coefficient must be zero, yielding the natural BC $\hat{n}_\alpha M_{\alpha\gamma}\hat{n}_\gamma = 0$, often written $M_{nn} = 0$. Physically, this BC means no external bending torque acts perpendicular to the boundary.

(Note that (37) is actually two scalar equations. It's often written as $\nabla \cdot N = 0$; it comes up constantly in both 3D and shell elasticity. It's often solved by introducing an 'Airy stress function' ψ that's strongly analogous to the electric potential in electrostatics: If the domain is simply connected (topology!) then (37) implies that $N_{\alpha\gamma} = \delta_{\alpha\gamma}\nabla^2\psi - \nabla_\alpha\nabla_\gamma\psi$ (it's easy to check that this form solves the equation — exercise!). Then the problem is simplified, because ψ is just a scalar field, which is easier to solve for than the vector $\boldsymbol{u}$. However, ψ isn't completely arbitrary: there's a 'compatibility' condition on it, encoding the requirement that you could back out a genuine $\boldsymbol{u}$ from it, i.e. a $\boldsymbol{u}$ whose mixed derivatives are symmetric. If this sounds a lot like the Gauss-Codazzi equations discussed earlier: yes, it's essentially the same thing.)

Anyway, we should now substitute in the definitions of N and M, and then ε and β, to get equations we can solve for the displacements. I'm not going to bother with the stress one, because it's clearly linear in derivatives of $\boldsymbol{u}$ so, (unsurprisingly) its solution in this problem is just $\boldsymbol{u} = \mathbf{0}$, due to the homogeneous BCs. However, (38) becomes

$$D\nabla^2\nabla^2 w + f = 0 \tag{39}$$

The 'biharmonic' operator $\nabla^2\nabla^2 w$ is sometimes written as $\nabla^4 w$ (or $\Delta^2 w$ if Δ is used for the Laplacian). (39) is the core equation of plate elasticity. It's not a million miles away from the Poisson equation, and similar solution techniques are possible.

We also have our BC $w = 0$ still, which implies that all along-boundary derivatives of w are zero. Bearing that in mind, our natural BC $M_{nn} = 0$ becomes a $\nabla^2 w = 0$ BC.

One way to solve our problem is to Fourier-expand w and f:

$$w(x, y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} w_{mn} \sin\left(\frac{m\pi x}{A}\right) \sin\left(\frac{n\pi y}{B}\right), \tag{40}$$

$$f(x, y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} f_{mn} \sin\left(\frac{m\pi x}{A}\right) \sin\left(\frac{n\pi y}{B}\right), \tag{41}$$

where I've assumed our plate has length A along x and B along y. You can easily check that such a w satisfies all of our BCs. Plugging into (39), differentiating termwise, and using orthogonality of the sine basis functions, one quickly finds

$$w_{mn} = \frac{f_{mn}}{\pi^4 D}\left(\frac{m^2}{A^2} + \frac{n^2}{B^2}\right)^{-2}. \tag{42}$$

As a result, we again find $w \propto 1/t^3$, coming from the fact that the plate bends but does not stretch (to leading order).

If, just for example, we have a point load at the center of the plate $f(x, y) =$

$F\delta(x - A/2)\delta(y - B/2)$, then one finds in the usual Fourier-series way that

$$f_{mn} = \frac{4F}{AB}\sin(m\pi/2)\sin(n\pi/2), \tag{43}$$

which can be plugged straight into (42). Then (40) gives the solution plotted in Fig. 10.

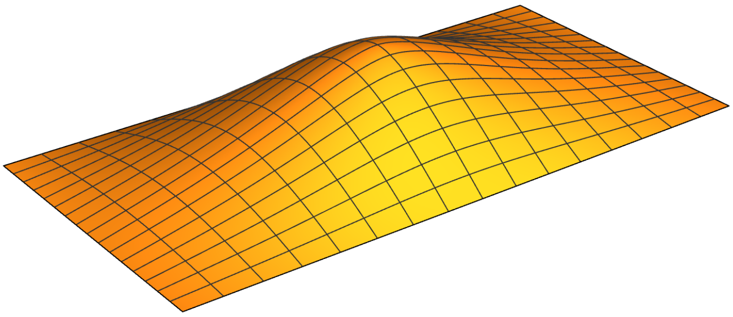

Figure 10: The normal displacement w (and thus the physical shape), of a simply supported rectangular plate subjected to a point load. The vertical scale was chosen arbitrarily.

Exercise: The rotationally symmetric Green function f for the 2D Laplacian satisfies $\nabla^2 f = \delta^{(2)}(\boldsymbol{r})$, and has the general form $f = \log(kr)/(2\pi)$, where the constant k is arbitrary. The rotationally symmetric Green function g for the 2D biharmonic satisfies $\nabla^2\nabla^2 g = \delta^{(2)}(\boldsymbol{r})$. Find the general form of g. Hint: The 2D Laplacian acts as $\nabla^2 g = (1/r)\partial_r(r\partial_r g)$.

1.3 Axisymmetrically deforming a cylinder

In the cantilever plate example, we applied a load at a free edge of the shell, in the direction normal to the shell. Let's now do the same, except with an initially-curved shell, rather than an initially-flat (plate) one. Specifically let's take the undeformed shell to be a cylinder, and let's apply a radial force $f\,\mathrm{d}s$ to each little element of one boundary; see Fig. 11. (The main reason I chose to use a cylinder for this example is that we can still use Cartesian coordinates (x, y). This means that all our derivatives can be partial derivatives, and we don't need to get into more advanced topics such as covariant differentiation.)

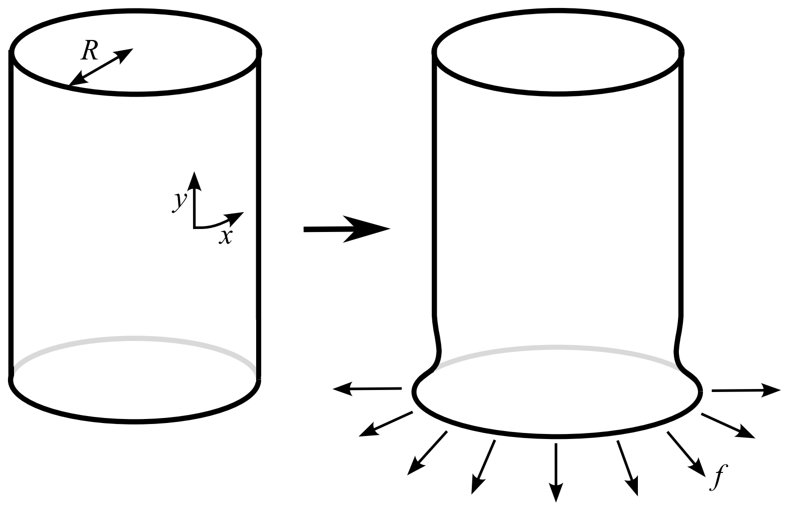

Figure 11: A cylinder (left) is deformed under a radial load f applied at a free boundary. Surprisingly, the resulting radial displacement is inwards in some regions and outwards in others!

The reference state will be a cylinder of radius R, inscribed inscribed with Cartesian coordinates $x^1 = x$ running azimuthally, and $x^2 = y$ running axially. Thus the azimuthal and axial displacements are $u_1 \equiv u$ and $u_2 \equiv v$ respectively, while the normal (radial) displacement is w again. The lower boundary of the cylinder will be at $y = 0$. As before, $\bar{a}$ is just the identity matrix, but now

$$\bar{\kappa} = \frac{1}{R} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{44}$$

This introduces a key new term in our linear-elastic strain expression, which becomes

$$\varepsilon_{\alpha\gamma} = \frac{1}{2} \left(\nabla_\alpha u_\gamma + \nabla_\gamma u_\alpha \right) + w \, \bar{\kappa}_{\alpha\gamma}. \tag{45}$$

That final term, which was zero for a plate, is the *key* difference between a plate and a curved shell: in a curved shell, normal displacements induce strain at leading order. This is pretty intuitive, if you think about a circle on the reference cylinder increasing its radius by an amount ΔR; the length of the circle has to increase by $2\pi\Delta R$, corresponding to an azimuthal ('hoop') strain of $\Delta R/R$.

What about our expression for β, is that also different now? In truth, yes, but we're going to keep the same expression (19), because the other terms would turn out to be negligible in our problem.

What about forcing? Well, we're just going to load radially at the free boundary $y = 0$, so the only load term in the energy is

$$- \oint_{y=0} f w \, \mathrm{d}s. \tag{46}$$

We're ready to crank the calculus-of-variations handle (exercise!). It's similar to the last time we did this. One extra term appears in the Euler-Lagrange equations, which become

$$\nabla_\alpha N_{\alpha\gamma} = 0, \tag{47}$$

$$-\nabla_\alpha \nabla_\gamma M_{\alpha\gamma} + N_{\alpha\gamma} \bar{\kappa}_{\alpha\gamma} = 0. \tag{48}$$

Reading off the BCs at the free boundary is a bit subtle, because the along-boundary component of ∇w is not independent of w on the boundary. We mentioned this subtlety earlier, but we have to do more to get the right answer this time, because the boundary displacements are unconstrained: We decompose $\nabla \delta w$ into its components, $\nabla \delta w = \hat{\boldsymbol{x}} \nabla_x \delta w + \hat{\boldsymbol{y}} \nabla_y \delta w$. The former term is involved in a 1D integration by parts that we must perform along the closed boundary curve:

$$\oint M_{yx} \nabla_x \delta w \, \mathrm{d}s = \oint M_{yx} \frac{\mathrm{d}\delta w}{\mathrm{d}s} \, \mathrm{d}s = \oint -\delta w \frac{\mathrm{d}M_{yx}}{\mathrm{d}s} \, \mathrm{d}s = \oint -\delta w \nabla_x M_{yx} \, \mathrm{d}s, \tag{49}$$

where M_{yx} is just a more readable way of writing M_{21}; accordingly there is no implicit Einstein summation happening anywhere in (49). After that the free-boundary BCs can all be read off straightforwardly:

$$N_{yx} = N_{yy} = 0, \tag{50}$$

$$M_{yy} = 0, \tag{51}$$

$$2\nabla_x M_{xy} + \nabla_y M_{yy} + f = 0. \tag{52}$$

At the other boundary we can have clamped BCs as before, but they aren't going to matter.

Let's now assume that the load force and the deformation are perfectly axisymmetric. Then all fields are functions of y only; we'll use primes to denote derivatives with respect to y. Furthermore, $u = 0$ by clockwise $\leftrightarrow$ anticlockwise symmetry. As a result, N and M are diagonal. With these simplifications, $\mathrm{div} N = 0$ becomes $N'_{yy} = 0$, so then (50) tells us $N_{yy} = 0$ everywhere! (This result is quite special to the free-boundary case. The other key features of our results are more general.) Thus, subbing the ε definition into N_{yy}, we find $v' = -\nu w/R$. With this relation, and the definition of M, (48) becomes

$$DR^2 w'''' + Yw = 0. \tag{53}$$

Compare to the cantilever equation $w'''' = 0$! Note: An easier way to get to this point would have been to impose translational invariance in the energy functional before minimising, as we did for the cantilever.

Now, (53) is linear and has constant coefficents, so is solved by exponentials. Subbing an exponential $\exp(iky)$ for w, we find that the possible wavevectors satisfy $k^4 = -Y/(DR^2)$. We thus see that our solution has a characteristic length scale $l = (DR^2/Y)^{1/4} \propto \sqrt{Rt}$; a key feature!

The possible k values sit at the corners of a square in the complex plane: $k = (\pm 1 \pm i)/\sqrt{2l^2}$. Thus we are left with

$$w = c_1\, e^{\frac{iy(1+i)}{\sqrt{2}\,l}} + c_2\, e^{\frac{iy(-1+i)}{\sqrt{2}\,l}} + c_3\, e^{\frac{iy(1-i)}{\sqrt{2}\,l}} + c_4\, e^{\frac{iy(-1-i)}{\sqrt{2}\,l}}, \tag{54}$$

for some arbitrary constants c_1, c_2, c_3, c_4. The final two terms grow into the bulk of the cylinder; we discard those, because we assume the non-loaded boundary is a distance $\gg l$ from the loaded boundary, and then any sensible BC applied at the non-loaded boundary would effectively set $c_3 = c_4 = 0$. Our two remaining free-end BCs (51) and (52) are just $w'' = 0$ and $w''' = f$. Imposing these yields

$$w = \frac{\sqrt{2} f l^3}{D} \exp\left(-\frac{y}{\sqrt{2}\,l}\right) \cos\left(\frac{y}{\sqrt{2}\,l}\right), \tag{55}$$

which is plotted in Fig. 12. Note: the solution doesn't just decay, it also *oscillates*! Some kind of 'flaring' occuring near the loaded end is very intuitive, but that the deformation is oscillatory rather than monotonic is pretty counter-intuitive I think. It means that even if f acts radially outwards, so the boundary circle moves outwards and goes into tension, nearby bits of shell move radially inwards and thereby go into compression!

I want to highlight how different the shell case is from the plate cases: Here the solution exhibits a characteristic short[6] length scale $l \propto \sqrt{t}$, decays away from the boundary (AKA has 'boundary-layer' character), is oscillatory, and has $w \propto t^{-3/2}$

[6]That the length scale $l \ll R$ is what self-consistently justifies our use of the Hessian form of β, (19): speaking loosely, a short length scale means gradients are large.

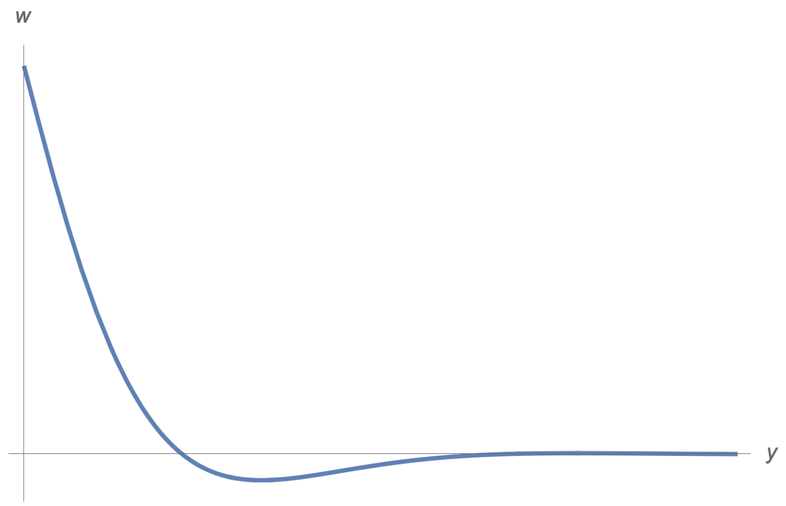

Figure 12: The normal displacement $w(y)$ of a cylinder loaded radially at a free boundary. The scales were chosen arbitrarily.

rather than t^{-3}. These differences are all consequences of that single extra term in the strain (45), which a shell has and a plate doesn't!

Note also that since t is small, the $w \propto t^{-3/2}$ scaling means this shell resists the load far more effectively than a cantilever would! This is exactly the curvature-induced rigidity we mentioned before in the context of pizza: the curved shell is much more rigid than a flat plate, because to deform it the load has to stretch it. Indeed, stretch and bend are competing ('trading off') in the boundary layer, which is the key to the thickness scalings that emerged. In contrast, the plates deformed isometrically via pure bend.[7]

Another, more subtle difference is that applying a load to a curved shell in the normal direction induces both normal and tangential displacements to leading (linear) order. We saw this earlier: v' was not zero!

All these differences make shells much more subtle and difficult to study than plates. On the flip side, it also makes them more interesting! In any case, their rigidity makes them more practically useful, so they must be studied whether we like it or not!

[7]Curvature-induced rigidity would eventually kick in *to some extent* in a plate problem like that of Sec. 1.2, for large enough deformations; we didn't see this because we were doing linear elasticity. For a plate, curvature-induced rigidity is a second-order (and hence rather weak) effect; Föppl–von Kármán plate theory captures it.

1.4 Cylinder buckling under axial compression

We'll cover a lot of ground quite fast in this section, with some glossing; if it feels that way, it's because it is that way!

For a moment let's consider an energy E that's a function of only two degrees of freedom (physical quantities), p and q, rather than the infinitely many degrees of freedom that a displacement *field* has. In an energy landscape $E(p,q)$, equilibria correspond to stationary points. *Stable* equilibria correspond to *minima* of E. Unstable equilibria correspond to maxima or saddle points.

Consider such an energy landscape that varies over time, (e.g. due to a load force increasing). Suppose a physical system lies in a minimum (stable equilibrium). It will stay trapped in that 'bowl' if the minima just moves around or changes shape (Fig. 13a-b). However, a mechanical instability occurs if the minimum ceases to be a minimum, and instead becomes a maximum or a saddle point; then the system will immediately 'roll downhill' to lower its energy (Fig. 13c-d).

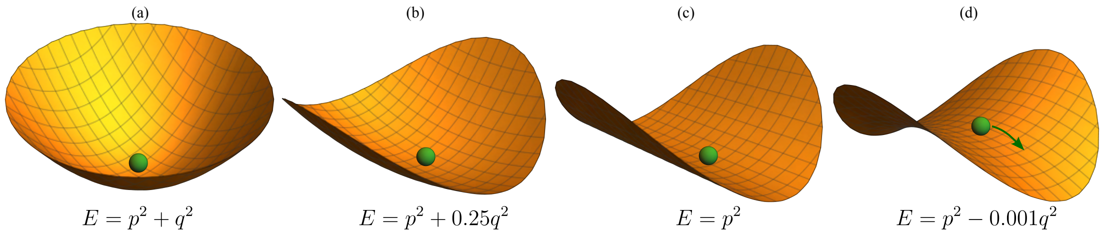

Figure 13: Example of energy landscape evolution. (a) An energy $E(p,q)$ initially has a minimum (stable equilibrium) $p = q = 0$. (b) The physical system (green ball) stays in the minimum, even as the energy landscape changes. Eventually the energy landscape reaches a critical/threshold case (c), just between stability and instability. Pushing even infinitesimally beyond this case, the landscape looks like (d) after zooming in, so the $p = q = 0$ equilibrium has become unstable, and the physical system will move to lower its energy. (Pictures not to scale.)

Let's get slightly more mathematical: At any equilibrium, $\nabla E = 0$, so perturbations δp and δq about that equilibrium cause energy changes only of quadratic-or-higher order in δp and δq. At a *stable* equilibrium, any such perturbations specifically cause energy *increases*. An instability is when the energy landscape changes such that (in at least one direction) perturbations can *decrease* the energy. The moment of crossover in behaviour — the threshold case, if you like — is the moment where perturbations can just manage to cause *zero* change in energy to quadratic order (Fig.13c). We say the instability occurs at this threshold.

This recipe for finding instabilities extends perfectly to the infinite-dimensional case, where the perturbations are deformation fields: If you have a family of equilibria ('base states'), perturb them, and write down the resulting quadratic-in-perturbations δE. Ask "do all nonzero perturbations make $\delta E > 0$?" When the answer changes from "yes" to "no, some make $\delta E = 0$", you have an instability.[8]

Now, here's something a little subtle: If you stare for a while at Fig. 13, or think for a while about quadratic forms, you'll realise that at the instability the quadratic δE must be *stationary* at the *perturbed* states that have $\delta E = 0$.[9] This means that at instability there's a whole 'direction' in perturbation space that you can move along at zero energy (the 'valley' in Fig. 13c)! That is in fact a more helpful recipe for finding an instability: *Find the first moment that the quadratic δE is stationary at a non-zero perturbation, i.e. its (functional) derivative is zero.* This way, in the infinite-dimensional case, finding the instability involves solving the Euler-Lagrange equations associated with the quadratic δE functional. Those equations are linear, so this is a 'linear stability analysis'.

So, let's find the instability that occurs when you squash a cylinder along its axis with a force f. In a careful experiment, the result might look something like Fig. 14a-b, but you've probably seen pretty similar things if you've ever squashed an aluminium can. Fig. 14c shows a sketch of the setup; the x coordinate runs azimuthally, while y runs axially. We'll assume the pre-instability base state to be a very simple and intuitive one: The stress is uniform and purely axial, and there are no bending forces.[10] This is always true away from the shell boundaries, and (I think) it's true everywhere when squashing between frictionless plates. Vertical force balance immediately gives the base-state membrane stress

$$N^0 = \frac{f}{2\pi R}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{56}$$

Furthermore, the actual *shape* of the cylinder is almost unchanged until instability

[8]Actually, in some cases there will be perturbations that *always* leave the energy unchanged to quadratic order (usually due to symmetries; see 'Goldstone modes'). Since, for these perturbations, the energy behaviour doesn't switch character as you vary (e.g.) the load on the system, their presence doesn't indicate an interesting instability.

[9]This is essentially the same as the fact that, if we have a positive semi-definite matrix M, the quadratic form $\boldsymbol{v}^{\mathrm{T}} M \boldsymbol{v}$ can only equal zero if $\boldsymbol{v} = \boldsymbol{0}$ or if $\boldsymbol{v}$ is an eigenvector of M with eigenvalue zero. And those possibilities are exactly the vectors that make the quadratic form stationary!

[10]A 'membrane' state is the name given to a state where bending is neglected.

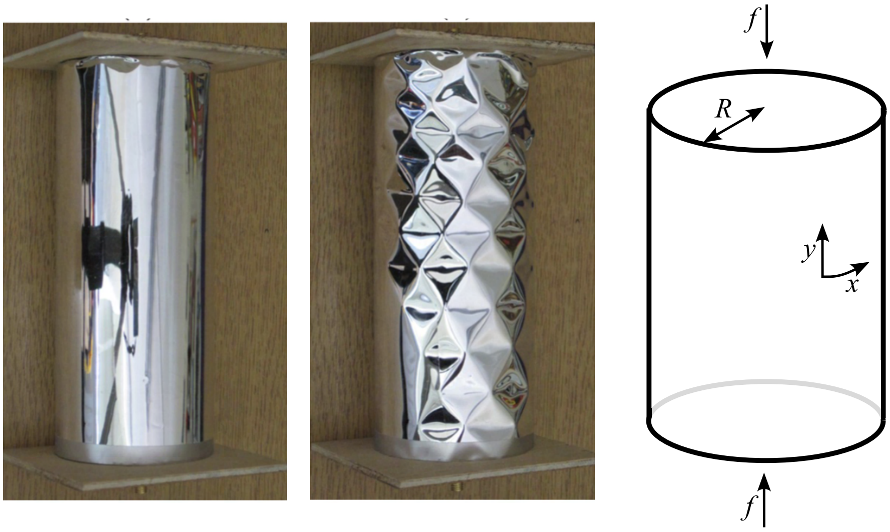

Figure 14: Left: A cylindrical shell, about to be squashed vertically (axially). Middle: The load force increases until a sudden instability occurs and the shell jumps to another shape. Usually the shape on the right would be seen only very briefly, en route to some less regular shape; here the nice pattern has been preserved with a trick (can you guess what's inside the cylinder to achieve this?). This 'Yoshimura' pattern also comes up in origami theory! Photos from Seffen+Stott, 2014. Right: Sketch of the setup.

occurs, so we can take the base-state curvature tensor to be

$$\kappa^0 = \bar{\kappa} = \frac{1}{R}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{57}$$

To probe the stability of this base state, we perturb with deformations $\delta \boldsymbol{u}(x, y)$ and $\delta w(x, y)$ in the tangential and normal directions respectively. What are the resulting changes in strain ε and bend β? Well, I'm going to ask you to take my word for it that the only *important* terms for this problem are the following:

$$\delta\varepsilon_{\alpha\gamma} = \overbrace{\left(\nabla_\alpha \delta u_\gamma + \nabla_\gamma \delta u_\alpha\right)/2 + \kappa^0_{\alpha\gamma}\,\delta w}^{\delta\varepsilon^1} + \overbrace{(\nabla_\alpha \delta w)(\nabla_\gamma \delta w)/2}^{\delta\varepsilon^2}, \tag{58}$$

$$\delta\beta_{\alpha\gamma} = -\nabla_\alpha \nabla_\gamma \delta w. \tag{59}$$

(The above expressions are just found by throwing away a bunch of unimportant terms from the exact (messy) expressions, which you can peruse in the Appendix.) The bend is just second derivatives of normal displacement, which is familiar for shells and plates. The strain has a linear in-plane contribution that's familiar from non-shell elasticity, and also a linear-in-normal-displacement term that's always present for curved shells; we saw it in Sec. 1.3. The final (nonlinear!) term in the strain is new, and encodes the *key* mechanism of shell buckling: *normal displacements can relieve tangential compression.* I reserve the word 'buckling' for instabilities driven by this mechanism.

This effect is quite intuitive: once the loading plates start squashing, the shell is a little too long to comfortably fit between them (hence compression), but it can 'waste' some of this 'excess length' with oscillatory normal deformations, allowing it to relax somewhat. However, other terms in the energy penalise such deformations, so there is competition, and the perturbations are only favoured when the compressive force f gets large enough; the moment this occurs is our buckling instability.

Now, we need to expand our energy to quadratic order in perturbations. This is just a matter of plugging $\varepsilon = \varepsilon^0 + \delta\varepsilon$ and $\beta = \beta^0 + \delta\beta$ into (16), using (32), (58), and (59), and throwing away higher-than-quadratic terms. Also, $YL(\varepsilon^0)/(1-\nu^2) = N^0$, the base-state membrane stress. Thus (exercise!), we arrive at

$$\delta E = \int \left[\frac{Y}{2(1-\nu^2)} Q(\delta\varepsilon^1) + \frac{D}{2} Q(\delta\beta) + \mathrm{tr}(N^0\, \delta\varepsilon^2) \right] \mathrm{d}A. \tag{60}$$

Now, remember that we're looking for the first moment (as f increases) that the functional derivative of δE can equal zero at nonzero $\delta\boldsymbol{u}$, δw. So we should solve the Euler-Lagrange equations! To find them, we perturb $\delta\boldsymbol{u}$ and δw (yes, perturbations of perturbations, e.g. $\delta\delta w$!) and do standard calculus-of-variations handle turning (exercise), yielding

$$\nabla_\alpha (\delta N^1)_{\alpha\gamma} = 0, \tag{61}$$

$$D\nabla^2\nabla^2 \delta w + \mathrm{tr}\left(\kappa^0\, \delta N^1\right) - \nabla\cdot\left(N^0\, \nabla\delta w\right) = 0, \tag{62}$$

where $\delta N^1 \equiv \frac{Y}{1-\nu^2} L(\delta\varepsilon^1)$.

Now we employ a common trick that's a *lot* like using a scalar potential in electromagnetism: Introduce a scalar 'Airy stress function' $\delta\psi$, and demand that $\delta N^1 = \Lambda\delta\psi$, where the differential operator $\Lambda_{\alpha\gamma} \equiv \delta_{\alpha\gamma}\nabla^2 - \nabla_\alpha\nabla_\gamma$. Doing so, (61) is immediately satisfied, because of symmetry of mixed partial derivatives. However, another equation emerges: because we're not solving for $\delta\boldsymbol{u}$ directly, we need to make sure that our $\delta\psi$ will yield a δN^1 that will yield a $\delta\varepsilon^1$ that will correspond to a genuine displacement $\delta\boldsymbol{u}$. In practice this means requiring that mixed partial derivatives of $\delta\boldsymbol{u}$ are symmetric, and after some tedium[11] that boils

[11] First, invert the stress-strain relation to get $Y\delta\varepsilon^1 = (1+\nu)\delta N^1 - \nu\mathrm{tr}(\delta N^1)I$. (To do so slickly, take the trace of the expression δN^1 to find $\mathrm{tr}\,\delta\varepsilon^1 = (1-\nu)\,\mathrm{tr}\,\delta N^1/Y$, and sub that back

down to

$$\nabla^2\nabla^2\delta\psi - Y\,\Lambda_{\alpha\gamma}(\kappa^0_{\alpha\gamma}\,\delta w) = 0. \tag{63}$$

The above requirement is sometimes called 'geometric compatibility'. Really it's just an incremental version of the Gauss compatibility equation discussed earlier for surfaces in general.

Now we have only two scalar equations to solve, (62) and (63). They're linear and have constant coefficients, so the solutions ('buckling modes') are just exponentials:

$$\begin{pmatrix}\delta w\\ \delta\psi\end{pmatrix} = \begin{pmatrix}A\\ B\end{pmatrix} e^{i(k_x x + k_y y)}. \tag{64}$$

Plugging this into (62) and (63), along with N^0 and κ^0, we get algebraic equations to solve for the ratio A/B and the force f. Interestingly, f only depends on k_x and k_y via the combination $k_\circ^2 \equiv (k_x^2 + k_y^2)^2/k_y^2$:

$$f = 2\pi\,\frac{Y + Dk_\circ^4 R^2}{k_\circ^2 R}. \tag{65}$$

So when does the instability occur? At the smallest force f that allows these solutions to exist; so we minimise the f expression over $k_\circ$, at long last finding the classic buckling threshold

$$f_\star = 4\pi\sqrt{YD}. \tag{66}$$

with the corresponding $k_{\circ\star}^4 = Y/(DR^2)$.

Some things to note: (1) The critical $k_{\circ\star}$ corresponds to infinitely many pairs (k_x, k_y); in fact those pairs form a 'Koiter circle' in $\boldsymbol{k}$-space, and the lovely diamond buckling pattern in Fig. 13 arises from several of these modes going unstable simultaneously. (2) The critical force $f_\star$ doesn't depend on R at all. (3) Instead, $f_\star \propto t^2$, and the wavevector $k_{\circ\star}$ implies a characteristic buckling wavelength $\sim \sqrt{Rt}$ — a length scale that also emerged in Sec. 1.3! These thickness scalings are characteristic of curved-shell buckling, and arise from the tradeoff between stretch and bend that is necessarily involved. This tradeoff is ubiquitous in shell mechanics, and is part of what makes it a rich and interesting subject :)

---

in to the δN^1 expression, then rearrange for $\delta\varepsilon^1$.) Now, Λ kills the $\boldsymbol{u}$ part of $\delta\varepsilon^1$, so applying Λ to $\delta\varepsilon^1$ and tracing gives one of the terms in (63) immediately. The other follows quickly from applying Λ to the inverted stress-strain relation we found, given $\nabla\cdot\delta N^1 = 0$. It is *not* obvious that (63) is a *sufficient* condition for compatibility.

I should come clean about glossing some things:

1. A tacit assumption we made was that our buckling modes had to be oscillatory, rather than decaying (i.e. we looked at real k_x, k_y only). Was that ok to assume? Not really, no. It is true that on a cylinder we will need to have a periodic boundary condition in the azimuthal (x) direction, which requires a real (and quantized) k_x, but this does not constrain k_y to be real. And actually, for a radially free boundary, buckling indeed occurs at an $f_\star$ that's much smaller than (66), with a buckling mode that oscillates but also decays exponentially away from the boundary; see Nachbar and Hoff's *The buckling of a free edge of an axially-compressed circular cylindrical shell* (1962).[12] However, for more conventional boundary conditions (e.g. clamped), that does not happen; E.g. see the cylinder-buckling chapters in Calladine's *Theory of Shell Structures*, and look again at Fig. 14. So one has to get stuck in the weeds with boundary conditions, which we avoided completely! What we found is a 'bulk' instability, by which I mean that the modes have a comparable amplitude over most of the shell, as opposed to being exponentially suppressed in some regions relative to others; in some setups this instability will occur first, and therefore be observed; in other setups other instabilities will occur first instead.

2. Real shells almost always buckle at a significantly lower force than (66) predicts, due to some combination of boundary effects and/or an acute sensitivity to imperfections; this is still not fully understood, despite a lot of work. Getting physically realistic buckling thresholds out of shell theory can sometimes require more care than we've put in!

Exercise (Euler buckling): Find the buckling force $f_\star$ for a flat plate ($\bar{\kappa} = \kappa^0 = 0$), with a compressive force f applied at each end in the x direction, so $N^0 = \mathrm{diag}(-f/S, 0)$ where S is the length in the y direction. Assume translational invariance in the y direction. Do not use an Airy stress function. You can use (61) and (62) if you find the appropriate BCs, but an easier approach will be to go back to (60), impose translational invariance there, and then crank the calculus-of-variations handle. Consider first simply-supported BCs, then clamped BCs. Which has the higher $f_\star$, and does that make sense intuitively? Compare the thickness scalings of the $f_\star$ values and the wavelengths to the cylinder case.

[12] The modes in Fig. 2c of *Lifting, Loading, and Buckling in Conical Shells* (2023) are similar.

2 Appendix: Changes in metric and curvature tensors up to second order in displacements

Consider an undeformed surface $\mathcal{S}$, and a surface $\mathcal{S}'$ that is the result of deforming $\mathcal{S}$. The surface $\mathcal{S}$ has some metric tensor a and second fundamental form (curvature tensor) κ. We also inscribe it with some coordinate system x^α, which we should think of as labels that material elements carry with them throughout deformation. The metric a' and curvature tensor κ' of $\mathcal{S}'$ can then be expressed with respect to the coordinate basis corresponding to x^α, which also form a coordinate system for $\mathcal{S}'$. When raising and lowering indices (which is just a notational shorthand remember!), we will in this appendix always implicitly be using a to do so, and as usual the upstairs components $a^{\alpha\beta}$ are defined via the matrix inverse: $a^{\alpha\gamma}a_{\gamma\beta} = \delta^\alpha_\beta$.

If $\mathcal{S}$ is described by the 3D position vector $\boldsymbol{R}(x^\alpha)$, then

$$\boldsymbol{e}_\alpha \equiv \partial_\alpha \boldsymbol{R}, \tag{67}$$

$$\hat{\boldsymbol{N}} \equiv \frac{\boldsymbol{e}_1 \times \boldsymbol{e}_2}{|\boldsymbol{e}_1 \times \boldsymbol{e}_2|} \tag{68}$$

are respectively the coordinate basis vectors corresponding to the x^α coordinates and the unit normal, both on $\mathcal{S}$. We have $a_{\alpha\beta} = \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta$ and $\kappa_{\alpha\beta} = \boldsymbol{e}_\beta \cdot \partial_\alpha \hat{\boldsymbol{N}} = -\hat{\boldsymbol{N}} \cdot \partial_\alpha \boldsymbol{e}_\beta$. The latter, combined with the fact that $\partial_\alpha \hat{\boldsymbol{N}}$ is perpendicular to $\hat{\boldsymbol{N}}$ because $\hat{\boldsymbol{N}}$ is unit-length, quickly yields the Weingarten relation $\partial_\alpha \hat{\boldsymbol{N}} = \kappa^\beta{}_\alpha \boldsymbol{e}_\beta$.

Now, it's useful to know how a' and κ' relate to a and κ. Let's describe the deformation by supposing that each material point in $\mathcal{S}$ is displaced by $\boldsymbol{u}$ and w, respectively in the directions tangential and normal to $\mathcal{S}$ at that material point. Thus, if $\mathcal{S}$ is described by the 3D position vector $\boldsymbol{R}(x^\alpha)$, then $\mathcal{S}'$ is described by

$$\boldsymbol{R}' = \boldsymbol{R} + u^\beta \boldsymbol{e}_\beta + w\hat{\boldsymbol{N}} \tag{69}$$

From (69), we can now calculate the coordinate basis vectors corresponding to the x^α coordinates on $\mathcal{S}'$:

$$\boldsymbol{e}'_\alpha \equiv \partial_\alpha \boldsymbol{R}' = \boldsymbol{e}_\alpha + \partial_\alpha(u^\beta \boldsymbol{e}_\beta) + \partial_\alpha(w\hat{\boldsymbol{N}}). \tag{70}$$

At this point, let's recall that if ∇ denotes the standard covariant derivative on $\mathcal{S}$, then $\nabla_\alpha u^\gamma = \boldsymbol{e}^\gamma \cdot \partial_\alpha(u^\beta \boldsymbol{e}_\beta)$. (The relevant Christoffel symbols are those calculated from a in the x^α coordinates.) We've thus packaged the two tangential components of $\partial_\alpha(u^\beta \boldsymbol{e}_\beta)$ nicely, but the normal component remains, and we have

$$\boldsymbol{e}'_\alpha = \boldsymbol{e}_\alpha + \boldsymbol{e}_\gamma \nabla_\alpha u^\gamma + \hat{\boldsymbol{N}} \left(\hat{\boldsymbol{N}} \cdot \partial_\alpha(u^\beta \boldsymbol{e}_\beta) \right) + \partial_\alpha(w\hat{\boldsymbol{N}}). \tag{71}$$

We can improve this expression by using the definition of κ and the Weingarten relation. After defining (for convenience) the linear-in-displacements tensors

$$d^\beta{}_\alpha \equiv \nabla_\alpha u^\beta + w\kappa^\beta{}_\alpha, \tag{72}$$

$$\phi_\alpha \equiv \kappa_{\alpha\beta} u^\beta - \partial_\alpha w, \tag{73}$$

it becomes

$$\boldsymbol{e}'_\alpha = \boldsymbol{e}_\alpha + d^\beta{}_\alpha \boldsymbol{e}_\beta - \phi_\alpha \hat{\boldsymbol{N}}. \tag{74}$$

Thus

$$\boxed{a'_{\alpha\beta} = \boldsymbol{e}'_\alpha \cdot \boldsymbol{e}'_\beta = a_{\alpha\beta} + d_{\alpha\beta} + d_{\beta\alpha} + d_{\sigma\alpha} d^\sigma{}_\beta + \phi_\alpha \phi_\beta.} \tag{75}$$

Note that this expression is exactly quadratic in the displacements. If you want a linear-in-displacements approximation, you just discard the last two terms. The implicit $\propto w$ linear terms represent the key difference between shells and plates: in shells a small normal displacement induces strain at leading order. Of all the quadratic terms, often the only important one for buckling is $(\partial_\alpha w)(\partial_\beta w)$, which captures the essence of shell buckling: out of plane deformation 'uses up' length to relieve compressive stress.

Now let's calculate

$$\hat{\boldsymbol{N}}' \equiv \frac{\boldsymbol{e}'_1 \times \boldsymbol{e}'_2}{|\boldsymbol{e}'_1 \times \boldsymbol{e}'_2|} \tag{76}$$

in the same kind of way. To begin, note that at each point the pair of vectors $\boldsymbol{e}_1 \mathrm{d}x^1$ and $\boldsymbol{e}_2 \mathrm{d}x^2$ define the standard parallelogram area element: $\mathrm{d}A = |\boldsymbol{e}'_1 \times \boldsymbol{e}'_2|\, \mathrm{d}x^1 \mathrm{d}x^2$. But we also have the ubiquitous expression $\mathrm{d}A = \sqrt{\det a'_{..}}\, \mathrm{d}x^1 \mathrm{d}x^2$, where $a'_{..}$ is the matrix of downstairs components. Comparing yields

$$|\boldsymbol{e}'_1 \times \boldsymbol{e}'_2| = \sqrt{\det a'_{..}}\,, \tag{77}$$

which can be checked via brute-force tedium if desired. By the same argument

$|\boldsymbol{e}_1 \times \boldsymbol{e}_2| = \sqrt{\det a_{..}}$, so (68) yields

$$\boldsymbol{e}_\alpha \times \boldsymbol{e}_\beta = \hat{\boldsymbol{N}}\, \epsilon_{\alpha\beta} \sqrt{\det a_{..}}. \tag{78}$$

where $\epsilon_{12} = -\epsilon_{21} = 1$, $\epsilon_{11} = \epsilon_{22} = 0$. The above equation quickly yields $|\boldsymbol{e}^1 \times \boldsymbol{e}^2| = 1/\sqrt{\det a_{..}}$ after defining $\boldsymbol{e}^\alpha = a^{\alpha\beta}\boldsymbol{e}_\beta$ as usual, which in turn means we can write $\hat{\boldsymbol{N}} = \boldsymbol{e}^1 \times \boldsymbol{e}^2 \sqrt{\det a_{..}}$. This form in turn yields

$$\boldsymbol{e}_\alpha \times \hat{\boldsymbol{N}} = \epsilon_{\sigma\alpha}\boldsymbol{e}^\sigma \sqrt{\det a_{..}} \tag{79}$$

directly.

Combining (74), (78), and (79), we can now find

$$\begin{aligned}
\boldsymbol{e}'_1 \times \boldsymbol{e}'_2 &= (\boldsymbol{e}_1 + d^\tau{}_1 \boldsymbol{e}_\tau - \phi_1 \hat{\boldsymbol{N}}) \times (\boldsymbol{e}_2 + d^\sigma{}_2 \boldsymbol{e}_\sigma - \phi_2 \hat{\boldsymbol{N}}) && (80)\\
&= ((\delta^\tau_1 + d^\tau{}_1)\boldsymbol{e}_\tau - \phi_1 \hat{\boldsymbol{N}}) \times ((\delta^\sigma_2 + d^\sigma{}_2)\boldsymbol{e}_\sigma - \phi_2 \hat{\boldsymbol{N}}) && (81)\\
&= \sqrt{\det a_{..}} \Big((\delta^\tau_1 + d^\tau{}_1)(\delta^\sigma_2 + d^\sigma{}_2)\epsilon_{\tau\sigma}\hat{\boldsymbol{N}} - (\delta^\tau_1 + d^\tau{}_1)\phi_2 \epsilon_{\gamma\tau}\boldsymbol{e}^\gamma + (\delta^\sigma_2 + d^\sigma{}_2)\phi_1 \epsilon_{\gamma\sigma}\boldsymbol{e}^\gamma \Big)\\
&= \sqrt{\det a_{..}} \Big(\det(I + d^{\cdot}_{\cdot})\hat{\boldsymbol{N}} + (\phi_\gamma + \phi_\gamma d^\beta{}_\beta - \phi_\beta d^\beta{}_\gamma)\boldsymbol{e}^\gamma \Big), && (82)
\end{aligned}$$

where $d^{\cdot}_{\cdot}$ is the matrix of d's mixed components. Combining with (77), we at last have

$$\hat{\boldsymbol{N}}' \equiv \frac{\boldsymbol{e}'_1 \times \boldsymbol{e}'_2}{|\boldsymbol{e}'_1 \times \boldsymbol{e}'_2|} = \frac{\sqrt{\det a_{..}}}{\sqrt{\det a'_{..}}} \Big(\det(I + d^{\cdot}_{\cdot})\hat{\boldsymbol{N}} + (\phi_\gamma + \phi_\gamma d^\beta{}_\beta - \phi_\beta d^\beta{}_\gamma)\boldsymbol{e}^\gamma \Big). \tag{83}$$

Note that the prefactor is the reciprocal of the deformation's area scale factor.

To compute $\kappa'_{\alpha\beta}$, we need to take another derivative of (74). It will be convenient to temporarily use local Cartesian coordinates, in which $\partial_\beta \boldsymbol{e}_\alpha = -\kappa_{\alpha\beta}\hat{\boldsymbol{N}}$. Then, remembering $\partial_\alpha \hat{\boldsymbol{N}} = \kappa^\beta{}_\alpha \boldsymbol{e}_\beta$ (Weingarten), we find

$$\begin{aligned}
\partial_\beta \boldsymbol{e}'_\alpha &= \partial_\beta \Big(\boldsymbol{e}_\alpha + d^\gamma{}_\alpha \boldsymbol{e}_\gamma - \phi_\alpha \hat{\boldsymbol{N}} \Big) && (84)\\
&= \partial_\beta \Big((\delta^\gamma_\alpha + d^\gamma{}_\alpha)\boldsymbol{e}_\gamma - \phi_\alpha \hat{\boldsymbol{N}} \Big) && (85)\\
&= \boldsymbol{e}_\gamma \partial_\beta (\delta^\gamma_\alpha + d^\gamma{}_\alpha) - (\delta^\gamma_\alpha + d^\gamma{}_\alpha)\kappa_{\gamma\beta}\hat{\boldsymbol{N}} - \hat{\boldsymbol{N}}\partial_\beta \phi_\alpha - \phi_\alpha \kappa^\gamma{}_\beta \boldsymbol{e}_\gamma\,. && (86)
\end{aligned}$$

Dotting this with $-\hat{\boldsymbol{N}}'$ yields

$$\kappa'_{\alpha\beta} = -\hat{\boldsymbol{N}}'\partial_\beta \boldsymbol{e}'_\alpha \tag{87}$$

$$= \frac{\sqrt{\det a_{..}}}{\sqrt{\det a'_{..}}}\Bigg(\det(I + d^{\cdot}_{\cdot})[(\delta^\gamma_\alpha + d^\gamma{}_\alpha)\kappa_{\gamma\beta} + \partial_\beta \phi_\alpha] + (\phi_\gamma + \phi_\gamma d^\sigma{}_\sigma - \phi_\sigma d^\sigma{}_\gamma)[-\partial_\beta d^\gamma{}_\alpha + \phi_\alpha \kappa^\gamma{}_\beta]\Bigg), \tag{88}$$

where we've used that δ^γ_α has zero derivative. In our local Cartesian coordinates, the partial derivatives above are also covariant derivatives, so we can write the above as

$$\kappa'_{\alpha\beta} = -\hat{\boldsymbol{N}}'\partial_\beta \boldsymbol{e}'_\alpha \tag{89}$$

$$= \frac{\sqrt{\det a_{..}}}{\sqrt{\det a'_{..}}}\Bigg(\det(I + d^{\cdot}_{\cdot})(\kappa_{\alpha\beta} + d^\gamma{}_\alpha \kappa_{\gamma\beta} + \nabla_\beta \phi_\alpha) + (\phi_\gamma + \phi_\gamma d^\sigma{}_\sigma - \phi_\sigma d^\sigma{}_\gamma)(\phi_\alpha \kappa^\gamma{}_\beta - \nabla_\beta d^\gamma{}_\alpha)\Bigg). \tag{90}$$

Now the above expression is fully tensorial, so since it holds in one (local Cartesian) coordinate system, it holds in any coordinate system! Remember that there are displacements also hidden in $\sqrt{\det a'_{..}}$, so κ' is not just cubic in displacements. In fact, it's probably best to use (75) to write the prefactor in (90) as

$$\frac{\sqrt{\det a_{..}}}{\sqrt{\det a'_{..}}} = \det\left(I + d^{\cdot}_{\cdot} + (d^{\cdot}_{\cdot})^{\mathrm{T}} + (d^{\cdot}_{\cdot})^{\mathrm{T}} d^{\cdot}_{\cdot} + \phi^{\cdot} \otimes \phi_{\cdot}\right)^{-1/2}. \tag{91}$$

Let's now expand κ' up to quadratic order in displacements. First, the above prefactor: The matrices $d^{\cdot}_{\cdot}$ and $(d^{\cdot}_{\cdot})^{\mathrm{T}}$ have the same trace and determinant. Moreover, for 2×2 matrices, $\det(I + M) = 1 + \operatorname{tr} M + \det M = 1 + \operatorname{tr} M + \left(\operatorname{tr}(M)^2 - \operatorname{tr}(M^2)\right)/2$. Using these identities, we get $\det(I + d^{\cdot}_{\cdot}) = 1 + \operatorname{tr} d^{\cdot}_{\cdot} + \det d^{\cdot}_{\cdot}$ and, from (91),

$$\frac{\sqrt{\det a_{..}}}{\sqrt{\det a'_{..}}} \approx \left(1 + 2\operatorname{tr} d^{\cdot}_{\cdot} + (\operatorname{tr} d^{\cdot}_{\cdot})^2 + 2\det d^{\cdot}_{\cdot} + |\boldsymbol{\phi}|^2\right)^{-1/2} \tag{92}$$

$$\approx 1 - \operatorname{tr} d^{\cdot}_{\cdot} + (\operatorname{tr} d^{\cdot}_{\cdot})^2 - \det d^{\cdot}_{\cdot} - |\boldsymbol{\phi}|^2/2 \tag{93}$$

to quadratic order in displacements. Thus

$$\kappa'_{\alpha\beta} \approx \kappa_{\alpha\beta} + d^{\gamma}{}_{\alpha}\kappa_{\gamma\beta} + \nabla_\beta \phi_\alpha - \frac{|\boldsymbol{\phi}|^2}{2}\kappa_{\alpha\beta} + \phi_\gamma(\phi_\alpha \kappa^{\gamma}{}_{\beta} - \nabla_\beta d^{\gamma}{}_{\alpha}) \tag{94}$$

to quadratic order in displacements.

We can now approximate the curvature tensor of S' similarly. In general for a matrix M, we have the expansion $(I+M)^{-1} = I - M + M^2 - M^3 + \ldots$, which leads to $(A+M)^{-1} = A^{-1} - A^{-1}MA^{-1} + A^{-1}MA^{-1}MA^{-1} - \ldots$, which we can apply to (75) to find

$$(a'^{-1})^{\alpha\beta} \approx a^{\alpha\beta} - d^{\alpha\beta} - d^{\beta\alpha} - \phi^\alpha\phi^\beta + d^{\alpha}{}_{\sigma}d^{\sigma\beta} + d^{\alpha}{}_{\sigma}d^{\beta\sigma} + d^{\beta}{}_{\sigma}d^{\sigma\alpha} \tag{95}$$

to quadratic order in displacements. Multiplying onto (94) yields the approximate primed shape operator

$$\boxed{(a'^{-1})^{\alpha\tau}\kappa'_{\tau\beta} \approx \kappa^{\alpha}{}_{\beta} + \nabla_\beta\phi^\alpha - \frac{|\boldsymbol{\phi}|^2}{2}\kappa^{\alpha}{}_{\beta} - \phi_\gamma\nabla_\beta d^{\gamma\alpha} - d^{\alpha\gamma}\kappa_{\gamma\beta} - (d^{\alpha\tau} + d^{\tau\alpha})\nabla_\beta\phi_\tau + d^{\alpha}{}_{\sigma}d^{\sigma\tau}\kappa_{\tau\beta},} \tag{96}$$

whose eigenvalues are the curvatures of S' to quadratic order in displacements etc.

Note: Throughout this appendix, the metric used to raise and lower indices and define the covariant derivative has been that of the un-perturbed shell. In general this might not be the elastic reference configuration.