Andrew Glassner's Notebook

http://www.glassner.com

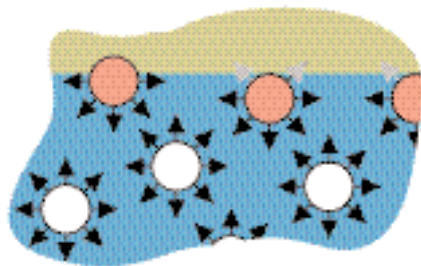## Soap Bubbles: Part 1

Andrew Glassner

I love soap bubbles. They're beautiful, delicate, and though they live only briefly, it's a glorious moment. They float on the air, brilliant colors wobbling over their surface, and when they pop into nothingness, it's just an opportunity to make more bubbles!

This issue and next I'll talk about the chemistry, physics, and computer graphics of soap bubbles. Here I'll focus on the physics of soap films, which are, after all, what bubbles are made of. We'll see what happens when soap dissolves in water and discuss some surprising properties of soap films. With this grounding in how films work, in the next issue I'll discuss how they give rise to the brilliant colors and beautiful 3D clusters that we associate with soap bubbles.

Soapy water

An important idea in soap films of all sorts is *surface tension*. For our purposes, we can think of surface tension as a force that lies in a very thin plane right at the surface of a liquid-air interface. It's a contractive force, trying to pull the surface into the smallest shape possible. Figure 1 shows a schematic view of how this works. Molecules in the liquid pull on each other uniformly, so that the forces upon them are basically neutralized. But for molecules near the surface, all the pull is in the plane of the surface or back into the liquid.

We can make bubbles with soapy water much more easily than with regular water. Why? Soap decreases the surface tension. Plain water has a surface tension of about 72.25 dynes per square centimeter at 20 degr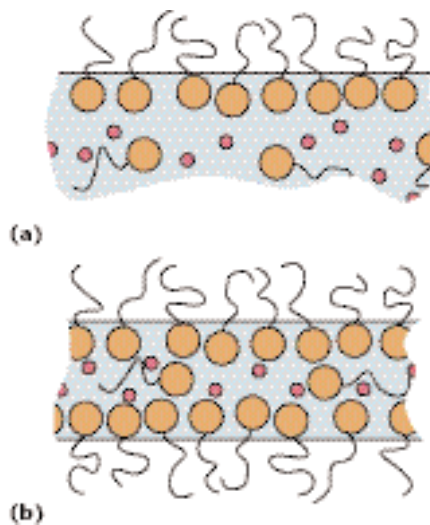ees Celsius. When you add soap, this surface tension drops by about 65 percent. It's this diminished surface tension that lets soap bubbles form and last. The higher tension in plain water causes it to pull against itself too tightly to form big, stable bubbles. That's why raindrops and dewdrops are little blobs of liquid (usually as small and spherical as the circumstances allow) rather than little bubbles or bits of froth.

Soap films have an interesting chemical structure. Soap molecules are the metal salts of long-chain fatty acid molecules. When dissolved in water, the soap molecules break apart and ionize. Let's look at this more closely.

A common soap, sodium stearate, has the chemical symbol $C_{17}H_{35}COO^-Na^+$. When added to water, this molecule breaks into two parts. The metal component splits off by itself and floats around as a little positively charged ion of sodium, written Na^+. The rest of the molecule structurally looks like a ball and chain, or a blob with a tail. The blob, or head, is the polar carboxyl group COO^-. The long tail consists of the hydrocarbon chain $C_{17}H_{35}$. These two pieces of the same molecule respond to water differently.

The head is called *hydrophilic*, which means water-loving, because this part of the molecule likes to be surrounded by water. But the tail is *hydrophobic*, or water-hating. It will dissolve in fat, oil, and grease, but

1 Surface tension at the boundary between water (blue) and air (green). The water molecules (white) attract each other. For molecules near the surface (pink), no outward forces exist, pulling the molecules sideways and back in to the liquid.

2 Soap molecules in water. (a) In solution, the heads stay in water and the tails poke out. The small purple circles represent sodium ions. (b) A thin film.
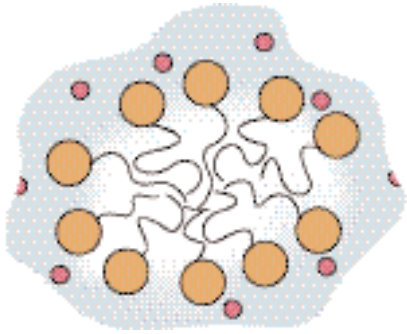
3 A micelle. Note that the heads keep the water molecules out of the center.

avoids water. This difference in preference results in each molecule migrating to the film's surface where the tail pokes out of the water and the head is submerged just inside, as in Figure 2a. Since it lives on the surface, this type of molecule is called a *surfactant*. The entire molecule is called *amphipathic*, or "both loving."

Soap molecules are pretty big—the tail is about 30Å (10 Ångstroms = 1 nanometer), and the head takes up about 40Å^2 of surface area. This explains why the surface tension of soapy water is less than that of water— the monomolecular layer of big carboxyl heads at the surface push away the much smaller water molecules, reducing the force they can exert on one another. Since the water molecules can't attract each other as strongly, the surface tension decreases.

In a soap film, two parallel sheets of soap molecules sandwich a layer of water (Figure 2b). Soap films are thin—usually between about 2×10^5 Å to 50 Å. In the thinnest films, the amphipathic molecules are almost head-to-head. We'll see next time that this range of thicknesses is what gives rise to the brilliant colors in soap bubbles.

Up until a certain point, as more soap is added to the solution, the density of the soap molecules at the surface increases until it reaches about one molecule per 50 Å^2. At this point the surface is pretty densely packed with soap molecules, and as more soap is added, the new molecules just float around inside the water. Things change when the concentration of the soap passes a point called the *critical miscellization concentration*, or CMC. Then the soap molecules spontaneously form little clusters called *micelles*, as shown in Figure 3. Micelles typically consist of 50 or more ions. The big heads on the outside of the cluster keep the water molecules out, effectively creating a little water-free bubble inside the solution.

Soap film geometry

There's an easy experiment you can do yourself that demonstrates some important properties of soap films. I'll describe the standard apparatus, and then we'll look at what we can discover using it.
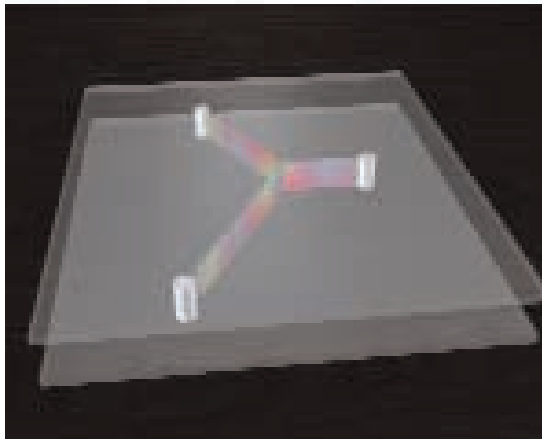
Take two clear plastic rectangles, about 6 inches on a side, and place some metal or plastic pins between them. The pins should be round and smooth and rather small in diameter, and they should hold the sheets so
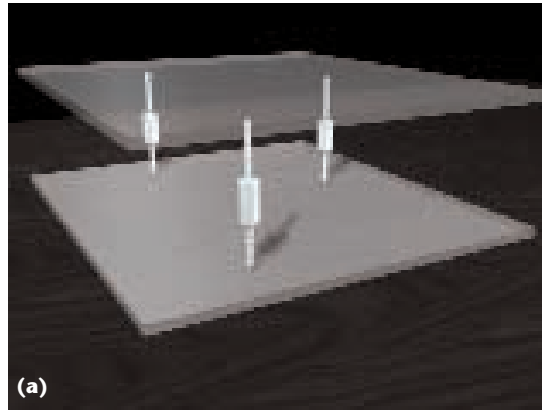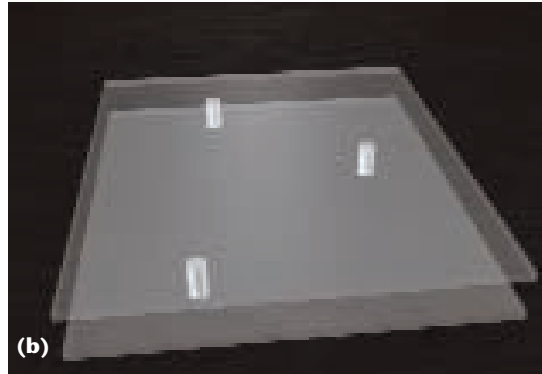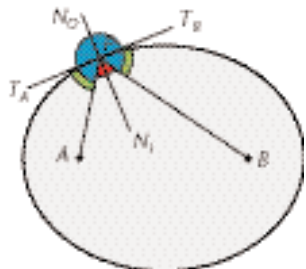
that they're as parallel as possible and about an inch apart. Figure 4 shows the idea. Here I've placed three pins at random. If we dip the apparatus into soapy water and then pull it out, we observe a soap film running between the sheets and joining up the pins, as in Figure 5. Every acute triangle formed by the pins will result in a film that looks like three sheets meeting together inside the triangle.

Let's look at two interesting properties shared by all soap films of this sort.

First, these films always form in flat segments. Consider a piece of film that connects two of the pins. The surface tension in the soap film is always trying to pull it into the shape with the least amount of surface area contained by its borders. When those borders are

6 An ellipse with focal points *A* and *B*, constant *k*, set of points *P*, the normal, and the tangent line. Ellipse property E2a is in green, property E2b in red, and property E2c in blue.

7 (a) Triangle *ABC* and point *P*. (b) A circle around *A* of radius *AP*. (c) Adding an ellipse around *BC* with constant *BP* + *CP*.

8 (a) The ellipse and circle meet at *P* with a common tangent. (b) The ellipse is too small and doesn't contact the circle. (c) The ellipse is too big and intersects the circle twice.

two parallel pins and two parallel sheets, that smallest area forms a flat rectangle.

The second property is that segments can either meet in pairs or in triples. Pairs occur only around the outside of the pins, if at all. Whenever segments meet three at a time, they always form three 120-degree angles, as in Figure 5. This is pretty surprising. It says that in any frothy soap film, all the edges meet at the same angle. Since this property also helps determine the geometry of how 3D soap bubbles join up, it's worth taking a brief detour to see why this happens. By the way, because the distance between the plastic sheets doesn't affect the film's shape (within reasonable limits), I'll adopt the common convention of discussing the geometry as though it were 2D. To get the 3D versions, just extrude the diagrams perpendicular to the page.

Before we get started, I'll point out a couple of useful properties of *ellipses*. An ellipse is a closed curve defined by two points (each a *focus*, or together the *foci*) and a constant. Let's call the two focal points *A* and *B*, and the constant *k*. Then the ellipse is that set of points *P* that lie

at an equal summed distance from the foci; that is, $AP + BP = k$. In Figure 6, I've also marked a random point on the ellipse as *P*, and at that point indicated the normal (with inside and outside points N_I and N_O), and the tangent line (with endpoints T_A and T_B).

For any point *P* on the ellipse, we can prove two well-known properties:

E1: $AP + BP = k$
E2a: $\angle APT_A = \angle BPT_B$
E2b: $\angle APN_I = \angle BPN_I$
E2c: $\angle APN_O = \angle BPN_O$

In words, E1 says that the sum of the distances from *P* to the foci is a constant. E2a says that the lines from the foci to *P* form equal angles with the tangent. E2b and E2c follow from E2a.

Let's take a look at why soap films always meet at 120 degrees. We'll do it by assuming that we know the point *P* and observing its properties. The trick is to realize that using *P*, we can build a circle and an ellipse that must be mutually tangent.

Let's begin with the triangle $\triangle ABC$ in Figure 7a. If we're also given a point *P*, then we know lengths *AP*, *BP*, and *CP*. (Formally, *AP* is the name of the segment, and $|AP|$ is its length, but I won't bother with the vertical bars here.) Let's draw a circle of radius *AP* around *A*, as in Figure 7b. Now use property E1 to draw an ellipse around foci *B* and *C* with $k = BP + CP$ as in Figure 7c. *P* is that point on the circle closest to the foci *B* and *C*.
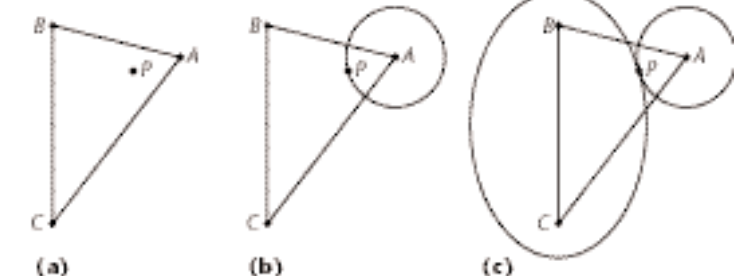
Geometrically, we can see that this circle and this ellipse must graze each other at point *P*, where they also share a tangent, as in Figure 8a. To show that this is true, suppose the opposite. Either the ellipse and circle miss each other completely as in Figure 8b, or they intersect twice as in Figure 8c. Look at point *Q* in Figure 8b. It can't possibly be *P*, because $QA + QB + QC < PA + PB + PC$. That is, the total path is smaller for point *Q*, but we assumed that point *P* was closest to the foci. This contradiction rules out any point on an ellipse that's too small to intersect the circle. Since we're aiming for a minimum length, this also rules out any ellipses so big that they contain the circle. Now consider point *Q* in Figure 8c. It also cannot be *P*, because we could reduce the length *QA* while keeping $QB + QC$ constant (that is, we could move along the ellipse to point Q', giving us $QA + QB + QC > Q'A + Q'B + Q'C$). This contradiction rules out any point on an ellipse that intersects the circle. Since Figures 8b and 8c are ruled out, we're left only with Figure 8a.

Since *P* lies on the ellipse, we know from property E2c that $\angle APB = \angle APC$, as shown in Figure 9a.
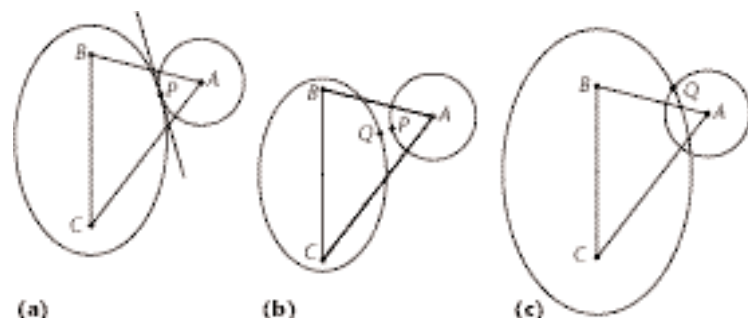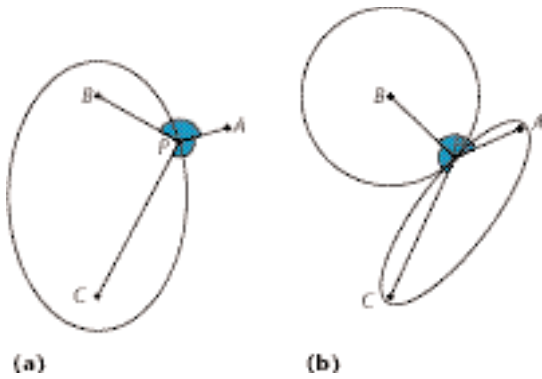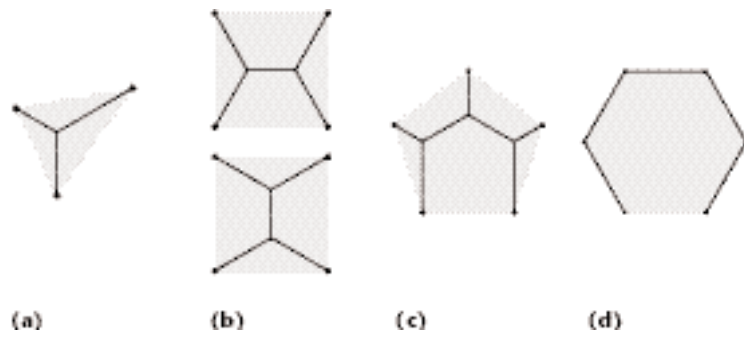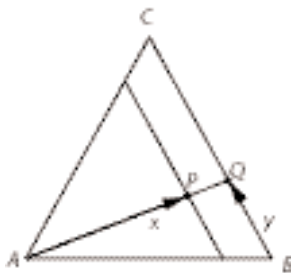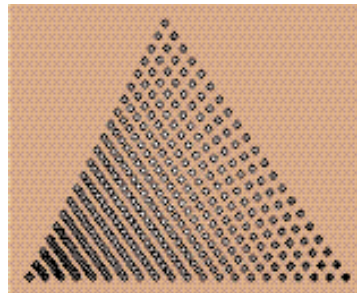
Now erase the circle and the ellipse. Repeat the process. This time draw the circle around point B with radius *BP* and the ellipse that just grazes it on foci *A* and *C*, as in Figure 9b. Property E2c tells us that $\angle APB = \angle CPB$. Since $\angle APB$ is equal to both of the others, we've found that

$$\angle APB = \angle CPB = \angle APC$$

Since all three angles are equal, and they obviously sum

(a)   **(b)**

9 (a) ∠APB = ∠APC. (b) ∠APB = ∠CPB.


(a)   **(b)**   **(c)**   **(d)**

10 Shortest-path (or Steiner) networks. (a) Triangle. (b) The square has two configurations. (c) Pentagon. (d) For hexagons and above, the perimeter minus one edge.

11 Configuration parameters for the triangle. Variable *y* chooses a point *Q* from *B* to *C*, then *x* chooses *P* from *A* to *Q*.

12 Total path length for different points in the triangle. The lengths are normalized, so black is longest and white shortest.

(a)

(b)
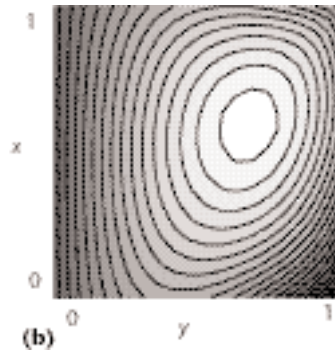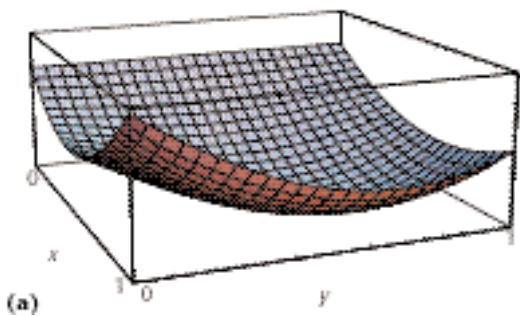
13 Total path length for the triangle. (a) Plotted as height. (b) The contours.

up to 360 degrees, each one is 120 degrees.

This is also called the three-point Steiner problem. The Steiner problem asks us to find the shortest network of straight lines that join up a collection of points. Whenever we're dealing with points in the plane, polygons are always interesting special cases (see Figure 10). We've just seen how to solve the problem for any triangle (we'll look at squares and rectangles in a moment). Notice that the regular pentagon's path is symmetrical, and of course all triples join at a 120-degree angle. For regular polygons with six or more sides, the shortest-path solution is simply the perimeter of the polygon with one side omitted. This kind of perimeter joining is the exceptional case when the films join in pairs rather than triplets and don't meet at 120-degree angles.

Now that we know the solution for a triangle, let's look at it a slightly different way. These other visualizations will be useful when we look at networks for the square and rectangle.

Figure 11 shows the basic idea. We have two variables, *x* and *y*, which we'll use to locate a point *P*. We'll use variable *y* to select a point *Q* along the line from *B* to *C*—the range [0, 1] maps *Q* to [*B*, *C*]. We'll then use variable *x* to locate a point *P* along the line from *A* to *Q*—the range [0, 1] maps to [*A*, *Q*]. These two variables are called the *configuration parameters* for this problem.

For each pair (*x*, *y*) we find *P* and then the total length of the soap film *AP* + *BP* + *CP*. Figure 12 shows this length for some points *P* swept out by this technique. I normalized the values, so a black dot represents the longest path and a white dot the shortest. It shouldn't be too surprising that the minimum looks to be about where the angles would all be 120 degrees.
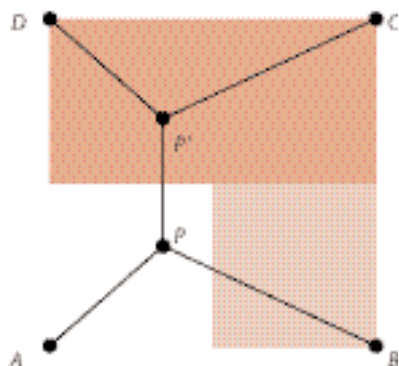
Figure 13 shows a 2D plot of the function, both as a surface and as contours. Notice that because of how I've set up the parameterization, all points with *x* = 0 lie at point *A*, so the value of *y* doesn't matter at *x* = 0. If we
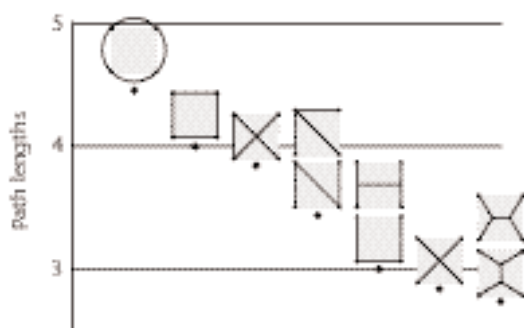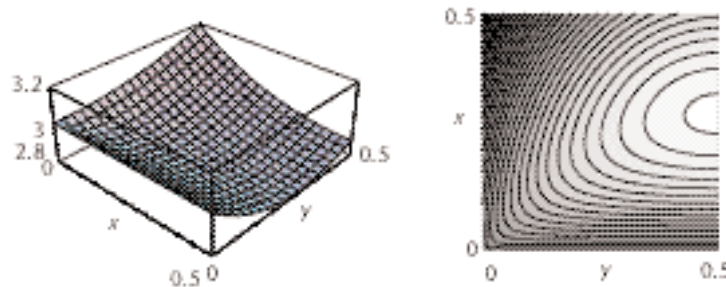
14 Path lengths for different networks on a square.

(a)   **(b)**
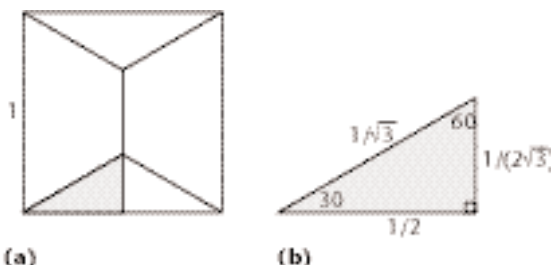
15 Finding the total path length $1 + \sqrt{3}$ for the best shaded triangle.

16 The configuration parameters place P in the lower left quadrant. The rest of the network is drawn as shown.

17 Total path length for the square. (a) Plotted as height.
(b) The contours.

were seriously investigating this space, we'd want to use a more uniform parameterization of the triangle.

Films on the square

Now that we've found the Steiner network for a triangle, let's look at a square. We can cook up lots of possibilities for straight-line networks that join the vertices of a unit square. Figure 14 shows a graph of some examples and the length of the network they generate. (I included the square's circumscribing circle as well, just for fun.)

The shortest path has a length of $1 + \sqrt{3}$. To confirm this, Figure 15 shows the geometry for this network, based on the 120-degree angles we know soap films always form. The shaded triangle in Figure 15a is a 30-60-90 triangle. The long leg has length $1/2$, meaning that the hypotenuse has length $1\sqrt{3}$ and the vertical leg is $1/2\sqrt{3}$. The total length is then the four hypotenuses plus the vertical leg: $4 (1/\sqrt{3}) + [1 - 2(1/(2\sqrt{3}))] = 1 + \sqrt{3} \approx 2.73$.

Let's look at how the length changes for different points in the square. Figure 16 shows how we'll set things up. The two configuration parameters x and y will sweep out the lower left corner of the square to identify point P. We'll join P to B, extend it vertically, and then reflect the result, letting symmetry simplify the problem. Figure 17 shows the surface and contour plots, and happily the minimum length appears where we just determined it should be.

Aspect ratios for films

Hang on a second. Figure 14 shows two configurations (on the far right) for the square that both meet our criterion for a shortest path. One is obviously nothing but a rotation of the other. Suppose we build a plastic and pin gadget using four pins that form the vertices of a square. If we dip this into a bucket of soapy water, when we pull it out the soap film will have only one configuration or the other. Which will it be?
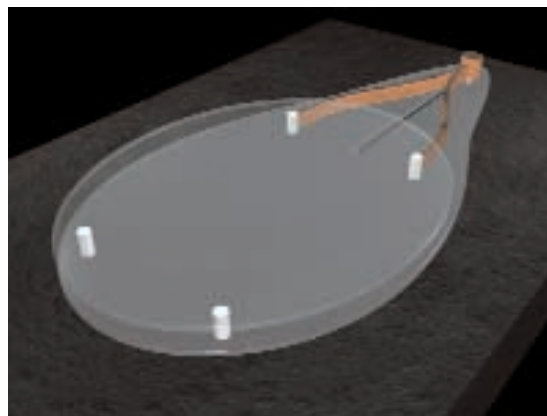

18 The rectangle device. The two pins in the bottom left are fixed at a distance of 1 unit from each other. The other two pins are attached to a handle, which holds them 1 unit from each other, and parallel to the first two. By moving the handle in and out, the four pins can form a variety of rectangles.
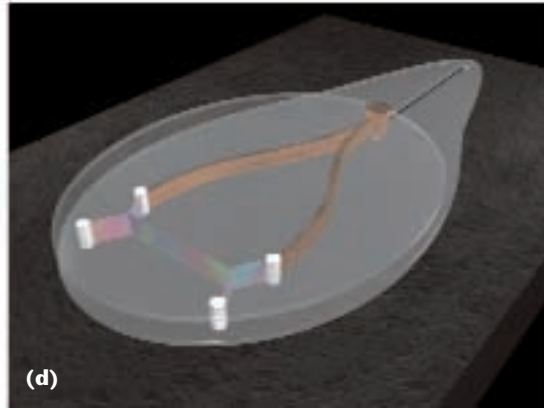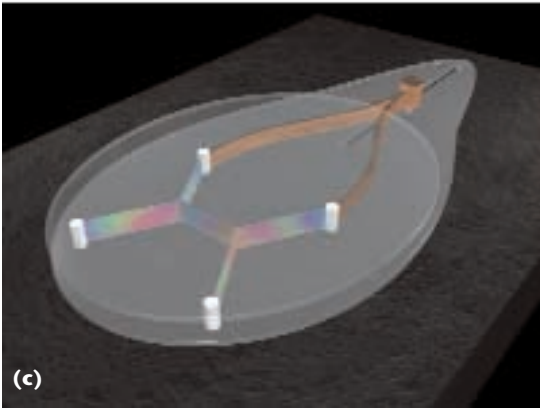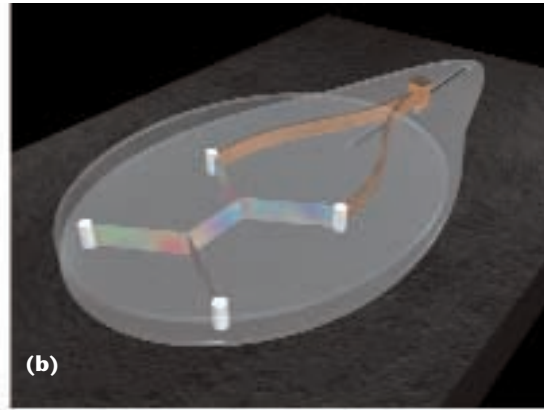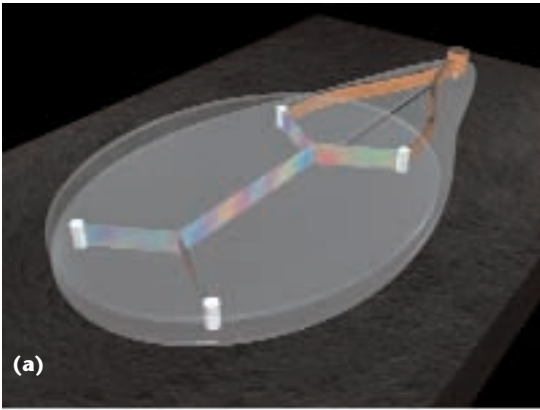
19 The soap films resulting from moving the handle. (a) A wide rectangle. (b) A square. (c) A square. (d) A narrow rectangle.

The easy answer is that it doesn't matter. As far as the plastic gadget is concerned, it can't tell whether it's been rotated 90 degrees. But this easy answer opens up a much more interesting question. Suppose that after we dip it in water, we place the gadget down on the table so that from our point of view, the middle bridge segment runs horizontally (that is, the film looks like a bent letter H). We know that the other configuration (a bent letter I) has exactly the same length. It turns out that our minimum-length criterion is equivalent to a minimum-energy criterion, meaning that the soap film contains the least amount of energy when it's of minimum length. Most physical systems try to adjust themselves to a configuration with minimum energy, and soap films are no exception. In these terms, it might be reasonable to expect the soap film to somehow constantly flip back and forth between these two low-energy states, since neither one is preferred. Of course, that doesn't happen, but why not?

Let's answer this with a new experiment using a new apparatus. Figure 18 shows the new gadget. Two pins are fixed in position, 1 unit apart. Another pair of pins, also 1 unit apart, are parallel to the first pair and connected to a handle. By moving the handle in and out we can adjust w, the length of one side of a rectangle
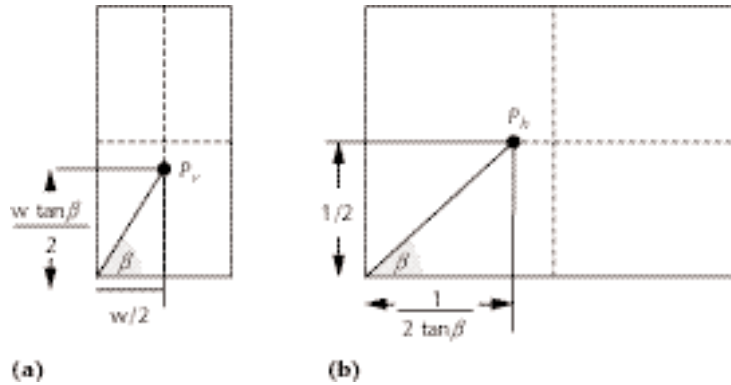

20 Geometry for points P. (a) $P_v = (w/2, (w \tan\beta)/2)$. (b) $P_h = (1/(2\tan\beta), 1/2)$.

formed by the four pins. Figure 19 shows the apparatus in four configurations and the soap films that result.

To begin discovering why the film doesn't spontaneously and frequently flip between the two middle configurations in Figure 19, let's first look at the configuration space for this film. We'll locate a point P in the rectangle, lying along either the vertical or horizontal centerline. For the configuration parameter I'll pick an angle β, which directs the bit of film coming out of the lower left corner. It hits the vertical centerline at point P_v, and the horizontal centerline at point P_h, as shown in Figure 20. Remember that the rectangle is 1 unit high by w units wide. To determine which of these points is closer (and therefore taken by the film), we compute a critical angle β_c that passes through the cen-

21 Finding the critical angle: β_c = arctan($1/w$).
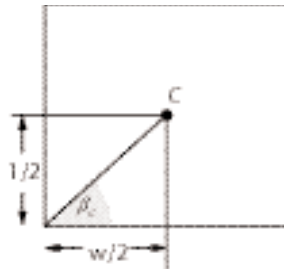
ter of the rectangle, as shown in Figure 21. When $\beta < \beta_c$, we use P_v for the meeting point of the films, and otherwise use P_h.

Since we know that the soap films want to meet at 120 degrees, we can find the value of β that corresponds to that configuration for any value of w. Figure 22 shows the geometry for a long, short rectangle and for a narrow, tall one. The important angles are $\beta = 30$ and 60 degrees, respectively.

Now we're ready to look at the soap film's behavior by plotting the length of the total film against the configuration parameter. Let's start with a narrow, tall rectangle by setting $w = 0.2$, as shown in Figure 23a. The plot of total length (or energy) versus the configuration parameter β appears at the left in Figure 23b (all the length curves in this section have been normalized to a range of [0, 1] before plotting), where I've highlighted the values at $\beta = 30$ and 60 degrees. Notice that $\beta = 30$ is a global minimum for this graph.

Imagine when we start pulling the apparatus from the soapy water, the film initially forms in some random state. That corresponds to some straight-line network for a random value of β.

Since the film wants to minimize its energy, it will move downhill on this graph as far as it can, looking for the network with the shortest path, and therefore the least energy. In this particular graph, no matter where you start, downhill motion always takes you to $\beta = 30$, so you'll always get the vertical configuration when you dip a rectangle with sides 1 and 0.2.

The rest of Figure 23 shows energy plots for six more values of w. Figure 23d plots $w = 1$. Notice that the ener-

gy profile has two distinct local minima. If the soap film randomly starts off with $\beta > 45$ degrees, then it will roll to the minimum on the right, otherwise it will go to the one on the left. Once in one of these minima, the film can't get to the other one without passing through a zone of higher energy (or longer length). That's why the film doesn't flip back and forth—it would have to take in energy to get over the hill to the other side. You can supply this energy yourself by blowing on the film carefully, thereby using the energy from your breath to nudge the film over the high-energy wall into the other valley.

Figure 23c shows a less extreme situation, which represents a value of w halfway between $1\sqrt{3}$ and 1. You can see that there's still a local minimum for the horizontal configuration, but it's not quite as wide. In random dippings, you're more likely to start in the left-hand valley and roll down to the vertical energy state. But this is still a minimum, and it's still stable— if the film takes on this configuration, it stays there.

Figure 23b shows what happens at $w = 1\sqrt{3}$. This is the value where we see a cusp for the horizontal configuration. This is exactly where this configuration ceases to be stable. If the rectangle is slightly wider than $1\sqrt{3}$, then the horizontal configuration can form and persist. If the rectangle is slightly narrower, then the horizontal configuration can no longer form in a stable way, and we always get the vertical form. In terms of the energy graph, no matter where you start, downhill always takes you to the vertical configuration; we've lost the little uphill barrier that formed a valley for the horizontal form.

The rest of Figure 23 shows what happens as w increases beyond 1. The cusp for the vertical configuration occurs at $w = \sqrt{3}$; any wider than that and the soap always comes up horizontally.

We've seen that when $w < 1/\sqrt{3}$ we always get the vertical form. When $w > \sqrt{3}$ we always get the horizontal one. What happens in between?

The colored lines on the right side of Figure 23 show the answer. Let's follow the gold path, starting at the top above a narrow rectangle. The shaded rectangle to its right just doesn't form—the angles aren't 120 degrees. If the film starts off in this position, it will immediately adjust itself to the stable configuration. This shaded rectangle corresponds to the blue dot in the energy graph, which isn't in a stable position.

If we increase w and pull the rectangle wider, following the gold path downwards in Figure 23, then the network adjusts, but stays in a vertical configuration. We can make the rectangle narrower or wider as we please, and the network stays vertical. But if we pull out the handle far enough to reach the critical value of $w = \sqrt{3}$, the vertical form is no longer stable. The film instantly snaps into the horizontal configuration, as shown by the right-pointing gold arrow. Now we're on the blue path. Increasing or decreasing the width keeps us on the blue path. In other
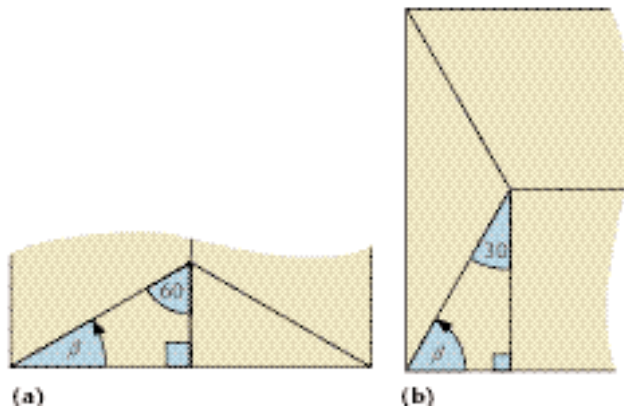

22 The ideal angle. All three lines meet in the colored rectangle at 120 degrees. (a) For narrow rectangles, $\beta = 30$ degrees. (b) For wide rectangles, $\beta = 60$ degrees.
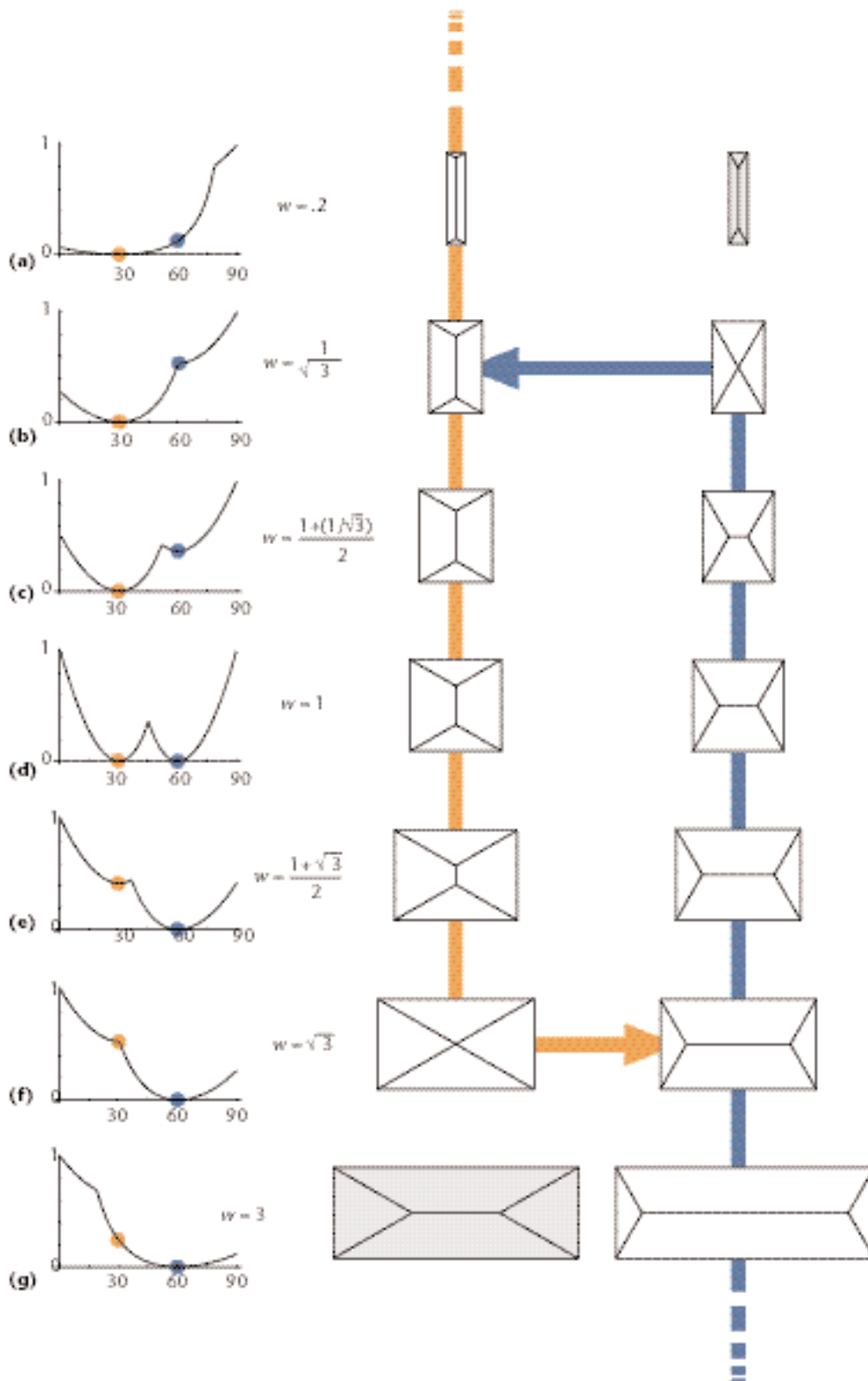
23 The soap film network for seven different values of the width *w*, for a rectangle with constant height 1. Each row shows the energy profile, the vertical configuration, and the horizontal configuration (the grayed out forms aren't stable). The gold and blue dots are at configuration parameter angles of 30 and 60 degrees. The paths on the right show how the network behaves as the width of the rectangle is adjusted. (a) *w* = 0.2, (b) *w* = 1√3, (c) *w* = (1 + (1√3))/2, (d) *w* = 1, (e) *w* = (1 + √3)/2, (f) *w* = √3, and (g) *w* = 3.

24 The energy profile for the rectangular gadget. The *w* axis runs from 0 to 3, and the *β* axis from 0 to 90. (a) The energy profile. Note that longer films naturally have a longer path and higher energy, making it difficult to see detail in the surface. (b) The same data, but each curve of constant *w* has been independently normalized to the range [0, 1]. (c) A contour plot of Figure 24b.

(a)   (b)   (c)

(a)   (b)
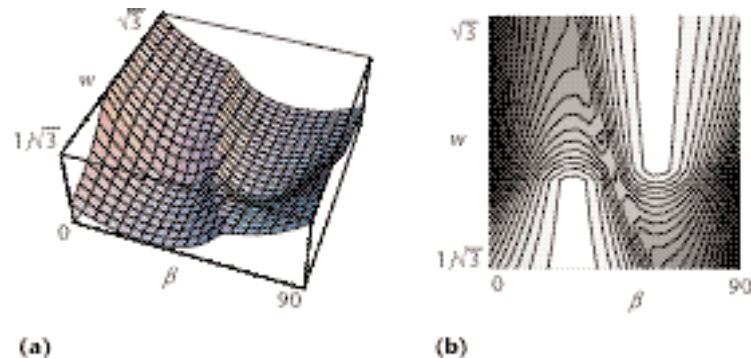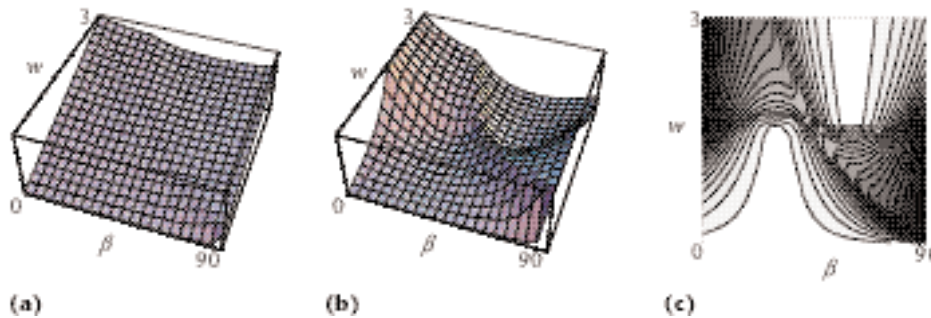
25 A close-up of Figure 23b in the hysteresis region $1/\sqrt{3} < w < \sqrt{3}$. (a) Height and (b) contour.
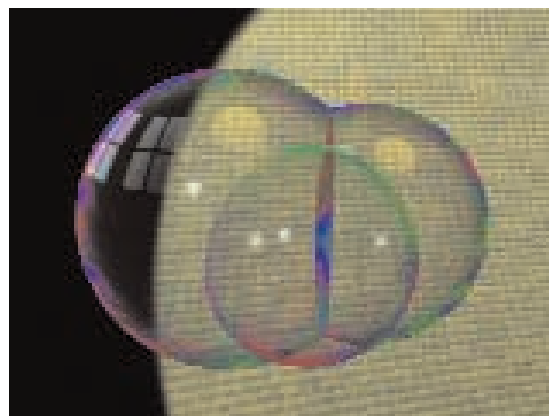
26 A cluster of three soap bubbles hovering in front of a cloth background.

Further Reading

The quintessential reading for soap bubbles is the book *Soap Bubbles: Their Colors and Forces Which Mold Them* by C.V. Boys (Dover Publications, New York, 1959), which was originally published in 1911. The book is a friendly and light introduction to the subject, aimed at the enthusiastic turn-of-the-century teenager with a scientific curiosity.

Two more recent books provide a more modern and complete introduction and served as my primary references. *Demonstrating Science with Soap Films* by David Lovett (Institute of Physics Publishing, Bristol, 1994) presents the idea of plotting the energy curves for a rectangle of changing shape, which I adapted for the discussion here. Another modern reference, *The Science of Soap Films and Soap Bubbles* by Cyril Isenberg (Dover Publications, New York, 1978), was my source for the chemistry of soap molecules.

words, there's no going right back—we can't recover the vertical form just by making the rectangle a little narrower. We say that this process is not reversible.

The horizontal form is stable from as wide as you want to as narrow as $w = 1/\sqrt{3}$, when it snaps into the vertical configuration, as shown by the left-pointing blue arrow.

Any rectangles narrower than $w = 1/\sqrt{3}$ are always in the vertical form, and any wider than $w = \sqrt{3}$ are in the horizontal form. Basically, any time you adjust the soap film, you stay in the valley you're already in until that valley ceases to exist. This phenomenon, called *hysteresis*, shows up in other situations, such as when you magnetize a piece of iron.

Figure 24 puts all this information together and shows the variation in the energy of the film with respect to the width *w* and the configuration parameter *β*. Figure 25 shows a close-up of the hysteresis region $1/\sqrt{3} < w < \sqrt{3}$.

Blowing bubbles

Now that we know all about soap films, we're ready to stretch our minds and our films into 3D. In the next issue we'll look at the geometry of bubbles and bubble clusters, discuss the source of the brilliant colors in the bubbles, and talk about making pictures like Figure 26. ∎

Readers may contact Glassner by e-mail at andrew_glassner@yahoo.com.