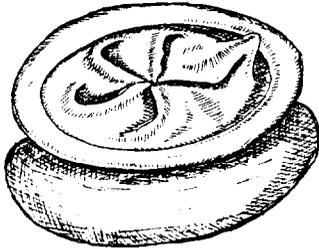


DRAWING SURFACES AND THEIR DEFORMATIONS: THE TOBACCO POUCH EVERSIONS OF THE SPHERE

GEORGE K. FRANCIS

University of Illinois
Urbana, Illinois 61801, USA



The various sensuously possible cases of figures are not, as Greek geometry, individually conceived and investigated, but all interest is concentrated on the manner in which they mutually proceed from each other. Insofar as an individual form is conceived, it never stands for itself alone but as a symbol of the system to which it belongs and as an expression for the totality of forms into which it can be transformed under certain rules of transformation.

Jean-Victor Poncelet

1. INTRODUCTION

Mathematicians communicate with pictures: a quick sketch on a coffee stained paper napkin, a saddle drawn on the blackboard before a calculus class, the totem figure of the discipline conjured up on the viewgraph before a learned audience of specialists. Illustrating some mathematical idea, one should be able to reconstruct the picture quickly, from memory, and with sufficient skill to ignite the desired idea in the viewer's mind. More elaborate pictures should be so drawn as to withstand the entropy of repeated mechanical reproduction. A truly heroic test is for the picture to survive a sequence of transfers, e.g., from design board to tracing paper, to xeroxed preprint, to acetate plastic film, to thermofaxed ditto master, and thus to handout before a lecture. Commonly, these pictures consist of elaborate typographical symbols or graphs of real valued functions of one variable. But when these pictures are not two-dimensional, one problem is to discover and organize some simple combinatorial rules for drawing pictures of mathematical surfaces extended in space. Any solution is subject to both geometrical and graphical constraints. The former are part of low-dimensional topology. The latter do not as yet belong to any particular mathematical discipline.

By *graphical constraints* we shall mean all the conditions of medium, recall, reproducibility, and artistic detail that just suffice to evoke a mental perception of the surface. These constraints are more utilitarian than aesthetic in nature. The graphical calculus we seek to develop should be easily learned and taught, and applicable to diverse areas of geometry and topology. The *geometrical constraints* are set, of course, by the abstract concepts one wishes to illustrate. To illustrate them is to encode these ideal mathematical objects in a pictorial language. The picture is itself a construction of the mind. What is on paper is a secondary set of graphical symbols that encodes the mental picture by a variety of conventions, tricks, and illusions. In the first part of this paper we summarize some of the graphical devices we have collected from many sources, including the works of Maurits Escher, and have adapted them to our limited artistic talents with much practice and experimentation. For now they are meant to be suggestions on how to draw effectively what is seen so clearly in the mind. A good set of pictures often forestalls a premature flight into algebraic abstraction at the expense of the interest and comprehension of the hearer. In the second part we "practice what we preach" by telling a picture story about a certain class of eversiones of the sphere.

2. TECHNIQUES AND TERMINOLOGY

Systematic geometry has always benefited from the axiomatic approach and we hope that our graphical calculus will eventually yield also to a formal treatment. For the present we proceed informally, borrowing rigorous definitions from differential topology as needed. Golubitsky and Guillemin's fine text [1] is more than adequate for this purpose. This is not to suggest, however, that differential topology is the correct discipline for the visual category. Many ideas in geometric topology and algebraic geometry, as they are first grasped intuitively and not as they are finally written up, make use of the same kind of visualizations.

Roughly speaking, a *surface* is what we draw a picture of . . . *lines* are what we draw the picture with . . . and a *point* is where we center our attention when looking at a detail of the surface. We regard a visible subset of 3-space as a *surface* if it can be presented locally by a certain class of smooth, model mappings, $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $(X, Y, Z) = F(x, y)$. By a *picture* of the surface we shall mean a projection of the image of F to the *viewing plane*, augmented by a set of graphical conventions indicating how the surface is to be imagined from the picture. All of our basic notions are illustrated in Fig. 1. The location of a picture is given by a three digit numeral specifying the figure number, row

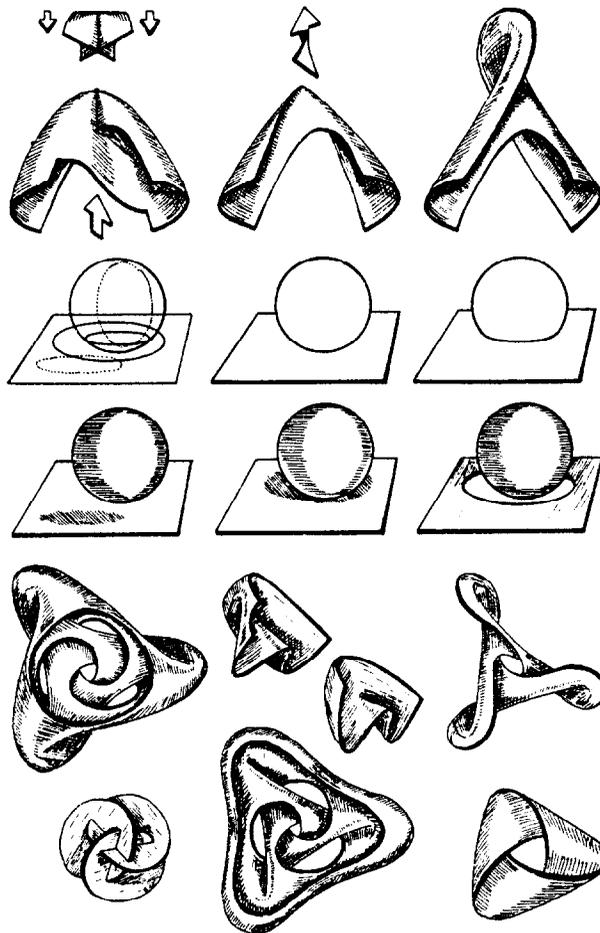


Fig. 1.

from top, and column from left. Thus, (153) refers to the Moebius band spanning the trefoil knot.

The parametric approach is more flexible and constructive than one in which we regard the surface as defined implicitly by constraints, $u(X, Y, Z) = 0$. For example, *Whitney's umbrella* [2] (upper detail of 111) is modelled on $X = x, Y = -y^2, Z = xy$, where XYZ are the observer's coordinates: left-to-right/down-to-up/far-to-near, with the origin at the center of attention. Setting $Z = \pm(-Y)^{1/2}X$, we think of the surface as generated by two horizontal lines, initially coincident along the X axis, rotating in opposite directions as they sink down the double line along the negative Y axis. Bending the X axis down, $Y' = Y - tX^2, 0 \leq t \leq 1$ (small arrows), produces the parabolic umbrella, $F = (x, -x^2 - y^2, xy)$ in (111) reminiscent of the top of *Steiner's crosscap*. A bend in the opposite direction produces a hyperbolic umbrella (not shown), which is a projection of the complex branchpoint $(x, y, x^2 - y^2, 2xy)$ from 4-space. The cubic form, $X^2Y + Z^2$, which vanishes on the ruled umbrella, is not nearly this flexible. Furthermore, it also vanishes on the extraneous, one-dimensional whisker extending above the surface along the positive Y axis.

Even when explicit formulas are inconveniently difficult to write down, we can pseudo-parametrically describe motions intended to be imagined between consecutive pictures. Thus, to pass from the bent umbrella (111) to the bent but otherwise nonsingular quadrilateral (112), retract a border segment transverse to the double line, somewhat like raising a window shade. Now stretch it out again with a left twisting motion suggested by the arrow (112) to reach *Whitney's cusp* (113) [3]. This surface, parametrized by $X = z^3 + yz, Y = y, Z = z$, graphs a cubic bifurcation with abscissa Z , ordinate X , and time axis Y . Rewriting,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -2z^3 \\ -3z^2 \\ z \end{bmatrix} + (y + 3z^2) \begin{bmatrix} z \\ 1 \\ 0 \end{bmatrix} \tag{*}$$

displays the surface as generated by straight lines parallel to the viewing plane, moving along the cubic spiral, which forms the apparent contour or horizon of the surface. Under projection, these lines are tangent to *Neil's cusp* (semicubical parabola) from which the surface takes its name. We refer to Chaps. 6 and 7 of [1], for further topological details.

Already, the reader will have observed a distinction between *intrinsic*, *apparent*, and *auxiliary* surface features, all of which we depict by drawing lines. The *border curves*, frequently introduced to delimit a surface detail, are almost always of the last kind. Sometimes though, as in surfaces spanning knots and links, the borders have an intrinsic meaning. *Intrinsic* to the surface are those features that are independent of the position it has been placed for projection into the viewing plane. Thus, a *pinchpoint* (the origin in the umbrella) and the *double line*, where two sheets of the surface cross, are intrinsic. The *apparent* features are accidents of the viewing position, but correspond to intrinsic features of the projection. Thus, a *contour line*, along which a visible sheet of the surface bends out of view, is induced on the surface by the projection. It is the locus of points where the surface-normal is perpendicular to the viewing direction or, more practically, where the surface-tangent-plane passes through the observer. A *cusp point* occurs on the contour when its tangent on the surface parallels the viewing direction.

Since these three kinds of lines are easily confused in a picture, we use several elementary tricks of perspective to force the eye to distinguish between them. The principles of linear perspective, that is, the foreshortening (especially) of surfaces with rectangular faces are well known from projective geometry. For curved surfaces we use *linear perspective* on auxiliary surfaces, such as coordinate planes, to induce the illusion of depth.

However, the surface pictured is also given finite thickness by drawing the borders heavy or light according to the angle made to the viewing direction. This, in turn, becomes the clue for a border line.

To improve the separation of contours from borders we use some rudimentary shading, which is part of *aerial perspective*. The surface bends away along the contour, so *shading* suggests decreasing reflection of light in the viewing direction. Secondly, a face of a surface further away should be darker than a nearer one. Finally, one part of the surface may throw a *shadow* on another. For technical simplicity and reliable reproducibility, we use a somewhat stylized shading technique based on a *field of lines* (or one-dimensional foliation). Differently oriented line fields help contrast different faces of a surface; like oriented line fields are useful to continue a face past an occlusion. A pair of transverse line fields usually suffices to deepen the tone, as for a shadow.

But, with only two grey tones (solid black regions are useless because they do not reproduce correctly on most copiers) we cannot obey all the rules of aerial perspective. The eye will probably scan each region of the picture separately to extract the information completely. Therefore, we choose for the direction of illumination the opposite of the viewing direction. [Sequence (131–133) is an exception, the light source here is in octant (+, +, -).] This convention permits us to use both shade and shadow to mark a partial occlusion when the occluding front face ends on a border curve. When the occluding face terminates along a contour, the conflict of shade and shadow can sometimes be resolved by stopping the shade short of the contour to produce the illusion associated with lateral inhibition in the retina, known as a *Mach line*. This thin highlight (absence of shading) close to the contour is exaggerated in (132).

For both visual and combinatorial purposes, the entire collection of lines to be used should be in as general a position as possible in the viewing plane. To simulate motion, changes from one picture to the next should be highly controlled. For these two reasons we usually begin with a *diagram* of all lines, such as (121). This serves as a tracing template for the subsequent pictures. It is also easier to memorize the diagram and then reconstruct the picture using the *occlusion rules* for border, contour, and double lines. The dotted ovals in (121) are shade and shadow delimiters. Thus, we obtain (122) for a sphere hovering above the plane and (123) for one intersecting the plane along a parallel of latitude. The various *occlusion rules* are all derived from simple observations, e.g., a surface path gains/loses one sheet-depth unit each time its projection crosses a line representing a border, etc. A field of horizontal lines seems adequate to shade the sphere (131). Curved lines would probably be more effective visually, but harder to draw. An elliptical shadow patch gives the illusion of altitude. Note that without shading, we cannot tell from (122) whether (131) or (132) is intended.

Even with shading we could not tell whether (123) is a sphere intersecting a plane or a soft ball resting on the plane. The elliptical double line, where the sphere crosses the plane, is tangent to the circular contour of the sphere. Double line/contour tangencies produce graphical ambiguities which are harder to resolve than those produced by border/contour tangencies. The latter were reduced by thickening the surface. The most effective device we have found to reduce the former is a window that eliminates the need to draw the double line altogether, (133). We define a *window* to be a fully visible topological disc which has been removed from the surface. That is, the restriction of a (local) parametrization F of the surface to a disc in the source of F is an embedding into space, as well as into the viewing plane under projection. A window is useful for showing how a surface continues behind an occlusion. More generally useful is the ability to place windows in such a way that no double lines need to be drawn. If the choice of the viewing direction is given some slight flexibility, such windows are always possible. The surface with windows is embedded, and its picture requires only border and contour lines.

For example, (141) shows a one-sided surface with four windows. The nearest window reveals three smaller windows with pairwise linked borders. If we span this link by discs intersecting as shown in (151), slide this complex into the surface, and cap the large window without intersecting the rest, we obtain *Boy's surface* [4,5]. This is an immersion of the projective plane in 3-space with one double curve crossing itself at a single triple point. The contour of this view is a right-handed trefoil knot with three cusps and crossovers. The crossovers can be eliminated by *rolling* the cusps *over* the contour, as in (142). This version of Boy's surface (152) has two disjoint, simple contours: a circle around a *deltoid* (three-cusped hypocycloid). The *ribbon neighborhood* (143) of the deltoid has three right-handed cusps, and thus is isotopic under the motion (113 \rightarrow 112) to the Moebius band with three half-twists (153). We suggest that the reader draw the immersed disc with a triple point, which together with (143) make up Boy's surface situated as in (152). A more difficult exercise would be to use the motion (112 \rightarrow 111) on (153) to produce three pinchpoints. Continue the three double lines through a triple point and terminate them at three further pinchpoints. The resulting surface will be topologically equivalent to *Steiner's Roman surface*. The challenge is to place the deltoid contour of this surface convincingly.

We close this section with a few remarks on designing the pictures. A slate blackboard is useful for laying out the *line diagram*. This is the set of curves in the viewing plane destined to become the projection of the set of contours and borders of the surface. The ease of erasing helps in achieving graphical general position of the diagram. Erasure facilitates also the task of finding the occlusions. White chalk for highlights and black chalk for accenting the deepest shades usually suffice to establish the illusion. Colored chalk, used to keep track of complicated convolutions of the surface, occasionally produces startling visual effects. The method of assemblages for surface-to-surface mappings [6–10] can be used to design a complete set of windows. Building up these combinatorics further, such as specifying the position of pinchpoints on the contours, assigning the handedness to the cusps and crossovers, etc., one might hope to solve the following problem. Given the line diagram, find all possible factorizations

$$M^2 \xrightarrow{F} \mathbf{R}^3 \xrightarrow{P} \mathbf{R}^2,$$

where M is a surface, F is a proper, stable mapping, P is a projection, and the composition has the given diagram as its singular locus. A combinatorial classification up to self-maps of M and regular homotopies in \mathbf{R}^3 that preserves the line diagram, would be very useful in the graphical calculus.

3. APPLICATION AND EXAMPLES

Nearly twenty years ago the late Arnold Shapiro rose to Raoul Bott's challenge to give a concrete description of how to evert the sphere. The possibility of turning the sphere inside out by a regular homotopy follows from an abstract theorem of Smale [11]. Shapiro's eversion [12,13] failed to satisfy the visual imagination of his audience. This happy fault led to a series of spectacularly artistic illustrations by Tony Phillips [14], Charles Pugh [15], Nelson Max [16], and Jean-Pierre Petit [17–19], all based on the brilliant eversions visualized by the blind topologist, Bernard Morin [17–20].

In Fig. 2 we present the essentials of the first ($n = 2$) of the countable sequence of highly symmetric *tobacco pouch eversions* [7], so named for their resemblance (for $n = 5, 6, 7$) to the closing operation of the common French *blague automatique*. These pictures were originally developed in collaboration with John Staudt, Timothy Daughters, and the

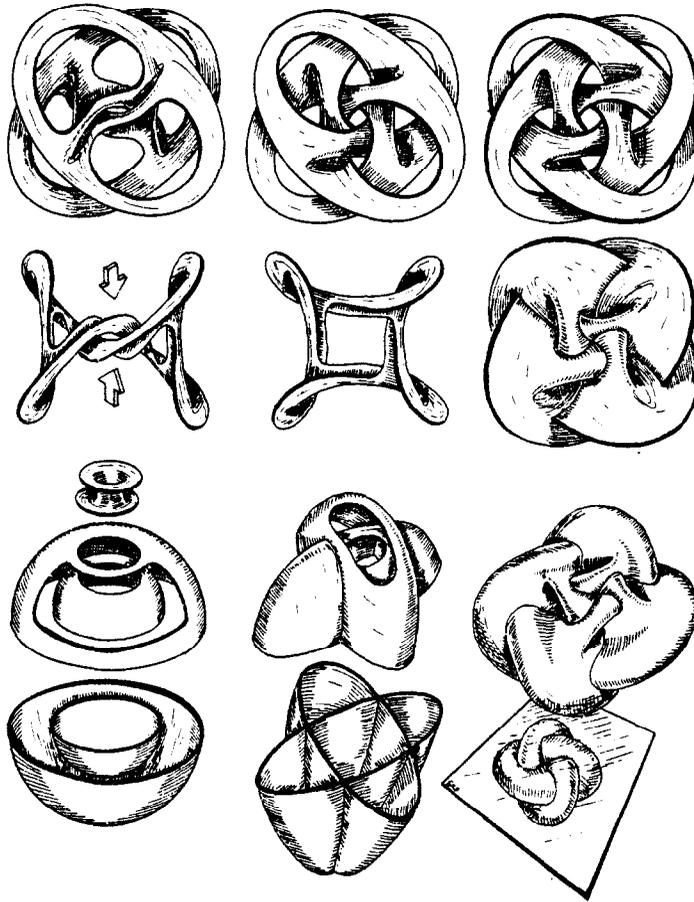


Fig. 2.

other members of my freshman topology seminar in 1977, after viewing Max's film. They undoubtedly owe a subliminal debt to Morin's models [20] and to the souvenir pouch he gave me in 1972.

An *immersed* sphere in 3-space is the image of a smooth map $F : S^2 \rightarrow R^3$ for which the Jacobean matrix DF of first partial derivatives is everywhere of rank 2. The sphere is *embedded* if F is also one-to-one. A deformation F_t , $t \in [0,1]$, through immersions is a *regular homotopy* if $t \rightarrow DF_t$ is also continuous. A regular homotopy through embeddings is an *isotopy*.

Graphically this means that each (sufficiently small) parameter patch on the surface looks like the graph of a smooth function on the plane tangent to the center of the patch, and that the patch moves isotopically under the deformation we imagine to occur between consecutive pictures. Since every patch centered at a pinchpoint has a double line, Whitney's umbrella (111) is not an immersed surface. An *eversion* of the sphere is a regular homotopy from the identity inclusion of S^2 in R^3 to the antipodal map of S^2 . The deformation

$$F_t = [X \cos(\pi t) + Y \sin(\pi t), -X \cos(\pi t) + Y \cos(\pi t), Z \cos(\pi t)],$$

$(X, Y, Z) \in S^2$, is not an eversion because it fails to be an immersion at $t = 1/2$ along the equator $Z = 0$.

Now begin by deforming the sphere isotopically into the shape of a *gastrula*. This embedded sphere consists of two concentric *spherical shells* connected by a *neck* shaped like the inner, negatively curved portion of a torus. We have further separated the shells into two concentric hemispheres (241) and zones (231). Seen from a point on its axis of revolution, the contour of the neck would simply be a circle. We may, however imagine how to twist two apparent swallow tails into this contour, like (221). In (211) we made the two rims of the so deformed neck elliptical for contrast and used four windows for visibility. We could produce such a shape from the *gastrula* by deforming each spherical shell separately into ellipsoidal shells, (232 and 242), with orthogonal major axes. This motion is a regular homotopy because a double curve is produced. Note further that, while the *gastrula* is a surface of revolution, the immersed sphere has merely twofold symmetry, i.e., it is invariant under a half-turn about its axis. This symmetry may be applied to each stage of the regular homotopy.

Now look again at the ribbon neighborhood (221) of the contour. A regular homotopy (arrows) moves the two longer segments across each other into the ribbon neighboring a four-cusped hypocycloid, or *astroid* (222). Observe that both (222) and (242) have a fourfold symmetry. If you were to color the sides of the surface, a quarter-turn about its axis moves these parts into themselves, but with colors exchanged. This is not yet the case for the entire surface, as may be seen from (212). This detail is the result of the homotopy (221 \rightarrow 222) applied to (211). To impose the fourfold symmetry, press down on the upper rim at nine and three o'clock to obtain (213). Placing windows in a symmetric way into (213) produces an immersed annulus that could look like (223). With some effort one can imagine that the double line of (223) forms a bouquet with four loops and four loose ends tied together at the quadruple point where the four windows cross the axis. Fitting (223) to (242) produces a version of *Morin's surface* (233). (243) is its reflection in a mirror. We leave to the reader the exercise of producing a more conventional view of Morin's surface by rolling the four cusps over the contour. To complete the eversion, apply the reverse of this homotopy to the other pair of opposite segments of the *astroid* contour.

The essential point, however, is that one really needs only the five embedded details (211–222) to argue convincingly for the existence of the homotopy, without worrying about the exact placement of the double curves. This idea generalizes to all symmetries of even order, $n = 2, 4, 6, \dots$. The contour of the middle stage, which is a generalization of Morin's surface, is an alternating torus knot that projects to a stellated hypocycloid of $2n$ cusps. This surface is invariant under a $1/2n$ turn about the axis, exchanging color; the homotopy is invariant under a $1/n$ turn.

For odd n , a different type of eversion is produced by this device. An odd number of swallow tails are twisted into the neck of a *gastrula*. The case $n = 3$ is illustrated in Fig. 3. Picture (311) is a stylized analog of (211) as it would result directly from the combinatorics of [9] applied to the contour of a distorted but still embedded sphere. The six cusp-ears of the ribbon neighboring the contour are joined by bridges to the annular rims of concentric spherical shells. In (312) the long arms have moved across the axis by a threefold (rotationally) symmetric regular homotopy. Note that the bent ribbon in (312) is a twofold cover of the Moebius band with three cusps (143) under normal projection. One should imagine a copy of (143) situated between the ribbon. Now move the ribbon normally through the Moebius band to position (321). While the colors have exchanged for the bent ribbon, they have not done so for the spherical shells, which have not moved yet.

Now uncross the long arms of the bent ribbon to (331) to obtain a ribbon like that of (311) but with colors exchanged. Unlike (311), however, the window borders of (331) are

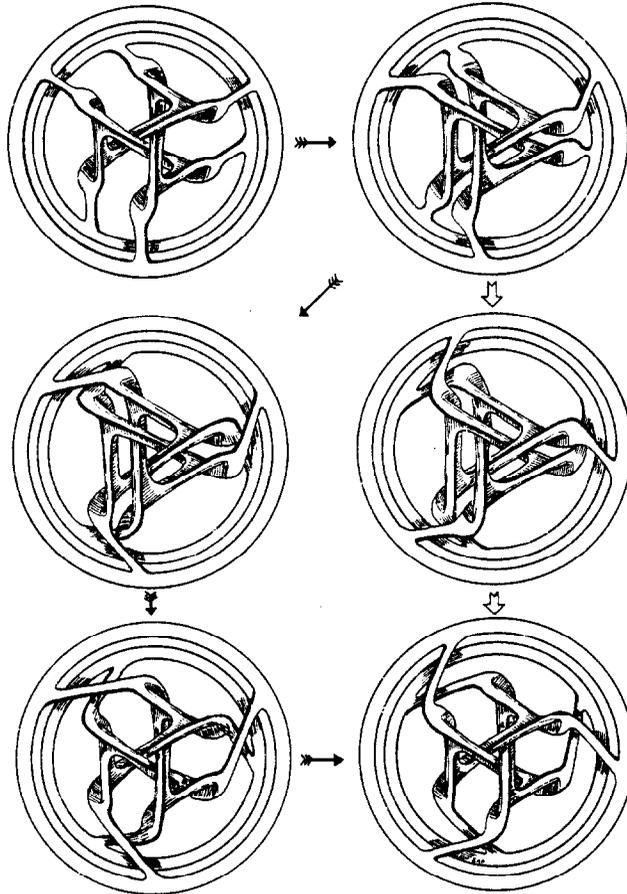


Fig. 3

linked and so the surface is only immersed. Exchange the spherical shells as suggested by their rims to obtain (333), which is isotopic to (311) with colors exchanged.

Although the essential part of this homotopy also proceeds in the (immersed) normal bundle of Boy's surface (or its generalization for odd $n > 3$) it differs in one curious respect from the original eversions of this type [12,14,22,23]. These take the standard sphere to a surface which parallels Boy's surface twice. Think of the skin of a thickened Boy's surface. This immersed sphere everts by a motion along the field of lines normal to Boy's surface, which it covers twice at half-time. Running the first part of the regular homotopy backwards returns the surface to the standard position with sides exchanged. Our homotopy (fleched arrows) does not possess this time symmetry, although it too passes through a parallel neighbor of Boy's surface (321). To recover temporal symmetry, exchange the spherical shells attached to the annular rims at the same time as the branches of the ribbon (312 \rightarrow 322). These two (isotopic) surfaces are not parallel to Boy's surface. Each is the image of a sphere immersed (not embedded) in the abstract normal bundle of the projective plane. This immersion crosses itself and the zero section along a double curve which is fixed, like a hinge, during the model of the motion (312 \rightarrow 322). Now, motion (322 \rightarrow 332) is essentially the same as (312 \rightarrow 311).

We close this paper with a brief remark on Morin's brilliant parametrization [21] of the tobacco pouch eversions. His formulas proceed analytically from the motion of the

contours in a way analogous to how we construct our pictures. The germ of the analogy is contained in the parametrization (*) of Whitney's cusp. However, computer graphical displays driven by his formulas do not produce our pictures: only the projection of the moving contour is the same. We leave a full discussion of this interesting discrepancy for another day.

Acknowledgements—We would like to thank Robin McLeod for organizing the eminently successful conference and workshop at Las Cruces, January 1980, which provided the opportunity to field-test this study. We are also grateful to Bill Abikoff, Tom Banchoff, Jim Callahan, Nelson Max, Tim Poston, Herman Rusius, and especially to Bernard Morin for critical comments and valuable suggestions at various stages of the project.

REFERENCES

1. M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities*, Grad. Texts in Math. No. 14, Springer-Verlag, New York (1973).
2. H. Whitney, The Singularities of a Smooth n -Manifold in $(2n-1)$ -Space. *Ann. Math.* **45**, 247–293 (1944).
3. H. Whitney, On Singularities of Mappings of Euclidean Spaces. I. Mappings of the Plane into the Plane. *Ann. Math.* **62**, 374–410 (1955).
4. W. Boy, Über die Curvatura Integra und die Topologie Geschlossener Flächen. *Math. Ann.* **57**, 151–184 (1903).
5. W. Boy, Bild der Projectiven Ebene Mit Nur Drei Extremen, Plaster of Paris Model, 30. Serie, Nr. 4, Verlag Martin Schilling, Halle, a.d. Saale (1902).
6. G. Francis, Assembling Riemann Surfaces With Given Boundary and Branchpoints. *Ill. J. Math.* **20**, 198–217 (1976).
7. G. Francis, Some Equivariant Eversions of the Sphere, privately circulated manuscript (1977).
8. G. Francis, Polymersions With Nontrivial Targets. *Ill. J. Math.* **22**, 161–170 (1978).
9. G. Francis and S. Troyer, Excellent Maps With Given Folds and Cusps. *Houston J. Math.* **3**, 165–194 (1977).
10. G. Francis and S. Troyer, Excellent Maps With Given Folds and Cusps, II. *Houston J. Math* (submitted).
11. S. Smale, A classification of Immersions of the Two-Sphere. *Trans. Am. Math. Soc.* **90**, 281–290 (1959).
12. A. Shapiro, Lectures (December 1960). See [13].
13. G. Francis and B. Morin, Arnold Shapiro's Eversion of the Sphere, *Math. Intelligen.* **2**(4) (1979).
14. A. Phillips, Turning a Surface Inside Out. *Sci. Am.* **214**, 112–120 (1966).
15. C. Pugh, Wire Mesh Models, Mathematics Department, University of California, Berkeley.
16. N. Max, Turning a Sphere Inside Out, Int. Film Bureau, Chicago (1977).
17. B. Morin and J. Petit, Problematique du Retournement de la Sphere. *C. R. Acad. Sc. Paris* **287**, 767–770 (1978).
18. B. Morin and J. Petit, Le Retournement de la Sphère. *C. R. Acad. Sc. Paris* **287**, 791–794 (1978).
19. B. Morin and J. Petit, Le Retournement de la Sphère. *Pour la Science* **15**, 34–49 (1979).
20. B. Morin, private collection of plasteline models (1967–1972).
21. B. Morin, Equations du Retournement de la Sphère. *C. R. Acad. Sc. Paris* **287**, 879–882 (1978).
22. B. De Witt, Eversions of the 2-Sphere, in *Battelle Recontres* 1967, pp. 546–557, Benjamin, New York (1968).
23. N. Kuiper, Convex Immersions of Closed Surfaces in E^3 . *Comm. Helv.* **35**, 85–92 (1961).