Continuous, Non-Singular Transformations from the Klein Bottle to the Torus

Rachel Lash
Department of Mathematics
Alfred University
Alfred, NY 14802
lashdb@alfredtech.edu

Juha Pohjanpelto, Advisor
Department of Mathematics
Oregon State University
Corvallis, OR 97330
juha@math.orst.edu

§1. A Question and an Observation.

Suppose we have two objects, very much alike, but not identical. What would it take to be able to transform one continuously into the other?

The answer to this question depends, of course, on the objects, and on the nature of their similarities and differences. As often happens, a simple example reveals much. Consider a dihedral group, for instance D_3, the group of symmetries of a triangle in the plane. If our two objects, alike but not identical, are the two triangles shown in Figure 1.1, then the first can be transformed continuously into the second by means of a rotation.

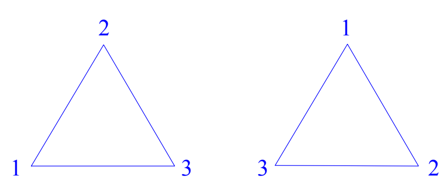


Fig. 1.1

We could plot each intermediate stage in the plane, and it would still be a triangle.

If, on the other hand, our two objects were those shown in Figure 1.2, things would be different.

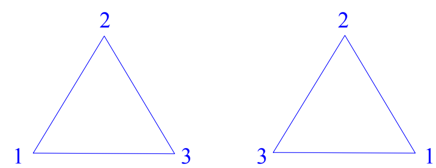


Fig. 1.2

No continuous non-singular motion in $\mathbb{R}^2$ will take us from the first triangle to the second. Our only hope is to collapse the triangle to a line and then gradually restore it, reflected. We must either accept the singularity induced by this mapping, or we must give up altogether and decide that reflections are simply discrete transformations.

For a Flatlander, the horns of this dilemma would be the precise resting place of our problem. Fortunately, however, we live not in the plane but in a spacious universe of three or four dimensions, and can see beyond the apparent difficulty. If we allow that the plane resides in $\mathbb{R}^3$, then we can interpolate non-degenerate triangles between the two shown in Fig.1.2, representing the stages of a continuous non-singular transformation. We fix an axis through vertex 2 and the opposite midpoint and rotate the triangle out of plane and through 3-space, passing through π radians and landing, reflected, back in the plane.

If such a simple matter of imbedding can change a reflection into a rotation, what else is possible? Obviously when we say that there is or is not a continuous deformation of one object into another, we are not necessarily talking about the objects intrinsically. We are talking about their relation to ambient space. How much freedom could an ambient space allow?

Myriad sub-questions suggest themselves: what about the symmetries of polyhedra? Why do we restrict ourselves only to those symmetry operations which can be performed in $\mathbb{R}^3$? What about two surfaces, alike in Euler characteristic but differing in orientability? Is there a space in which a "rigid" transformation carries us from one to the other? What about surfaces like the sphere and the projective plane, which have the same local geometry, but differ in orientability and in Euler characteristic (if only by 1)? How much can we get away with, what ambient space do we need, and how many "niceness" conditions can we impose?

The following subsections explore briefly two of the above questions: that of polyhedra, and that of a transformation between the sphere and the projective plane. These, among other aspects of the project, are ones which I intend to explore in greater depth at a later date. The main focus of the present paper is on the detailed construction of several types of transformation from the Klein bottle to the torus. These two surfaces were a natural choice, because they are the simplest pair having the same Euler characteristic. Any work done on the Klein bottle and the torus bids fair to generalize to Klein bottles and tori of higher genus, another goal for work I will do in the future. Possibilities of such generalization are hinted at in this paper, but as yet my main focus has been to lay as firm a foundation as possible at the level of the transformation $K^2 \to T^2$.

A. Polyhedra. In this short discussion, let us restrict ourselves to polyhedra with triangular faces. The reason for this choice will shortly become apparent. Let **S** be the set of all possible arrangements of the faces[1] of a polyhedron P, subject to the requirement that for each member Q of **S,**

1. Any two faces of Q intersect only at one vertex or one edge.
2. Any two faces of Q are connected by a path of faces having a common edge.
3. Each edge belongs to exactly two faces.
4. The same number of faces meet at each vertex of Q.

Clearly the set of symmetry operations on P in $\mathbb{R}^3$ is a subset of **S**. Also contained in **S** are all possible reflections of P as well as numerous less familiar types of rearrangements. A reasonable way to interpret the members of **S** might be as the simplectic approximations to compact surfaces. If we further restrict to a subset **T** of **S** such that the number of faces joining at each vertex is constant from member to member of **T**, we then have a set of simplectic approximations to surfaces with the same Euler characteristic, such as the torus and the Klein bottle, as well as operations like rotations and reflections on these complices.

It would be interesting to construct some sort of ambient space in which the members of **T** or even of **S** could all be considered as symmetries of each other. Referring once more to our first example of the group D_3, it is intriguing to note that all six permutations of three objects are represented as symmetries in D_3, a state of affairs which does not obtain in any other hedral symmetry group of operations in $\mathbb{R}^3$, whether we are considering rearrangements of faces or vertices. In some sense, $\mathbb{R}^3$ is a maximal imbedding for the triangle. Are there maximal imbeddings for other simplectic complices? For example, the tetrahedron would be an interesting case to examine, since it generalizes the triangle in the sense that it is the simplex for $\mathbb{R}^3$.

For the hedra or gons which are not simplices, the story is probably more complicated. If we have a square and we exchange, say, the bottom two vertices, this transformation corresponds to neither a rotation nor a reflection, but a kind of twisting which is impossible for the triangle. This suggests that for the square, built of two 2-d simplices, or the octahedron, built of four 3-d simplices, or for any other figure reducible to more than one simplex, the maximal imbedding space is more difficult to find, and is probably not Euclidean. However the issue seems interesting and possibly useful. Looking at the problem from a slightly different vantage point, it is even possible that we could explain the absence of a non-orientable surface of Euler characteristic 2 by invoking some lack of twistability in the tetrahedron analogous to that in the triangle.

B. The Sphere and the Projective Plane. The direction of the previous discussion leads us naturally to more abstract types of transformations between

[1] Note that this might not necessarily be the best definition; we could also have used the vertices.

topological spaces, transformations not necessarily representable as permutations of hedral faces. There is neither a non-orientable surface of Euler characteristic 2 nor an orientable surface of Euler characteristic 1, so it is reasonable to pair S^2 and P^2 as surfaces we would like to transform into one another.

Our ambient space in the following construction is not a concrete physical space, but one which helps our transformation along by including some conventions in its structure. The basic space is itself a sphere having flowlines on its surface. The flowlines are a family of great circles all intersecting at the north and south poles, as shown in Figure 1.3:

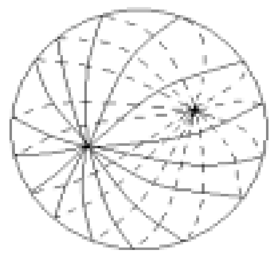


Fig. 1.3

In this ambient space surfaces are represented by figures drawn on the surface of the sphere, with the understanding that when the boundary of a figure intersects the same flowline in two different places, these two boundary points are to be identified. Thus the projective plane is represented as a 2-gon containing one of the regions of intersection of the great circles:

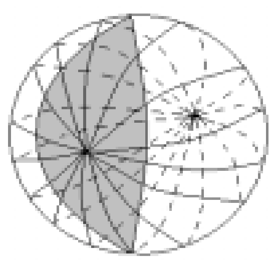


Fig. 1.4

The resulting identifications yield the usual schema for the projective plane. To obtain the sphere, we simply rotate this 2-gon so that it no longer contains any region of intersection:

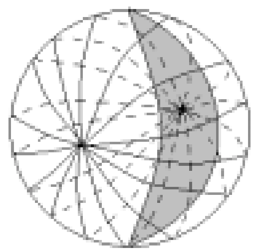


Fig. 1.5

I include this discussion of the transformation $P^2 \to S^2$ partly because it will be necessary in future research for generalizing my investigations to transformations of surfaces other than K^2 and T^2, but also because it serves as an introduction to the formalism used in the §2. The next five sections will be devoted to building a succession of representations of the transformation $K^2 \to T^2$.

§2. A Transcendental Transformation from K^2 to T^2

In what follows, we present an abstract construction of the transformation $K^2 \to T^2$. As mentioned above, the flavor of this construction is very similar to that of our work in §1B building $P^2 \to S^2$. This time our ambient space is a torus $S^1 \times S^1$, parametrized by φ and t, φ being the small circles and t the large ones. The flowlines are given by a transformation of t, dependent on φ:

$$s = \tan^{-1}(\cos\varphi \tan t), \quad 0 \leq \varphi < 2\pi, \quad -\pi \leq t < \pi,$$

where s is understood to be in the same quadrant as t when $\cos\varphi$ is positive, and in the negative of t's quadrant when $\cos\varphi$ is negative. Figure 2.1 shows the flowlines for several discrete values of t. The vertical direction represents t going from $-\frac{\pi}{2}$ to $\frac{3\pi}{2}$ at intervals of 0.2, while the horizontal direction represents φ between 0 and 2π.

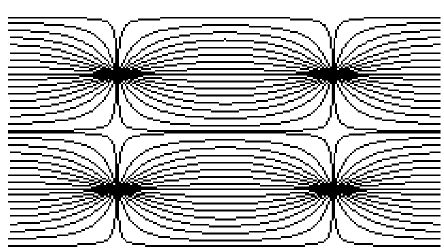


Fig. 2.1 Flowlines on flat torus

The interpretation behind s is the following: consider a unit circle C in $\mathbb{R}^3$, initially in the xy-plane. Fix an axis A through two antipodal points of C, say $(-1, 0)$ and $(1, 0)$, and begin rotating C about A. Parametrize its rotation by an angle φ measuring the from the xy-plane, and let φ run from 0 to 2π, so that C makes a full rotation about A. At each stage in the rotation, project C into the xy-plane.

Assume that the circle C is parametrized by a parameter t which is simply $-\pi \leq t < \pi$. This is the parametrization of C within whatever plane it happens to inhabit at some given stage φ_0 in its rotation. When we project C into the xy-plane, however, the parametrization t induces for the resulting ellipse will no longer move at a uniform speed; it will be slow near $(-1, 0)$ and $(1, 0)$ and fast in between.

These are the ingredients from which we get s. Specifically, suppose C is partway through its rotation, and we are in the plane described by the given $\varphi = \varphi_0$. Take the line L in this plane which makes an angle of t with A. Now project C into the xy-plane and take s to be the angle the projection of L makes with A. We now have a new parametrization of the unit circle in the xy-plane, given by $-\pi \leq s < \pi$. Rotation of C through 2π encompasses the whole horizontal length $\varphi \in [0, 2\pi)$ of Figure 2.1, since C is edge-on at two times during a complete rotation, corresponding to the fact that there are two regions of intersection in the flowlines.

The transformation $K^2 \to T^2$ is represented in this space by a cylindrical band drawn on the torus and allowed to move over its surface. The boundaries of the band are t-parameter lines, and the band moves in the direction pointed by the φ-parameter lines. One caveat applies: the coordinates of the boundaries must differ in φ by less than π radians. As in the ambient space for $P^2 \to S^2$, boundary points get identified when they intersect the same flowline in two different places. Thus when the band contains one of the regions where the flowlines intersect, the quotient space indicated by our convention is the Klein bottle, and when the band contains no such region, the quotient space is the torus.

It is the significance of s which seems to me the most illuminating aspect of this space. The intersections of the flowlines appear to be singular regions, but actually they are just caused by a projection of the circle C rotating in 3-dimensional space, a non-singular process. Translations of the boundaries of the cylindrical band through $\Delta\varphi$ along the φ-parameter lines correspond to rotation of the circle C through the same angle. In a very real sense, the rotation of C through π represents one of the boundaries of the cylindrical band being flipped over before it is identified with the other. This corresponds with the picture we have in our heads when we think of the construction of a Klein bottle. Picture a cylinder in $\mathbb{R}^3$, and imagine rotating one of its boundaries through an angle of π about an axis through two antipodal points. By doing this, we produce a region of self-intersection in the cylinder in $\mathbb{R}^3$, and when we identify the boundaries, we obtain a Klein bottle:

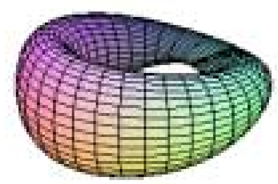


Fig. 2.2 Klein bottle immersion

What is interesting is that the boundary of the cylindrical band was somehow "stored" in a set of dimensions beyond the ambient space, where it was free to rotate. It became associated with the circle C, which had an abstract but nevertheless relevant and non-singular existence "beyond" the ambient transformation space, hence the nomenclature introduced in the title of this section.

The storage of some part of the Klein bottle/torus in a meta-space is one which we shall encounter again, but in a more concrete setting. This is the primary difficulty with the transcendental transformation: it is hard to pin down. It obviously has connections to covering spaces and monodromy groups, but does not exactly speak their language. Therefore it is our next aim to approach the transformation $K^2 \to T^2$ from a very concrete perspective, starting in $\mathbb{R}^4$ with an actual parametrization of the Klein bottle and examining precisely what must happen to it as we change it into T^2.

§3. Parametric Transformation from K^2 to T^2.

The following parametrization imbeds the Klein bottle in $\mathbb{R}^4$ (Do Carmo 436):

$$G(u,v) = \left\langle (r\cos v + a)\cos u,\ (r\cos v + a)\sin u,\ r\sin v\cos\frac{u}{2},\ r\sin v\sin\frac{u}{2} \right\rangle,$$

where $u, v \in [-\pi, \pi]$. Ignoring the fourth component yields the same $\mathbb{R}^3$ immersion as shown above in Figure 2.2, which was obtained by setting $r = 1$ and $a = 3$. The self-intersection occurs at $u = \pi$.

Our strategy in transforming K^2 into T^2 will be to make our parametrization "forget" about the self-intersection. This the only thing stopping our Klein bottle from being a torus, and we must devise a transformation that loses the information about how $u = \pi$ behaves. After the information is lost, the transformation will simply assume that $u = \pi$ behaves like all the other u-circles, and we will have a torus.

In order to lose information, we construct a singularity. The region of self-intersection must shrink down to a point, leaving the rest of the surface

2-dimensional. Since we are picturing this process happening in time, we add a third parameter named t to the function $G(u,v)$, where t runs from 0 to 1 and represents a smooth transformation of K^2 into the pinched Klein bottle. The value $t=0$ will correspond to K^2; $t=1$ to the pinched Klein bottle. The parameter t operates by making r a function varying with u and t as follows:

$$r(u,t) = (1-t)r_0 + tr_0 \cos^2 \frac{u}{2},$$

where r_0 replaces the constant r in Do Carmo's parametrization. At $t=1$, we ignore the fourth component to get the surface shown below in $\mathbb{R}^3$, again taking $r=1$ and $a=3$:

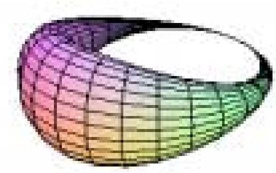


Fig. 3.1

At this point, however, our surface has lost its status as a Klein bottle; we have turned the self-intersection in $\mathbb{R}^3$ into a singular point. The information is lost. A reparametrization makes it clear that the surface in Figure 3.1 could just as easily be regarded as a pinched torus. Consider the patch

$$F(u,v) = \left\langle (r(u,1)\cos v + a)\cos u,\ (r(u,1)\cos v + a)\sin u,\ r(u,1)\sin v \left|\cos\frac{u}{2}\right|,\ r(u,1)\sin v \sin\frac{u}{2} \right\rangle,$$
$$u, v \in [-\pi, \pi],$$

which also parametrizes the surface in Fig. 3.1, but in such a way as to suggest that it is a pinched torus. Note that if preferred, the absolute value may be omitted from the parametrization by taking two patches instead of one:

$$F_1(u,v) = \left\langle (r(u,1)\cos v + a)\cos u,\ (r(u,1)\cos v + a)\sin u,\ r(u,1)\sin v \cos\frac{u}{2},\ r(u,1)\sin v \sin\frac{u}{2} \right\rangle,$$
$$u \in [0,\pi],\ v \in [-\pi,\pi]$$
$$F_2(u,v) = \left\langle (r(u,1)\cos v + a)\cos u,\ (r(u,1)\cos v + a)\sin u,\ -r(u,1)\sin v \cos\frac{u}{2},\ r(u,1)\sin v \sin\frac{u}{2} \right\rangle,$$
$$u \in [\pi, 2\pi],\ v \in [-\pi,\pi].$$

Just as we defined a smooth transformation from the Klein bottle to the pinched surface, we can define a smooth transformation from the present reparametrization of the pinched surface to a torus by extending our variation of r

with u and t. We have t run from 1 to 2, with 1 representing the pinched surface and 2 the torus. We replace $r(u,1)$ in the above patches with

$$r(u,t) = (t-1)r_0 + (2-t)r_0 \cos^2 \frac{u}{2}, \ t \in (1,2].$$

Now we have a more definite picture to work from. We know what happens to an actual imbedding, we know that there is a singularity, and we know exactly where and how this singularity appears. Of course, ideally we want a non-singular way to talk about the transformation $K^2 \to T^2$, which will inform our next approach. The present section, however, serves the same purpose as our observation in §1 that in the plane, the transformation between the two triangles in Figure 1.2 must be singular if it is to be continuous. The location and character of the singularity were closely related to the resolution of the transformation as a rotation in $\mathbb{R}^3$.

The next three sections will work toward constructing two different types of non-singular representations of the transformation $K^2 \to T^2$.

§4. A Fibration of Covering Spaces.

Our functions $F(u,v,t)$ and $G(u,v,t)$ treat t like a time variable. When we constructed the transformation $K^2 \to T^2$, the idea which guided us was, loosely speaking, an image of a Klein bottle cinching its region of self-intersection into a point, and then puffing this pinched place out into a tube so that it became a torus. It is reasonable, then, to make t a time-dimension and think of the functions $F(u,v,t)$ and $G(u,v,t)$ as coordinate patches on a 3-dimensional space immersed in $\mathbb{R}^4$. We will be interested in the level surfaces $t = t_0$ of the manifold, since these snapshots tell us exactly what happened as we passed from the Klein bottle to the torus.

Instead of regarding the actual surfaces as the $t = t_0$ slices, however, it is simpler to unwrap each stage into its universal covering space and use these as our $t = t_0$ slices. We are essentially describing a fibration $\mathbf{X}$ whose base space is t and whose fibers are the universal covering surfaces. From this vantage point, the transformation $K^2 \to T^2$ can be located at the level of the monodromy groups. It is very simple to state in this language how the $t = 0$ and $t = 2$ fibers differ: one of our monodromy group generators changes from a glide reflection at $t = 0$ to a translation at $t = 2$. Equivalently, we could say that along the canonical curves $u = (4n+1)\pi$, $n \in \mathbb{Z}$ in our tessellation, the orientation of the quotient-space identifications reverses. (Of course, our choice of $u = (4n+1)\pi$ was arbitrary; we could just as easily have taken $u = -(4n+1)\pi$).

What happened between $t = 0$ and $t = 2$? Let us look at how the fundamental region $u = [-\pi,\pi]$, $v = [-\pi,\pi]$ is affected as t runs from 0 to 2. As t increases from 0, the side edges of our fundamental region begin to shrink. To see this, suppose $t = t_0$ in $G(u,v,t)$, with t_0 strictly between 0 and 1. Then we can cut around the Klein bottle along the circle $u = \pi$ and then around $v = \pi$

to "unpeel" K^2 into a fundamental region we can draw on the plane, preserving the surface area of the Klein bottle. We lay the line $v = 0$ along the x-axis and $u = -\pi$ along the y-axis. For any given $t \in [0, 1]$, the following lines are then the upper and lower boundaries of our fundamental region, the left and right boundaries being the lines $x = 0$ and $x = 1$ ($u = -\pi$ and $u = \pi$) :

$$\text{Upper boundary}: \quad y = \pi\left((1-t)\,r_0 + tr_0\cos^2\left(\frac{2\pi x - \pi}{2}\right)\right)$$
$$\text{Lower boundary}: \quad y = -\pi\left((1-t)\,r_0 + tr_0\cos^2\left(\frac{2\pi x - \pi}{2}\right)\right).$$

A graph of the fundamental region appears below for $r_0 = 1$ and $t = 0.5$:

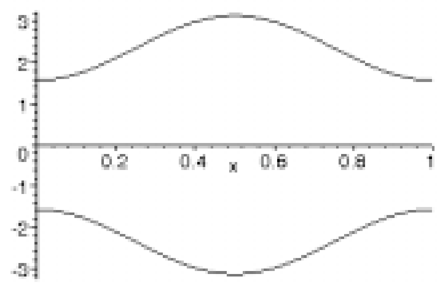


Fig. 4.1

We have also defined a coordinate transformation of the plane in general by specifying that the unit square be sent to the fundamental region in Figure 4.1. The coordinate transformation can be visualized by thinking of tiling a surface with fundamental regions like this one. The transformation is

$$T_t\,(x, y) = \left\langle x, \quad \pi(2y-1)\left[(1-t)\,r_0 + tr_0\cos^2\left(\frac{2\pi x - \pi}{2}\right)\right]\right\rangle.$$

We have constructed $T_t(x, y)$ in such a way that it represents an unrolling of the universal covering space at $t = t_0$ onto the plane, preserving area and length in the vertical direction. Since the Euclidean plane is not the covering space when t is not equal to 0 or 2, the fundamental regions appear distorted for $t \in (0, 2)$. Actually, the universal covering space is imbedded in $\mathbb{R}^3$ as the surface of revolution of the curve $y = (1-t)r_0 + tr_0\cos\left(\frac{2\pi x - \pi}{2}\right)$ about the x-axis, except that we must consider it to be an infinite-sheeted surface of revolution, so that the translation in the y-direction continues to be of infinite order.

When $t = 1$, the side edges of the fundamental region have shrunk altogether to zero, and our covering space is homeomorphic to an infinite chain of infinite-sheeted spheres. It is tessellated by 2-gons as its fundamental regions, and the orbit space of the action of our generators is a pinched torus, which is topologically equivalent to a sphere joined at the north and south poles. Indeed, one of our generators, the one that slides each 2-gon onto the next moving in the y-direction, is equivalent to the monodromy group generator operating on the sphere, with orbit space also the sphere. However, we also have another generator, the glide reflection in the x-direction, which now simply identifies the points to which the y-parameter lines have converged. Therefore it doesn't really matter whether it is a glide reflection or a translation at $t = 1$, and we can change it to a translation.

As far as radius $r(u, t)$ and surface area of the torus are concerned, the function $F(u, v, t)$ is just $G(u, v, t)$ "run in reverse" over $t \in (1, 2]$. Hence it takes over exactly where $G(u, v, t)$ left off, and transforms our fundamental region into a square again, in a manner completely analogous to that just described for the transformation from square to 2-gon induced by $G(u, v, t)$. Since we changed the glide reflection to a translation at $t = 1$, the end result of the transformation induced on the covering space by $F(u, v, t)$ is a monodromy group whose generators are two perpendicular translations and whose orbit space is a torus.

It is interesting to analyze in parallel the effect of the transformation $K^2 \to T^2$ on the covering spaces and on the monodromy groups of the intermediate stages. Several nice properties emerge, as well as some interesting discrepancies. The most notable of these latter is the fact that while the covering spaces became singular at $t = 1$, the monodromy groups did not. At $t = 1$, the determinant of the Jacobian of $T_1(x, y)$ vanishes at $x = n$, $n \in \mathbb{Z}$, but throughout the transformation, the monodromy groups are always free groups of rank 2. They are not all Abelian, so they cannot all be isomorphic, but within the part of the transformation induced by G they are isomorphic, and the monodromy group for the $t = 1$ covering space is isomorphic to the monodromy groups for $t \in (1, 2]$. Also, the fibration $\mathbf{X}$ is nearly a fibration with fiber $\mathbf{X}_{t=0}$, for with the exception of the slice $t = 1$, all the fibers are homeomorphic to one another.

From this discussion it emerges that there are two main defects in the fibration transformation as it stands at present: first of all, we still have singularities, and second, the monodromy groups make an abrupt change from non-Abelian to Abelian. The first problem will be solved in §5; the second will have to wait until §6. Despite the fact that singularities are the problem we have focused on, they are relatively easy to deal with compared to the issue of the monodromy groups, which turns out to be deeper. In a sense, the abrupt change in the monodromy groups does not matter, because it occurs at the time when the universal covering space is singular, and as we have already observed, it makes no difference at that stage whether individual points are identified by means of a glide reflection or a translation. But if we are about to resolve the singularities, soon we will no longer have this excuse to lean on, and we will need to confront the difficulty head-on, devising a way for the monodromy groups somehow to change gradually. Obviously we will have to look at things from new perspective to do this, for it is difficult to see at this point exactly how we can make the change be gradual in any sense.

However our current concern is the singularities. We can be fairly sure they are not necessary, since the singularities in the transcendental transformation in §2 come with their own built-in resolution as a result of the interpretation of the flowlines. Our task then is to find a non-singular representation closer to our current constructions.

§5. Resolving the Singularities in X.

In what follows we show that the singularities in a space based on **X** can be resolved in $\mathbb{R}^4$. The space **Y** we construct will have no singularities, and will run over all values of t, whereas **X** was only defined for $t \in [0, 2]$. The projection of **Y** into $\mathbb{R}^3$ will represent **X** and the fibration transformation $K^2 \to T^2$ of §4.

Before proceeding, however, we make some convenient changes to the coordinates u, v, and t. Both u and v are rescaled by $x = \frac{u}{2\pi}$, $y = \frac{v}{2\pi}$. Additionally we will adjust our coordinate patch so that t running from 0 to 1 represents the entire transformation $K^2 \to T^2$, and we will build **Y** in such a way that only one patch is necessary. Let us call this patch $P(x, y, t)$, its components being P^1, P^2, P^3, and P^4. For simplicity, we talk about **X** as well as **Y** in terms of the adjusted coordinates x, y, and t described above, even though we actually did not formulate it in this way. Thus we speak as though the t-slice in **X** where G and F join is at $t = \frac{1}{2}$, and as though the fundamental region at $t = 0$ is described by $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, $y \in \left[-\frac{1}{2}, \frac{1}{2}\right]$.

In order to resolve our singularities, we look at what happened to the $t = t_0$ slices in the previous section. We saw there that the lines described in our current coordinates by $x = \frac{2n-1}{2}$, $n \in \mathbb{Z}$ shrank, eventually reaching a singular stage, and then grew again. We must now view this shrinking as the result of a projection. The $t = \frac{1}{2}$ slice we have been visualizing in **X**, the one whose fundamental region is a 2-gon, is not the real $t = \frac{1}{2}$ slice, but merely its shadow in an unflattering light. The real $t = \frac{1}{2}$ slice, in **Y**, is not tesselated by 2-gons at all. It is tesselated by squares, but it is twisted so that at $x = \frac{2n-1}{2}$, it does not live in the P^2-dimension. We need to use the P^4-dimension in order to see where the seemingly lost length went when the $x = \frac{2n-1}{2}$ edges of our fundamental region appeared to be shrinking.

Having now an intuitive picture of what happened, we wish to begin parametrizing what we have seen. First of all, we organize everything we know so far about what the parametrization should be like. We know that we want it to extend over all values of x, y, and t, which will force us to deviate slightly from the direction pointed out by our work in §3 and §4. Currently, our function $r(u, t)$ poses a problem for any natural extrapolation in the t-direction, owing to the manner in which it suggests y and t are related in the $x = \frac{2n-1}{2}$ ($u = (2n-1)\pi$) slices of **X**. If we were to pursue the implications of $r(u, t)$, we would send the line $x = -\frac{1}{2}$, $y = y_0$ to $P^1 = -\frac{1}{2}$, $P^2 = -2y_0t + y_0$, $P^4 = 0$, and the line $x = \frac{1}{2}$, $y = y_0$ to $P^1 = \frac{1}{2}$, $P^2 = -2y_0t + y_0$, $P^4 = 0$. Such an assignment would mean that as t grew, the image of the fundamental region $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, $y \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ would expand without bound in the y-direction. As the t-slices of our image space are supposed to be universal covering spaces of the torus, the Klein bottle, and the pinched surface, this extrapolation of the domain would have no relevance to the present problem. It would be best, therefore, if we sent

$x = \frac{2n-1}{2}$, $y = y_0$ to something periodic in t with period 2, so that t running from 0 to 1 would represent exactly one transformation $K^2 \to T^2$. Thus we want to send the line $x = \frac{2n-1}{2}$, $y = y_0$ to some variety of cosine curve in t with period 2 in the plane $P^1 = \frac{2n-1}{2}$ and, for now, $P^4 = 0$. As yet, we have done nothing to resolve the singularity at $t = \frac{1}{2}$, but sending $y = y_0$ to a cosine curve will allow us to run t forever, with the effect that on top of the $P^3 = 1$ face of the image of the cube $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, $y \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, $t \in [0, 1]$, we will set another "cube" representing a transformation from T^2 to K^2, and on top of that one from K^2 to T^2, and so on and so on. (Of course, the "cubes" extend below $t = 0$ as well.) Note that we can also make this t-extrapolation serve to extrapolate the $x = \frac{n}{2}$ slices over all y-values as well by giving our cosine curve an amplitude of y.

Such an extrapolation of the space $\mathbf{X}$ is beneficial for two primary reasons, the first being that it simply makes $\mathbf{Y}$ more general than $\mathbf{X}$, and the second being that its periodicity would even allow us identify a fundamental region in $\mathbf{Y}$ and take the quotient space on this level as well as the original level. While this latter avenue is one we will not pursue in this paper, it may be useful in future research. However the most important fact about $\mathbf{Y}$ is what we know about the $t = n$, $n \in \mathbb{Z}$ slices: they must be planes. In particular, $P(x, y, 0) = \langle x, y, 0, 0\rangle$ and $P(x, y, 1) = \langle x, y, 1, 0\rangle$, or, more generally, $P(x, y, n) = \langle x, y, n, 0\rangle$. We also know something about the $t = \frac{1}{2}$ slice of $\mathbf{Y}$; namely, that it must behave something like a helicoid of period 2 in P^1, P^2, and P^4, at least between $x = -\frac{1}{2}$ and $x = \frac{1}{2}$. For we know that it twists fully out of the P^2-dimension and into P^4 at $x = -\frac{1}{2}$ and $x = \frac{1}{2}$, and that it lies flat in P^2 at $x = 0$. However, we quickly realize that the $t = \frac{1}{2}$ slice cannot behave like a helicoid for x-values less than $-\frac{1}{2}$ or greater than $\frac{1}{2}$, because if it did, we would find it difficult or impossible to arrange things so that $\mathbf{Y}$ was periodic in the x-direction. In particular, we would have problems introducing a smooth interpolation between the helicoid and the planes at $t = 0$ and $t = 1$ while still maintaining a periodicity of 2 over x. So we settle instead on a surface similar to a helicoid, but which doubles back on itself when it has made a half rotation instead of continuing to corkscrew around in the same direction. Call this surface a ruffleoid. It is shown in Fig. 5.1, and can be parametrized by $P^1 = x$, $P^2 = y\cos(\frac{\pi}{2}\sin\pi x)$, $P^3 = \frac{1}{2}$, and $P^4 = y\sin(\frac{\pi}{2}\sin\pi x)$.

How shall we interpolate between the planes and the ruffleoid? This is not difficult. The factor of $\frac{\pi}{2}$ in front of $\sin\pi x$ is what controls the angle from the P^2-axis the ruffleoid rotates through before it doubles back, because it is the amplitude of the argument and therefore constrains the argument within $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ from the P^2-axis. Thus it is possible merely by varying this amplitude to construct a ruffleoid ruffling through any angle we choose. Let us call such a general ruffleoid a θ-ruffleoid, where θ is the amplitude of the argument. Since a plane is a 0-ruffleoid, we want to make θ a periodic function of t with amplitude $\frac{\pi}{2}$, and we want it to be zero at $t = 0$, maximized at $t = \frac{1}{2}$, and zero at $t = 1$.

We also would like it to have period 1, so we take $\theta = \frac{\pi}{2}\sin^2 \pi t$.

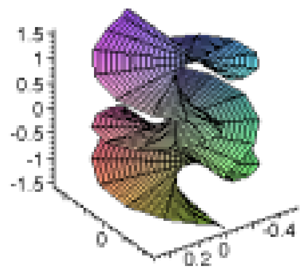


Fig. 5.1 Ruffleoid.

Putting all the pieces together, we can write our coordinate patch

$$P(x,y,t) = \left\langle x, \quad y\cos\left(\frac{\pi}{2}(\sin^2 \pi t)(\sin \pi x)\right), \quad t, \quad y\sin\left(\frac{\pi}{2}(\sin^2 \pi t)(\sin \pi x)\right)\right\rangle.$$

This resolution and imbedding can furnish us with a representation of the transformation $K^2 \to T^2$ free of singularities. We have a space $\mathbf{Y}$ which is non-singular and which projects into $\mathbb{R}^3$ as a smoothed-out and extended version of $\mathbf{X}$; now the key is in the way we get the orbit space from the covering surface. The method we will use, which I call decomposition, takes its inspiration from the transcendental transformation and the idea of a meta-space. It keeps the same monodromy group we used on $\mathbf{X}$, a glide reflection and a translation up until $t = \frac{1}{2}$, and then two translations after that. The group actions at each value of t are understood to reside in that particular t-slice, although as yet we make no identifications. First we must perform some projections, because now that we have resolved the singularities, we must once again achieve a state of affairs in which it does not matter that the glide reflection makes an abrupt change to a translation. Since we know how $\mathbf{Y}$ projects into the first three dimensions of $\mathbb{R}^4$, let us perform this projection first. We ignore P^4, obtaining our smoothed version of $\mathbf{X}$, and at each value of t take the orbit space. The result is the same succession of surfaces we had in §§3 and 4.

Meanwhile, we also project $\mathbf{Y}$ into the 3-dimensional space spanned by P^1, P^3, and P^4. At $t = 0$ and $t = 1$, we only get a line, but as t increases, we get something which looks rather like a bow-tie between $x = -\frac{1}{2}$ and $x = \frac{1}{2}$, and is periodic outside that range (see Fig. 5.2, graphed for $t = \frac{1}{2}$). Our interpretation is to be that this second projection exactly recaptures the length and surface area which appear to be lost in the first projection. In this sense the two projections are dual to each other. This being the purpose, let us define that in the $P^1P^3P^4$ projection, we remove any point or set of points where the projection is not rank 2, since at these places, no length or area is lost in our first projection. Therefore we remove the entire projection at $t = 0$ and $t = 1$, saying that it is null at these t-values, and at every stage we remove the points $x = n$, $n \in \mathbb{Z}$. Then we perform the identifications indicated by the monodromy groups as projected into the present three dimensions, and we are left with a pinched surface missing its pinch-point. The advantage we gain is that in the $P^1P^3P^4$ projection, our orbit space is topologically a cylinder, no matter what the value of t (except at $t = 0$ and $t = 1$). Thus it doesn't matter at all how the identifications work, or whether they are the result of glide reflections or translations. The change in the monodromy group will not materially affect either component of the decomposition. On the other hand, the character of the components as dual to each other allows us reasonably to claim that this transformation cannot strictly be considered singular.

What happened here was that we used a new way of losing information in the process of our transformation. Rather than simply eating up the information in a singularity, we stored it somewhere where it was irrelevant. This technique is one which looks promising for surfaces of higher genus. However we would still like to talk about the transformation in terms of covering spaces and monodromy groups. If we could do this in a non-singular, non-discrete way, we would not even need to worry very much about the orbit spaces in the intermediate stages of the transformation; they might have to be very strange and probably singular at points, but we would have a complete description of the transformation without referring to them, and so this need not unduly trouble us.

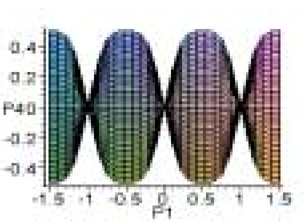


Fig. 5.2

§6. Non-Vanishing Monodromy Groups on a Non-Singular Covering Space.

In this section, our goal will be to lay the framework for representing our transformation at each stage by a freely generated rank 2 monodromy group acting on a non-singular 2-dimensional covering surface. We want all identifications that occur to be the result of taking the quotient space; in other words, we do not want monodromy groups which identify any points in the covering space beforehand. Our covering space will be assumed always to be a plane, and the monodromy groups between $t = 0$ and $t = 1$ must represent a continuous deformation of the monodromy group for the Klein bottle into that for the torus.

At $t = 0$, our monodromy group M is generated by two maps: the glide reflection $g : (x, y) \to (x+1, -y)$ and the translation $l : (x, y) \to (x, y+1)$. At $t = 1$, both our generators are translations; call them g^*: $(x, y) \to (x+1, y)$ and $l^* = l$, and, in accordance with this notation, call the group they generate M^*. The basic problem we face in transforming M into M^* is that at $t = 0$, the Jacobian of g is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, while at $t = 1$, g^* has Jacobian $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. How can we transform a mapping whose Jacobian has determinant -1 smoothly into a mapping with $\det_J = 1$? We must pass through a map with Jacobian 0. But as things stand now, this means that our generator is at some stage no longer a bijection. Our task must be to circumvent this issue in some way.

What we have set out to do looks impossible. It would seem at first glance that either the transformation from g to g^* must have a singularity, or it must be discrete. But consider our example in §1, a reflection in the plane. What the Flatlanders declared to be impossible turned out to be very easy in our universe. How can we apply such an idea to the transformation from g to g^*? This transformation also appears to be necessarily either discrete or singular. Can we somehow "rotate" g into g^* in some space?

Before we proceed, we must be careful of what we are saying. The fact that the transformation from g to g^* appears discrete or singular is not directly related to the fact that g itself is a reflection and therefore a discrete or singular transformation. Of course, there is a relationship in the sense that it is g's status as a reflection which prevents it from being easily transformed into g^*, but the particulars of why g and g^* are dissimilar do not concern us. It is the mere fact of their difference in which we are interested. We are moving up a level in the transformations we construct. Presently it appears that we need a transformation between transformations, and this must necessarily take place in the space of transformations rather than in any more concrete space. Thus in order to transform g into g^*, we do not suggest that a rotation of the covering surfaces in some arcane ambient space will solve our problems. Rather, we must find a way to describe mappings of the plane in such a way that g and g^* can be considered to be parallel but pointing in opposite directions on a line in transformation space. There must be another class of mappings which can be regarded as being in some sense perpendicular to g and g^*, and we must rotate g towards g^* through the space spanned by g and by one of the perpendicular transformations.

The key to constructing such perpendicular transformations lies in higher-order x and y dependence. We must generalize the ideas encapsulated in translations and glide reflections to make them hold for non-linear actions. First of all, each of g, g^*, and l are rank 2 mappings of the plane; this is the most important feature to generalize. Also, each component of each of these mappings can be considered as a function of both variables, x and y, where it happens that in g, g^*, and l, the first component depends only on x and the second only on

y. At this point, let us change our notation to the more appropriate $g(x, y, t)$, $l(x, y, t)$, where $g(x, y, 0)$ is our original g, $g(x, y, 1)$ our original g^*, and similarly with $l(x, y, t)$. To help our intuition along, let us regard each component of each of the mappings as a surface in 3-dimensional space. For example, if we are considering our original g, then $g_1(x, y) = x + 1$ and $g_2(x, y) = -y$, both planes. Each level slice of the plane described by g_i corresponds to a set of points in the xy-plane to which g assigns the same ith coordinate. Obviously, every point in the xy-plane must belong to some level slice of g_1 and to some level slice of g_2. In order to be sure that our monodromy mapping g does not identify points before we take the orbit space, we must check that no level slice of g_1 intersects any level slice of g_2 more than once. This is because a level slice of g_i represents a set of points g_i cannot distinguish between. If we do not want points identified, we must be sure that g_2 can distinguish between any pair of points g_1 fails to differentiate, and vice versa.

Looking at things in this way, opportunities for generalization are obvious. For instance, there is no reason why we have to restrict the components to planar surfaces; perhaps one or more of our components is a non-linear function, and has a curved surface as its graph. Further, what if we want one of our components to be a surface which is rank 2 but not representable by a single coordinate patch? Essentially this would mean that one of the components of our monodromy mapping was a multivalued function, but multivalued in a nice and well-behaved way, and in a way that could reasonably be claimed as a generalization of conventional monodromy mappings. It is interesting to note that, loosely speaking, we add dimensions to the Jacobian of our generator by making one or both of its components multivalued. The Jacobian is no longer a single number.

Exactly how much generalization is necessary to construct $g(x, y, t)$ I do not yet know, although I suspect that multivalued functions may be necessary. The following proposition narrows down our field of inquiry somewhat:

Proposition 6.1. Suppose that $g(x, y, t)$ is differentiable in x and y, g_1 depends only on x and g_2 depends only on y, g_1 and g_2 are piecewise monotonic in x and y, respectively, and, fixing (x, y) equal to any (x_0, y_0), $\frac{dg_1}{dx}(x, t)|_{x=x_0}$ and $\frac{dg_2}{dy}(y, t)|_{y=y_0}$ are continuous and piecewise monotonic in t. Then there must be some t_0 for which $g(x, y, t_0)$ is not 1-1.

We prove this statement by first showing that if g_1 or g_2 has an extremum, then g cannot be 1-1, and then showing that g_1 or g_2 must have an extremum for some value of t.

Proof. 1. (If g_1 or g_2 has an extremem, g is not 1-1). We can see why an extremum causes problems by looking at the level slices. Suppose that at some $t = t_0$, g_i has a local extremum at a point a. If it is a maximum, choose two points p and q for which $p < a < q$, $g_i(p, t_0) = g_i(q, t_0) < g_i(a, t_0)$, and the slope of g_i is positive in the interval (p, a) and negative in the interval (a, q). (The intermediate value property insures that this is possible.) If it is a minimum, choose p and q such that $p < a < q$, $g_i(p, t_0) = g_i(q, t_0) > g_i(a, t_0)$, and the slope of g_i is negative in the interval (p, a) and positive in the interval (a, q). From now on, however, assume that the extremum is a maximum. The proof is analogous for a minimum.

We then use the intermediate value property of g_i to assert that for any value c_1 such that $g_i(p, t_0) < c_1 < g_i(a, t_0)$, there exists a point p_1 for which $p < p_1 < a$ and $g_i(p_1, t_0) = c_1$. Again by the intermediate value property, there exists a point q_1 such that $a < q_1 < q$ and $g_i(q_1, t_0) = c_1$.

If we repeat the process and choose c_2 such that $g_i(p_1, t_0) < c_2 < g_i(a, t_0)$ and apply the intermediate value theorem again to get points p_2 and q_2 such that $p_1 < p_2 < a < q_2 < q_1$ and $g_i(p_2, t_0) = g_i(q_2, t_0) = c_2$, we will have begun a succession of c_k whose corresponding p_k and q_k grow arbitrarily close to a as k increases. The result is that an uncrossable line develops in our pattern of level slices of g_i. We know that p and q belong to the same level slice of g_i, as do p_1 and q_1, and p_k and q_k in general. The schematic diagram below shows intuitively what is happening near our extremum, with each level slice represented by a different type of dotted line:

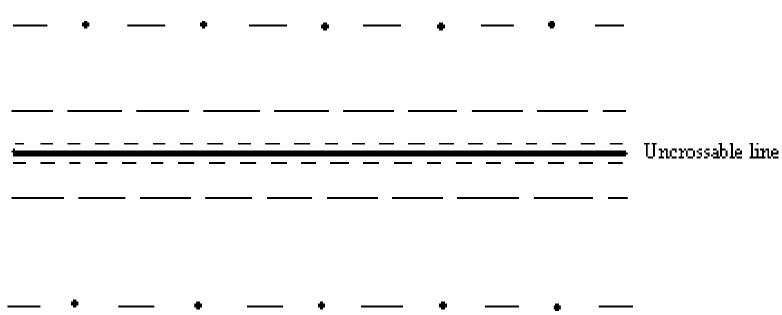


Fig. 6.2

The vertical direction represents whichever variable (x or y) g_i depends on; the remaining variable runs in the horizontal direction. The instant g_j crosses the heavy line in Fig. 6.2, it begins intersecting level slices of g_i it has already intersected, and g is no longer 1-1. But since g_1 depends only on x and g_2 only on y, the level slices of each are sets of infinite straight lines, with g_2's level slices perpendicular to g_1's. Thus there is no alternative but for g_j to cross g_i's uncrossable line, and g cannot be 1-1.

2. (There is an extremum): Since $\frac{dg_1}{dx}(x,t)|_{x=x_0}$ and $\frac{dg_2}{dy}(y,t)|_{y=y_0}$ are continuous and piecewise monotonic in t, $\det_{J(g)}(x_0, y_0, t)$ is continuous and piecewise monotonic in t for any (x_0, y_0). For all $(x, y) \in \mathbb{R}^2$, $\det_{J(g)}(x, y, 0) = -1$ and $\det_{J(g)}(x, y, 1) = 1$. Therefore for every (x_0, y_0), we can find some t_0 for which $\det_{J(g)}(x_0, y_0, t_0) = 0$, and for which there is some $\varepsilon > 0$ s.t. if $t \in (t_0 - \varepsilon,$ $t_0)$, $\det_{J(g)}(x_0, y_0, t) < 0$, and if $t \in (t_0,\, t_0 + \varepsilon)$, $\det_{J(g)}(x_0, y_0, t) > 0$.

As we pass through t_0, one or both of $\frac{dg_1}{dx}(x,t)|_{x=x_0,\ t=t_0}$ and $\frac{dg_2}{dy}(y,t)|_{y=y_0,\ t=t_0}$ is equal to zero, and one of them changes sign. Since we know that at some point $\frac{dg_2}{dy}(y,t)|_{y=y_0}$ changes sign, let us suppose that it does so now, although the proof would be analogous if we assumed that $\frac{dg_1}{dx}(x,t)|_{x=x_0}$ changed sign. Additionally let us suppose that $\frac{dg_2}{dy}(y,t)|_{y=y_0}$ changes from negative to positive. However this choice also is not essential to the structure of the proof.

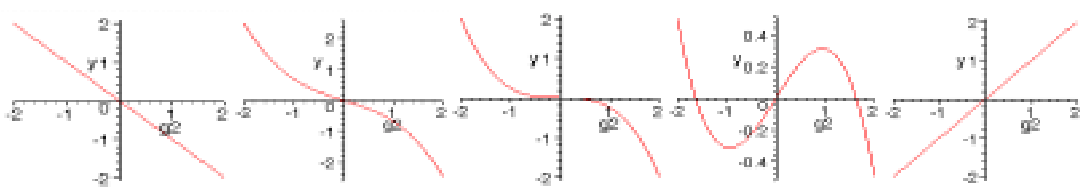
With regard to the constant-t slices of $\frac{dg_2}{dy}(y,t)|_{t\in(t-\varepsilon,\ t+\varepsilon)}$, one of two things happens. Either the slices for $t \in (t_0-\varepsilon,\ t_0)$ all have $\frac{dg_2}{dy}(y,t) \le 0\ \forall y$ and those for $t \in (t_0,\ t_0+\varepsilon)$ all have $\frac{dg_2}{dy}(x,t) \ge 0\ \forall y$, or there is at least one $t_1 \in (t_0-\varepsilon,\ t_0+\varepsilon)$ such that $\frac{dg_2}{dy}(y,t)|_{t=t_1}$ is negative for some y and positive for some y. The first alternative is impossible. The reason for this is that at most, a t-slice $\frac{dg_2}{dy}(y,t)|_{t=\text{constant}}$ can be zero at isolated values of y; if $\frac{dg_2}{dy}(y,t)|_{t=\text{constant}}$ were zero for an interval (a,b) of y-values, then g would send the 2-dimensional region $x \in (-\infty,\infty)$, $y \in (a,b)$ to a 1-dimensional region, violating our requirement that no identifications be performed except those involved in taking the orbit space. Thus when t passes through t_0, the first of our two alternatives would violate the intermediate value theorem, since for infinitely many values of y, $\frac{dg_2}{dy}(y,t)$ would have to jump from negative to positive values.

Hence t_1 exists, and because g_1 and g_2 are differentiable, their derivatives have the intermediate value property, so there is an extremum on the $t = t_1$ slice. ■

On the other hand, we can show by example that a multivalued mapping of the kind described above will work. The multivalued mapping I used for $g(x,y,t)$ is shown below, written in terms of y since this is the most convenient form:

$$y = (2t-1)g_2 + (t^2-t)g_2^3$$

A few graphs are sufficient to show what happens to g_2 over t. The t-value of each graph is listed beneath it.



0 0.25 0.5 0.75 1

Fig. 6.1

§7. Goals for the Future.

Obviously this project leaves much room for much further exploration. Some aspects I intend to add or to expand upon are briefly discussed in this section.

The first priority is of course to finalize the results of §6; namely, to determine the least possible amount of generalization necessary in order to construct continuously changing, non-singular monodromy groups for the transformation $K^2 \to T^2$. I am currently in the process of generalizing Proposition 6.1 to cases when there is cross-dependence; this would either show that multivalued functions are necessary, or at least give us a new lower bound on the complexity of dependence necessary in order to acheive the desired end. If multivalued functions are not necessary, then the next step would be to find an example of a single-valued function which works, and, if possible, to characterize the class of all single-valued functions we could have used.

Another goal would be to clarify the relations among various different representations of $K^2 \to T^2$ constructed in this paper. For example how, precisely, is the abstract approach typified in the transcendental transformation related to the monodromy approach, or to the decomposition approach? Also, it would be interesting to analyze more closely the character of the orbit space, in particular under the generalized monodromy approach. To do this, we would need to introduce concepts from dynamical systems, since the orbit of a point now consists of the iterates of a non-linear function.

The largest goal, as has already been mentioned, is for the generalization of this project to other compact surfaces than the torus and the Klein bottle. The immediate next step would be to tackle the two-holed torus and two-holed Klein bottle. In the monodromy approach, this would necessitate the use of Fuchsian groups on a hyperbolic covering space tesselated with octagons meeting eight at a vertex. Other related possibilities for expansion and extension of scope have already been discussed in §1.

Works Cited.

Do Carmo, Manfredo P. *Differential Geometry of Curves and Surfaces.* Englewood Cliffs: Prentice-Hall, Inc., 1976.