RICHARD EVAN SCHWARTZ

# Pappus's theorem and the modular group

*Publications mathématiques de l'I.H.É.S.*, tome 78 (1993), p. 187-206
<http://www.numdam.org/item?id=PMIHES_1993__78__187_0>

# PAPPUS'S THEOREM AND THE MODULAR GROUP

*by* Richard SCHWARTZ

## 1. Introduction

Pappus's theorem is as old as the hills. It refers to the configuration of points and lines shown in Figure 1.1.

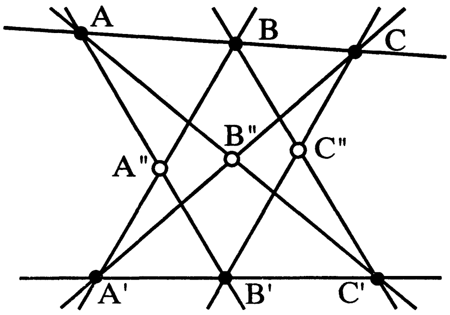


Fig. 1.1

*Pappus's Theorem. — If the points* A, B, *and* C *are collinear and the points* A′, B′ *and* C′ *are collinear, then the points* A″, B″ *and* C″ *are collinear.*

A slight twist makes this theorem new again. The twist is to iterate, and thereby treat Pappus's theorem as a dynamical system.

Pappus's theorem may be considered as a dynamical system defined on objects called *marked boxes*. Essentially a marked box is a collection of points and lines in the projective plane **P** which comprises the initial data for Pappus's theorem. (See § 2.2 for a precise definition.) When Pappus's theorem is applied to a marked box, more lines and points are produced. These new lines and points may be used to form new initial data for the theorem, and so on.

In fact, one can construct an entire group G of operations to perform on marked boxes. Algebraically G is the modular group, $\mathbf{Z}/2 * \mathbf{Z}/3$, and is generated by the simple operations alluded to above. (For a precise definition of the operations, see § 2.3.) The orbit $\Omega$ of a marked box is an infinite collection of marked boxes, nested inside and outside of each other in a pattern encoded by the modular group.

The modular group makes a second entrance into Pappus's theorem, as the group of projective symmetries of $\Omega$. Projective transformations are analytic diffeomorphisms of $\mathbf{P}$ which take collinear points to collinear points. Projective dualities are analytic diffeomorphisms between $\mathbf{P}$ and its dual space of lines $\mathbf{P}^*$ which take collinear points to coincident lines. Projective transformations and dualities together generate the Lie Group of projective symmetries of $\mathbf{P}$. It turns out that the projective symmetry group $\overline{\mathrm{M}}$ of $\Omega$ (those projective symmetries which permute the marked boxes of $\Omega$) is again algebraically the modular group. The group of projective symmetries and the group of operations commute, and so there are two commuting modular group actions on $\Omega$. As the original marked box varies, the orbit varies, along with the two group actions. Everything varies smoothly, thereby producing two commuting modular group actions on the space of all marked boxes.

There is a certain geometric condition one can put on marked boxes, called *convexity*. The set of convex marked boxes forms an open subset of the set of all marked boxes. For these marked boxes, there is a certain amount of geometry and topology associated to the group actions described above.

In the convex case, there is a fractal curve associated to the orbit $\Omega$. Certain " distinguished " points of the marked boxes in $\Omega$ are dense in a topological circle $\Lambda$, which is generally non-smooth. Dually, certain " distinguished " lines of the marked boxes in $\Omega$ are dense in a topological circle L of lines. The circles $\Lambda$ and L are " self-projective " curves in the sense that they are preserved (or swapped) by the modular group of projective symmetries of $\Omega$. In a certain sense, they are analogues of quasi-circles in the projective plane. Figure 1.2 shows certain of the convex marked boxes in an orbit. Both $\Lambda$ and L are " hinted at " in the picture.

Projective symmetries induce analytic diffeomorphisms of the projective tangent bundle $\mathscr{P}$ of the projective plane. When the original marked box is convex, the symmetry group $\overline{\mathrm{M}}$ gives a discrete group action on $\mathscr{P}$. In fact this group action has a domain of discontinuity and a corresponding quotient 3-manifold. More precisely, let $\mathscr{N} \subset \mathscr{P}$ denote those pairs $(p, l)$ such that $p \notin \Lambda$ and $l \notin \mathrm{L}$. The fixed points of $\overline{\mathrm{M}}$ are dense in, and contained in, the complement of $\mathscr{N}$, and $\mathscr{N}$ itself is the domain of discontinuity for $\overline{\mathrm{M}}$. The quotient $\mathscr{E} = \mathscr{N}/\overline{\mathrm{M}}$ is a three-dimensional analytic manifold with the homotopy type of the trefoil knot complement in the three sphere.

The manifold $\mathscr{E}$ is a " $(\mathscr{G}, \mathscr{X})$-manifold " in the sense of [T]. Here $\mathscr{G}$ is the Lie Group of projective symmetries, and $\mathscr{X}$ is the universal to the projective tangent bundle $\mathscr{P}$. There is a classical geometric structure on the trefoil knot complement coming from its description as $SL_2(\mathbf{R})/SL_2(\mathbf{Z})$. In the case of a " totally symmetric " marked box, one essentially recovers this structure. However, as the marked box varies, one obtains a 2-parameter family of " exotic " geometric structures on the (homotopy) trefoil knot complement.

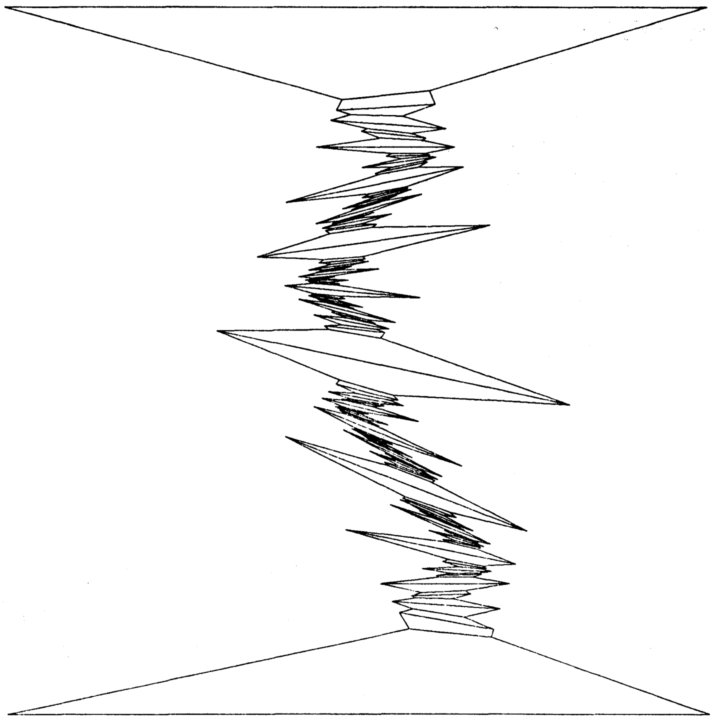

FIG. 1.2

In sum, each initial choice of lines and points involved in Pappus's theorem (subject to the convexity condition) gives rise to the following objects:

1. an infinite collection $\Omega$ of marked boxes, which is indexed by the modular group, and which has the modular group as its group of projective symmetries;
2. a pair of (usually fractal) curves $\Lambda \in \mathbf{P}$ and $L \in \mathbf{P}^*$ which are invariant (or swapped) under a modular group of projective symmetries;
3. a representation $\overline{M}$ of the modular group into the Lie Group of symmetries of the projective tangent bundle $\mathscr{P}$;
4. a (usually " exotic ") $(\mathscr{G}, \mathscr{X})$-structure $\mathscr{E} = \mathscr{N}/\overline{M}$ on the (homotopy) trefoil knot complement.

This paper is self-contained, but does assume some basic hyperbolic and projective geometry. The hyperbolic geometry can be found in [B], and the projective geometry, as well as an excellent treatment of Pappus's theorem, can be found in [H].

Here is a list of contents:



I would like to thank Hsin-Ta (Frank) Cheng, Peter Doyle, Bill Goldman, Jun Hu, Linda Keen, and Dennis Sullivan for helpful and interesting conversations concerning this work. I would also like to thank the IHES, and the Einstein Chair at CUNY, for their support during the genesis of this paper.

## 2. Iterating Pappus's Theorem

### 2.1. The Projective Plane

As usual, we will take the projective plane to be the space $\mathbf{P}$ of one-dimensional subspaces of $\mathbf{R}^3$. The ordinary plane $\mathbf{R}^2$ sits naturally as a subset of $\mathbf{P}$, and consists of all one dimensional subspaces which do not lie on the $xy$-plane of $\mathbf{R}^3$.

Projective transformations of $\mathbf{P}$ are just projectivized linear transformations of $\mathbf{R}^3$. The group of projective transformations is naturally isomorphic to $SL_3(\mathbf{R})$. There is a unique projective transformation which takes one specified set of 4 (ordered) general position points to another such set.

$\mathbf{P}^*$ will denote the space of lines of $\mathbf{P}$. This set is naturally in bijection with $\mathbf{P}$. A duality is just a projectivized linear transformation from $\mathbf{R}^3$ to its dual space. More concretely, a duality is a homeomorphism $\mathbf{P} \to \mathbf{P}^*$ which takes collinear points to coincident lines. As a special case, a duality which is induced by a positive definite inner product on $\mathbf{R}^3$ is called a polarity. Each duality $\Delta$ induces a dual map $\Delta^* : \mathbf{P}^* \to \mathbf{P}$, defined by the formula $\Delta^*(\overline{pq}) = \Delta(p) \cap \Delta(q)$.

A set $Q \subset \mathbf{P}$ is convex if there is a projective transformation $T$ such that $T(Q)$ is a convex subset of the ordinary plane.

## 2.2. Marked Boxes

Informally, a *convex marked box* is a convex quadrilateral in $\mathbf{P}$ with a distinguished top edge, a distinguished bottom edge, a distinguished top point and a distinguished bottom point.

Here are some definitions, which culminate in a formal definition of a convex marked box, the main object of study. An *overmarked box* is a pair of 6 tuples

$$((p, q, r, s; t, b), (P, Q, R, S; T, B))$$

having the incidence relations shown in Figure 2.2. There is an involution on the set of overmarked boxes:

$$((p, q, r, s; t, b), (P, Q, R, S; T, B)) \to ((q, p, s, r; t, b), (Q, P, S, R; T, B))$$

A *marked box* is an equivalence class of overmarked boxes under this involution. Let $\Theta$ be the marked box labelled as in Figure 2.2. The *top* of $\Theta$ is the pair $(t, T)$. The *bottom* of $\Theta$ is the pair $(b, B)$. The *distinguished edges* of $\Theta$ are T and B. The *distinguished points* of $\Theta$ are $t$ and $b$.

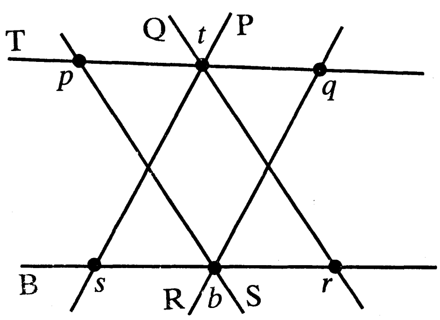


Fig. 2.2

So far, the definitions make sense in a projective plane over any field. (The exception is the projective plane over $\mathbf{Z}/2$, which doesn't have enough points to contain a marked box.) The remainder of the definitions depend on the ordering properties of $\mathbf{R}$. The marked box $\Theta$ is *convex* if the following 4 conditions hold:

1. $p$ and $q$ separate $t$ and $T \cap B$ on the line T.
2. $r$ and $s$ seperate $b$ and $T \cap B$ on the line B.
3. P and Q seperate T and $\overline{bt}$ on the pencil of lines through $t$, in the cyclic ordering of lines through $b$.
4. R and S seperate B and $\overline{bt}$ on the pencil of lines through $b$.

The first two conditions imply the last two, and vice versa.

The *convex interior* of $\Theta$ is the open convex quadrilateral whose vertices, in cyclic order, are $p$, $q$, $r$ and $s$.

Projective symmetries act on marked boxes in a way which preserves convexity: Let $\hat{x} = T(x)$ for a projective transformation T. We define:

$$T(\Theta) = ((\hat{p}, \hat{q}, \hat{r}, \hat{s}; \hat{t}, \hat{b}), (\hat{P}, \hat{Q}, \hat{R}, \hat{S}; \hat{T}, \hat{B})).$$

Similarly, given a duality $\Delta$, let $x^* = \Delta(x)$ when $p$ is a point, and $X^* = \Delta^*(X)$ when X is a line. We define:

$$\Delta(\Theta) = ((P^*, Q^*, S^*, R^*; T^*, B^*), (q^*, p^*, r^*, s^*; t^*, b^*)).$$

The reordering of elements of $\Delta(\Theta)$ is deliberate.

## 2.3. Operations on Marked Boxes

There are three natural *box operations* one can perform on marked boxes. Before defining these operations, we introduce some notation. If $p, q \in \mathbf{P}$, then $pq$ is the line containing $p$ and $q$. Likewise, if P and Q are distinct lines in $\mathbf{P}$, then PQ is the intersection of P and Q. This notation may be applied recursively; for example, $(ab)(cd)$ is the intersection point of the lines $ab$ and $cd$. Given

$$\Theta = ((p, q, r, s; t, b), (P, Q, R, S; T, B))$$

as above, we define:

$$i(\Theta) = ((s, r, p, q; b, t), (R, S, Q, P; B, T)).$$
$$\tau_1(\Theta) = ((p, q, QR, PS; t, (qs)(pr)), (P, Q, qs, pr; T, (QR)(PS))).$$
$$\tau_2(\Theta) = ((QR, PS, s, r; (qs)(pr), b), (pr, qs, S, R; (QR)(PS), B)).$$

Finally, for completeness, we define the identity operation $1(\Theta) = \Theta$. The three box operations are shown schematicaly in Figure 2.3, for a convex marked box.

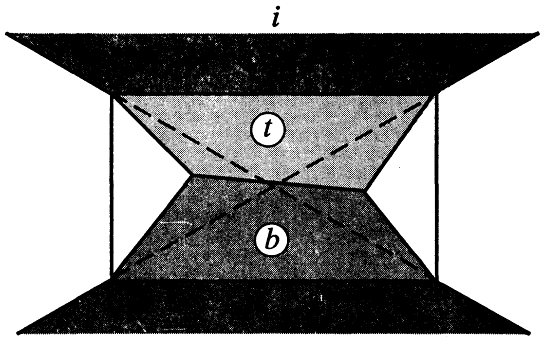


Fig. 2.3

The box operations may be applied iteratively, to form a semigroup G of box operations. We will use the notation $ab$ to mean: " first apply $b$, then apply $a$ ". It is clear that $i$ and $\tau_j$ commute with projective transformations. Two simple computations, one for $i$ and one for (say) $\tau_1$, show that dualities commute with these box operations as well. Hence *the semigroup of box operations commutes with projective symmetries.*

*Lemma* **2.3.** — *The following relations always hold:*

$$i^2 = 1; \quad \tau_1 i \tau_2 = i; \quad \tau_2 i \tau_1 = i; \quad \tau_1 i \tau_1 = \tau_2; \quad \tau_2 i \tau_2 = \tau_1.$$

*Proof.* — The first relation is obvious. By symmetry, we only have to verify the third and fourth relations. We compute:

$$i\tau_1(\Theta) = ((\mathrm{PS}, \mathrm{QR}, p, q; (qs)\,(pr), t), (qs, pr, \mathrm{Q}, \mathrm{P}; (\mathrm{QR})\,(\mathrm{PS}), \mathrm{T})).$$

$$\tau_1 i\tau_1(\Theta) = ((\mathrm{PS}, \mathrm{QR}, (pr)\,\mathrm{Q}, (qs)\,\mathrm{P}; (\mathrm{QR})\,(\mathrm{PS}), (q(\mathrm{QR}))\,(p(\mathrm{PS}))), (\ldots)).$$

$$\tau_2 i\tau_1(\Theta) = (((pr)\,\mathrm{Q}, (qs)\,\mathrm{P}, q, p; ((\mathrm{PS})\,p)\,((\mathrm{QR})\,q), t), (\ldots)).$$

Looking at Figure 2.2, we see that

$$(pr)\,\mathrm{Q} = r; \quad (qs)\,\mathrm{P} = s; \quad (q(\mathrm{QR}))\,(p(\mathrm{PS})) = \mathrm{RS} = b.$$

Substituting these identities in, we see that the 6 points of $\tau_1 i\tau_1(\Theta)$ match the 6 points of $\tau_2(\Theta)$, and the 6 points of $\tau_2 i\tau_1(\Theta)$ match those of $i(\Theta)$. To complete the proof, we note that the 6 points of a marked box clearly determine the entire marked box. ∎

The relations in Lemma 2.3 imply that the semigroup G is in fact a *group*. Clearly G is generated by $\alpha = i$ and $\beta = i\tau_1$. Note that

$$\beta^3 = i\tau_1\, i\tau_1\, i\tau_1 = i\tau_1\, i\tau_2 = 1.$$

Looking at the nesting properties of the convex interiors of the marked boxes in the orbit of $\Theta$, in the case $\Theta$ is convex, we see that the two elements $\alpha\beta = \tau_1$ and $\alpha\beta^2 = \tau_2$ generate a free semigroup. This is enough to identify G as the *modular group*:

$$\mathrm{G} = \langle\, \alpha, \beta : \alpha^2 = 1, \beta^3 = 1 \,\rangle.$$

## 2.4. Two Commuting Group Actions

Let $\Omega = \Omega(\Theta)$ be the orbit of $\Theta$ under G. It is not hard to see, particularly in the convex case, that G acts freely, faithfully, and transitively on $\Omega$.

The structure of $\Omega$ can be expressed by an *incidence graph* $\Gamma$. The edges of $\Gamma$ correspond to marked boxes in $\Omega$, the vertices correspond to tops and bottoms of boxes, and each edge is directed from the top to the bottom. Vertices on distinct edges are identified if the corresponding distinguished sides coincide. A picture of $\Gamma$ is shown in Figure 3.1.1. Technically, $\Gamma$ is just an abstract graph, but we will always embed $\Gamma$ in the hyperbolic plane, as the tiling associated to the hyperbolic reflection group generated by reflections in an ideal hyperbolic triangle.

The group G of box operations has a concrete description as a permutation subgroup of edges of $\Gamma$. The operation $i$ reverses the orientations on each edge. The element $\tau_1$ " rotates " each edge counterclockwise one " click " about the tail point. The element $\tau_2$ " rotates " each edge one " click " clockwise about the head point. Note that G is *not* a group of homeomorphisms of the hyperbolic plane. Each marked box induces a faithful representation $\overline{G}$ of G into the permutation subgroup of marked boxes in the orbit $\Omega$.

Besides G, there is a second, commuting, modular group action M on $\Gamma$. This action M is generated by order 2 isometric rotations about centers of edges of $\Gamma$ together with order 3 isometric rotations about centers of triangles in $\Gamma$. The marked box $\Theta$ induces a representation $\overline{M}$ into the Lie Group $\mathscr{G}$ of projective symmetries of $\mathbf{P}$.

*Theorem* **2.4.** — *Let $\Theta$ be a marked box, with orbit $\Omega$. There is a faithful representation $\overline{M} : M \to \mathscr{G}$ which takes isometries of $\Gamma$ to projective symmetries of $\Omega$ in a way which is natural with respect to the labelling of $\Gamma$.*

*Proof.* — We will show that, for each marked box $\Psi \in \Omega$:

1. There is an order 3 projective transformation having the cycle
$$i(\Psi) \to \tau_1(\Psi) \to \tau_2(\Psi).$$
2. There is a polarity having the cycle $\Psi \to i(\Psi)$.

The theorem then follows from the fact that everything in sight commutes.

Let $\Psi = ((b_1, b_3, a_3, a_1; a_2, b_2), (\dots))$. Figure 2.4.1 shows a particular normalization of $\Psi$. We compute

$$\begin{aligned} i(\Psi) &= ((a_1, a_3, b_1, b_3; a_2, b_2), (\dots)), \\ \tau_1(\Psi) &= ((b_1, b_3, c_1, c_3; b_2, c_2), (\dots)), \\ \tau_2(\Psi) &= ((c_1, c_3, a_1, a_3; c_2, a_2), (\dots)). \end{aligned}$$

The obvious Euclidean rotation satisfies (1).

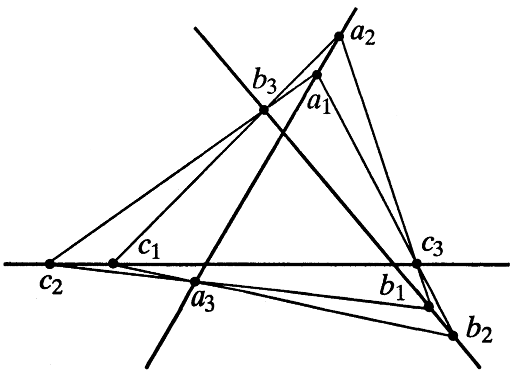


Fig. 2.4.1

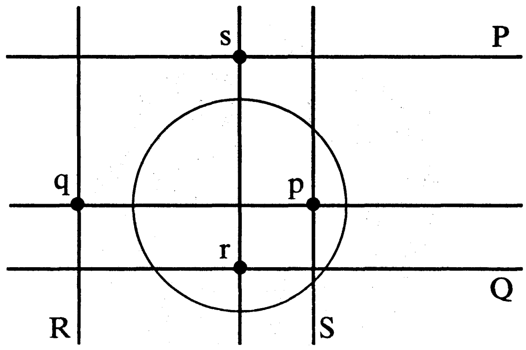


FIG. 2.4.2

Now let $\Psi = ((p, q, r, s; t, b), (P, Q, R, S; T, B))$. Figure 2.4.2 shows a normalization of $\Psi$ in which $t$ and $b$ are at infinity, and $\|p\| \, \|q\| = \|r\| \, \|s\| = 1$ in the standard norm on the Euclidean plane. The polarity induced by the standard inner product on $\mathbf{R}^3$ satisfies (2). ∎

All in all, we have produced two different commuting modular group actions $\overline{G}$ and $\overline{M}$ on the orbit $\Omega$. As the original marked box varies, the orbit varies, and these group actions vary as well. Thus we produce two different commuting group actions on the space of marked boxes.

## 3. Geometry of Marked Boxes

Henceforth, we make the blanket assumption that all our marked boxes are convex.

### 3.1. Depth of a Box

The labeling of the incidence graph $\Gamma$ is determined by the choice of $\Theta$ in the orbit $\Omega$. It can be arranged that the edges corresponding to $i(\Theta)$, $\tau_1(\Theta)$, and $\tau_2(\Theta)$ bound the hyperbolic triangle which contains the origin $0 \subset \mathbf{R}^2$. Given this labelling, we make the following definitions concerning the placement of other boxes in $\Omega$ with respect to $\Theta$.

Let $v$ be a vertex of $\Gamma$. Say that the *depth* of $v$ is the minimum number of edges of $\Gamma$ which must be crossed by any path from the origin to $v$. The depths of various vertices are shown in Figure 3.1.1.

Each directed edge $e \in \Gamma$ has a half space $H_e$ associated to it. The assignment $e \to H_e$ is made so that $H_e \subseteq H_f$ if and only if the convex interior of the marked box labelled by $e$ is contained in the convex interior of the marked box labelled by $f$. Let $H_1$, $H_2$ and $H_3$ be the three half spaces shown in Figure 3.1.1. Assuming that $H_e \subset H_j$ for some $j \in \{1, 2, 3\}$, define the *major depth* of the edge $e$ to be the maximum depth of the two vertices bounding $e$. Define the *minor depth* of $e$ to be the minimum depth of its two

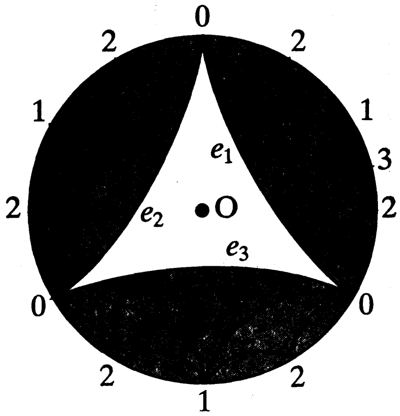


FIG. 3.1.1

vertices. There are only finitely many edges at every major depth, and infinitely many at each minor depth.

The directed edges of $\Gamma$ are in bijection with the marked boxes of $\Omega$, so we can transfer the notion of major and minor depth to marked boxes. We will use the notation $md(\Upsilon)$ and $Md(\Upsilon)$ to denote the minor, and major depth of $\Upsilon$.

The purpose of this section is to make precise the statements that *boxes having high depth are small and thin.*

The projective plane inherits a Riemannian metric $\rho$ from its description as the sphere modulo antipodal points. The metric is not very natural with respect to the full group of projective transformations, but it will nonetheless be useful to us here and in § 4. Unless otherwise stated, all metric measurements (distances and angles) will be made with respect to $\rho$. On the other hand, when we speak of convexity, we will always refer to the projective geometric notion, described in § 2.1.

Given a marked box $\Upsilon \in \Omega$, let $\overline{\Upsilon}$ denote the convex interior of $\Upsilon$, let $\hat{\Upsilon}$ denote the (smaller) convex quadrilateral shown in Figure 3.1.2, and let $\alpha(\Upsilon)$ denote the angle shown in Figure 3.1.2. $|X|$ will denote the $\rho$-diameter of the set $X \subset \mathbf{P}$.

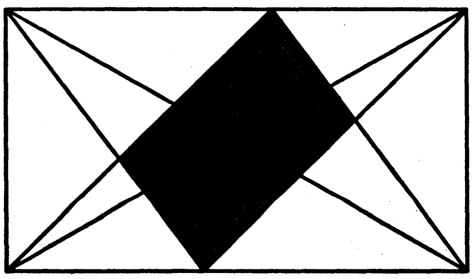

FIG. 3.1.2

*Depth Lemma.* — *Let* $\varepsilon > 0$ *be fixed. Then there is a constant* $N = N(\varepsilon, \Theta)$ *such that, for any marked box* $\Psi \in \Omega$

1. $md(\Psi) > N \Rightarrow |\overline{\Psi}| < \varepsilon$;
2. $Md(\Psi) > N \Rightarrow |\hat{\Psi}| < \varepsilon$;
3. $Md(\Psi) > N \Rightarrow \alpha(\Psi) < \varepsilon$.

*Proof.* — Without loss of generality, we assume $|\Psi| \in |E_1|$, and we normalize so that $\overline{E}_1$ is the unit square. The spherical metric $\rho$ restricted to the unit square is only boundedly different from the Euclidean metric there, so for ease of computation we will work with the Euclidean metric. This bounded change, and the original normalization, merely alter the constants in the Lemma.

We may write $\Psi = w(E_1)$, where $w$ is a word in the box operations $\tau_j$. By throwing out uniformly small words, and using symmetry (between $\tau_1$ and $\tau_2$) we can assume that $w$ has one of the following properties:

1. The string $\ldots \tau_1 \tau_2 \ldots$ occurs $n$ times in $w$, for large $n$.
2. $w = \tau_2 \tau_1^n$, for large $n$.
3. $w = \tau_1^n$, for large $n$.

In the first two cases, both $md(\Psi)$ and $Md(\Psi)$ are large. In the third case, only $Md(\Psi)$ is large.

*Case* **1**. — Recall that $E_1$ is the unit square. Suppose $F \in \Omega$ is any marked box with $\overline{F} \subset \overline{E}_1$. We will show that there is a constant $\eta = \eta(\Theta) < 1$ such that $|\overline{\tau_1 \tau_2(F)}| \leqslant \eta\, |\overline{F}|$.

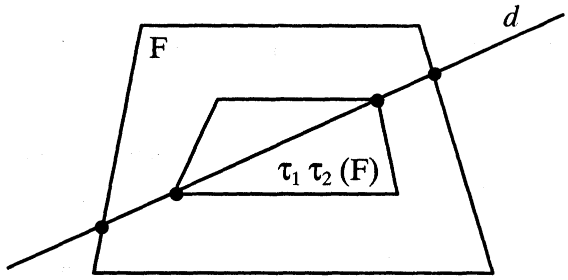


FIG. 3.1.3

Consider the diagonal $d$ shown in Figure 3.1.3. If $d$ is chosen correctly, $|\overline{\tau_1 \tau_2(F)}| = |d|$. Also, $|\overline{F}| \geqslant |d \cap \overline{F}|$. All the boxes in $\Omega$ are equivalent under a projective symmetry, so the cross ratio of the 4 points shown in Figure 3.1.3 takes on one of finitely many values.

Here is a useful trick: Let $a_n < b_n < c_n < d_n \subset \mathbf{R}$ be an infinite sequence of 4-tuples of points. Then the ratio $|c_n - b_n| / |a_n - d_n|$ cannot converge to 1 if the cross ratio

$\xi_n = [a_n, b_n, c_n, d_n]$ takes values in a finite set. This trick bounds the diameter of the smaller quadrilateral away from the diameter of the bigger one. Thus $|(\hat{\Psi})| \leqslant |\bar{\Psi}| \leqslant \eta^n$.

The argument for $\alpha(\Psi)$ is similar: one uses a similar trick to show that $\alpha(\tau_1 \tau_2(F)) \leqslant \eta' \alpha(F)$, for some other constant $\eta' < 1$.

*Case* **2.** — From Theorem 2.4, there is a projective transformation T which has the infinite orbit $E_1 \to \tau_1^2(E_1) \to \tau_1^4(E_1) \ldots$ This map T fixes the distinguished point $p$ at the top of $E_1$, and takes the open unit square into itself. It follows easily that T has no fixed points in the open unit square. Since T preserves the two lines $x$ and $y$ shown in Figure 3.1.4, $\tau_2 \tau_1^{2k}(E_1) = T^k(\tau_2(E_1))$ shrinks as a set to the point $p$ as $k$ grows. Hence $\bar{\Psi}$ and $\hat{\Psi}$ are shrinking to points as well.

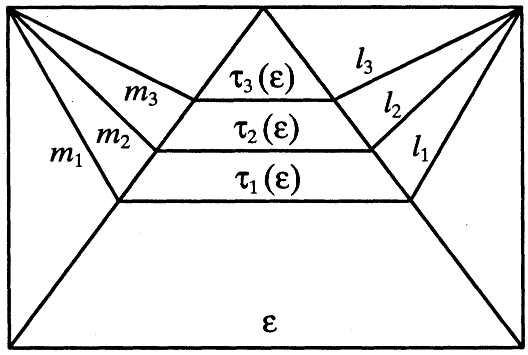


Fig. 3.1.4

The lines $l_n$ and $m_n$, shown in Figure 3.1.4, converge to the top line of $E_1$, squeezing the angle $\alpha$ down to zero.

*Case* **3.** — The argument here is the similar to that in Case 2, except that the top edge of $\tau_1^n(E_1)$ does not shrink to a point. The only difference in this case is that $|\bar{\Psi}|$ remains large.

Stringing the three cases together proves the Lemma. ■

## 3.2. The Pappus Curve

The marked box $\Theta$ determines a natural map $\lambda$ from vertices of the incidence graph $\Gamma$ into the projective plane. The rule is that $\lambda(v)$ is the common distinguished point on all the marked boxes labelled by edges in $\Gamma$ emanating from $v$. Part of the image of $\lambda$ is " hinted at " in Figure 1.2. Recall from § 2.4 that $\overline{M}$ is the representation of the hyperbolic isometry (modular) group M into the Lie Group of projective symmetries. The index 2 subgroup of $M' \subset M$ consisting of those group elements taken to projective *transformations* preserves the image $\Lambda$ of $\lambda$. Furthermore, the vertices of $\Gamma$ are dense in $S^1$, and it is natural to ask whether or not $\lambda$ extends continuously. If this is true, then the whole image $\Lambda = \lambda(S^1)$ is invariant under $\overline{M}'$.

*Theorem* **3.2.** — *The map extends to give a continuous homeomorphism from* $S^1$ *onto its* $M'$*-equivariant image* $\Lambda \subset \mathbf{P}$. *This homeomorphism conjugates the action of* $M'$ *to that of* $\overline{M}'$.

*Proof.* — Recall that $\hat{\Psi}$ is the small convex quadrilateral shown in Figure 3.1.2. The image vertices of $\lambda$ are arranged according to Figure 3.2. Since the nesting properties of the edges of $\Gamma$ exactly reflect the nesting properties of the convex interiors of boxes in $\Omega$, the following fact is sufficient to prove the theorem: let $e_1, e_2, \ldots$ be any infinite sequence of nested edges in $\Gamma$. (This means that $H_{e_j} \supset H_{e_{j+1}}$.) Let $\Psi_j$ be the corresponding marked box. Then $|\hat{\Psi}_j| \to 0$ as $j \to \infty$. But this is an immediate consequence of the Depth Lemma. ■

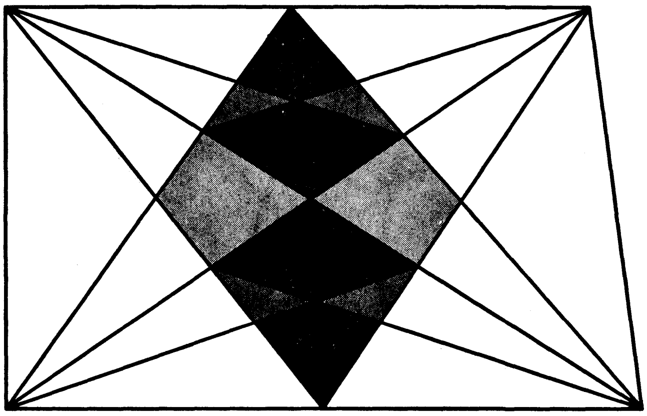

Fig. 3.2

## 3.3. Curve and Linefield

The dualities of $\overline{M} - \overline{M}'$ take the curve $\Lambda$ to a curve of lines L, which contains the tops and bottoms of the marked boxes in the orbit $\Omega$. The curve L is " hinted at " in Figure 1.2.

The two objects L and $\Lambda$ not only are dual to each other, but also have a special geometric relationship. A *transverse linefield* to $\Lambda$ is a continuous curve X of lines such that each line of X intersects $\Lambda$ in exactly one point, and that each point of $\Lambda$ is contained in some line of X. Dually, a *section* of L is a curve X such that each point of X is contained in exactly one line of L, and each line of L contains exactly one point of X.

From the nesting properties of the marked boxes in the orbit $\Omega$, it follows that L is a transverse linefield to $\Lambda$, and dually $\Lambda$ is a section to L.

*Theorem* **3.3.** — *If* $\Lambda$ *is not a straight line, then* L *is the unique transverse linefield to* $\Lambda$. *Dually, if* L *is not a curve of coincident lines, then* $\Lambda$ *is the unique section to* L.

*Proof.* — We will prove the first statement. By Theorem 3.2, the continuous extension $\bar{\lambda}$ of $\lambda$ conjugates the action of $M'$ on $S^1$ to the action of $\overline{M}'$ on $\Lambda$. In particular, $\overline{M}'$ has a dense set of fixed points in $\Lambda$. By continuity, we just have to produce a single point $p \in \Lambda$ at which any two transverse linefields must agree.

Say that a pair of points $(p, q) \in \Lambda \times \Lambda$ is a hyperbolic pair if there is a nontrivial element $T \in \overline{M}'$ which fixes both $p$ and $q$. Such pairs are dense in $\Lambda \times \Lambda$ because the hyperbolic translations of $M'$ have fixed points which are dense in $S^1 \times S^1$. Such an element T attracts about one of the fixed points, say $p$, and expands around the other one.

Since $\Lambda$ is not a straight line (and not convex either) there are two points $p, q \in \Lambda$ such that any line sufficiently close to $\overline{pq}$, emanating from $p$, intersects $\Lambda$ in some point other than $p$. By density, we can assume that $(p, q)$ is a hyperbolic pair. The situation is shown in Figure 3.3.

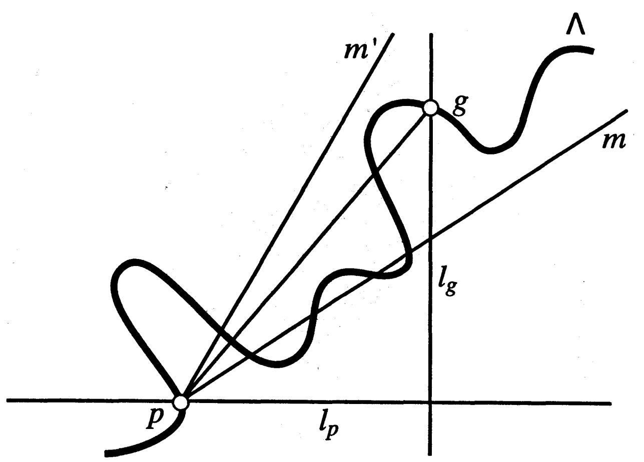


FIG. 3.3

Let $l_p$ be the line of L through $p$ and let $l_q$ be the line of L through $q$. Then T preserves $\overline{pq}$, $l_p$, $l_q$, and fixes the point $l_p \cap l_q$. Let $m$ be any fourth line through $p$. We will prove below (Lemma 3.3) that $T^n(m)$ converges to $l_p$.

Now, if $m$ is sufficiently close to $\overline{pq}$, then $m$ will intersect $\Lambda$ elsewhere. By equivariance, $T^n(m)$ will intersect $\Lambda$ elsewhere as well. By continuity, all lines containing $p$ between $m$ and $T^n(m)$ will intersect $\Lambda$ elsewhere. Letting $n \to \infty$, and $m \to \overline{pq}$, we see that all lines between $\overline{pq}$ and $l_p$ which contain $p$ will intersect $\Lambda$ elsewhere. The same argument may be repeated for lines $m'$ on the other side of $\overline{pq}$. Hence, every line through $p$ except $l_p$ intersects $\Lambda$ in at least two points. ■

*Lemma* **3.3.** — Suppose $T \in \overline{M}'$ *fixes two distinct points* $p, q \in \Lambda$, *and attracts about* $p$. *Suppose* $m$ *is a line through* $p$ *distinct from* $l_p$ *and* $\overline{pq}$. *Then* $T^n(m)$ *converges to* $l_p$.

*Proof.* — The line $m$ intersects $l_q$ at a point $x \in l_q$. This intersection point is neither $q$ or $l_p \cap l_q$. If $T^n(m)$ failed to converge to $l_p$, then T would fix $x$. Fixing three distinct points on $l_q$, T would be the identity on $l_q$. By looking at the conjugated action of T back on $S^1$, we observe the following fact: if $\Psi$ is any marked box such that $q \in \overline{\Psi}$, then the open convex interior of $T(\overline{\Psi})$ contains the closure of the convex interior of $\overline{\Psi}$. This is not consistent with T fixing $L_q$ pointwise. ∎

## 4. The Three-Manifold

### 4.1. The Projective Tangent Bundle

The projective tangent bundle $\mathscr{P}$ to the projective plane is the closed analytic 3-manifold of pairs $(p, \mathscr{P})$, such that $p \in \mathbf{P}$, $P \in \mathbf{P}^*$, and $p \in P$. We will call points of $\mathscr{P}$ *flags*. The Lie Group $\mathscr{G}$ of projective symmetries acts as a group of analytic diffeomorphisms of $\mathscr{P}$, as follows: $T((p, P)) = (T(p), T(P))$ for a projective transformation T, and $\Delta((p, P)) = (\Delta^*(P), \Delta(p))$ for a duality $\Delta$.

We will say that a *pencil* of $\mathscr{P}$ is a set of flags, such that one of the two coordinates varies while the other one takes on all possible values. The objects with which we will deal, in § 4.3 below, are *ruled surfaces* in the sense that they can be expressed as the union of continuously varying disjoint pencils.

### 4.2. Domain of Discontinuity

The representation $\overline{M}$ constructed in Theorem 2.4 gives an action on $\mathscr{P}$ by analytic diffeomorphisms. Let $\mathscr{N} \subset \mathscr{P}$ denote the set of pairs $(p, P)$ such that $p \notin \Lambda$ and $P \notin L$. The representation $\overline{M}$ preserves both $\mathscr{N}$ and its complement.

*Lemma* **4.2.1.** — *The fixed points (flags) of* $\overline{M}$ *are dense in, and contained in, the complement of* $\mathscr{N}$.

*Proof.* — We will prove density first, then containment. Consider the set $\mathscr{M}$ of pairs $(p, P)$, such that $P \in L$. This set makes up essentially "half" of the complement of $\mathscr{N}$. It is sufficient to show that the fixed flags are dense in $\mathscr{M}$. There is a continuous map

$$\varphi : \Lambda \times \Lambda - \text{Diagonal} \to \mathscr{M}$$

defined as follows. Let $l_p$ and $l_q$ be the lines of L which contain the two points $p, q \in \Lambda$. Define $\varphi(p, q) = (l_p \cap l_q, l_p)$.

A proof very similar to the one given in Theorem 3.3 implies that the image of $\varphi$ contains the pencil determined by $l_q$ provided $q$ is part of a hyperbolic pair $(p, q)$. These pencils are dense in $\mathcal{M}$; hence so is the image of $\varphi$. Since the hyperbolic pairs are dense in the domain of $\varphi$, their image is dense in the range as well, which is to say that fixed flags are dense in $\mathcal{M}$.

Now for containment. Recall that M is the modular group of hyperbolic isometries acting on $\Gamma$. There are four possibilities for elements of M: order 2, order 3, parabolic and hyperbolic. Consider the corresponding elements of $\overline{\mathrm{M}}$.

1. By the proof of Theorem 2.4, these are polarities, and have no fixed flags.
2. By the proof of Theorem 2.4, these are conjugate to a Euclidean rotation. Again, no fixed flags.
3. The fourth power of such an element fixes pointwise the distinguished line of some marked box, and also has infinite order. By duality, this fourth power preserves all the lines through the distinguished point of the same marked box. All of these fixed points and lines give rise to two fixed pencils, both of which belong to the complement of $\mathcal{N}$. These are the only fixed flags: any more would make the element trivial.
4. The proof of Lemma 3.3 implies that the square of this type of element has exactly 3 fixed points and exactly 3 fixed lines, making 6 flags fixed in all. All of these flags belong to the complement of $\mathcal{N}$. ∎

One consequence of Lemma 4.2.1 is that $\overline{\mathrm{M}}$ acts freely on $\mathcal{N}$. To show that $\mathcal{N}$ is a domain of discontinuity for $\overline{\mathrm{M}}$, we need to show that the action here is also *proper*. This is to say that the set $\{ \mathrm{T} \in \overline{\mathrm{M}} : \mathrm{T}(\mathcal{K}) \cap \mathcal{K} \neq \emptyset \}$ is finite for any compact $\mathcal{K} \subset \mathcal{N}$. Before proving this, we state the following equivalent version, which is easier to prove:

*There is a finite subset* $\overline{\mathrm{H}} \in \overline{\mathrm{M}}$ *such that, for any compact* $\mathcal{K}$ *and infinite* $\overline{\mathrm{S}} \subset \overline{\mathrm{M}}$ *there is some element* $\mathrm{T} \in \overline{\mathrm{S}}$ *and two elements* $h_1, h_2 \in \overline{\mathrm{H}}$ *such that* $h_1 \mathrm{T} h_2(\mathcal{K})$ *does not intersect* $\mathcal{K}$.

To recover the original statement, one simply applies the new criterion to compact subsets of the form $\overline{\mathrm{H}}\mathcal{K}$. The reason for this restatement is the following Lemma about the hyperbolic isometry group M. We will refer to the notion of depth used in § 3.1, and also to the fixed edge $e_1$ described in § 3.1.

*Lemma* **4.2.2.** — *There is a finite subset* $\mathrm{H} \in \mathrm{M}$ *having the following property: let* $\mathrm{S} \subset \mathrm{M}$ *be an infinite sequence of distinct elements. Then for each integer* $\mathrm{N} > 0$ *there are elements* $h_1, h_2 \in \mathrm{H}$ *and an element* $\mathrm{T} \in \mathrm{S}$ *such that* $h_1 \mathrm{T} h_2(e_1) \subset \tau_1 \tau_2(e_1)$ *and* $md(h_1 \mathrm{T} h_2(e_1)) > \mathrm{N}$.

*Proof.* — It is an easy exercise in hyperbolic geometry to show that there is a finite set of elements $\mathrm{H}_1 \subset \mathrm{M}$ such that all but finitely many elements of S are hyperbolic after composition with some element in H. In fact, it is easy to guarantee that the fixed points of these new hyperbolic elements remain far away from each other in the round metric on the circle. To be more concrete, we can assume that each element of S has fixed points in regions bounded by distinct and non-adjacent edges shown in Figure 4.2. Finally, after conjugation with another finite set $\mathrm{H}_2$ of elements of M, we can guarantee

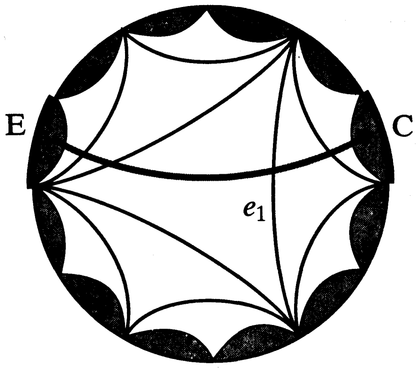


FIG. 4.2

that the expanding fixed point of every element of S is contained in the region labelled E, and that the contracting fixed point is contained in the region labelled C. Composing at most twice on the left and on the right by elements and inverses of elements in $H_j$, we can assume that every element in S expands about a point in E and contracts about a point in C. Most of these elements will have very large translation lengths, and such elements will clearly satisfy the conclusion of the Lemma. ∎

Based on Lemma 4.2.2, and the reformulation of the properness criterion, we can without loss of generality assume for the rest of this section that each transformation $T \in \overline{M}$ which we consider has the following properties:

1. T has a repelling fixed point in $\Lambda - E_1$.
2. T has an attracting fixed point in $\Lambda \cap \tau_1 \tau_2(E_1)$.
3. $T(E_1) \subset \tau_1 \tau_2(E_1)$.

Recall that $\rho$ is the " spherical metric " on $\mathbf{P}$. Let $N_\varepsilon(X)$ be the $\varepsilon$ neighborhood about the set $X \in \mathbf{P}$. If $p \in \Lambda$, then $l_p$ will denote the line of L which contains $p$.

*Lemma* **4.2.3.** — *Let* $\{T_k\}$ *be any infinite sequence of distinct projective transformations in* $\overline{M}$. *For any* $\varepsilon > 0$ *there is some* $T_n \in \{T_k\}$, *and points* $p, q \in \Lambda$ *(not necessarily distinct) such* $T_n(\mathbf{P} - N_\varepsilon(l_q)) \subset N_\varepsilon(p)$ *and* $T_n(\mathbf{P} - N_\varepsilon(q)) \subset N_\varepsilon(l_p)$.

*Proof.* — Since the fixed points of $T_k$ are contained in two disjoint compact subsets, we can normalize so that

1. $p_k$ is the origin and $q_k$ is at infinity, on the $y$-axis;
2. $l_{p_k}$ is the $x$-axis, and $l_{q_k}$ is the line at infinity;
3. the convex interior of $\tau_1 \tau_2(E_1)$ contains a fixed small square centered at the origin, and is contained in a fixed larger square centered at the origin.

Under this normalization, a sufficiently large ball about the origin contains the complement of $N_\varepsilon(l_q)$. Likewise, the set of lines having absolute slope less than a sufficiently large (but fixed) number contains the complement of $N_\varepsilon(q)$. Finally, we have

$$T_k = \begin{pmatrix} \lambda_k & 0 \\ 0 & \mu_k \end{pmatrix}.$$

Let $E' = \tau_1 \tau_2(E_1)$. For most values of $k$, $T_k(E')$ will have large minor depth, and hence by the Depth Lemma of § 3.1, $T_k(E')$ will be " small and thin ". Since $E'$ is close in size and shape to the unit square, $\lambda_n$, $\mu_n$, and $\mu_n/\lambda_n$ are all very close to zero for large $n$. In particular, $T_n$ contracts a large ball about the origin into a very small one, and contracts lines of bounded large absolute slope to lines having very small absolute slope. ∎

*Theorem* **4.2.** — *The set $\mathcal{N}$ is the domain of discontinuity for the group action* $\overline{M}$.

*Proof.* — Let $\pi_1$ and $\pi_2$ be the obvious projections from $\mathcal{P}$ to $\mathbf{P}$ and $\mathbf{P}^*$ respectively. Let $\mathcal{K}_1$ denote the set of flags in $\mathcal{K}$ whose projections under $\pi_1$ are not contained in $N_\varepsilon(l_q) \cup N_\varepsilon(p)$. Dually let $\mathcal{K}_2$ denote the set of flags in $\mathcal{K}$ whose projections under $\pi$ do not intersect $N_\varepsilon(q) \cup N_\varepsilon(l_p)$. Since $\mathcal{K}$ is compact and is contained in $\mathcal{N}$, the two sets $\mathcal{K}_j$ cover $\mathcal{K}$ once $\varepsilon$ is small enough. Let $\{ T_k \}$ be an infinite sequence of transformations which (putatively) does not move $\mathcal{K}$ completely off itself. By Lemma 4.3.3, $\pi_1(T_n(\mathcal{K}_1)) \subset N_\varepsilon(p)$ and $\pi_2(T_n(\mathcal{K}_2)) \subset N_\varepsilon(l_p)$. But then $T_n(\mathcal{K}_j) \cap \mathcal{K} = \emptyset$. Since these two sets cover $\mathcal{K}$, we have $T_n(\mathcal{K}) \cap \mathcal{K} = \emptyset$, a contradiction. We have shown that the fixed points of $\overline{M}$ are dense in, and contained in the complement of $\mathcal{N}$. We have now just seen that $\overline{M}$ acts properly and freely on $\mathcal{N}$. Hence $\mathcal{N}$ is a domain of discontinuity for $\overline{M}$. ∎

We emphasize that this domain $\mathcal{N}$ varies with the initial choice of convex marked box. Usually, it is a set with a " fractal boundary ".

## 4.3. Trefoil Knot Complement

Since $\mathcal{N}$ is a domain of discontinuity for $\overline{M}$, the quotient manifold $\mathcal{E} = \mathcal{N}/\overline{M}$ is a three dimensional analytic manifold. $\mathcal{E}$ has a $(\mathcal{G}, \mathcal{X})$ structure in the sense of [T]. Here $\mathcal{X}$ is the simply connected cover of the projective tangent bundle. In this section, we will prove

*Theorem* **4.3.** — *The manifold $\mathcal{E}$ is homotopy equivalent to the complement of the trefoil knot in the three sphere.*

Before proving Theorem 4.3, we will analyze the bilaterally symmetric case. The analysis here is related to the classical fact that $SL_2(\mathbf{R})/SL_2(\mathbf{Z})$ is homeomorphic to the trefoil knot complement. To prove the general case, we will show that nothing on the level of homotopy changes as we deform away from the symmetric case.

If the marked box has bilateral symmetry, then the curve $\Lambda$ will be a line of $\mathbf{P}$, and the curve L will consist of coincident lines. We can conjugate by a projective transformation so that $\Lambda$ is the line at infinity and L is the set of lines going through the origin in $\mathbf{R}^2$. Then the set $\mathscr{N}$ consists of flags $(p, \mathrm{P})$, where $p \in \mathbf{R}^2 - \{0\}$ and P is a line through $p$ which misses the origin. Thus $\mathscr{N}$ is homeomorphic to $\mathbf{R}^2 \times \mathbf{S}^1$.

The modular group $\overline{\mathrm{M}} = \langle \alpha, \beta : \alpha^2 = \beta^3 = 1 \rangle$ acts on $\mathscr{N}$. $\beta$ is an order three linear map of $\mathbf{R}^2$ and $\alpha$ can be taken as the polarity induced by the standard inner product on $\mathbf{R}^3$. The universal cover $\hat{\mathscr{N}}$ has a description as the set of "flags" $(p, \mathrm{P})$, where $p$ is a point in the universal cover of the punctured plane, and $l$ is a line through $p$ missing the origin. Let $c$ denote the covering transformation of $\hat{\mathscr{N}}$. The "rotation" $\beta \in \overline{\mathrm{M}}$ acts on $\hat{\mathscr{N}}$ by setting $\beta^3 = c$.

The polarity $\alpha$ has the following concrete description: take the point $p$ on the unit circle to the tangent line through the antipodal point $p'$ in the unit circle. From this description, we can clearly lift $\alpha$ to an action on $\hat{\mathscr{N}}$ which satisfies $\alpha^2 = c$. Thus, the $c$-extension of the modular group acts on $\hat{\mathscr{N}}$. The quotient of this lifted action is the original manifold $\mathscr{E}$. Hence $\mathscr{E}$ is the quotient of $\mathbf{R}^3$ by the $c$-extension of the modular group, which is the trefoil knot complement fundamental group.

To analyze the general case, we introduce some notions from point set topology. Let M be a topological 3-manifold. An embedded surface $\mathrm{S} \subset \mathrm{M}$ is *tamely embedded* if each point on the S has a small neighborhood so that the pair $(\mathrm{N} \cap \mathrm{S}, \mathrm{N})$ is homeomorphic to the standard pair $(\mathrm{B}^2, \mathrm{B}^3)$, where $\mathrm{B}^j$ is the open unit $j$ ball. Two tamely embedded surfaces $\mathrm{S}_j$ *intersect tamely* if each point of the intersection has a neighborhood N such that the triple $(\mathrm{S}_1 \cap \mathrm{N}, \mathrm{S}_2, \cap \mathrm{N}, \mathrm{N})$ is homeomorphic to the triple $(\mathrm{B}_1^2, \mathrm{B}_2^2, \mathrm{B}^3)$, and if these homeomorphisms are all compatible with the homeomorphisms to standard pairs which take place away from the intersection. Here $\mathrm{B}_1^2$ and $\mathrm{B}_2^2$ are orthogonal open unit disks. Here is a technical fact:

*Suppose* M *is a closed topological* 3-*manifold and* $\mathrm{A}_t$ *and* $\mathrm{B}_t$ *are two continuous families of tamely embedded and tamely intersecting closed surfaces. Then the homeomorphism type of* $\mathrm{M} - \mathrm{A}_t - \mathrm{B}_t$ *remains unchanged as* $t$ *varies.*

We omit the proof because it is technical, and would lengthen the exposition considerably.

*Proof of Theorem 4.3.* — Let $\mathscr{A} \subset \mathscr{P}$ consist of those flags whose first coordinate is contained in $\Lambda$. Let $\mathscr{B} \subset \mathscr{P}$ consist of those flags whose second coordinate is contained in L. Then $\mathscr{N} = \mathscr{P} - \mathscr{A} - \mathscr{B}$. The surface $\mathscr{A}$ is a bundle of projective lines sitting over $\Lambda$; it is closed, and ruled in the sense of section § 4.1. Dually for $\mathscr{B}$. In fact, in the local coordinates shown in Figure 4.3, the two surfaces are locally both ruled surfaces of $\mathbf{R}^3$ in the ordinary sense. Furthermore, they intersect exactly in a curve. It is a standard exercise to show that locally such surfaces are tamely embedded, and intersect tamely, in $\mathbf{R}^3$. Hence the same is true globally, in $\mathscr{P}$. Taking a family of marked boxes varying

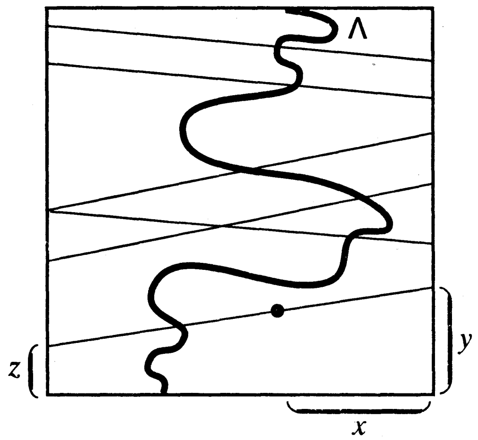


FIG. 4.3

continuously from a bilaterally symmetric one to the one under study, we see that $\mathcal{N}$ is homeomorphic to $\mathbf{S}^1 \times \mathbf{R}^2$, as in the symmetric case. Finally, the two perturbed elements $\alpha$ and $\beta$ may be lifted to the universal cover, since $\beta$ is conjugate to a rotation and $\alpha$ is still a polarity. The identity $\alpha^2 = \beta^3 = c$ cannot change either. Hence the trefoil knot fundamental group acts on the universal cover of $\mathcal{N}$, which is just $\mathbf{R}^3$. The quotient is $\mathscr{E}$, as desired. ■

REFERENCES

[B] A. BEARDON, *The Geometry of Discrete Groups*, Springer-Verlag, 1983.
[H] HILBERT and COHN-VOSSEN, *Geometry and the Imagination*, Chelsea Publishing Company, 1950.
[T] W. THURSTON, *The Geometry and Topology of Three-Manifolds*, Princeton University Lecture Notes, 1978.

Mathematics Department
UCSD
La Jolla, California 92093
USA

*Manuscrit reçu le 11 juin 1992.*