The Banach–Tarski paradox for the hyperbolic plane

by

Jan Mycielski (Boulder, CO)

Abstract. Every two bounded sets with nonempty interiors in the hyperbolic plane are equivalent by finite decomposition.

1. Introduction. In 1914 Hausdorff proved that there is no finite, nonzero, finitely additive and rotation-invariant measure over the family of all subsets of the two-dimensional sphere S^2. In 1923 Banach showed that there exists a finitely additive measure μ over all bounded subsets of the plane $\boldsymbol{R}^2$ which is invariant under all isometries of $\boldsymbol{R}^2$ and satisfies $\mu(I^2) = 1$, where I^2 is the unit square. In 1924 Banach and Tarski improved Hausdorff's theorem by showing that any two bounded sets A and B with nonempty interios in $\boldsymbol{R}^n$ $(n \geqslant 3)$ or in S^n $(n \geqslant 2)$ are equivalent by finite decomposition, i.e., there exist finite partitions of A and B into disjoint sets $A = A_1 \cup \ldots \cup A_m$, $B = B_1 \cup \ldots \cup B_m$ such that A_i is isometric to B_i for $i = 1, \ldots, m$. (Moreover the underlying isometries can be assumed to preserve the orientation of $\boldsymbol{R}^n$ or S^n respectively.) Of course the number m depends on A and B. For an attractive presentation of this subject see the book of S. Wagon [2].

It is also known that the theorem of Banach and Tarski is true for the hyperbolic space H^n for $n \geqslant 3$, and the proof is the same as for $\boldsymbol{R}^n$ (since the group of rotations of H^n around a fixed point act upon H^n in the same way as SO_n acts upon $\boldsymbol{R}^n$).

It is the purpose of this paper to extend the theorem of Banach and Tarski to the hyperbolic plane H^2 (Theorem 2.8). Again the underlying isometries will be sense-preserving. (This problem for H^2 was raised independently by Andrew Haas and the author.) I am indebted to Stan Wagon for his stimulating influence.

In Section 2 we state the preliminaries. All arguments given in that section are essentially known and stated only for convenience of the reader. Only Lemma 2.7 requires an original proof, which is given in Section 3. Additional remarks and problems are stated in Section 4.

2. Preliminaries and the main theorem. We shall use the Poincaré disk model for H^2, so $H^2 = \{z \mid |z| < 1\}$ and all its sense-preserving isometries are generated by rotations $z \mapsto e^{i\alpha}z$, where $0 \leqslant \alpha < 2\pi$, and translations $z \mapsto (z+a)/(\bar{a}z+1)$, where $|a| < 1$. We put $\tau_a(z) = (z+a)/(\bar{a}z+1)$ and recall that the diameter line of the unit circle passing through a is invariant under τ_a and circular arcs passing through the ends of this diameter are also invariant under τ_a.

The following lemma is due to Klein and Fricke (see [0]).

2.1. LEMMA. *If* $1 > a \geqslant \frac{1}{\sqrt{2}}$, *then the two isometries* τ_a *and* τ_{ia} *are free generators of a free subgroup of the group of isometries of* H^2. *Moreover no element of this subgroup except the identity has any fixed points in* H^2.

Proof. Consider Fig. 1. The translation τ_a maps the hyperbolic straight line L_1 onto L_2 and τ_{ia} maps L_3 onto L_4. Since $a \geqslant 1/\sqrt{2}$, it is easy to check that the four closed regions A_1, A_2, B_1, and B_2 are disjoint (the A_i's touch the B_i's at infinity if $a = 1/\sqrt{2}$). Let C be the central region with the lines L_1 and L_3 but without L_2 and L_4, A_1 and B_1 be open and A_2 and B_2 be closed. Then τ_a maps $C \cup B_1 \cup B_2$ into A_2 and τ_{ia} maps $C \cup A_1 \cup A_2$ into B_2. Let w be a nontrivial reduced group word in τ_a and τ_{ia}. It is clear from the above, by induction on the length of w, that $w(C) \subseteq A_1 \cup A_2$ or $w(C) \subseteq B_1 \cup B_2$ (depending on the first term of w). Thus w has no fixed points in C. Since C is a fundamental domain this proves Lemma 2.1.

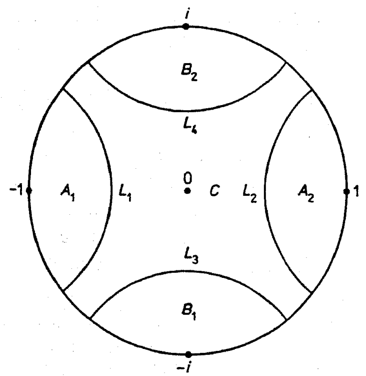

Fig. 1

2.2. LEMMA. *Almost all pairs* (a_1, a_2) *of reals in the open square* $(-1, 1)^2$, *in the sense of measure or category, are such that for every finite sequences of integers* p_{ij}, q_{ij}, $i = 1, \ldots, n$, $j = 1, 2$, *where* $n \geqslant 2$ *and* $p_{k1} + p_{k2} \neq 0$ *for all* $k > 1$ *and* $q_{k1} + q_{k2} \neq 0$ *for all* $k < n$, *the function*

$$f(z) = \tau_{a_1}^{p_{11}} \tau_{a_2}^{p_{12}} \tau_{ia_1}^{q_{11}} \tau_{ia_2}^{q_{12}} \ldots \tau_{a_1}^{p_{n1}} \tau_{a_2}^{p_{n2}} \tau_{ia_1}^{q_{n1}} \tau_{ia_2}^{q_{n2}}(z)$$

is not the identity map of H_2 *onto itself.*

Proof. By Lemma 2.1, the equation $f(0) = 0$ defines a proper subset of the diagonal $a_1 = a_2$. Hence $f(0) = 0$ defines a proper subset of the square $(-1, 1)^2$. Moreover this set is algebraic and hence it is meager and of measure zero.

The following lemma is obvious.

2.3. LEMMA. *An orientation-preserving isometry of* H^2 *different from the identity has at most one fixed point.*

$\equiv$ will denote equivalence of sets by finite decomposition.

2.4. LEMMA. *If* $A \subseteq H^2$, A *has interior points,* $C \subseteq A$ *and* C *is countable and bounded, then* $A \equiv A - C$.

Proof. Let D be a circular disc such that $D \subseteq A$, c is the center of D and $c \notin C$. Since C is countable and bounded there exists a $C_1 \subseteq D$ such that $c \notin C_1$, $C_1 \cap C = \emptyset$ and $C - D \equiv C_1$. Now let $C_2 = (C \cap D) \cup C_1$. It is clear that $C \equiv C_2$. Let ϱ be a rotation of D such that C_2, $\varrho(C_2)$, $\varrho^2(C_2)$, ... are disjoint from each other. Since $\varrho(C_2 \cup \varrho(C_2) \cup \ldots) = \varrho(C_2) \cup \varrho^2(C_2) \cup \ldots$ we see that $D \equiv D - C_2$. This implies $A \equiv A - C$.

2.5. LEMMA. *If* F *is a free nonabelian group of transformations of a set* S *that acts without fixed points upon* S *(i.e. for each* $\varphi \in F$, $\varphi \neq e$ *and each* $s \in S$, $\varphi(s) \neq s$), *then* S *can be split into two disjoint parts such that each of them is equivalent to* S *by finite decomposition using the transformations of* F.

See [2] Corollary 4.3 for a proof of this lemma. Many stronger related facts are proved there and in [1] and [3].

2.6. LEMMA. *If* $A \subseteq B \subseteq C$ *and* $A \equiv C$, *then* $A \equiv B$.

See [2] again.

2.7. Lemma. *For every* $\varepsilon > 0$ *there exists a set* $B \subseteq H^2$ *with nonempty interior, diameter* $\leqslant \varepsilon$ *such that* B *is equivalent by finite decomposition to two disjoint copies of* B.

This lemma is the main new step made in this paper and it will be shown in the next section. It yields easily our main result:

2.8. THEOREM. *If* X, $Y \subseteq H^2$ *are bounded and have nonempty interiors, then* $X \equiv Y$.

Proof. We can find a set B satisfying the conclusion of Lemma 2.7 such that both X and Y contain isometric copies of B. Of course $B \equiv nB$ (n disjoint copies of B in H^2). Since X is bounded, there exists an X' such that $X \equiv X'$ and $X' \subseteq nB$ for some n. So, by Lemma 2.6, $B \equiv X$. By the same reason $B \equiv Y$, and $X \equiv Y$ follows. ■

3. Proof of Lemma 2.7. We define B as the shaded part in Fig. 2, where the points α, β, γ, δ are the corners of a square of diameter $< \varepsilon$ centered at the origin, but the sides of B are not hyperbolic straight segments but are segments of equidistants to the real or imaginary axis respectively. Since the diameter of B is small, B is close to an ordinary Euclidean square. The point α and the half open sides $[\alpha, \beta)$ and $[\alpha, \delta)$ are in B. The closed sides $[\beta, \gamma]$ and $[\gamma, \delta]$ are in the complement of B.

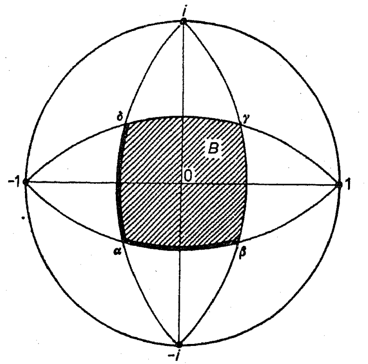

Fig. 2

Now we will define a bijection φ of B onto itself (see Fig. 3). Let

$$-s < -u \leqslant -v < v \leqslant u < s$$

be such that φ translates the lens $\beta, s, \gamma, u, \beta$ onto the lens $\alpha, -u, \delta, -s, \alpha$; φ translates the curved quadrilateral $\alpha, \eta, v, \sigma, \delta, -s, \alpha$ onto the curved quadrilateral $\varepsilon, \beta, s, \gamma, \tau, -v, \varepsilon$; finally φ translates the curved quadrilateral $\eta, \beta, u, \gamma, \sigma, v, \eta$ onto the curved quadrilateral $\alpha, \varepsilon, -v, \tau, \delta, -u, \alpha$. Thus the bijection φ consists of three translations, and hence it is a piecewise isometry.

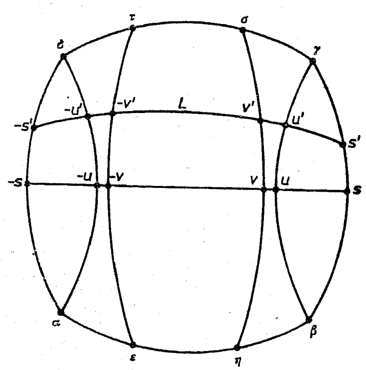

Fig. 3

Let us examine the action of φ upon a curve L equidistant from the real axis. First we parametrize L by arc length. Let $-s' < -u' \leqslant -v' < v' \leqslant u' < s'$ be real numbers representing points of L such that their differences are the lengths of the corresponding segments of L (see Fig. 3). Hence, in terms of this parametrization,

φ acts upon L in the following way

$$\varphi(x) = x+(s'-v') \quad \text{for } x \in [-s', v'),$$
$$\varphi(x) = x-(v'+u') \quad \text{for } x \in [v', u'), \text{ and}$$
$$\varphi(x) = x-(u'+s') \quad \text{for } x \in [u', s'),$$

where + and − are ordinary addition and subtraction of real numbers.

3.1. LEMMA. *For every positive integer n and every $x \in L$ there exists nonnegative integers $n_1+n_2 \geqslant n$ such that*

$$\varphi^n(x) = x+n_1(s'-v')-n_2(u'+s').$$

Proof. By the previous discussion for every $x \in L$ there exist three nonnegative integers $k_1+k_2+k_3 = n$ such that

$$\varphi^n(x) = x+k_1(s'-v')-k_2(v'+u')-k_3(u'+s')$$
$$= x+(k_1+k_2)(s'-v')-(k_2+k_3)(u'+s').$$

So the integers $n_1 = k_1+k_2$ and $n_2 = k_2+k_3$ satisfy the conclusion of 3.1.

3.2. LEMMA. *If $(s'-v')/(u'+s')$ is irrational, then $\varphi^k(x) \neq x$ for every integer $k \neq 0$ and every $x \in L$.*

Proof. An obvious consequence of 3.1.

Let d denote the distance from L to the real axis. Then $(s'-v')/(u'+s')$ is a function of d, s and v.

3.3. LEMMA. *We can fix s and v such that the ratio $(s'-v')/(u'+s')$ is irrational except possibly for a countable set of values of d.*

Proof. We can choose s, and this determines u (since the curve β, u, γ is a translate of the curve $\alpha, -s, \delta$). Then we can choose $v \in (0, u]$ such that $(s-v)/(u+s)$ is irrational. In this case $(s'-v')/(u'+s')$ is an analytic function of d whose value for $d = 0$ is irrational. Hence 3.3 follows.

Now we shall conclude the proof of 2.7.

Let s and v be such that $v \leqslant u < s$. $\mathrm{Re}(\sigma) > 0$, the diameter of B is less than ε and the conclusion of Lemma 3.3 holds. Let φ be defined by those parameters as above. By Lemma 3.1 there exist two translations, namely

$$\varphi_1 = \tau_s\tau_{-v} \quad \text{and} \quad \varphi_2 = \tau_{-s}\tau_{-v}$$

such that for every positive integer n there exists a finite partition of B into sets S_i such that φ^n restricted to any S_i is of the form $\varphi_1^{n_1}\varphi_2^{n_2}$, where $n_1, n_2 \geqslant 0$ and $n_1+n_2 \geqslant n$.

Notice that

$$\varphi_1\varphi_2 = \varphi_2\varphi_1$$

and that

$$\varphi_1 = \tau_{a_1} \quad \text{and} \quad \varphi_2 = \tau_{a_2},$$

where

$$a_1 = \tau_r(-v) = \frac{s-v}{-sv+1} \quad \text{and} \quad a_2 = \tau_{-s}(-u) = \frac{-s-u}{su+1}.$$

By the definition of B and φ, u is a function of s, $u \to 0$ as $s \to 0$ and $u \to \infty$ as $s \to \infty$. It follows that $a_2 \to 0$ as $s \to 0$ and so a_2 is a nonconstant function of s. And, of course, for a fixed s, a_1 is a nonconstant function of v.

So we may still refine our choice of s and v such that the pair (a_1, a_2) is not in the exceptional set of measure zero of Lemma 2.2.

Then, using the same values s and v, we define a bijection ψ of B onto itself in the same way as φ except that ψ will consist of translations preserving the imaginary axis of B. Then we have two translations ψ_1 and ψ_2 with the same relationships to ψ and φ_1 and φ_2 to φ.

Therefore, by Lemma 3.1, for every nontrivial reduced word W of the form

$$\varphi^{p_1}\psi^{q_1} \dots \varphi^{p_n}\psi^{q_n} \quad (p_k \neq 0 \text{ for } k > 1,\ q_k \neq 0 \text{ for } k < n),$$

there exists a finite partition of B into sets S_i such that W restricted to any S_i is of the form

$$\varphi_1^{p_{11}}\varphi_2^{p_{12}}\psi_1^{q_{11}}\psi_2^{q_{12}} \dots \varphi_1^{p_{n1}}\varphi_2^{p_{n2}}\psi_1^{q_{n1}}\psi_1^{q_{n2}},$$

where p_{kj} and q_{kj} depend on S_i but are of the same sign as p_k or q_k respectively and

$$|p_{k1}+p_{k2}| \geqslant |p_k| \quad \text{and} \quad |q_{k1}+q_{k2}| \geqslant |q_k|.$$

So by our choice of the parameters s and v in terms of which φ_j and ψ_j are defined and by the Lemmas 2.2, 2.3, 3.2 and 3.3, W restricted to S_i has at most one fixed point. Hence W has only finitely many fixed points in B.

It follows that φ and ψ are free generators of a free nonabelian group F of piecewise isometries of B onto itself. Also the set C of all fixed points of all elements of F except the identity is countable. Of course $C \subseteq B$ and F acts without fixed points on $B-C$. Hence, by Lemma 2.5, $B-C \equiv 2(B-C)$ and by Lemma 2.4, $B-C \equiv B$. Therefore $B \equiv 2B$, which concludes the proof of Lemma 2.7.

4. Remarks and problems. 1. The above proof yields a free nonabelian group of piecewise isometries of B onto itself. Answering a question of S. Wagon and the author, Craig Squier found a free nonabelian group of piecewise translations of an infinite cyclic group onto itself. However many elements of such groups must have many fixed points since, by the theorem of Banach stated in the introduction, there exists no paradoxical decomposition of $[0, 1]$ and since abelian groups are amenable.

2. The proofs of the lemmas 2.5 and hence 2.7 use the Axiom of Choice. This cannot be avoided since the parts of a paradoxical decomposition of B cannot be Lebesgue measurable.

3. Marczewski's problem (see [2]): Is the Banach–Tarski Paradox possible if the parts of the underlying decompositions are required to have the property of Baire? This is unknown for all the spaces R^n $(n \geqslant 3)$, S^n $(n \geqslant 2)$ and H^n $(n \geqslant 2)$.

4. We can prove that $H^2 \equiv$ (a half-plane of H^2) and that $H^2 \equiv$ (any solid angle of H^2). This follows from a cancellation law of type $nA \equiv nB \Rightarrow A \equiv B$ (see [2]) and a theorem proved in [1] which tells that H^2 can be split into two sets each of which is equivalent to H^2 by decomposition into two parts. Moreover the proof of [1] does not use the Axiom of Choice and all the parts are countable unions of disjoint convex sets. Can one establish the above theorem about hyperbolic half-planes and solid angles restricting the decomposition to Borel sets? (The problem is open since the cancellation law is based on the Axiom of Choice and we do not know if the parts of the finite decomposition which it requires can be made measurable or have the property of Baire. We know only (by a theory developed in [3]) that, if equivalence by decompositions into countably many parts is considered, then Borel parts can be secured.)

References

[0] R. C. Lyndon and J. L. Ullman, *Pairs of 2-by-2 matrices that generate free products*, Michigan Math. J. 15 (1968), 161–166; and *Groups generated by two parabolic linear fractional transformations*, Canadian J. Math. 21 (1969), 1388 1403.

[1] J. Mycielski and S. Wagon, *Large free groups of isometries and their geometrical uses*, l'Enseignement Mathématique, 30 (1984), 247–267.

[2] S. Wagon, *The Banach–Tarski Paradox*, Cambridge Univ. Press, 1985.

[3] A. Tarski, *Cardinal Algebras*, New York, Oxford 1949.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF COLORADO
Boulder, CO 80309-0426

Received 9 July 1987;
in revised form 21 September 1987