PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 101, Number 4, December 1987

k-TO-1 FUNCTIONS ON AN ARC

HIDEFUMI KATSUURA AND KENNETH R. KELLUM

(Communicated by Doug W. Curtis)

ABSTRACT. Recently Jo W. Heath [6] has shown that any 2-to-1 function from an arc onto a Hausdorff space must have infinitely many discontinuities. Here we investigate extending Heath's result to k-to-1 functions for $k > 2$. Examples show that in general Heath's theorem cannot be extended even for functions from an arc into itself. However, if f is a k-to-1 function ($k \geq 2$) from an arc *onto* an arc, then we prove that f has infinitely many discontinuities.

Introduction. A function f is called k-to-1 for some integer k if $f^{-1}(y)$ contains exactly k points for each y in the image of f. Since Harrold [4] proved that no continuous 2-to-1 function exists on $[0, 1]$, several mathematicians have considered k-to-1 functions. It is known that no 2-to-1 continuous function exists on an n-cell ($n = 1, 2, 3$) (see Harrold [4], Roberts [10] and Civin [2]). This problem remains open for $n > 3$. Other relevant papers are listed in the references. Recently Heath [6] has shown that any 2-to-1 function from an arc onto a Hausdorff space must have infinitely many discontinuities. Here we investigate extending Heath's result to k-to-1 functions for $k > 2$. It follows from Harrold [5] that no continuous k-to-1 function exists from an arc to an arc for $k \geq 2$. Examples show that in general Heath's theorem cannot be extended even for functions from an arc into itself. However, if f is a k-to-1 function ($k \geq 2$) from an arc *onto* an arc, then we prove that f has infinitely many discontinuities. Our proofs are elementary, making use only of the sequential compactness of an interval, the intermediate value theorem, the extreme value theorem and simple divisibility facts.

Examples. In [4], Harrold gives an example, which he attributes to G. E. Schweigert, indicated in Figure 1, of a 3-to-1 continuous function from the interval $[0, 1]$ onto a circle. By appending the half-open interval $(1, 2]$ and wrapping it once around the circle one can obtain a 4-to-1 continuous function from $[0, 2]$ onto a circle and this can clearly be extended to show that for any $k > 2$ there exists a continuous k-to-1 function from an arc onto a circle.

In Figure 2 we have "unwound" Harrold's example to get an example of a 3-to-1 function f from $[0, 1]$ into itself with only one discontinuity. If we extend the domain of f to include $(1, 2]$ by defining $f(x) = 2 - x$ for $x \in (1, 2]$, we get a 4-to-1 function with one discontinuity. Clearly we can continue this process to show that for each $k \geq 4$ there exists a k-to-1 function from an arc *into* an arc with only $k - 3$ discontinuities. This naturally leads to

QUESTION 1. For each $k > 2$ what is the minimum number of discontinuities for a k-to-1 function from an arc into itself?

Received by the editors May 14, 1986 and, in revised form, August 12, 1986.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 26A15, 26A03, 54C30, 54C10.
Key words and phrases. k-to-1 function, 2-to-1 function.

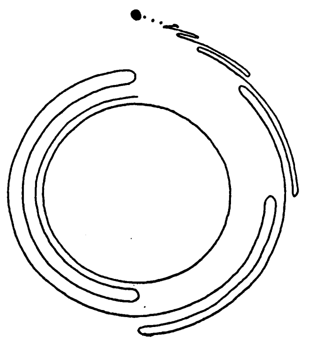

FIGURE 1

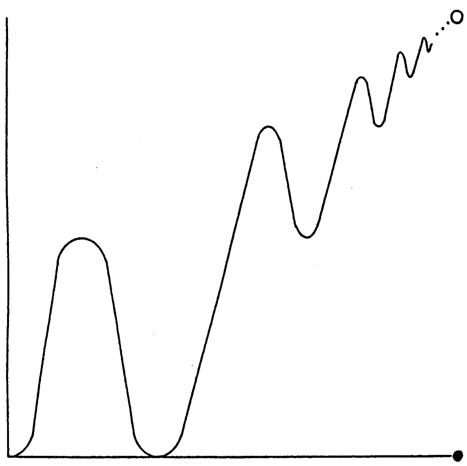

FIGURE 2

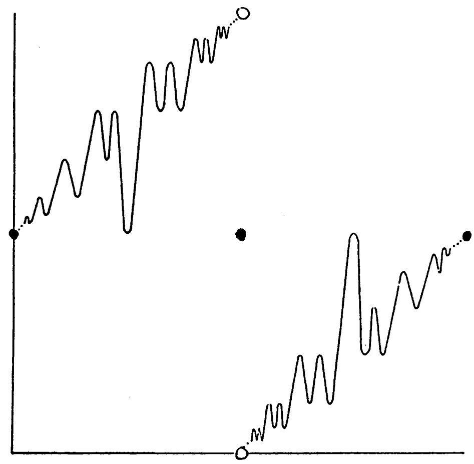

FIGURE 3

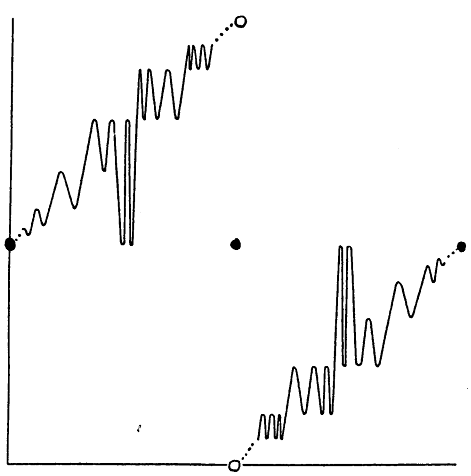

FIGURE 4

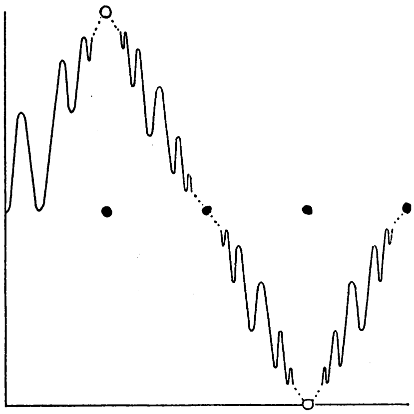

FIGURE 5

In Figure 3, we have an example of a 5-to-1 function with one discontinuity. If one carefully replaces some line segments in this example with "N's", the example can be made to be 7-to-1 (see Figure 4), and this can be repeated to see that for odd k the answer to Question 1 is 1. For even k we do not know. Figure 5 is an example of a 6-to-1 function with two discontinuities.

Theorems. The following theorem shows that the above examples cannot be greatly simplified. The technique of the proof is the basis for the proof of our main theorem.

THEOREM 1. *If $f\colon [0,1] \to [0,1]$ is k-to-1 ($k \geq 2$) and has finitely many discontinuities, then f must have infinitely many local extrema.*

PROOF. Let D be the set of discontinuities, let E be the set of points at which local extrema occur, and assume that each of these sets is finite. Write $f^{-1}(f(D \cup E \cup \{0,1\}))$ as $0 = x_0 < x_1 < \cdots < x_m = 1$. Then for $i = 1, 2, \ldots, m$, f restricted to (x_{i-1}, x_i) is continuous and one-to-one. Suppose U and V are two open intervals of the form (x_{j-1}, x_j), and suppose that $f(U) \cap f(V) \neq \emptyset$. Now, $f(U)$ and $f(V)$ are open intervals, say (p,q) and (r,s). If we assume that $f(U) \neq f(V)$, then the end-point of one open interval is in the other, say $p \in (r,s)$. But since $f^{-1}(p)$ consists of exactly k points at which f is continuous and locally monotone, for sufficiently small $\varepsilon > 0$, $f^{-1}(p+\varepsilon)$ would have $k+1$ elements. Thus if $f(U) \cap f(V) \neq \emptyset$, then $f(U) = f(V)$. Thus, for each (x_{i-1}, x_i), $f^{-1}(f(x_{i-1}, x_i))$ is the union of exactly k of these open intervals. Hence, the number of intervals on the x-axis, m, must be divisible by k. However, f is also k-to-1 on $x_0, x_1, \ldots, x_m$, which implies that $m+1$ is also divisible by k, a contradiction.

That a function is *finite-to-one* means that the preimage of each point is finite.

The proofs of Lemmas 1, 2 and 3 are left to the reader.

LEMMA 1. *Suppose f: $(0,1) \to (0,1)$ is finite-to-1 and continuous. If $x \in (0,1)$ then:*

(1) *There exists p such that $0 < p < x$ and either $f[p,x] \subseteq (0, f(x)]$ or $f[p,x] \subseteq [f(x), 1)$.*

(2) *There exists q such that $x < q < 1$ and either $f[x,q] \subseteq (0, f(x)]$ or $f[x,q] \subseteq [f(x), 1)$.*

LEMMA 2. *Let f be as in Lemma 1. Then the limit from the right of f at 0 and the limit from the left of f at 1 both exist.*

LEMMA 3. *Let f be as in Lemma 1. If s is the limit from the right of f at 0, and s is neither 0 nor 1, then there exists d, with $0 < d < 1$, such that either $f(0,d) \subseteq (0,s)$ or $f(0,d) \subseteq (s,1)$. Also, the analogous condition holds at 1.*

LEMMA 4. *Let f be a k-to-1 continuous function from an open set U of real numbers onto $(0,1)$. If (a,b) is a component of U then the limit from the right of f at a is either 0 or 1 and the limit from the left of f at b is also either 0 or 1.*

PROOF. By Lemma 2 the limit from the right of f at a exists, call it y and assume that y is neither 0 nor 1. By Lemma 3 there exists d such that $(a,d) \subseteq U$ and $f(a,d)$ is a subset of $(y,1)$ or of $(0,y)$, say the latter. Let $f^{-1}(y)$ be written as $c_1 < c_2 < \cdots < c_k$ and note that none of the c_i's is in (a,d). For each $j = 1, 2, \ldots, k$, let d_j and e_j be numbers, guaranteed by Lemma 1, such that $d_j < c_j < e_j$, $[d_j, e_j] \subseteq U$, $f[d_j, c_j]$ is a subset of $(0, y]$ or of $[y, 1)$ and $f[c_j, e_j]$ is a subset of $(0, y]$ or of $[y, 1)$. Without loss of generality we may assume that $d < d_1 < c_1 < e_1 < d_2 < c_2 < e_2 < \cdots < d_k < c_k < e_k$. For each $j = 1, 2, \dots, k$, $f[d_j, c_j]$ and $f[c_j, e_j]$ are nondegenerate intervals. Since f is k-to-1, at most k of these intervals are subsets of $(0, y]$, and at most k of them are subsets of $[y, 1)$. Hence, exactly k intervals of the $2k$ intervals $f[d_j, c_j]$ and $f[c_j, e_j]$ are subsets of $(0, y]$ and each of these k intervals contains y. Now, we can choose a number u in $f(a, d)$ so that u is large enough to be in each of the above k intervals which are subsets of $(0, y]$. This means that $f^{-1}(u)$ has at least k different members outside of (a, d). But since $f^{-1}(u)$ must also have a point in (a, d), this contradicts the fact that $f^{-1}(u)$ has k elements.

THEOREM 2. *Suppose f is a k-to-1 continuous function from an open set U of real numbers onto $(0,1)$. Then the number of components of U is not more than k.*

PROOF. Assume, on the contrary, that U has components C_i for $i = 1, 2, \dots, k+1$. By Lemma 4, for each $i = 1, 2, \dots, k+1$, $f(C_i)$ is of one of the forms $(0,1)$, $[y, 1)$ or $(0, y]$. Let P be the set of integers i such that $f(C_i)$ is of the form $(0, 1)$, Q the set of integers i such that $f(C_i)$ is of the form $[y, 1)$, and R the set of integers i such that $f(C_i)$ is of the form $(0, y]$. Let p, q and r denote the number of elements in P, Q and R, respectively. Then $p + q + r = k + 1$ and we may assume without loss that $q \leq r$. (If $r = q = 0$, the result follows easily.) For each $i \in R$, let $f(C_i) = (0, y_i]$ for some $0 < y_i < 1$. Choose a number u such that $0 < u < \min\{y_i : i \in R\}$. Then

$$\begin{aligned} k = |f^{-1}(u))| &\geq \sum_{i \in P} |f^{-1}(u) \cap C_i| + \sum_{i \in R} |f^{-1}(u) \cap C_i| \\ &\geq p + 2r \geq p + q + r = k + 1, \end{aligned}$$

a contradiction.

THEOREM 3. *If $f\colon [0,1] \to [0,1]$ is a k-to-1 surjection with $k \geq 2$, then f has infinitely many discontinuities.*

PROOF. Assume f is only discontinuous at $c_1, c_2, \dots, c_p$, for some integer p. Let

$$\{0, 1, f(0), f(1)\} \cup \{f(c_i) : i = 1, 2, \dots, p\} = \{y_1, y_2, \dots, y_n\}$$

for some integer n, where the y_i's are labeled so that $0 = y_1 < y_2 < \cdots < y_n = 1$. Write $\bigcup_{i=1}^{n} f^{-1}(y_i)$ as $0 = x_1 < x_2 < \cdots < x_{kn} = 1$. For each $i = 2, 3, \dots, n$, let $f^{-1}(y_{i-1}, y_i) = U_i$ and let $g_i = f|_{U_i} \colon U_i \to (y_{i-1}, y_i)$. Then for each $i = 2, 3, \dots, n$, U_i is the union of finitely many open intervals of the form (x_{j-1}, x_j), say m_i many. Hence

$$\sum_{i=2}^{n} m_i = kn - 1. \tag{$*$}$$

On the other hand, for each $i = 2, 3, \dots, n$, $g_i \colon U_i \to (y_{i-1}, y_i)$ is a k-to-1 continuous surjection. Thus Theorem 2 applies, and it implies that, for each $i = 2, 3, \dots, n$, $m_i \leq k$. Hence

$$\sum_{i=2}^{n} m_i \leq \sum_{i=2}^{n} k = k(n-1) = kn - k.$$

But since $k \geq 2$, this contradicts $(*)$.

QUESTION 2. Harrold [**5**] showed that no arc is the image of a continuum under a continuous k-to-1 map. Nadler and Ward [**8**] have generalized this result to some spaces other than arcs. If X is a continuum and f is a k-to-1 function from X onto $[0, 1]$, must f have infinitely many discontinuities?

NOTE. After this paper was accepted for publication, Professor Jo W. Heath settled Question 1 by proving that for even $k > 4$, the answer is 2. She also answered Question 2 in the affirmative.

REFERENCES

1. K. Borsuk and R. Molski, *On a class of continuous maps*, Fund. Math. **45** (1958), 84–98.
2. P. Civin, *Two-to-one mappings of manifolds*, Duke Math. J. **10** (1943), 49–57.
3. P. Gilbert, *n-to-one mappings of linear graphs*, Duke Math. J. **9** (1942), 475–486.
4. O. G. Harrold, *The non-existence of a certain type of continuous transformation*, Duke Math. J. **5** (1939), 789–793.
5. O. G. Harrold, *Exactly* $(k, 1)$ *transformations on connected linear graphs*, Amer. J. Math. **62** (1940), 823–834.
6. J. W. Heath, *Every exactly* 2-*to*-1 *function on the reals has an infinite set of discontinuities*, Proc. Amer. Math. Soc. **98** (1986), 369–373.
7. V. Martin and J. H. Roberts, *Two-to-one transformations on* 2-*manifolds*, Trans. Amer. Math. Soc. **49** (1941), 1–17.
8. J. Mioduszewski, *On two-to-one continuous functions*, Dissertationes Math. (Rozprawy Mat.) **24** (1961).
9. S. B. Nadler, Jr. and L. E. Ward, Jr., *Concerning exactly* $(n, 1)$ *images of continua*, Proc. Amer. Math. Soc. **87** (1983), 351–354.
10. J. H. Roberts, *Two-to-one transformations*, Duke Math. J. **6** (1940), 256–262.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, SAN JOSÉ STATE UNIVERSITY, SAN JOSÉ, CALIFORNIA 95192