APPLICATIONS OF ALMOST ONE-TO-ONE MAPS

ALEXANDER BLOKH, LEX OVERSTEEGEN, AND E. D. TYMCHATYN

Abstract. A continuous map $f : X \to Y$ of topological spaces X, Y is said to be *almost 1-to-1* if the set of the points $x \in X$ such that $f^{-1}(f(x)) = \{x\}$ is dense in X; it is said to be *light* if pointwise preimages are 0-dimensional. In a previous paper we showed that sometimes almost one-to-one light maps of compact and σ-compact spaces must be homeomorphisms or embeddings. In this paper we introduce a similar notion of an almost d-to-1 map and extend the above results to them and other related maps. In a forthcoming paper we use these results and show that if f is a minimal self-mapping of a 2-manifold then point preimages under f are tree-like continua and either M is a union of 2-tori, or M is a union of Klein bottles permuted by f.

1. Introduction

In our previous paper [BOT02] we study almost 1-to-1 maps and prove that under some assumptions they are homeomorphisms. We are motivated by a few factors of which an important one is that certain classically studied maps (such as minimal maps, see [KST01]) are in fact almost 1-to-1 so that our tools would apply to them. The main result of [BOT02] is the following theorem.

Theorem 1.1. *Suppose that $f : M \to N$ is a light and almost 1-to-1 map from an n-manifold M into a connected n-manifold N. Then*

$$f|_{M\setminus\partial M} : M \setminus \partial M \to N$$

is an embedding. In particular, if M is a closed manifold, then f is a homeomorphism.

Together with some dynamical arguments it allows us to prove the following theorem which will appear in a forthcoming paper [BOT03]. By a *minimal* map we mean a map $f : X \to X$ of a topological space X into itself such that all its orbits are dense in X (the *orbit* of a point x

Date: June 20, 2003.

The first author was partially supported by NSF Grant DMS-0140349, the second author by NSF grant DMS-0072626, and the third author by NSERC grant OGP00561.

is the set of points $\{x, f(x), f^2(x), \dots\}$). Also, a continuum X is *tree-like* provided for every $\varepsilon > 0$ there exists an open cover of mesh less than ε whose nerve is a finite tree (alternatively, X is homeomorphic to the inverse limit of finite trees)

Theorem 1.2. *Suppose that $f : M \to M$ is a minimal map of a 2-manifold. Then f is a monotone map with tree-like point inverses and M is either a finite union of tori, or a finite union of Klein bottles which are cyclically permuted by f.*

In this paper we extend results of [BOT02] onto *(branched) covering* maps. Let us fix some general notation and terminology. All spaces are separable and metric. For a subset Y of a topological space X, we denote the *boundary* of Y by $\mathrm{Bd}(Y)$ and the *interior* of Y by $\mathrm{Int}(Y)$. A *continuum* is a compact and connected space. A locally connected continuum containing no subsets homeomorphic to the circle is called a *dendrite.* By a *dendroid* we mean an arcwise connected continuum containing no subsets homeomorphic to the circle. We rely upon the standard definition of a *closed n-manifold* (compact connected manifold without boundary); by a *manifold* we mean a manifold of any sort (closed, manifolds with boundary, open). In the case of a compact manifold M *with* boundary its boundary is called the *manifold boundary* of M and denoted by ∂M. Thus, if D is homeomorphic to the closed unit ball in a Euclidean space then $\mathrm{Bd}(D)$ is empty while ∂D is the corresponding unit sphere.

A map $f : X \to Y$ of a locally connected space X onto a locally connected space Y is said to be a *covering map* if every point $y \in Y$ has a connected neighborhood U that is *evenly covered* by f, i.e. such that each component of the preimage $f^{-1}(U)$ is mapped by f homeomorphically onto U. If the number of components of the preimage of U is the same for all points and is equal to k, the map is said to be *k-fold covering.* By a *branched covering* map $f : M \to N$ between two locally connected continua M and N we mean an open map f for which there exists a finite set C such that $f|_{M \setminus C} : M \setminus C \to N \setminus f(C)$ is a covering map. Finally, a set X is said to be *ε-separated* if the distance between every two points of X is at least ε. Then a map $f : X \to Y$ is said to be *almost separated* if there exists ε and a dense in $f(X)$ set of points $A \subset f(X)$ such that for each point $x \in A$ the set $f^{-1}(x)$ is ε-separated.

Our main results deal with covering and branched covering maps. More precisely, we will show that almost separated maps between closed manifolds are covering maps. Subsequently we will obtain similar results for branched covering maps of manifolds. An interesting fact which we establish along the way is that a continuous light map of a compact n-manifold X into a compact n-manifold Y, whose critical value set is of codimension at least 2, is open. In particular, in the S^2-case, if the critical value set is totally disconnected, then the map is open and, hence, by the Stoilow theorem it is a branched covering map with finite critical set. Thus the critical value set of a light map $f : S^2 \to S^2$ is either finite or has non-trivial components. Finally, we extend these results to arbitrary (not only light) maps relying upon the monotone-light decomposition of maps.

We would like to add here that a great source of information and methods in the field is a well-known Whyburn's book [Why42].

2. Results concerning almost 1-to-1 maps

To simplify reading we summarize the results and the definitions from [BOT02] in this section. We begin with some technical results. We say that a map $f : X \to Y$ is *quasi-interior* provided for each open subset $U \subset X$, the interior of $f(U)$ is non-empty. Suppose that $f : X \to Y$ is a map of a metric σ-compact space X into a metric σ-compact space Y. Denote by R_f the set of all points $y \in Y$ which have unique preimages under f and by D_f the set of all points $x \in X$ such that $f^{-1}(f(x)) = x$. The following lemma summarizes some technical results obtained in [BOT02].

Lemma 2.1. *Suppose that $f : X \to Y$ is a continuous map of metric σ-compact spaces. Then D_f is a G_δ subset of X while the set R_f is a G_δ-subset of $f(X)$. Moreover, if f is a closed map (e.g., if X is compact) then the following properties are equivalent:*

1. *D_f is dense in X,*
2. *R_f is dense in $f(X)$ and f is quasi-interior as a map from X to $f(X)$.*

The next lemma from [BOT02] shows that light maps of n-manifolds are quasi-interior (essentially, this is so because such maps do not lower dimension). The converse is false as can be seen from the map of the 2-sphere onto itself with just one non-degenerate arc fiber.

Lemma 2.2. *If X and Y are n-manifolds then any light map of X to Y is quasi-interior. Thus, the interior of $f(X)$ in Y is dense in $f(X)$ and all relatively open subsets of $f(X)$ have non-empty interior in Y.*

The next results from [BOT02] deal with one-dimensional continua such as dendrites and dendroids (defined in Introduction). To state it we need the following definition: a subset A of a topological space Y is said to be *con-dense* if each non-degenerate continuum in Y contains a point of A.

Theorem 2.3. *Suppose that $f : X \to Y$ is a light map, X is a continuum and Y is a dendrite. Moreover, suppose that the set R_f is con-dense in $f(X)$. Then f is an embedding (so if R_f is also dense in Y then f is a homeomorphism of X onto Y).*

The following analog of Theorem 2.3 deals with dendroids.

Proposition 2.4. *Suppose that $f : X \to Y$ is a light map from an arcwise connected continuum X into a dendroid Y. Moreover, suppose that the set R_f is con-dense in $f(X)$. Then f is an embedding.*

To state the next theorem we need the following definition: a map $f : X \to Y$ from a continuum X to a continuum Y is *weakly-confluent* provided for every continuum $K \subset Y$ there exists a component C of $f^{-1}(K)$ such that $f(C) = K$. Now we are ready to state Theorem 2.5 ([BOT02]) which provides a mechanism for proving that a map is a homeomorphism if it is a 1-to-1 map on a dense set of points.

Theorem 2.5. *Suppose that $g : X \to Y$ is a weakly confluent, light and almost 1-to-1 mapping of a continuum X onto a locally connected continuum Y. Then g is a homeomorphism.*

These results and some extra-arguments allow us to prove Theorem 1.1 ([BOT02]) quoted in Introduction. Finally, we state one more result from [BOT02] which deals with general properties of light maps and which we will generalize to arbitrary maps below.

Theorem 2.6. *Suppose that $f : X \to Y$ is a light mapping from a continuum X onto a continuum Y.*

Then either:

1. *there exists an open set U and a dense G_δ-subset D' of $f(U)$ such that for all $y \in D'$,*
$$|f^{-1}(y) \cap U| = 1, \text{ or}$$
2. *there exists a dense G_δ-subset D of Y such that for each $y \in D$, $f^{-1}(y)$ is homeomorphic to the Cantor set.*

3. Light maps which are not almost 1-to-1

The main idea behind the results of [BOT02] is that, depending on the properties of topological spaces, if a map is 1-to-1 on a set which is either dense or is in some other way "widespread", then the map is actually a homeomorphism or at least an embedding. Our next task is to see how these results can be extended onto other classes of maps. There are two classes of non-invertible (non-injective) maps which can be considered more or less analogous to homeomorphisms/embeddings here, namely covering maps and, to a lesser degree, branched covering maps. Since we want to be able to conclude that a map in question is a covering or a branched covering, it is natural to change the assumption that the map is almost 1-to-1 and replace it by the assumption that points with finite preimages of the same cardinality are dense. Thus, a natural notion extending that of an almost 1-to-1 map seems to be the following one: a map $f : M \to N$ is said to be *almost d-to-1* if it is quasi-interior and there exists a dense subset R of $f(M)$ each of whose point-preimages have cardinality d. It will also be convenient to call a set $A \subset Y$ is said to be *massive* (in Y) if it contains a dense G_δ-subset of Y.

Let $f : M \to N$ and denote the set of points in N with exactly i preimages by R^i_f, the set of points in N with more than i preimages by E^i_f, and the set of points in N with at most i preimages by $G^i_f = N \setminus E^i_f$. Then $G^i_f = \cup^i_{j=0} R^j_f$. The following lemma is analogous to parts of Lemma 2.1.

Lemma 3.1. *Suppose that $f : M \to N$ is a continuous map of metric σ-compact spaces. Then for any $k \geq 0$ the set G^k_f is a G_δ-set.*

Proof. Without loss of generality we may assume that $f(M) = N$. It is enough to show that the set E^k_f is an F_σ-subset of N. Since M is σ-compact we can represent it as a countable union of metric compacta $X_i, i = 0, 1, \dots$. It suffices to prove that $E^k_f \cap f(X_i)$ is an F_σ-set. Thus it is enough to prove the lemma assuming that both M and N are compact.

Now, for every m let B_m be the set of points in N such that there are $k+1$ points in their preimages which are $1/m$-separated. It is easy to show that the set B_m is closed for every m. Since $E^k_f = \cup^\infty_{m=0} B_m$, we are done. □

To describe the properties of the set R^k_f is more complicated. However for almost k-to-1 maps of n-manifolds the set R^k_f contains a dense open subset of $f(M)$ as the lemma below shows. To state it we need the following definition. Suppose that $f : X \to Y$ is a map between two locally connected continua. We say that a point $y \in Y$ is *k-covered* if there exists a connected neighborhood U of y such that its preimage set $f^{-1}(U)$ can be represented as the union of k pairwise disjoint open sets $V_1, \dots, V_k$ such that $f|_{V_j} : V_j \to U$ is a homeomorphism between V_j and U for each $1 \leq j \leq k$. By the definition the set of all k-covered points is open.

Lemma 3.2. *Let $f : M \to N$ is an almost k-to-1 light map of two n-manifolds. Then the set R^k_f contains an open dense subset A of $f(M)$ such that each point of A is k-covered.*

Proof. First we prove that R_f^k is massive in $f(M)$ and that G_f^{k-1} is nowhere dense in $f(M)$. Observe that the set $R_f^1 = R_f$ is a G_δ-subset of $f(M)$ by Lemma 2.1 even in the general case of σ-compact spaces. We prove the rest by induction.

Suppose that f is almost k-to-1. By the assumption and Lemma 3.1 G_f^k is massive in $f(M)$. It would suffice if we proved that G_f^{k-1} is nowhere dense in $f(M)$ because then R_f^k would contain the complement of the closure of G_f^{k-1} intersected with G_f^k, and since both these sets are massive in $f(M)$, we would be done.

By way of contradiction assume that G_f^{k-1} is dense in a ball in $f(M)$. By Theorem 2.2 we can then find a ball U in the interior of $f(M)$ in which G_f^{k-1} is dense too. By Lemma 3.1 $G_f^{k-1} \cap U$ is massive in U. On the other hand it is well-known that the set Y of points, on which taking f-preimage sets is continuous, is massive in $f(M)$. Hence the set $C = U \cap G_f^{k-1} \cap Y$ is massive in U (so, it has no isolated points).

Take a point $y \in C$ such that it has maximal number of preimages possible for points in C. Denote its preimages by $y_1, \dots, y_r$. Then any point $z \in C$ close to y has the full preimage $f^{-1}(z) = \{z_1, \dots, z_r\}$ consisting of $r < k$ points which are close to $y_1, \dots, y_r$ respectively (the preimages are continuous in C, yet there cannot exist extra-preimages of z because r is chosen to be maximal for points of C).

Let us show that for a small neighborhood W_i of y_i the map $f|W_i$ is an embedding. Indeed, by Theorem 2.2 the interior of $f(W_i)$ is dense in $f(W_i)$, and hence $C \cap f(W_i)$ is dense in $f(W_i)$. On the other hand by the previous paragraph all points of $C \cap f(W_i)$ have unique preimages in W_i. Hence $f|W_i$ is almost 1-to-1 which implies by Theorem 1.1 that $f|W_i$ is an embedding. We can choose neighborhoods $W_1, \dots, W_r$ so that $f(W_i) = V$ is the same open neighborhood of y for all $i = 1, \dots, r$.

It is easy to see that points of V cannot have preimages other than those in the sets W_i. Indeed, otherwise there exists $a \notin \cup W_i$ such that $f(a) = b \in V$. Since f is quasi-interior we can find a neighborhood T of a whose image contains an open subset V' of V. Then any point of $C \cap V'$ will have at least $r+1$ preimages, a contradiction. So, all points of V have $r \le k-1$ preimages while by the assumption there is a dense set of points with exactly k preimages, a contradiction. We conclude that the set G_f^{k-1} is nowhere dense in $f(M)$ and so the set R_f^k is massive in $f(M)$ as explained above.

Now, consider the intersection C' of R_f^k with the set of points of continuity of preimage. For every point $z \in C'$ the number of preimages is k, and mimicking the above arguments we can show that the map around each point of $f^{-1}(z)$ is an embedding. So we have k balls mapped onto a neighborhood of z homeomorphically. Again mimicking the arguments from the previous paragraph we see that points of this neighborhood of z have no other preimages because otherwise we will be able to find an open set of points with at least $k+1$ preimages, a contradiction. Thus, if the map is light and almost k-to-1 then in fact there exists an open dense set of points each of which has a neighborhood W with $f^{-1}(W) = W_1 \cup W_2 \cup \cdots \cup W_k$ where the W_i are pairwise disjoint open sets each homeomorphic by f to W. □

Lemma 3.2 seems to be the best possible without any extra-assumptions. Indeed, we claim that there exist examples of continuous maps which are neither covering maps nor branched covering maps and which have the following properties.

(1) A map of the line to itself such that each point has exactly 3 preimages;
(2) a map of the circle such that all points except one have 3 preimages;
(3) a map of the interval such that all points except for the endpoints have 3 preimages;
(4) a map of the plane such that each point has 3 preimages.

First, consider a "zigzag" map of the line defined as follows:

(1) $f(2n) = n$ for any integer n;
(2) $f(2n+1) = n+2$ for any integer n;
(3) f is defined as a linear map on every interval with integer endpoints.

It is easy to see that under this map each point x of the line has exactly 3 preimages. Compactifying the line with one point we get a map of the circle into itself such that all points have three preimages except for the one point at infinity. Compactifying the line with two points we will get a similar interval map, now with two points having a unique preimage. In all these cases the maps in question are almost 3-to-1, but they are not open and hence not branched covering maps.

Moreover, we can extend the above defined map $f : \mathbb{R} \to \mathbb{R}$ of the line onto itself to a map $F : \mathbb{R}^2 \to \mathbb{R}^2$ defining F as a skew product over f which acts, say, as the identity on the extra-coordinate. Then $F = f \times \mathrm{id}$ is 3-to-1 everywhere. Note that F is not even a branched covering map of the plane. We can extend it to a map $\hat{F} : S^2 \to S^2$ of the 2-sphere by mapping the point at infinity to itself. Then $\tilde{F}$ is exactly 3-to-1 except at the point at infinity which has a one point preimage.

Developing this example further we can show that there exists a light map of the plane such that the origin has an uncountable preimage while all other points have exactly 3 preimages. Indeed, let $f : \mathbb{R} \to \mathbb{R}$ be the zigzag map from the previous paragraph. The map g is obtained from f by adding a point at $-\infty$ and defining $g(-\infty) = -\infty$. Hence the domain and range of g are homeomorphic to the half line $[0, \infty) = \mathbb{R}^+$, and we can consider g is a map from $[0, \infty)$ onto $[0, \infty)$ such that $g^{-1}(0) = 0$ and $|g^{-1}(x)| = 3$ for each $x > 0$. Let C denote the set of all critical points of g. Without loss of generality we may assume that g has a local maximum at 1 and that 1 is a fixed point of g. Suppose that the adjacent to 1 local extrema of g are the points $a < 1 < b < c < d$; without loss of generality we may assume that $g(d) = 1$. Observe that then by the construction $g(a) < g(b) < g(1) = g(d) = 1 < g(c)$. Again without loss of generality (composing our map with an appropriate homeomorphism of $\mathbb{R}^+$) we may assume that there exists a g-fixed point $u \in (0, 1)$ greater than both a and $g(b)$. Then because of the properties of zigzag functions $g([0, u]) = [0, u]$ and $g([u, 1]) = [u, 1]$.

Define a continuous one-parameter family of maps $\{C_t : [0, 1] \to [0, 1]\}_{t=0}^{t=1}$ so that the following properties hold:

(1) C_t is a homeomorphism for every $0 \le t < 1$;
(2) C_0 is the identity map:
(3) C_1 maps $[0, u]$ into 0 and $[u, 1]$ homeomorphically onto $[0, 1]$.

In other words, maps C_t contract $[0, u]$ to the point of collapsing it onto 0 and expand $[u, 1]$ to the point of mapping it onto the entire $[0, 1]$.

Now, consider the one parameter family of maps $g_t = C_t \circ g \circ C_t^{-1}$ for all $0 \le t < 1$. By the construction $g_0 = g$ while g_t converges to a continuous map g_1 such that the following holds:

(1) $g_1(0) = 0, g_1(1) = 1, g_1$ is an increasing homeomorphism on $[0, 1]$ mapping the interval $[0, 1]$ onto itself;
(2) g_1 is a decreasing homeomorphism on $[1, b]$ mapping the interval $[1, b]$ onto $[0, 1]$ so that $g_1(b) = 0$;
(3) g_1 is increasing on $[b, c]$ and maps $[b, c]$ onto $[0, g(c)]$;
(4) $g_1|_{[c,\infty)} = g|_{[c,\infty)}$ and so $g_1(d) = 1$;
(5) as follows from this, all positive integers have exactly 3 preimages under g_1 while 0 has only 2 preimages, 0 and b.

Let us now construct the map $G : \mathbb{R}^2 \to \mathbb{R}^2$ such that except for the origin all points have exactly 3 preimages while the preimage of the origin is uncountable. To this end choose a Cantor set $\hat{C}$ of arguments and define G on the corresponding rays as g_1 mapping each ray onto itself. On the other hand for every complementary interval of arguments

$U_n = (a_n, b_n)$ $(n = 1, 2, \dots)$ we pick its midpoint c_n and define G on the corresponding ray with argument c_n as $g_{1-1/n}$. Finally, the map G on the rays with arguments x between a_n and c_n (and c_n and b_n) is defined as the map g_t with the appropriate choice of $t \in [1-1/n, 1]$ so that $t \to 1$ as $x \to a_n$ $(x \to b_n$, respectively). The resulting map G is such that all points of the plane except for the origin have 3 preimages while the preimage of the origin is the union of the origin and a Cantor set of points with the arguments from $\hat{C}$ which are b-distant from the origin.

All this shows that some extra-assumptions must be made in order to come up with the desired conclusions that a map is covering or branched covering. To introduce them we need some definitions which mimic those from [BOT02]. Namely, a map $f : X \to Y$ is said to be *weakly d-confluent* if every sufficiently small connected closed set $A \subset Y$ has at least d components of its preimage, each of which has the image A. Now we are ready to prove the following theorem which extends Theorem 2.5.

Theorem 3.3. *Suppose that $f : X \to Y$ is a weakly k-confluent, light and almost k-to-1 mapping of a continuum X onto a locally connected continuum Y. Then each y in Y has exactly k preimages and f is a k-fold covering map.*

Proof. Since f is weakly k-confluent, $|f^{-1}(y)| \geq k$ for each $y \in Y$. By way of contradiction suppose that there exists a point $y \in Y$ such that $f^{-1}(y)$ contains at least $k+1$ points. Since f is light and Y is locally connected there exists a sufficiently small open connected neighborhood V of y such that if $C = \overline{V}$, then $f^{-1}(C)$ has at least $k+1$ components which meet $f^{-1}(y)$. Since f is weakly k-confluent there exist k components K_i of $f^{-1}(C)$ such that $f(K_i) = C$. Then $f^{-1}(C \cap R^k_f) \subset \cup K_i = K$ (recall that R^k_f is the set of all points in the range with exactly k preimages). Let x be a point in $f^{-1}(y) \setminus K$ and let U be an open neighborhood of x such that $U \cap K = \emptyset$ and $f(U) \subset V$. Since f is quasi-interior (recall that by the definition of an almost k-to-1 map any such map is quasi-interior), $f(U) \cap R^k_f \neq \emptyset$ contradicting the fact that $f^{-1}(V \cap R^k_f) \subset K$. Thus, each point of Y has exactly k preimages. The fact that then f is a k-fold covering map easily follows and completes the proof. □

The main problem with Theorem 3.3 is that the hypotheses for it are hard to verify. Of course the same applies to Theorem 2.5 from [BOT02] (see Section 2). To improve Theorem 2.5 we may need to prove an analog of Theorem 1.1 of [BOT02] in the setting of d-to-1 maps. As the proof of Theorem 1.1 in [BOT02] shows, one important ingredient of the arguments there is the verification of the fact that in the case of manifolds light and almost 1-to-1 maps are weakly confluent. An analogous result in the setting of d-to-1 maps would have to have a conclusion that under certain assumptions the map is weakly d-confluent. As the above examples show, if the assumptions are that a map is almost d-to-1 and light, this is not necessarily the case.

Thus we need another approach. The one we adopt here is based upon the notion of an almost separated map given in Introduction. The aim is to show that almost separated maps are covering maps, or at least that they are covering maps on big subsets of their domains. Observe that claims like that would be in the spirit of Lemma 3.2 where we prove that a light almost k-to-1 map has an open dense set of k-covered points in the range. Now we specify what kind of set it is if the map is not just almost 1-to-1, but almost separated.

Theorem 3.4. *Suppose that $f : M \to N$ is an almost separated light map between two n-manifolds. Then f is a local embedding at any point of $M \setminus \partial M$. In particular, if M is compact then every point $y \in N \setminus (\partial N \cup f(\partial M))$ is k-covered for some $k \in \{0, 1, \dots\}$. Moreover, if M is a union of closed manifolds then f is a covering map.*

Proof. We will rely upon the tools developed in [BOT02] to study almost 1-to-1 maps and apply them to study almost separated maps. Let $A \subset f(M)$ be the dense set of points whose preimages are ε-separated for some $\varepsilon > 0$. Consider a point $x \in M \setminus \partial M$ and choose a small neighborhood U of x of diameter less than ε. Then every point y from the set $A \cap f(U)$ has exactly one preimage in U because all other preimages of y are too far. On the other hand, $f(U)$ has interior, and moreover by Theorem 2.2 the interior of $f(U)$ is dense in $f(U)$. Since A is dense in $f(M)$, it is dense in the interior of $f(U)$, so finally we conclude that $A \cap f(U)$ is dense in $f(U)$. By Theorem 1.1 $f|U$ is an embedding which proves the first claim of the theorem. This implies that if M is compact then every point $y \in N \setminus (\partial N \cup f(\partial M))$ is k-covered for some k and proves the second claim of Theorem 3.4.

Thus we extend Theorem 1.1 which deals with almost 1-to-1 maps and proves that they are global embeddings outside the boundary of the manifold. Note that in general a local homeomorphism is not necessarily a covering map (i.e., consider the map $f : \mathbb{R}^+ \to S^1$ defined by $f(x) = e^{2\pi i x}$ and pull back neighborhoods of the point with the argument 0). To prove the third claim of the theorem we need the following well-known fact: a local homeomorphism g of a closed n-manifold M

into a connected n-manifold N is a covering map. Applying this to the components of M we prove the third claim of the theorem. $\square$

The result of Theorem 3.4 cannot be strengthened unless we make additional assumptions about the map. Indeed, consider a closed 2-disk and visualize it as a long and thin strip. Map it onto a closed unit 2-disk by making it go around to begin with and then making various twists and turns as if the image follows some 1-dimensional path inside the unit disk. Clearly this can be done in such a way that the map will be almost separated (in fact even separated in the sense that for some ε all point preimages are ε-separated sets) while the set of k-covered points with various k's is exactly the one described in the above theorem.

As we say, this result can be extended by making additional assumptions. A map $f : X \to Y$ is *confluent* provided it is onto and for each subcontinuum $K \subset Y$ and each component C of $f^{-1}(K)$, $f(C) = K$. It is well known that all open onto mappings between compacta are confluent.

Theorem 3.5. *Suppose that $f : M \to N$ is a confluent local homeomorphism between manifolds M and N (e.g., f can be almost separated, light and confluent map of an n-manifold M without boundary). Then f is a covering map.*

Proof. Suppose that f is given as above and $y \in N$. Choose a closed ball $B \in N$ containing y in its interior. Let C be a component of $f^{-1}(B)$. Since f is a local homeomorphism, C is arcwise connected.

We claim that $f|_C$ is one-to-one. Suppose not, then there exist $x \neq y \in C$ such that $f(x) = f(y)$. Let $A = \alpha([0, 1])$ be an arc in C joining x to y. Then $f(A)$ is a closed path in B. Let H be a homotopy in B which contracts $f(A)$ to a point while fixing the point $f(x)$. Since f is a local homeomorphism the homotopy H lifts to C while fixing the points x and y. This allows us to find an arc in C with the endpoints at x and y mapped into a point, a contradiction. Hence B is evenly covered by the components of $f^{-1}(B)$ and f is a covering map. $\square$

Obviously, if the map maps an n-manifold M with boundary onto an n-manifold N without boundary it cannot be covering because points which belong to the image of the boundary of M will not have neighborhoods U with all components of the preimage of U homeomorphic to U. It is easy to suggest examples of such maps. Indeed, consider the closed unit disc identified with the complement T to the open unit disc in the closed complex sphere $\overline{\mathbb{C}}$. Define the map f of T onto S^2 such that f identifies every point with polar coordinates $(1, \theta)$ with the point having polar coordinates $(1, -\theta)$ (so the identification simply flattens the unit circle), and f is one-to-one elsewhere. Clearly, this is a closed, onto, almost 1-to-1 map $f : T \to S^2$, however at points of the unit circle it is not a local homeomorphism, and at points with polar coordinates $(1, 0)$ and $(1, \pi/2)$ the map is not even a local embedding. The map is clearly not confluent either.

Of course the idea of the previous example comes from Caratheodory theory. Indeed, suppose that J is a dendrite in the complex sphere. Then the Riemann map maps the open unit disc in the complex plane onto the complement of J in the complex sphere. Since J is locally connected, this maps extends to a map of the closed unit disc onto the entire sphere such that the boundary of the unit disc is mapped onto J. This extended map is one-to-one on the open disc and, hence, almost 1-to-1. However, it does not have nice properties everywhere because of its behavior at points of S^1, i.e. at points of the boundary of the unit disk.

These examples can be extended onto the entire 2-sphere. Indeed, we can identify the inside and the outside of the unit circle by means of the inversion map $\hat{I}$, and then extend the maps onto the entire S^2 so that the points y and $\hat{I}(y)$ are mapped into the same point. The resulting map of the sphere has its "peak" at the unit circle which gets squashed into the dendrite whereas each of the two complementary components of the unit circle is mapped homeomorphically onto the complement of J (the outside and the inside of the unit circle are actually mapped onto the complement of J homeomorphically). Observe that in this case the image of the set of all points at which the map is not a local embedding is 1-dimensional.

Now, let us see how these examples help study arbitrary light maps. We need the following notation and definitions. Let $f : M \to N$ be a light map of closed n-manifolds. For $\varepsilon > 0$ denote by S_ε the set of all points such that their preimages are ε-separated, let B_ε be the interior of the set $\overline{S_\varepsilon}$, and let $E_\varepsilon = f^{-1}(B_\varepsilon)$. Also, let $B = \cup_{\varepsilon>0} B_\varepsilon$ and $E = f^{-1}(B)$. Furthermore, a point x is said to be a *critical* point of a map $f : M \to N$ if the map f is not a local embedding in any neighborhood of x. Denote by C the set of all critical points of f, call C the *critical set* of f and call $f(C)$ the *critical value set* of f. Clearly, when M is compact, both sets C and $f(C)$ are compact.

We establish several properties of the just introduced sets in the lemma below.

Lemma 3.6. *In the notation introduced above the following holds.*

(1) *$f|E_\varepsilon$ a local embedding (in particular there are no critical points of f in E_ε);*
(2) *the entire set B_ε is contained in S_δ for each $0 < \delta < \varepsilon$;*
(3) *$C \cap E = \emptyset$;*
(4) *$f(C) \cap B = \emptyset$ and $f(C) \cup B = f(M)$.*

Proof. (1) Clearly, S_ε is dense in B_ε, therefore the map f restricted onto the set $f^{-1}(B_\varepsilon) = E_\varepsilon$ is almost separated. By Theorem 3.4 we conclude that $f|E_\varepsilon$ a local embedding. In particular there are no critical points of f in E_ε.

(2) Indeed, otherwise there is a point $y \in B_\varepsilon$ and $x_1, x_2 \in f^{-1}(y)$ such that $d(x_1, x_2) < \delta$. By (1) we can choose disjoint open neighborhoods U_i of x_i of diameter less than $r = 0.25(\varepsilon - \delta)$ such that $f(U_1) = f(U_2) = V$ is an open neighborhood of y. Every point $w \in V$ has the preimage $w' \in U_1$ and the preimage $w'' \in U_2$ of which we can say that $w' \neq w''$ (because U_1 and U_2 are disjoint) while by the triangle inequality $d(w', w'') < \delta + 0.5(\varepsilon - \delta) < \varepsilon$. This implies that there are no points of S_ε in V, a contradiction with $y \in B_\varepsilon$ and V being a neighborhood of y (recall that S_ε is dense in B_ε).

(3) It follows from (1) that $C \cap \cup_{\varepsilon>0} E_\varepsilon = C \cap E = \emptyset$.

(4) The sets $f(C)$ and B are disjoint by (2). On the other hand, consider a point $y \in f(M) \setminus B$. If the set $f^{-1}(y)$ is infinite then clearly there are critical points in it (recall that M is compact) and so $y \in f(C)$. If the set $f^{-1}(y) = \{y_1, \dots, y_l\}$ is finite and is disjoint from C then we can choose small neighborhoods $\{W_1, \dots, W_l\}$ of points of $f^{-1}(y)$ whose closures are pairwise disjoint and whose homeomorphic images coincide with a small neighborhood V of y. The set $R = f(M \setminus (\cup_{i=1}^l W_i))$ is a compact subset of $f(M)$ not containing y, hence we can shrink sets W_i to smaller sets W'_i with the images coinciding with a small neighborhood V' of y such that $V' \cap R = \emptyset$. Then any point z of V' has the preimage $f^{-1}(z) = \{z_1, \dots, z_l\}$ such that $z_i \in W'_i$ for each i. By the choice of the neighborhoods W_i this implies that for some $\varepsilon > 0$ the set $f^{-1}(z)$ is ε-separated, a contradiction with the assumption that $y \notin B$. Hence $f(C) \cup B = f(M)$ as desired. $\square$

Observe that by Lemma 3.6 the set B coincides with the union of the interiors of sets S_ε themselves. Also, as we see from the same lemma, our map f is rather nice on the set E, so we now need to study the map on the set C. It turns out that we can make interesting conclusions about the map relying only upon information about the topology of the set $f(C)$, namely upon the fact that it is not "big". More precisely, we require that $f(C)$ cannot separate open sets in the following sense: for any open connected set U the set $U \setminus f(C)$ remains connected. In the case of n-manifolds it is known ([Eng78, Theorem 1.8.13]) that this is equivalent to the fact that $\dim N - \dim f(C) \geq 2$ (i.e., $f(C)$ has codimension at least 2). Now, the following theorem holds.

Theorem 3.7. *Suppose that $f : X \to Y$ is a light map of a closed n-manifold X into a closed n-manifold Y. Suppose that the critical value set $f(C)$ is of codimension at least 2. Then the map f is open. In particular, if $f : S^2 \to S^2$ is a light map such that the critical value set is totally disconnected then f is a branched covering map (and so in fact the critical set is finite).*

Proof. To begin with observe that since f is light, the assumption that the critical value set is of codimension at least 2 implies that the critical set is of codimension at least 2 as well [Eng78, Theorem 1.12.4].

Consider a point $x \in X$ and its neighborhood U. By way of contradiction suppose that $f(x)$ does not belong to the interior of $f(U)$. Then x does not belong to the interior of the image of any subset of U. Consider an open connected set V such that $x \in V \subset U$ and $f(x) \notin f(\mathrm{Bd}(V))$. Such sets exist because the set $f^{-1}(f(x))$ is totally disconnected (recall that f is light) and so x has a basis of neighborhoods whose boundaries contain no points of $f^{-1}(f(x))$.

Then $f(x)$ belongs to one of the components of the open set $Y \setminus f(\mathrm{Bd}(V))$. Let us denote this component W and prove that in fact $W \subset f(V)$. The idea of the proof is to find (by way of contradiction) points of $f(V)$ and of its complement inside W, connect them with the appropriately chosen arc and show that this arc cannot exit $f(V)$ thus getting a contradiction.

Indeed, assume by way of contradiction that W is not a subset of $f(V)$. Then there exists a point $y \in W$ such that $y \notin f(V)$. Since no point of W can belong to $f(\mathrm{Bd}(V))$ we see that in fact $y \notin f(\overline{V})$.

Let us show that there exists a point $z \in f(V) \cap W$ which is not a critical value. Indeed, since f is light it is quasi-interior by Lemma 2.2. Hence there is an open set in $f(V)$ arbitrarily close to $f(x)$. Clearly this open set contains a point z which are not a critical value because the critical value set $f(C)$ is of codimension at least 2.

Connect y and z by an arc B which is contained in W and avoids the set of critical values; this is possible because the codimension of the set $f(C)$ is at least 2. Consider the set $B' = B \cap f(V)$. Since points of B are neither critical values nor points of $f(\mathrm{Bd}(V))$ we see that any point belongs to B' together with its small neighborhood, and therefore B' is open in B. On the other hand, consider the set $B'' = B \cap (X \setminus \overline{f(V)})$; this set is also open in B. Since $B \subset W$ and W is a component of $X \setminus f(\mathrm{Bd}(V))$ we conclude that $B \cap f(\mathrm{Bd}(V)) = \emptyset$.

Therefore, $B = B' \cup B''$ which contradicts the fact that B is connected and completes the proof of the first part of the theorem.

In dimension 2 we can use Stoilow theorem according to which an open map of S^2 into itself has to be a branched covering map. This proves the second part of the theorem. □

As one can see on applications of Theorem 1.1 such as Theorem 3.7, some of them deal with the question as to what kind of set can the critical value set be for a map of an n-manifold into an n-manifold. In our view, this is an interesting question which should be thoroughly studied. Leaving such thorough study for the future, we would like to give below another example of a similar application of Theorem 1.1. By *perfect* we mean a map such that any compact set in the range has a compact full preimage in the domain.

Theorem 3.8. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous light perfect map with a critical set $C \neq \emptyset$. Then the critical value set $f(C)$ cannot by bounded and nowhere dense with connected and simply connected complement. In particular, if $n \geq 3$ then the set C cannot be finite.*

Proof. By way of contradiction let us assume that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a perfect light continuous map such that the critical value set $f(C)$ is bounded and nowhere dense while its complement is connected and simply connected. Indeed, it is easy to see that preimages of bounded sets under a perfect map are bounded too. On the other hand, the set C is closed because if $c_1, c_2, \dots$ is a converging sequence of critical points of f then its limit c is also a critical point of f. Since $f(C)$ is bounded, then C is bounded, hence C (and so $f(C)$) are compact.

Now, let $A = f^{-1}(C)$ and take a component B of $\mathbb{R}^n \setminus A$ and prove that it maps onto $\mathbb{R}^n \setminus f(C)$. Clearly, $f(B)$ is disjoint from $f(C)$, so it is contained in a connected set $\mathbb{R}^n \setminus f(C)$. Also, $f(B)$ is open because f is a local homeomorphism on B. Let us show that $f(B)$ is closed in $\mathbb{R}^n \setminus f(C)$. Indeed, suppose that $x_1, x_2, \dots$ is a sequence of points in $f(B)$ which converges to x in $\mathbb{R}^n \setminus f(C)$. If $x \in f(B)$ there is nothing to prove, so let us assume that $y \notin f(B)$. The preimage of the compact set $X = \{x_i\}_{i=1}^{\infty} \cup \{x\}$ is a compact set Y which is disjoint from A. Choose (by passing to a subsequence if necessary) points $y_i \in B$ so that $f(y_i) = x_i$ and so that $y_i \to y$. Then by continuity $f(y) = x$. By the assumption $x \notin f(B)$, hence $y \notin B$. Therefore $y \in \mathrm{Bd}(B)$ is contained in A, a contradiction. So, $f(B)$ is closed and open in $\mathbb{R}^n \setminus f(C)$ which implies that $f(B) = \mathbb{R}^n \setminus f(C)$ since $\mathbb{R}^n \setminus f(C)$ is connected by the assumption.

This actually implies that $\mathbb{R}^n \setminus A$ is connected. Indeed, since C is compact, there exists only one component of $\mathbb{R}^n \setminus A$ which extends to infinity. On the other hand, by the previous paragraph the image of each component of $\mathbb{R}^n \setminus A$ is not bounded. Each such component must extend to infinity. Hence there exists only one component of $\mathbb{R}^n \setminus A$. Thus, the set $\mathbb{R}^n \setminus A = B$ is connected.

Since $\mathbb{R}^n \setminus f(C)$ is simply connected and $f : B \to \mathbb{R}^n \setminus f(C)$ is clearly a covering map, then f maps B homeomorphically onto $\mathbb{R}^n \setminus f(C)$. Now, since f is light and $f(C)$ is nowhere dense then so is A. So, $f : \mathbb{R}^n \to \mathbb{R}^n$ is a map which is a homeomorphism on a dense set B in the domain which by Theorem 1.1 implies that in fact this is an embedding. In paragraph 2 of this proof we proved that f is onto $\mathbb{R}^n$. Hence, f is an onto homeomorphism contradicting that C is non-empty. □

4. Applications to arbitrary continuous maps

In this section we discuss applications of our results to the study of arbitrary, not just light, maps. In fact, some of our results can be in a way "translated" for the arbitrary continuous maps thanks to the existence of the *monotone-light decomposition* of an arbitrary continuous map of topological spaces. To state the appropriate result we need the following definition: a continuous map $f : X \to Y$ of topological spaces is said to be *monotone* if point preimages are connected. The following theorem is well-known.

Theorem 4.1. *Let $f : X \to Y$ be a continuous map of compact, metric spaces. Then there exists a compact metric space Z, a monotone map $g : X \to Z$ and a light map $h : Z \to Y$ such that $f = hg$.*

The idea of the proof is to define the map g by collapsing all components of preimages of points under f and thus to create a quotient space Z for which g is the factor map. Clearly, the map h determined then by the functional equation $f = hg$ is well-defined. Now, suppose that in the monotone-light decomposition of a map f the light map h is an embedding; then the map f is monotone. Conversely, if f is monotone then h has to be an embedding. Thus, it is natural to speak of monotone maps as analogous to embeddings in the framework of arbitrary continuous maps. Therefore below in extensions of our results monotone maps will often replace embeddings.

Let us now show how one can restate some of our main results for arbitrary maps using Theorem 4.1. The problem with having them all generalized for the general case is that when we collapse the domain by a monotone factor map the resulting quotient space may not inherit all the properties of the domain. In particular, it may stop being a manifold so that some of our results applicable to manifolds cannot be extended. Still, some facts do extend, and we list them in this section.

To begin with we need appropriate versions of some notations and definitions. Denote by H^1_f the set of all points x in Y such that $f^{-1}(x)$ is connected. Then we say that a map $f : X \to Y$ is *almost monotone* if the set $H^1_f \subset f(X)$ is dense in $f(X)$ and f is quasi-interior. The next lemma is a version of Lemma 2.1; it easily follows from Lemma 2.1 and Theorem 4.1 so the proof is left to the reader.

Lemma 4.2. *Suppose that $f : M \to N$ is a continuous map of metric σ-compact spaces. Then the set H^1_f is a G_δ-subset of $f(M)$.*

The next theorem follows from Theorem 2.3 and Theorem 4.1 and is given here without proof.

Theorem 4.3. *Suppose that $f : X \to Y$ is a continuous map, X is a continuum and Y is a dendrite. Moreover, suppose that the set H^1_f is con-dense in $f(X)$. Then f is monotone.*

The next theorem follows from Proposition 2.4 and Theorem 4.1 and is given here without proof.

Proposition 4.4. *Suppose that $f : X \to Y$ is a continuous map from an arcwise connected continuum X into a dendroid Y. Moreover, suppose that the set H^1_f is con-dense in $f(X)$. Then f is monotone.*

The following theorem follows from Theorem 2.5 and Theorem 4.1.

Theorem 4.5. *Suppose that $g : X \to Y$ is a weakly confluent and almost monotone mapping of a continuum X onto a locally connected continuum Y. Then g is monotone.*

Theorem 1.1 can be extended slightly in dimension 2 as follows.

Theorem 4.6. *Let $f : L \to N$ be a mapping of a closed 2-manifold L to a connected 2-manifold N such that f is almost monotone and for each $y \in N$ each component of $f^{-1}(y)$ is tree-like. Then f is monotone and onto.*

Proof. Let $f = g \circ m$ be the monotone-light factorization of f and put $m(L) = M$. Then M is a closed 2 manifold [RS38] while g is light and almost 1-to-1. By Theorem 1.1 g is a homeomorphism. Therefore f is monotone and onto. □

The following theorem follows from the monotone-light decomposition Theorem and extends Theorem 2.6 to arbitrary continuous maps.

Theorem 4.7. *Suppose that $f : X \to X$ is a continuous mapping from a continuum X onto itself. Then one of the following holds:*

(1) *there exists an open set U and a dense G_δ subset D' of $f(U)$ such that for all $y \in D'$ the set $f^{-1}(y) \cap U$ is connected;*

(2) *there exists a dense G_δ subset D of X such that for each $y \in D$, $f^{-1}(y)$ can be monotonically projected onto the Cantor set.*

Before we finish this section we would like to suggest the following example which shows that Theorems 2.5 and 4.6 are the best possible.

Example. There exists an almost 1-to-1 map from S^2 onto itself which is not monotone. This map is not weakly confluent, factors through the universal dendrite and has point inverses which contain circles. More precisely, Ward (see [W76]) has constructed a map $f : D \to S^2$ of a dendrite onto S^2. It is easy to see from the construction that D has a dense set of endpoints, f is light and is almost 1-1 (f identifies only sets of endpoints of D). It is an easy exercise to construct an almost 1-1 map $g : S^2 \to D$ onto D whose fibers (point preimages) are points, simple closed curves, finite wedges of simple closed curves and Hawaiian earrings. If E is a subcontinuum of D then E is locally connected and so $f(E)$ is also locally connected. Since S^2 contains non-locally connected subcontinua it follows that f is not weakly confluent. Hence, $f \circ g$ is not weakly confluent either. In particular, $f \circ g$ is not monotone. However, $f \circ g$ is almost 1-1 as it is a composition of almost 1-1 maps of compacta.

References

[BOT02] A. M. Blokh, L. G. Oversteegen, and E.D. Tymchatyn, *On almost one-to-one maps*, Preprint (2002).

[BOT03] A. M. Blokh, L. G. Oversteegen, and E.D. Tymchatyn, *On minimal maps of 2-manifolds*, Preprint (2003).

[Eng78] R. Engelking, *Dimension theory*, North-Holland and PWN, 1978.

[KST01] S. Kolyada, L. Snoha and S. Trofimchuk, *Noninvertible minimal maps*, Fundamenta Mathematicae, **168** (2001), 141–163.

[RS38] J. H. Roberts and N. E. Steenrod, *Monotone transformations of two-dimensional manifolds*, Annals of Math. **39** (1938), 851–862.

[W76] L. E. Ward, *A generalization of the Hahn-Mazurkiewicz Theorem*, Proc. Amer. Math. Soc. **58** (1976), 369-374.

[Why42] G. T. Whyburn, *Analytic topology*, vol. 28, AMS Coll. Publications, Providence, RI, 1942.

(Alexander Blokh and Lex Oversteegen) Department of Mathematics, University of Alabama at Birmingham, Birmingham, AL 35294-1170

(E. D. Tymchatyn) Department of Mathematics, Saskatchewan University

E-mail address, Alexander Blokh: `ablokh@math.uab.edu`

E-mail address, Lex Oversteegen: `overstee@math.uab.edu`

E-mail address, Ed Tymchatyn: `tymchat@snoopy.math.usask.ca`