Márton Elekes, Department of Analysis, Eötvös Loránd University, Kecskeméti u. 10-12, Budapest, 1053, Hungary; e-mail: `emarci@cs.elte.hu`
Tamás Keleti, *Department of Analysis, Eötvös Loránd University, Kecskeméti u. 10-12, Budapest, 1053, Hungary; e-mail: `elek@cs.elte.hu`
Vilmos Prokaj,†Department of Probability and Statistics, Eötvös Loránd University, Kecskeméti u. 10-12, Budapest, 1053, Hungary; e-mail: `prokaj@cs.elte.hu`

The Fixed Point of the Composition of Derivatives ‡

Abstract

We give an affirmative answer to a question of K. Ciesielski by showing that the composition $f \circ g$ of two derivatives $f, g : [0,1] \to [0,1]$ always has a fixed point. Using Maximoff's theorem we obtain that the composition of two $[0,1] \to [0,1]$ Darboux Baire-1 functions must also have a fixed point.

1 Introduction

In [4] R. Gibson and T. Natkaniec mentioned the question of K. Ciesielski, which asks whether the composition $f \circ g$ of two derivatives $f, g : [0,1] \to [0,1]$ must always have a fixed point. Our main result is an affirmative answer to this question. (An alternative proof has also been found by M. Csörnyei, T. C. O'Neil and D. Preiss, see [3].)

Theorem 1.1. *Let f and $g : [0,1] \to [0,1]$ be derivatives. Then $f \circ g$ has a fixed point.*

Key Words: fixed point, derivative, composition, gradient, level set, connected
Mathematical Reviews subject classification: Primary 26A99; Secondary 26A15, 26B99, 54H25
*Partially supported by OTKA grant F 029768
†Partially supported by OTKA grant T 032662
‡This research was completed during the Workshop on Geometric and Dynamical aspects of Measure Theory in Révfülöp, which was supported by the Erdős Center.

Since any bounded approximately continuous function is a derivative (see e.g. [1]) we get the following:

Corollary 1.2. *Let f and $g : [0,1] \to [0,1]$ be approximately continuous functions. Then $f \circ g$ has a fixed point.*

Now suppose that $f, g : [0,1] \to [0,1]$ are Darboux Baire-1 functions. Then, by Maximoff's theorem ([5], see also in [6]), there exist homeomorphisms $h, k : [0,1] \to [0,1]$ such that $f \circ h$ and $g \circ k$ are approximately continuous functions. Then clearly $\tilde{f} = k^{-1} \circ f \circ h$ and $\tilde{g} = h^{-1} \circ g \circ k$ are also $[0,1] \to [0,1]$ approximately continuous functions and $f \circ g$ has a fixed point if and only if $\tilde{f} \circ \tilde{g}$ has a fixed point. Thus from Corollary 1.2 we get the following:

Corollary 1.3. *Let f and $g : [0,1] \to [0,1]$ be Darboux Baire-1 functions. Then $f \circ g$ has a fixed point.*

Preliminaries

We shall denote the square $[-1,2] \times [-1,2]$ by Q. The partial derivatives of a function $h : \mathbb{R}^2 \to \mathbb{R}$ with respect to the i-th variable ($i = 1,2$) will be denoted by $\partial_i h$. The sets $\{h = c\}$, $\{h \neq c\}$, $\{h > c\}$, etc. will denote the appropriate level sets of the function h. The notations $\mathrm{cl}\, A, \mathrm{int}\, A$ and ∂A stand for the closure, interior and boundary of a set A, respectively. By *component* we shall always mean connected component.

2 Skeleton of the proof of Theorem 1.1

Let $f, g : [0,1] \to [0,1]$ be the derivatives of F and G, respectively. We can extend F and G linearly to the complement of the unit interval such that they remain differentiable and such that the derivatives of the extended functions are still between 0 and 1. Therefore we may assume that f and g are derivatives defined on the whole real line and $0 \leq f, g \leq 1$.

The key step of the proof is the following. Let

$$H(x,y) = F(x) + G(y) - xy \quad (\ (x,y) \in \mathbb{R}^2 \ ), \tag{2.1}$$

where $F' = f, \ G' = g : \mathbb{R} \to [0,1]$ are derivatives.

Then H is differentiable as a function of two variables, and its gradient is

$$H'(x,y) = (f(x) - y, g(y) - x). \tag{2.2}$$

Thus the gradient vanishes at (x_0, y_0) if and only if x_0 is a fixed point of $g \circ f$ and y_0 is a fixed point of $f \circ g$. Moreover, the gradient cannot be zero outside the closed unit square, since $0 \leq f, g \leq 1$. Therefore it is enough to prove that there exists a point in $\mathbb{R}^2$ where the gradient vanishes.

We argue by contradiction, so throughout the paper we shall suppose that the gradient of H nowhere vanishes.

Let us now examine the behavior of H on the edges of Q. By (2.2) the partial derivative $\partial_1 H$ is clearly negative on the top edge of Q and positive on the bottom edge, and similarly $\partial_2 H$ is positive on the left edge and negative on the right edge. This implies the following.

Lemma 2.1. *H is strictly decreasing on the top and right edges, and strictly increasing on the bottom and left edges of Q.*

Consequently,

$$\max\left(H(\text{-}1,\text{-}1), H(2,2)\right) < \min\left(H(\text{-}1,2), H(2,\text{-}1)\right).$$

This inequality suggests that there must be a kind of 'saddle point' in Q. Therefore define

$$c = \inf\{d : (\text{-}1,\text{-}1) \text{ and } (2,2) \text{ are in the same component of } \{H \leq d\} \cap Q\}$$

as a candidate for the value of H at a saddle point. At a typical saddle point q of some smooth function $K : \mathbb{R}^2 \to \mathbb{R}$ such that $K(q) = d$, the level set $\{K = d\}$ in a neighborhood of q consists of two smooth curves intersecting each other at q. In Section 3 we shall prove the following theorem, which states that the level set $\{H = c\}$ behaves indeed in a similar way.

Theorem 2.2. *$\{H = c\} \cap Q$ intersects each edge of Q at exactly one point (which is not an endpoint of the edge). Moreover, this level set cuts Q into four pieces, that is $\{H \neq c\} \cap Q$ has four components, and each of these components contains one of the corners of Q. The values of H are greater than c in the components containing the top left and the bottom right corners and less than c in the components containing the bottom left and the top right corners.*

This theorem suggests the picture that the level set $\{H = c\} \cap Q$ looks like the union of two arcs crossing each other, one from the left to the right and one from the top to the bottom. We are particularly interested in the existence of this 'crossing point', because at such a point four components of $\{H \neq c\} \cap Q$ should meet. Indeed, we shall prove the following (see Section 4).

Theorem 2.3. *There exists a point $r \in \{H = c\} \cap \operatorname{int} Q$ that is in the closure of more than two components of $\{H \neq c\} \cap Q$.*

Remark. In fact, it is also true that there exists a point $r \in \{H = c\} \cap \operatorname{int} Q$ in the intersection of the closure of exactly four components but we do not need this here.

But on the other hand we shall also prove (see Section 5) the following theorem about the local behavior of the level set.

Theorem 2.4. *Every $r \in \{H = c\} \cap \operatorname{int} Q$ has a neighborhood that intersects exactly two components of $\{H \neq c\} \cap Q$.*

However, the last two theorems clearly contradict each other, which will complete the proof of Theorem 1.1.

3 Proof of Theorem 2.2

Recall that we defined

$$c = \inf\{d : (\text{-}1,\text{-}1) \text{ and } (2,2) \text{ are in the same component of } \{H \le d\} \cap Q\}.$$

Lemma 3.1. *In fact, c is a minimum; that is, $(\text{-}1,\text{-}1)$ and $(2,2)$ are in the same component of $\{H \le c\} \cap Q$.*

Proof. Let K_n be the component of $\{H \le c + 1/n\} \cap Q$ containing both $(\text{-}1,\text{-}1)$ and $(2,2)$. Then K_n is clearly a decreasing sequence of compact connected sets. A well known theorem (see e.g. [7, I. 9. 4]) states that the intersection of such a sequence is connected. Hence $\cap_n K_n \subset \{H \le c\} \cap Q$ is connected and contains both $(\text{-}1,\text{-}1)$ and $(2,2)$, which shows that $(\text{-}1,\text{-}1)$ and $(2,2)$ are in the same component of $\{H \le c\} \cap Q$. □

Lemma 3.2. $\max\{H(\text{-}1,\text{-}1), H(2,2)\} < c < \min\{H(\text{-}1,2), H(2,\text{-}1)\}$.

Proof. To show $H(2,2) < c$ consider the polygon P with vertices $(\text{-}1,2)$, $(1,2)$, $(2,1)$ and $(2,\text{-}1)$. Since $\partial_1 H(x,y) = f(x) - y < 0$ if $y > 1$ and $\partial_2 H(x,y) = g(y) - x < 0$ if $x > 0$, it is easy to check that H is strictly bigger than $H(2,2)$ on the whole P. Therefore P connects $(\text{-}1,2)$ and $(2,\text{-}1)$ in $\{H > H(2,2)\} \cap Q$, thus $(\text{-}1,\text{-}1)$ and $(2,2)$ cannot be in the same component of $\{H \le H(2,2)\} \cap Q$. By the previous lemma this implies that $H(2,2) < c$. Proving that $H(\text{-}1,\text{-}1) < c$ is similar.

To show that $c < H(\text{-}1,2)$ consider the polygon P' with vertices $(\text{-}1,\text{-}1)$, $(\text{-}1,1)$, $(0,2)$ and $(2,2)$. Like above one can show that H is strictly less than $H(\text{-}1,2)$ on the whole P', so denoting the maximum of H on P' by d, we have $d < H(\text{-}1,2)$. Since P' connects $(\text{-}1,\text{-}1)$ and $(2,2)$ in $\{H \le d\} \cap Q$ we also have $c \le d$, therefore we have $c < H(\text{-}1,2)$. Proving that $c < H(2,\text{-}1)$ would be similar. □

The previous lemma and the strict monotonicity of H on the edges of Q (Lemma 2.1) clearly implies the first statement of Theorem 2.2: $\{H = c\} \cap Q$ intersects each edge of Q at exactly one point (which is not an endpoint of the edge). Since four points cut ∂Q into four components we also get that $\{H \neq c\} \cap \partial Q$ has four components and each of them contains a vertex of Q. Therefore for completing the proof of Theorem 2.2 we have to show that any component of $\{H \neq c\} \cap Q$ intersect ∂Q and that the vertices of Q belong to different components of $\{H \neq c\} \cap Q$.

The first claim is clear since if C were a component of $\{H \neq c\} \cap Q$ inside Q then one of the (global) extrema of H on the (compact) closure of C could not be on the boundary of C (where $H = c$), so this would be a local extremum but a function with non-vanishing gradient cannot have a local extremum.

To prove the second claim first note that the components of $\{H \neq c\} \cap Q$ are the components of $\{H < c\} \cap Q$ and the components of $\{H > c\} \cap Q$. By the previous lemma, $(-1,-1)$ and $(2,2)$ are in $\{H < c\} \cap Q$ while $(-1,2)$ and $(2,-1)$ are in $\{H > c\} \cap Q$, so it is enough to prove that the opposite vertices belong to different components.

Assume that $(-1,-1)$ and $(2,2)$ are in the same component of $\{H < c\} \cap Q$. Then there is a continuous curve in $\{H < c\} \cap Q$ that connects $(-1,-1)$ and $(2,2)$. Let d be the maximum of H on this curve. Then on one hand $d < c$; on the other hand, $(-1,-1)$ and $(2,2)$ are in the same component of $\{H \leq d\} \cap Q$, so $c \leq d$, which is a contradiction.

Finally, assume that $(-1,2)$ and $(2,-1)$ are in the same component of $\{H > c\} \cap Q$. Then there is a continuous curve in $\{H > c\} \cap Q$ that connects $(-1,2)$ and $(2,-1)$, so it also separates $(-1,-1)$ and $(2,2)$. But this contradicts Lemma 3.1.

4 Proof of Theorem 2.3

Let us denote by C_p the component of $\{H \neq c\} \cap Q$ that contains the corner p of Q (Theorem 2.2 shows that $p \in \{H \neq c\} \cap Q$). Again by Theorem 2.2 we have $\{H \neq c\} \cap Q = C_{(-1,-1)} \cup C_{(2,2)} \cup C_{(-1,2)} \cup C_{(2,-1)}$, and that H is less than c in the first two and greater than c in the last two components. Thus $\mathrm{cl}\, C_{(-1,-1)}$ and $\mathrm{cl}\, C_{(2,2)}$ are closed subsets of $\{H \leq c\} \cap Q$. Since (by Lemma 3.1) $(-1,-1)$ and $(2,2)$ are in the same component of $\{H \leq c\} \cap Q$, this implies that $\mathrm{cl}\, C_{(-1,-1)}$ and $\mathrm{cl}\, C_{(2,2)}$ cannot be disjoint. Let $r \in \mathrm{cl}\, C_{(-1,-1)} \cap \mathrm{cl}\, C_{(2,2)}$.

Our next aim is to show that r cannot be on the edges of Q. The other cases being similar we only consider the case of the left edge.

Lemma 4.1. *The component $C_{(2,2)}$ does not intersect the strip $[-1,0) \times [-1,2]$.*

Proof. We have to prove that if $(x,y) \in [-1,0) \times [-1,2]$ is such that $H(x,y) < c$, then $(x,y) \in C_{(-1,-1)}$. In order to do this, it is sufficient to show that the vertical segment S_1 between (x,y) and $(x,-1)$ and the horizontal segment S_2 between $(x,-1)$ and $(-1,-1)$ together form a polygon in $\{H(x,y) < c\}$ connecting (x,y)

to (-1, -1). But by (2.2), the partial derivative $\partial_2 H$ is positive on S_1, thus all values of H are less then $H(x, y)$ here, hence less then c, and $\partial_1 H$ is positive on S_2, thus the values of H here are less then $H(x, \text{-}1)$, which is again less than c. □

If we repeat the argument of the previous lemma for the other three edges, then we get that the point r we obtained above must be in the interior of Q.

To complete the proof of Theorem 2.3 it is enough to show that there exists a component in which H is greater than c (that is either component $C_{(\text{-}1,2)}$ or $C_{(2,\text{-}1)}$) such that the closure of the component contains r. But $r \in \operatorname{int} Q$, so this statement follows from the fact that the gradient of H nowhere vanishes, so H attains no local extremum.

5 Proof of Theorem 2.4

Lemma 5.1. *Let $x_1, x_2, y_1, y_2 \in \mathbb{R}$ be such that $x_1 < x_2$ and $y_1 < y_2$. Then $H(x_1, y_2) + H(x_2, y_1) - H(x_1, y_1) - H(x_2, y_2) > 0$.*

Proof. Using the definition (2.1) of H, a straightforward computation shows that the value of the above sum in fact equals $(x_2 - x_1)(y_2 - y_1) > 0$. □

Lemma 5.2. *Let $(x_0, y_0) \in \mathbb{R}^2$ be such that $H(x_0, y_0) = 0$ and $\partial_2 H(x_0, y_0) \neq 0$. Then there exists $\varepsilon > 0$ such that if $x_1 \in [x_0 - \varepsilon, x_0 + \varepsilon]$ and $y_1, y_2 \in [y_0 - \varepsilon, y_0 + \varepsilon]$ and $H(x_1, y_1) = H(x_1, y_2) = 0$, then one of the followings holds:*

(i) $y_1 = y_2 = y_0$,

(ii) $y_1 < y_0$ *and* $y_2 < y_0$,

(iii) $y_1 > y_0$ *and* $y_2 > y_0$.

Proof. From $H(x_0, y_0) = 0$ and (2.1) we get $G(y_0) = x_0 y_0 - F(x_0)$, while $\partial_2 H(x_0, y_0) \neq 0$ means $G'(y_0) \neq x_0$. These imply that for a sufficiently small ε if $y_0 - \varepsilon < y_1 \leq y_0 \leq y_2 < y_0 + \varepsilon$ and $y_1 \neq y_2$ then the slope of the segment between $(y_1, G(y_1))$ and $(y_2, G(y_2))$ is not in $[x_0 - \varepsilon, x_0 + \varepsilon]$. But if $H(x_1, y_1) = H(x_1, y_2) = 0$ then $G(y_1) = x_1 y_1 - F(x_1)$ and $G(y_2) = x_1 y_2 - F(x_1)$, so the slope of the segment between $(y_1, G(y_1))$ and $(y_2, G(y_2))$ is x_1. Therefore, with this ε, the conditions of the lemma imply that one of (i),(ii) or (iii) must hold. □

Before turning to the proof of Theorem 2.4 observe that if H is given by equation (2.1) then $H_1(x, y) = -H(1 - x, y)$ is also of the form (2.1). Indeed, taking $F_1(x) = -F(1 - x)$ and $G_1(y) = y - G(y)$, we get $f_1(x) =$

$F_1'(x) = f(1-x)$ and $g_1(y) = G_1'(y) = 1 - g(y) : [0,1] \to [0,1]$ derivatives and $H_1(x,y) = F_1(x) + G_1(x) - xy = -H(1-x,y)$. It is also clear that if H is of the form (2.1) then $H_2(x,y) = H(y,x)$ is also of the form (2.1). This means that when we study the local behavior of the function H around an arbitrary point $(x,y) \in \{H = 0\} \cap Q$ we can assume that $\partial_2 H(x,y) > 0$ as this is true (at the appropriate point) for at least one of the functions we can get using these symmetries. By adding a constant to H we can also assume that $c = 0$. Therefore for proving Theorem 2.4 we can assume these restrictions, so it is enough to prove the following.

Claim 5.3. *Let $(x_0, y_0) \in \operatorname{int} Q$ be such that $H(x_0, y_0) = 0$ and $\partial_2 H(x_0, y_0) > 0$. Then there exists a neighborhood of (x_0, y_0) that intersects exactly two components of $\{H \neq 0\} \cap Q$.*

Proof. Every neighborhood of (x_0, y_0) has to meet at least two components of $\{H \neq 0\} \cap Q$, otherwise (x_0, y_0) would be a local extremum but we assumed that the gradient of H is nowhere vanishing.

In order to show that a small neighborhood of (x_0, y_0) meets at most two components of $\{H \neq 0\} \cap Q$ it is enough to construct a polygon in Q around the point which intersects $\{H = 0\}$ at exactly two points. Indeed, in this case there can be at most two components of $\{H \neq 0\} \cap Q$ that intersect the polygon (since two points cut a polygon into two components). On the other hand, every component of $\{H \neq 0\} \cap Q$ that intersects the interior of the polygon has to meet the polygon itself, since by Theorem 2.3 each component of $\{H \neq 0\} \cap Q$ contains a vertex of Q.

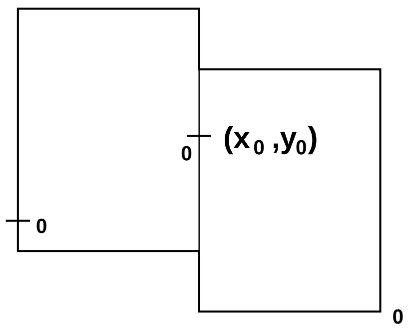

Figure 1: The polygon around (x_0, y_0)

To construct such a polygon it is enough to find two rectangles in Q, one to the left and one to the right from (x_0, y_0), which contain this point on vertical edges (but not as a vertex) and which both intersect $\{H = 0\}$ in just one additional point. (See Fig. 1)

Since $\partial_2 H(x_0, y_0) > 0$ we can find a small open sector in Q with vertex (x_0, y_0) with vertical axis upwards, where H is positive and a similar negative sector downwards. (See Fig. 2. and Fig. 3.) Choose the radius of the sectors smaller than the ε we get in Lemma 5.2. Choose $x_1 < x_0$ and $x_2 > x_0$ such that the vertical lines through them intersect the two sectors.

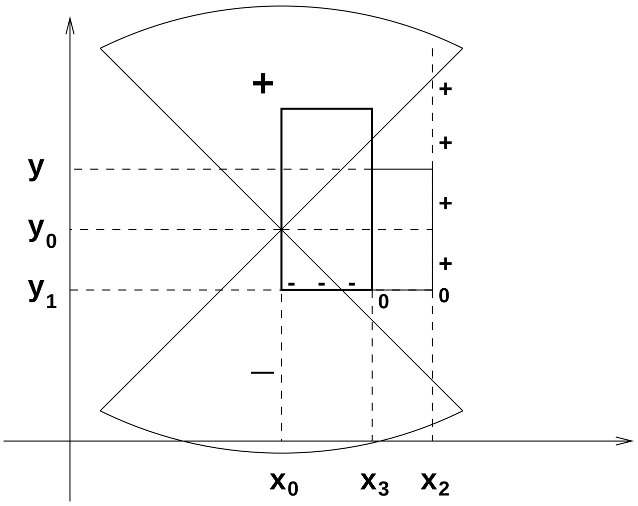

Figure 2: The rectangle on the right-hand side

First we construct the rectangle on the right-hand side. (See Fig. 2.) If (i) holds in Lemma 5.2 for the point (x_0, y_0) and for x_2 then we are done, so as the other two cases are similar we assume that (ii) holds. This means that even for the maximum y_1 of the zero set of H between the two sectors on the vertical line $(x = x_2)$ we have $y_1 < y_0$. We choose the leftmost point (x_3, y_1) of the zero set on the segment $[x_0, x_2] \times \{y_1\}$ as the lower right corner of the rectangle and we choose the upper right corner from the upper sector. Clearly there are only negative values on the bottom edge, so what remains to show is that all values are positive above (x_3, y_1) on the vertical line until it reaches the upper sector. If $x_3 = x_2$ then we are done, otherwise this is an easy consequence of Lemma 5.1 once we apply it to x_3, x_2, y_1 and any number $y > y_1$ such that (x_3, y) is still between the two sectors.

Now we turn to the left-hand side. (See Fig. 3.) We assume again that (ii) of Lemma 5.2 holds for x_1, and choose a maximal y_3 such that (x_1, y_3)

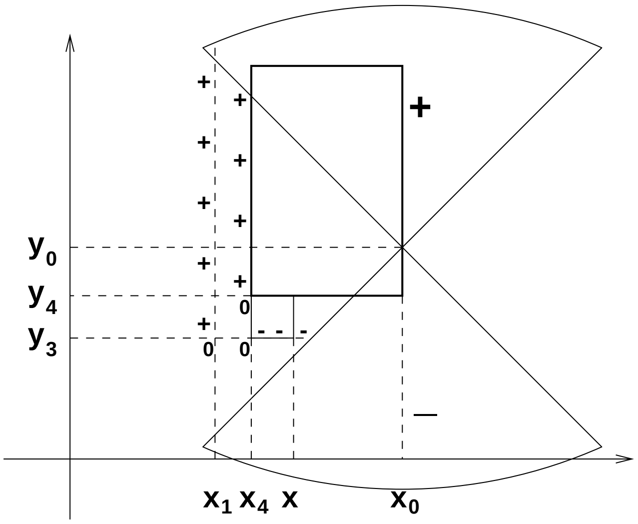

Figure 3: The rectangle on the left-hand side

is between the two sectors and $H(x_1, y_3) = 0$. Note that, by (ii), we have $y_3 < y_0$. Let (x_4, y_3) be the rightmost point of the zero set of H on the segment $[x_1, x_0] \times \{y_3\}$. Let us now define the lower left corner (x_4, y_4) of the rectangle as the point of maximal y coordinate on the vertical line $x = x_4$ between the sectors, where H vanishes. By Lemma 5.2, $y_4 < y_0$ and clearly all values on the vertical half-line from (x_4, y_4) are positive until it reaches the upper sector. Thus we only have to show that $H(x, y_4) < 0$ for all $x_4 < x < x_0$, which easily follows from Lemma 5.1 once we apply it to x_4, x, y_3 and y_4. $\square$

Remark. Theorem 2.4 is not true for every $\mathbb{R}^2 \to \mathbb{R}$ differentiable function with nowhere vanishing gradient. For an example see [2].

References

[1] A. Bruckner, *Differentiation of Real Functions*, CRM Monograph Series, 1994.

[2] Z. Buczolich, *Level sets of functions $f(x, y)$ with non-vanishing gradient*, Journ. Math. Anal. Appl. **185** (1994), 27-35.

[3] M. Csörnyei, T. C. O'Neil, D. Preiss, *The composition of two derivatives has a fixed point*, to appear in Real Anal. Exchange.

[4] R. Gibson, T. Natkaniec, *Darboux like functions. Old problems and new results*, Real Anal. Exchange **24** (2) (1998-99), 487-496.

[5] I. Maximoff, *Sur la transformation continue des fonctions* Bull. Soc. Phys. Math. Kazan **12** (1940), no. 3, 9-41 (Russian, French summary).

[6] D. Preiss, *Maximoff's Theorem*, Real Anal. Exchange **5** (1) (1979-80), 92-104.

[7] G. T. Whyburn, *Analytic Topology*, AMS Colloquium Publications Vol. 28, 1942.