Natural pseudo-distances between closed curves

Pietro Donatini[*1] and **Patrizio Frosini**[**1,2]

[1] Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato, 5
I-40126 Bologna, ITALY

[2] Advanced Research Centre for Electronic Systems "Ercole De Castro", Università di Bologna, via Toffano, 2/2
I-40135 Bologna, ITALY

Key words Natural pseudo-distance, measuring function

2000 Mathematics Subject Classification Primary 58E35, 58D99, 53C99; Secondary 58C05, 58C25, 53A04

Let us consider two closed curves $\mathcal{M}$, $\mathcal{N}$ of class C^1 and two functions $\varphi : \mathcal{M} \to \mathbb{R}$, $\psi : \mathcal{N} \to \mathbb{R}$ of class C^1, called measuring functions. The natural pseudo-distance d between the pairs $(\mathcal{M}, \varphi)$, $(\mathcal{N}, \psi)$ is defined as the infimum of $\Theta(f) \stackrel{def}{=} \max_{P \in \mathcal{M}} |\varphi(P) - \psi(f(P))|$, as f varies in the set of all homeomorphisms from $\mathcal{M}$ onto $\mathcal{N}$. The problem of finding the possible values for d naturally arises. In this paper we prove that under appropriate hypotheses the natural pseudo-distance equals either $|c_1 - c_2|$ or $\frac{1}{2}|c_1 - c_2|$, where c_1 and c_2 are two suitable critical values of the measuring functions. This equality shows that the relations between the natural pseudo-distance and the critical values of the measuring functions previously obtained in higher dimensions can be made stronger in the particular case of closed curves. Moreover, the examples we give in this paper show that our result cannot be further improved, and therefore it completely solves the problem of determining the possible values for d in the 1-dimensional case.

Introduction

In recent years, considerable research has been carried out to formalize the concept of shape for topological spaces and manifolds, and to provide an efficient comparison of shapes. Part of this research (cf., e.g., [1, 9, 17, 22]) concerns the study of the geometrical/topological properties of a suitable real function defined on the considered space. In this approach, a main issue is the stability under perturbations of the proposed methods, meaning that small changes in the real functions produce small changes in the geometrical descriptors (cf., e.g., [2, 3, 4]). This has led to the study of the concept of *natural pseudo-distance*, as introduced in Size Theory. In this formalization, the pairs $(\mathcal{M}, \varphi)$ are considered, where $\mathcal{M}$ is a topological space and $\varphi : \mathcal{M} \to \mathbb{R}$ is a continuous function. These pairs are called *size pairs* and each function φ is called a *measuring function*. The usual setting assumes that $\mathcal{M}$ is a closed manifold, and that both $\mathcal{M}$ and φ are C^1. If two size pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$ are given and $\mathcal{M}$, $\mathcal{N}$ are homeomorphic, then the *natural pseudo-distance* δ is defined by setting $\delta((\mathcal{M}, \varphi), (\mathcal{N}, \psi)) \stackrel{def}{=} \inf_{f \in H(\mathcal{M}, \mathcal{N})} \max_{P \in \mathcal{M}} |\varphi(P) - \psi(f(P))|$, where $H(\mathcal{M}, \mathcal{N})$ denotes the set of all homeomorphisms from $\mathcal{M}$ onto $\mathcal{N}$. Clearly, in cases in which the natural pseudo-distance is zero, we can obtain homeomorphisms for which the difference between the values taken by the measuring functions at corresponding points is arbitrarily small. On the other hand, if the infimum is large, we find that every homeomorphism between the considered manifolds must change the values taken by our measuring function considerably.

Besides being a key theoretical tool for applications involving comparison of shapes, the natural pseudo-distance is challenging from the mathematical point of view and there are several open questions concerning its properties. This paper is the last of three papers devoted to this subject. In [6] it was proved that in general dimensions a suitable multiple of $\delta((\mathcal{M}, \varphi), (\mathcal{N}, \psi))$ by a positive integer k coincides with the distance between two critical values of the functions φ, ψ. In [8] the particular case of surfaces was considered and it was proved that the number k can be assumed to equal 1, 2 or 3, using some results concerning Riemannian metrics (cf. [12, 14]) and a theorem by Jost and Schoen about harmonic maps (cf. [16]). In this final paper we succeed in

* e-mail: donatini@dm.unibo.it, Phone: ++ 39 051 209 4421, Fax: ++ 39 051 209 4490

** Corresponding author: e-mail: frosini@dm.unibo.it, Phone: ++ 39 051 209 4478, Fax: ++ 39 051 209 4490

proving that an even stronger statement holds in the 1-dimensional case, and the integer k can be assumed to be either 1 or 2 for curves, using a linearization technique. Moreover, the examples we give in this paper show that our result cannot be further improved, and therefore it completely solves the problem of determining the possible values for the natural pseudo-distance in the 1-dimensional case.

The subject of this paper fits in the current mathematical research and interest about simple closed curves, motivated by problems concerning shape comparison in Pattern Recognition and Computer Vision (cf., e.g., [18, 19]).

In Section 1 we sketch the main ideas in our proof. In Section 2 we give the main definitions and some examples, while in Section 3 the concepts of train and minimal d-approximating sequence are illustrated, together with some related results. In Section 4 we prove our main result (Theorem 4.5) about the natural pseudo-distance between closed curves.

1 The point of this paper

As described in the previous section, in [6] it was proved that the natural pseudo-distance between size pairs always equals $|c_1 - c_2|/k$, where c_1, c_2 are two suitable critical values of the measuring functions and k is an appropriate integer number. The minimum possible value for k is called the *analytic folding number*.

Interestingly enough, in every known example, the analytic folding number is 1 or 2.

Two questions naturally arise. Are there any examples showing an analytic folding number strictly greater than 2? Is this question related to the dimension of our manifolds?

In this paper we take another step towards answering these questions, completely solving the 1-dimensional case.

The key fact in the study of the natural pseudo-distance is that the attempt to minimize the change $\Theta(f) \stackrel{def}{=} \max_{P\in\mathcal{M}} |\varphi(P)-\psi(f(P))|$ in the measuring functions under the action of f does not, in general, lead to a homeomorphism, as we are going to show in the following section. The possible emergence of degeneracies prevents us from studying a single optimal homeomorphism. Therefore, approximating sequences of homeomorphisms must be considered, instead of a single homeomorphism. However, we can show that "optimal" approximating sequences (f_i) of homeomorphisms exist, converging to relations that represent the best way to take one manifold to another with respect to the change in the measuring functions. The study of these relations leads us to define the concept of train of "limit d-jumps", describing some degeneracies corresponding to the sequence (f_i). As we are going to see later on, the properties of these structures imply the properties of the analytic folding number.

How can we study these properties in the case of curves?

In [6] local deformations were used, based on the flow diffeomorphism of the gradient of the measuring functions but, unfortunately, this approach does not seem to be sufficient to answer the questions we posed. The main idea of this paper is to use a linearization procedure to confront the 1-dimensional case.

We shall proceed in this way. We shall consider each "optimal" sequence (f_i) of homeomorphisms between the manifolds $\mathcal{M}$ and $\mathcal{N}$ that we are examining, where optimal means that $\inf_i \max_{P\in\mathcal{M}} |\varphi(P) - \psi(f_i(P))|$ equals $\delta((\mathcal{M},\varphi),(\mathcal{N},\psi))$. Then we shall describe the degeneracies relating to (f_i) using a train of "limit d-jumps", and assume that the degeneracies of (f_i) are minimal with respect to a suitable order $\preceq$ we are going to define.

Finally, we shall apply a local linearization procedure to each f_i far from the critical points. The key remark will be that the change we are going to apply produces a new sequence that is "smaller" than (f_i) with respect to $\preceq$. Since (f_i) will already be minimal, some further information about the length of the trains of d-jumps for (f_i) will be derived, implying the main result obtained in this paper.

The outline of the proof will be similar to the one exposed in [8], but the technique that we shall use will be different, since a linearization procedure will be involved, in place of the theory of harmonic maps. This choice will allow us to get a stronger result than the one we obtained for surfaces.

Obviously, some technicalities will be necessary in order to use our ideas in practice, but the key point is simply the possibility of reducing the change of the measuring functions by locally decreasing the energy of the transformations we use between our curves. The following sections will formalize the ideas we have just described.

1.1 Some remarks concerning our mathematical setting

Before proceeding in our exposition, it may be useful to make some points clear. First of all the use of the word "shape" in this paper derives from the application of Size Theory in Pattern Recognition. Indeed, from the beginning of the 90's the idea of comparing topological spaces (or manifolds) endowed with measuring functions has been used to compare shapes of objects in Computer Vision (cf., e.g., [5, 10, 11, 13, 20, 21]).

The reader may wonder why we study size pairs instead of a unique pair (X, φ) where X is a finite disjoint union of circles. After all the manifolds $\mathcal{M}$ and $\mathcal{N}$ are the same manifold, from a topological point of view. Formally, we could do that, but it would be useless and a little misleading as regards applications. The point is that the homeomorphism used to identify $\mathcal{M}$ and $\mathcal{N}$ *is not canonical*. Indeed, finding the "best homeomorphisms" to identify $\mathcal{M}$ and $\mathcal{N}$ (supposing that they exist) is just the main motivation of our research. Assuming to know them from the beginning is equivalent to look for the solution of a problem that has been already solved. Obviously, we could also consider a generic homeomorphism h between our manifolds, but in this case this subjective choice would introduce some degree of indetermination in our model. Furthermore, while this approach would give us just a slight formal advantage in exposition, it would make the theory more difficult to apply, since the "normalizing" homeomorphism h should be chosen each time.

Another question could concern the motivation of studying the functional $\inf_{f \in H(\mathcal{M},\mathcal{N})} \|\varphi - \psi \circ f\|_\infty$ instead of $\inf_{f \in H(\mathcal{M},\mathcal{N})} \|\varphi - \psi \circ f\|_p$. The reason of this choice is that the latter functional is *not* a pseudo-distance between size pairs (cf. [4]).

Finally, the reader may wonder why we use measuring functions taking values in $\mathbb{R}$ rather than in another space Y. Indeed, the main definitions considered in Size Theory can be adapted to the case of Y being a partially ordered metric space. However, in this case the treatment becomes much more difficult and, in particular, it does not seem easy to extend Theorems 2.4 and 4.5. For these reasons, we postpone the study of this kind of generalization to further papers.

2 The natural pseudo-distance

2.1 The main definition

The definition of natural size pseudo-distance can be introduced for n-dimensional manifolds. Let us consider the set $Size_n$ of all pairs $(\mathcal{M}, \varphi)$, where $\mathcal{M}$ is a closed n-manifold of class C^k and $\varphi : \mathcal{M} \to \mathbb{R}$ is a function of class C^k. We shall call $(\mathcal{M}, \varphi)$ an (n-dimensional) *size pair* of class C^k and φ a *measuring function*.

Assume that $(\mathcal{M}, \varphi), (\mathcal{N}, \psi)$ are two size pairs. $H(\mathcal{M}, \mathcal{N})$ will denote the set of all homeomorphisms from $\mathcal{M}$ to $\mathcal{N}$.

Definition 2.1 If $H(\mathcal{M}, \mathcal{N}) \neq \emptyset$, the function $\Theta : H(\mathcal{M}, \mathcal{N}) \to \mathbb{R}$

$$\Theta(f) = \max_{P \in \mathcal{M}} |\varphi(P) - \psi(f(P))|$$

is called *natural size measure* with respect to the measuring functions φ and ψ.

In other words, Θ measures how much f changes the values taken by the measuring functions at corresponding points.

Definition 2.2 We shall call *natural size pseudo-distance* the pseudo-distance $\delta : Size_n \times Size_n \to \mathbb{R} \cup \{+\infty\}$ so defined:

$$\delta((\mathcal{M}, \varphi), (\mathcal{N}, \psi)) = \begin{cases} \inf_{f \in H(\mathcal{M},\mathcal{N})} \Theta(f) & \text{if } H(\mathcal{M}, \mathcal{N}) \neq \emptyset \\ +\infty & \text{otherwise.} \end{cases}$$

In the following, the symbol d will denote the value of the natural size pseudo-distance computed between the pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$ that we are considering. As we previously explained, this pseudo-distance gives a method for comparing two manifolds with respect to the measuring functions chosen.

We point out that δ is not a distance, since two size pairs can have a vanishing pseudo-distance without being equal. On the other hand, the symmetry property and the triangle inequality can be trivially proved.

Remark 2.3 The presence of the word "size" in our definitions is due to the link existing between the pseudo-distance δ, size functions and size homotopy groups (cf. [11, 15]). However, for the sake of simplicity, we shall often drop the word "size" in the expressions "natural size measure" and "natural size pseudo-distance". The term "natural" is used in order to distinguish the pseudo-distance studied here from other pseudo-distances we can define between submanifolds of the Euclidean space and between manifolds paired with measuring functions (cf. [10]).

In spite of the considerable difficulty in computing natural size pseudo-distances, the following result holds for the general dimension n (cf. [6]):

Theorem 2.4 *Assume that $\mathcal{M}$ and $\mathcal{N}$ are two homeomorphic closed manifolds of class C^1 and that $\varphi : \mathcal{M} \to \mathbb{R}$ and $\psi : \mathcal{N} \to \mathbb{R}$ are two functions of class C^1. Then, if d denotes the natural pseudo-distance between the size pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$, a positive integer k exists for which one of the following properties holds:*

$i)$ *k is odd and kd equals the distance between a critical value of φ and a critical value of ψ;*

$ii)$ *k is even and kd equals either the distance between two critical values of φ or the distance between two critical values of ψ.*

The smallest positive integer k for which either $i)$ or $ii)$ of Theorem 2.4 holds will be called *analytic folding number* for the pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$.

In this paper we shall prove that in the case of closed curves the analytic folding number always equals either 1 or 2. This fact, besides showing a particular property of the 1-dimensional case, allows us to make a direct computation of natural pseudo-distances for closed curves easier.

However, the hypothesis $n = 1$ will be not used until Section 4.

In the following subsection 2.2, we shall show that the infimum of $\Theta(f)$ varying f in $H(\mathcal{M}, \mathcal{N})$ is not always a minimum. When such an infimum is also a minimum, we shall say that each homeomorphism $f \in H(\mathcal{M}, \mathcal{N})$ with $d = \Theta(f)$ is an *optimal homeomorphism*.

In the case where an optimal homeomorphism exists, the following result holds (Theorem 6.3 in [6]).

Theorem 2.5 *Assume that $\mathcal{M}$ and $\mathcal{N}$ are two C^1 closed homeomorphic manifolds and that $\varphi : \mathcal{M} \to \mathbb{R}$ and $\psi : \mathcal{N} \to \mathbb{R}$ are of class C^1. If an optimal homeomorphism $f \in H(\mathcal{M}, \mathcal{N})$ between the size pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$ exists, then the natural pseudo-distance $d = \delta((\mathcal{M}, \varphi), (\mathcal{N}, \psi))$ equals the distance between a critical value of φ and a critical value of ψ.*

N.B.: For the sake of conciseness, all through this paper we shall use the expression "closed curve" to mean a closed 1-manifold (we shall *not* require this manifold to be connected).

In order to simplify our notations, we shall assume that the manifolds $\mathcal{M}$ and $\mathcal{N}$ do not meet, and that the corresponding measuring functions are obtained by restriction of a function $\omega : \mathcal{M} \cup \mathcal{N} \to \mathbb{R}$, so that $\varphi = \omega_{|\mathcal{M}}$ and $\psi = \omega_{|\mathcal{N}}$. In this way we can use just one symbol to denote both the measuring functions. These hypotheses are not restrictive, since we can always replace the size pair $(\mathcal{N}, \psi)$ with a new size pair $(\overline{\mathcal{N}}, \overline{\psi})$, having vanishing pseudo-distance from the previous one and such that $\mathcal{M} \cap \overline{\mathcal{N}} = \emptyset$. Sometimes, when not confusing, we shall use the symbol ω to denote both $\omega_{|\mathcal{M}}$ and $\omega_{|\mathcal{N}}$.

Moreover, it is easy to prove that, for every 1-dimensional size pair $(\mathcal{M}, \omega)$ of class C^k, an embedding $h : \mathcal{M} \to \mathbb{R}^3$ of class C^k exists such that $z(P) = \omega(h^{-1}(P))$ for each point $P \in h(\mathcal{M})$. If ω is Morse (i.e., smooth and having invertible Hessian at each critical point), we can assume that z is Morse on $h(\mathcal{M})$, too. In other words, there is no lack of generality in assuming that the measuring functions associated with the studied closed curves $\mathcal{M}$, $\mathcal{N}$ are obtained by restriction of the z-coordinate in $\mathbb{R}^3$. Sometimes, when not confusing, we shall use the symbol z to denote both $z_{|\mathcal{M}}$ and $z_{|\mathcal{N}}$ and use the expression "height of a point". For the sake of clarity, in our examples and figures we shall often assume that our measuring function is the z-coordinate.

2.2 Two examples

Now we give two simple examples in order to make our definitions clear.

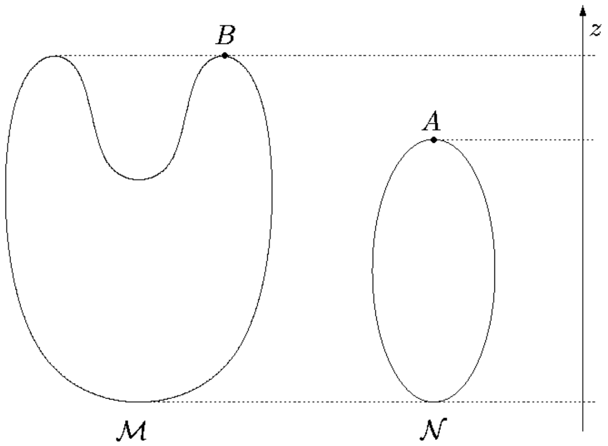

Fig. 1 The natural pseudo-distance between the size pairs $(\mathcal{M}, z)$ and $(\mathcal{N}, z)$ is $z(B) - z(A)$.

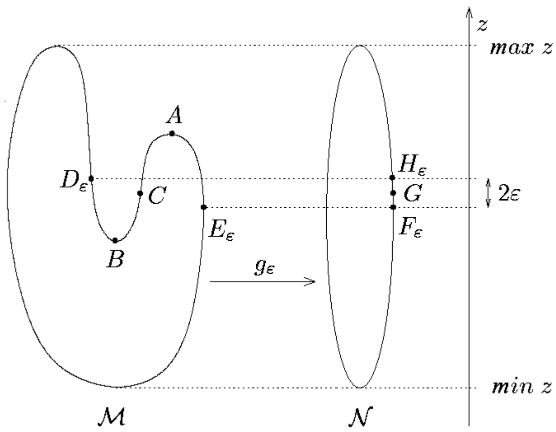

Fig. 2 Construction of the homeomorphism g_ε for which $\Theta(g_\varepsilon) \leq d + \varepsilon$.

Example 2.6 The first example we give is shown in Figure 1. $\mathcal{M}$ and $\mathcal{N}$ are smooth closed curves in $\mathbb{R}^3$, embedded in the xz-plane. It is clear that the natural pseudo-distance d between the size pairs $(\mathcal{M}, z)$ and $(\mathcal{N}, z)$ equals $z(B) - z(A)$, that is, the distance between a critical value of $z_{|\mathcal{M}}$ and a critical value of $z_{|\mathcal{N}}$.

In this example no optimal homeomorphism exists, since it ought to take both the maximum points for $z_{|\mathcal{M}}$ to A, against injectivity.

Example 2.7 Let us consider the smooth closed curves $\mathcal{M}$ and $\mathcal{N}$ in Figure 2. The points A and B are critical points of the function z and $z(C) = \frac{1}{2}(z(A) + z(B)) = z(G)$. We want to prove that the natural pseudo-distance between the size pairs $(\mathcal{M}, z)$ and $(\mathcal{N}, z)$ takes the value

$$d = \frac{1}{2}(z(A) - z(B))$$

and that no optimal homeomorphism exists. In order to do that we shall construct a sequence of homeomorphisms (f_n) for which $\lim_n \Theta(f_n) = \frac{1}{2}(z(A)-z(B))$, and show that $\Theta(f) > \frac{1}{2}(z(A)-z(B))$ for every homeomorphism $f \in H(\mathcal{M}, \mathcal{N})$.

Let us start by proving that, for every $\varepsilon > 0$, a homeomorphism $g_\varepsilon : \mathcal{M} \to \mathcal{N}$ exists, such that $\Theta(g_\varepsilon) \leq \frac{1}{2}(z(A) - z(B)) + 2\varepsilon$. Consider the points D_ε, E_ε, H_ε and F_ε in Figure 2, verifying $z(D_\varepsilon) = z(H_\varepsilon) = z(C) + \varepsilon$

and $z(E_\varepsilon) = z(F_\varepsilon) = z(C) - \varepsilon$. We choose a homeomorphism g_ε, taking the arc $D_\varepsilon C E_\varepsilon$ to the arc $H_\varepsilon G F_\varepsilon$ in such a way that $g_\varepsilon(D_\varepsilon) = H_\varepsilon$ and $g_\varepsilon(E_\varepsilon) = F_\varepsilon$. Outside the arc $D_\varepsilon C E_\varepsilon$ in $\mathcal{M}$ we define g_ε by taking every point P to a point $g_\varepsilon(P)$, verifying $z(P) = z(g_\varepsilon(P))$.

For every $n \in \mathbb{N} - \{0\}$ we set $f_n = g_{1/n}$. It is easy to prove that

$$\lim_n \Theta(f_n) = \frac{1}{2}(z(A) - z(B)).$$

Now we have only to verify that no homeomorphism between $\mathcal{M}$ and $\mathcal{N}$ exists for which $\Theta(f) \leq \frac{1}{2}(z(A) - z(B))$. If such a homeomorphism existed, for every $P \in \mathcal{M}$ we would have

$$|z(P) - z(f(P))| \leq \frac{z(A) - z(B)}{2}$$

and hence $z(f(A)) \geq z(G) \geq z(f(B))$. Therefore we could easily find points $P \in \mathcal{M}$, for which $|z(P) - z(f(P))| > \frac{1}{2}(z(A) - z(B))$, contradicting our assumption.

3 Some technical tools and definitions

3.1 The concept of train of "limit d-jumps"

In order to prove our main theorem, we need some new definitions and technical results. Assume two size pairs $(\mathcal{M}, \omega)$, $(\mathcal{N}, \omega)$ are given.

The symbol $S_H(\mathcal{M}, \mathcal{N})$ will denote the set of all sequences of homeomorphisms (f_n) in $H(\mathcal{M}, \mathcal{N})$ such that $\Theta(f_n) \to d$. Every sequence in $S_H(\mathcal{M}, \mathcal{N})$ will be called a *d-approximating sequence*.

Let us consider a sequence $(f_n) \in S_H(\mathcal{M}, \mathcal{N})$. We shall say that a pair of points $(P, Q) \in \mathcal{M} \times \mathcal{N}$ is in relation with respect to (f_n) if a sequence (P_r) in $\mathcal{M}$ exists, together with a strictly increasing sequence (i_r) in $\mathbb{N}$ such that

$$(P, Q) = \lim_r (P_r, f_{i_r}(P_r)).$$

In this case we shall write either $P \rho Q$ or $Q \rho P$, indifferently.

In the following part of this section we shall assume that $0 < d < +\infty$. The following compact sets are defined with respect to each d-approximating sequence (f_n):

$$\begin{aligned}
\mathbf{N}^+_\mathcal{M} = \mathbf{N}^+_\mathcal{M}\big((f_n)\big) &= \{P \in \mathcal{M} \mid \exists Q \in \mathcal{N} : P\,\rho\,Q,\ \omega(Q) - \omega(P) = d\} \\
\mathbf{N}^-_\mathcal{M} = \mathbf{N}^-_\mathcal{M}\big((f_n)\big) &= \{P \in \mathcal{M} \mid \exists Q \in \mathcal{N} : P\,\rho\,Q,\ \omega(P) - \omega(Q) = d\} \\
\mathbf{N}^+_\mathcal{N} = \mathbf{N}^+_\mathcal{N}\big((f_n)\big) &= \{Q \in \mathcal{N} \mid \exists P \in \mathcal{M} : P\,\rho\,Q,\ \omega(P) - \omega(Q) = d\} \\
\mathbf{N}^-_\mathcal{N} = \mathbf{N}^-_\mathcal{N}\big((f_n)\big) &= \{Q \in \mathcal{N} \mid \exists P \in \mathcal{M} : P\,\rho\,Q,\ \omega(Q) - \omega(P) = d\}\,.
\end{aligned}$$

In other words, the points P in $\mathbf{N}^+_\mathcal{M}$ are those for which a point $Q \in \mathcal{N}$ exists, such that the pair (P, Q) can be approximated arbitrarily well by a pair $(P_n, f_{i_n}(P_n))$ whose "jump" $\omega(f_{i_n}(P_n)) - \omega(P_n)$ is arbitrarily close to d. Hence, if we think of ω as a "height" function (cf. the examples in the previous section), the points P_n have images with height approximated by $\omega(P_n) + d$. In $\mathbf{N}^+_\mathcal{M}$, the symbol $\mathcal{M}$ recalls the manifold to which P belongs, while the symbol $+$ recalls that, by taking P to Q, we increase the value of the measuring function, i.e. the "jump" starting from the node in $\mathcal{M}$ is "upwards". The notations used for the other three sets are quite analogous. The symbol $-$ is used as a sign for denoting nodes from which "downwards jumps" start (the starting node belonging to the manifold shown as subscript).

It is clear that, for every point $P \in \mathbf{N}^+_\mathcal{M}$, a point $Q \in \mathbf{N}^-_\mathcal{N}$ exists such that $P\,\rho\,Q$ (and vice versa), and that an analogous relation holds for the sets $\mathbf{N}^-_\mathcal{M}$ and $\mathbf{N}^+_\mathcal{N}$. For every sequence of homeomorphisms in $S_H(\mathcal{M}, \mathcal{N})$ the sets $\mathbf{N}_\mathcal{M} = \mathbf{N}^+_\mathcal{M} \cup \mathbf{N}^-_\mathcal{M}$ and $\mathbf{N}_\mathcal{N} = \mathbf{N}^+_\mathcal{N} \cup \mathbf{N}^-_\mathcal{N}$ are non-empty because of the compactness of the manifolds and the continuity of the measuring functions.

Now we shall define the concept of "train" for a d-approximating sequence:

Definition 3.1 Let $(N_0, N_1, \ldots, N_k)$ be an ordered $(k+1)$-tuple of points in $\mathcal{M} \cup \mathcal{N}$ with $k \geq 1$ such that, for $i = 0, \ldots, k-1$ the following properties hold:

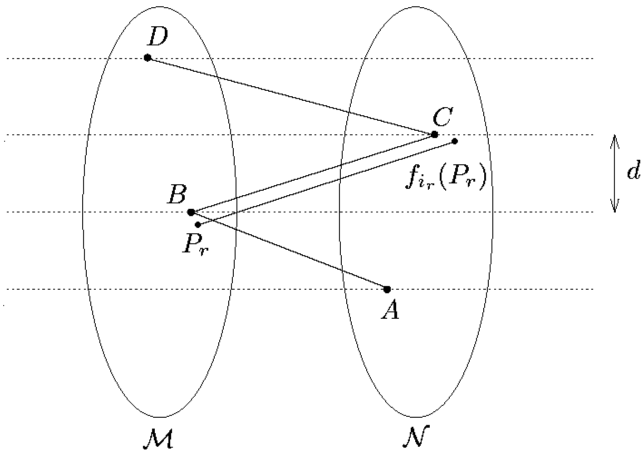

Fig. 3 A train of limit d-jumps given by the quadruple (A, B, C, D).

a) $\omega(N_{i+1}) = \omega(N_i) + d$;

b) $N_i \,\rho\, N_{i+1}$.

In this case the ordered set $(N_0, N_1, \ldots, N_k)$ will be called a *train of limit d-jumps for the sequence (f_n)* (or, in short, a *train*) and its points will be called *nodes*. The pairs (N_i, N_{i+1}) will be known as the *wagons* of the train. The number k will be called *length of the train* and each train that is not included in any other train will be said to be *maximal*. If $(N_0, \ldots, N_k)$ is a maximal train, its wagons (N_0, N_1) and (N_{k-1}, N_k) will be called *initial* and *final train wagons* (respectively), while N_0 and N_k will be the *initial* and *final train nodes*. The remaining nodes will be called *internal nodes*. The symbol $W\big((f_n)\big)$ will denote the set of all the train wagons (for all the existing trains).

Since each point belonging either to $\mathbf{N}_{\mathcal{M}}$ or to $\mathbf{N}_{\mathcal{N}}$ is a node for at least one train, the set of all trains is not empty. Notice that the point P is an initial node for at least a maximal train if and only if either $P \in \mathbf{N}^+_{\mathcal{M}} - \mathbf{N}^-_{\mathcal{M}}$ or $P \in \mathbf{N}^+_{\mathcal{N}} - \mathbf{N}^-_{\mathcal{N}}$, whereas it is a final node if and only if either $P \in \mathbf{N}^-_{\mathcal{M}} - \mathbf{N}^+_{\mathcal{M}}$ or $P \in \mathbf{N}^-_{\mathcal{N}} - \mathbf{N}^+_{\mathcal{N}}$.

In Figure 3 we provide a graphic representation of a maximal train (A, B, C, D). In this particular case, we have that $A \in \mathbf{N}^+_{\mathcal{N}}$, $B \in \mathbf{N}^+_{\mathcal{M}} \cap \mathbf{N}^-_{\mathcal{M}}$, $C \in \mathbf{N}^+_{\mathcal{N}} \cap \mathbf{N}^-_{\mathcal{N}}$ and $D \in \mathbf{N}^-_{\mathcal{M}}$. Hence A is the initial node and D is the final train node, while B and C are internal nodes. The three ordered pairs $(A, B), (B, C), (C, D)$ are the three wagons in the train; (A, B) and (C, D) are its initial and final wagon, respectively.

In Figure 4 we find the maximal train (B, G, A) associated to the d-approximating sequence we described in Example 2.7. In fact we can easily prove that $B\rho G$, $G\rho A$, $z(G) - z(B) = d$ and $z(A) - z(G) = d$. Hence $B \in \mathbf{N}^+_{\mathcal{M}}$, $G \in \mathbf{N}^+_{\mathcal{N}} \cap \mathbf{N}^-_{\mathcal{N}}$ and $A \in \mathbf{N}^-_{\mathcal{M}}$.

Remark 3.2 The example described in Figure 4 shows that the existence of a train of length 2, such that its initial node (in this case B) and its final node (in this case A) are critical points of the measuring function z, guarantees that the natural pseudo-distance equals half the distance between two critical values of the measuring function.

Our main goal will be to show that in the case of closed curves it is always possible to construct a sequence of d-approximating homeomorphisms for which we can demonstrate the existence of a train of length 1 or 2, beginning and ending at critical heights for the measuring functions. We shall do that in the next subsection 3.2 and in Section 4. The example we have just seen justifies our task, since it points out a simple relation between d and the critical values of z.

Now, in order to attain our goal, we need to introduce the concept of *minimal d-approximating sequence*.

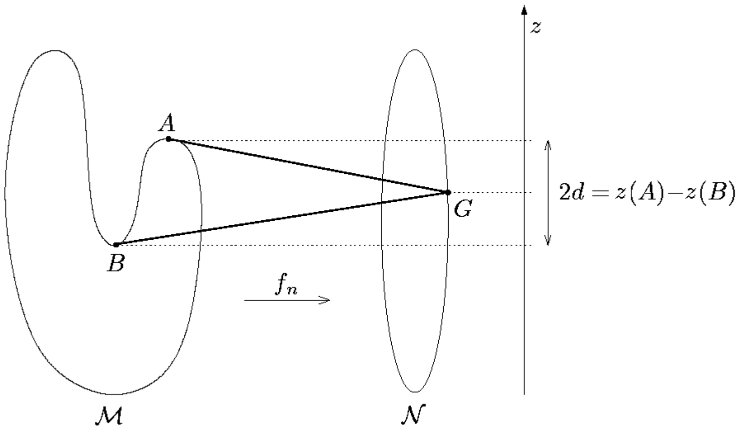

Fig. 4 An example of a train of limit d-jumps given by the triple (B, G, A).

3.2 Minimal d-approximating sequences

The concept of train that we have just mentioned allows us to prove Theorem 2.4 cited in Section 2, and will be central in the following sections, devoted to the proof of the main result in this paper (Theorem 4.5).

In this subsection we shall assume that $\mathcal{M}, \mathcal{N}$ are smooth homeomorphic closed manifolds and φ, ψ are Morse functions.

As we explained in the introduction, the main goal of this paper is to show that the analytic folding number is either 1 or 2 in the case of closed curves.

The idea is proving that we can always get the existence of a train like the one shown in Remark 3.2. In order to obtain that, from a constructive point of view we need to take a d-approximating sequence and improve it by shortening its trains as much as possible, until we get a train of length 1 or 2, beginning and ending at critical heights for the measuring functions.

This procedure will be carried out in two steps. The former will consist in a reduction of the trains applicable in any dimension, which has been developed and applied in [6] (Lemma 3.6 in this paper) in order that only trains beginning and ending at critical points for the measuring functions remain.

The latter will be a reduction process, expressly developed for the case of curves, allowing us to get a further shortening of trains.

Our goal requires a formal definition of "improving" a d-approximating sequence.

Hence we need to define the following preordering $\preceq$ on the set $S_H(\mathcal{M}, \mathcal{N})$ of the d-approximating sequences.

Definition 3.3 If (f_n) and (g_n) are two d-approximating sequences, we set

$$(g_n) \preceq (f_n) \ \ (\text{or, equivalently,} \ \ (f_n) \succeq (g_n))$$

if $\varphi\big(\mathbf{N}^+_{\mathcal{M}}\big((g_n)\big)\big) \subseteq \varphi\big(\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)\big)$ and $\varphi\big(\mathbf{N}^-_{\mathcal{M}}\big((g_n)\big)\big) \subseteq \varphi\big(\mathbf{N}^-_{\mathcal{M}}\big((f_n)\big)\big)$.

Definition 3.4 Let (f_n) and (g_n) be two d-approximating sequences. We say that $(g_n) \prec (f_n)$ (or, equivalently, $(f_n) \succ (g_n)$) if $(g_n) \preceq (f_n)$ and either $\varphi\big(\mathbf{N}^+_{\mathcal{M}}\big((g_n)\big)\big) \neq \varphi\big(\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)\big)$ or $\varphi\big(\mathbf{N}^-_{\mathcal{M}}\big((g_n)\big)\big) \neq \varphi\big(\mathbf{N}^-_{\mathcal{M}}\big((f_n)\big)\big)$ (i.e., at least one of the two inclusions in Definition 3.3 is proper).

We shall say that $(f_n) \in S_H(\mathcal{M}, \mathcal{N})$ is a *minimal sequence* if no sequence $(g_n) \in S_H(\mathcal{M}, \mathcal{N})$ exists such that $(g_n) \prec (f_n)$.

Remark 3.5 The relations $\preceq$ and $\prec$ could be defined by referring to the nodes in $\mathcal{N}$ in place of the nodes in $\mathcal{M}$. In fact, our definitions immediately imply that the inclusion $\varphi\big(\mathbf{N}^+_{\mathcal{M}}\big((g_n)\big)\big) \subseteq \varphi\big(\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)\big)$ is equivalent to the inclusion $\psi\big(\mathbf{N}^-_{\mathcal{N}}\big((g_n)\big)\big) \subseteq \psi\big(\mathbf{N}^-_{\mathcal{N}}\big((f_n)\big)\big)$ and the inclusion $\varphi\big(\mathbf{N}^-_{\mathcal{M}}\big((g_n)\big)\big) \subseteq \varphi\big(\mathbf{N}^-_{\mathcal{M}}\big((f_n)\big)\big)$ is equivalent to the inclusion $\psi\big(\mathbf{N}^+_{\mathcal{N}}\big((g_n)\big)\big) \subseteq \psi\big(\mathbf{N}^+_{\mathcal{N}}\big((f_n)\big)\big)$. An analogous statement holds for proper inclusions.

We observe that, in our definition, $(g_n) \preceq (f_n)$ *does not* mean that either $(g_n) \prec (f_n)$ or $(g_n) = (f_n)$.

The minimal sequences for $\prec$ are, in some ways, the best sequences of homeomorphisms whose measure approximates the natural size pseudo-distance, since they minimize the sets $\varphi\big(\mathbf{N}_{\mathcal{M}}^{+}\big)$ and $\varphi\big(\mathbf{N}_{\mathcal{M}}^{-}\big)$ (and hence the sets $\psi\big(\mathbf{N}_{\mathcal{N}}^{+}\big)$ and $\psi\big(\mathbf{N}_{\mathcal{N}}^{-}\big)$, too, i.e. the sets of node heights for the four types of node we have considered). Afterwards, we shall see that it is always possible to construct a d-approximating sequence of homeomorphisms such that the sets $\varphi\big(\mathbf{N}_{\mathcal{M}}\big)$ and $\psi\big(\mathbf{N}_{\mathcal{N}}\big)$ are finite, and that this can be done by using minimal sequences, too.

The existence of minimal sequences with respect to the preordering $\prec$ will be important in the following Section 4.

The main tool used in [6] for proving Theorem 2.4 is the following

Lemma 3.6 *Assume that* $0 < d < +\infty$ *and the measuring functions* φ, ψ *are Morse. For every sequence of homeomorphisms* (f_n) *in* $S_H(\mathcal{M},\mathcal{N})$ *a new sequence* (g_n) *exists in* $S_H(\mathcal{M},\mathcal{N})$ *such that all maximal trains begin and end at critical points of the measuring functions and* $W\big((g_n)\big) \subseteq W\big((f_n)\big)$.

Remark 3.7 We observe that in Lemma 3.6 the relation $(g_n) \preceq (f_n)$ is easily implied by the inclusion $W\big((g_n)\big) \subseteq W\big((f_n)\big)$.

The following proposition (also used in [8]) shows some properties of the minimal sequences we are going to exploit, under the hypotheses that our measuring functions are Morse. We recall the proof for the reader's convenience.

Proposition 3.8 *Assume that* $0 < d < +\infty$ *and the measuring functions* φ, ψ *are Morse, and set* $\mathcal{A} = \{z \in \mathbb{R} | \exists c_1, c_2 \in \varphi(K_\varphi) \cup \psi(K_\psi), r, s \in \mathbb{N} : z - c_1 = rd, c_2 - z = sd\}$. *Then the following statements hold:*

(a) *If a train for a* d*-approximating sequence begins and ends at critical points of the measuring functions, the heights of its nodes belong to the finite set* $\mathcal{A}$.

(b) *For every* d*-approximating sequence* (f_n)*, a minimal sequence* $(h_n) \preceq (f_n)$ *exists whose maximal trains begin and end at critical points of the measuring functions.*

(c) *If a* d*-approximating sequence* (g_n) *is minimal, the height of every node of its trains belongs to the finite set* $\mathcal{A}$.

Proof. (a) It trivially follows from the definition of train. The finiteness of $\mathcal{A}$ follows from the finiteness of the sets K_φ and K_ψ (here we are using the hypothesis that the measuring functions are Morse).

(b) Lemma 3.6 ensures that we can take a sequence $(g_n) \preceq (f_n)$ whose maximal trains begin and end at critical points of the measuring functions. The previous statement (a) and the definition of the relation $\prec$ imply that no infinite descending chain $\big((g_n) \succ (g_n^1) \succ (g_n^2) \succ \dots\big)$ beginning at (g_n) can exist. Let us consider the last term (g_n') in a maximal descending chain beginning at (g_n). Obviously, (g_n') is a minimal d-approximating sequence. Unfortunately, statement (b) is still not proved, since some maximal train of (g_n') could either begin or end at regular points of the measuring functions, contrary to what happens for (g_n). However, by applying Lemma 3.6 to (g_n') we get a new d-approximating sequence (h_n) that is still minimal and has the required property about maximal trains.

(c) Because of previous statement (b), a minimal sequence $(h_n) \preceq (g_n)$ exists whose maximal trains begin and end at critical points of the measuring functions. Since (g_n) is already minimal, it follows that $\varphi\big(\mathbf{N}_{\mathcal{M}}^{+}\big((h_n)\big)\big) = \varphi\big(\mathbf{N}_{\mathcal{M}}^{+}\big((g_n)\big)\big)$ and $\varphi\big(\mathbf{N}_{\mathcal{M}}^{-}\big((h_n)\big)\big) = \varphi\big(\mathbf{N}_{\mathcal{M}}^{-}\big((g_n)\big)\big)$ (and hence $\psi\big(\mathbf{N}_{\mathcal{N}}^{+}\big((h_n)\big)\big) = \psi\big(\mathbf{N}_{\mathcal{N}}^{+}\big((g_n)\big)\big)$ and $\psi\big(\mathbf{N}_{\mathcal{N}}^{-}\big((h_n)\big)\big) = \psi\big(\mathbf{N}_{\mathcal{N}}^{-}\big((g_n)\big)\big)$). Statement (a) ensures that $\varphi\big(\mathbf{N}_{\mathcal{M}}(h_n)\big) \cup \psi\big(\mathbf{N}_{\mathcal{N}}(h_n)\big)$ is included in the finite set $\mathcal{A}$, and therefore the same happens to the set $\varphi\big(\mathbf{N}_{\mathcal{M}}\big((g_n)\big)\big) \cup \psi\big(\mathbf{N}_{\mathcal{N}}\big((g_n)\big)\big)$. □

4 Our main result

In Section 2 we have recalled (Theorem 2.4) that the natural pseudo-distance between two size pairs is always related to the critical values of their measuring functions.

However, the examples we have displayed suggest that our results can be improved. In fact our examples show an analytic folding number k that is never greater than 2. In the first part of this section we shall prove (Theorem 4.3) that this condition always holds, under the assumption that $\mathcal{M}$ and $\mathcal{N}$ are two homeomorphic smooth closed curves and the measuring functions φ, ψ are Morse.

These hypotheses will make our proofs easier from a technical point of view. In subsection 4.1 we shall weaken our assumptions and come back to the case of class C^1 (Theorem 4.5).

So, from now to subsection 4.1, assume that $\mathcal{M}$, $\mathcal{N}$ are smooth homeomorphic closed curves and φ, ψ are Morse functions.

Now we introduce two useful lemmas. The former clarifies the local nature of the concept of node.

Lemma 4.1 *Assume $0 < d < +\infty$. Let U be an open subset of $\mathcal{M}$ and (f_n) and (g_n) be two d-approximating sequences such that, for every $n \in \mathbb{N}$, f_n coincides with g_n in U. Then $\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big) \cap U = \mathbf{N}^+_{\mathcal{M}}\big((g_n)\big) \cap U$ and $\mathbf{N}^-_{\mathcal{M}}\big((f_n)\big) \cap U = \mathbf{N}^-_{\mathcal{M}}\big((g_n)\big) \cap U$.*

Proof. It immediately follows from the definitions of the sets $\mathbf{N}^+_{\mathcal{M}}$ and $\mathbf{N}^-_{\mathcal{M}}$. $\square$

A similar result obviously holds for an open subset V of $\mathcal{N}$, and can easily be obtained by interchanging the roles of the sequences (f_n), (g_n) and (f_n^{-1}), (g_n^{-1}) in Lemma 4.1.

The useful property described by the following lemma justifies the introduction of the concept of minimal sequence in the case of closed curves.

Lemma 4.2 *Let $\mathcal{M}$ and $\mathcal{N}$ be two homeomorphic closed curves for which $0 < d < +\infty$, and (f_n) be a minimal d-approximating sequence. Then, for every $N \in \mathbf{N}^+_{\mathcal{M}}$, either $\varphi(N)$ is a critical value for the function φ, or the number $\varphi(N) + d$ is a critical value for the function ψ.*

In other words, under the hypotheses of the lemma (possibly by exchanging the roles of the two curves), if we consider the heights of two consecutive nodes in a train of a minimal sequence, at least one of them is a critical value.

Proof. We shall prove that, if $\varphi(N)$ is not a critical value for φ and $\varphi(N) + d$ is not a critical value for the function ψ, then we can get a new d-approximating sequence $(\hat{f}_n)$ such that $(\hat{f}_n) \prec (f_n)$, contradicting the assumption that (f_n) is minimal.

Let us call D^φ_ε the open set $\{P \in \mathcal{M} : |\varphi(P) - \varphi(N)| < \varepsilon\}$ and $D^\psi_{2\varepsilon}$ the open set $\{Q \in \mathcal{N} : |\psi(Q) - (\varphi(N) + d)| < 2\varepsilon\}$, and choose $\varepsilon > 0$ so small that

i) $\overline{D^\varphi_\varepsilon}$ does not contain critical points for φ;

ii) $\partial D^\varphi_\varepsilon$ does not contain nodes belonging to $\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)$;

iii) $D^\psi_{2\varepsilon}$ does not contain critical points for ψ.

The existence of such an ε is ensured by the assumption that $\varphi(N)$ is not a critical value for φ and $\varphi(N) + d$ is not a critical value for the function ψ, and by the fact that the set of heights of the nodes is finite (see Proposition 3.8(c)).

Condition $ii)$ implies that an $\eta > 0$ exists such that, for any large enough n, $\psi(f_n(P)) - \varphi(P) \le d - \eta$ for every P in the set $\partial D^\varphi_\varepsilon$. In fact, if that inequality did not hold, at least one of the points in the set $\partial D^\varphi_\varepsilon$ should be a node belonging to $\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)$, contradicting $ii)$. This is a consequence of the compactness (actually finiteness, since ψ is Morse) of the set $\partial D^\varphi_\varepsilon$ (cf. [6], Lemma 4.1).

Now, before constructing the required sequence $(\hat{f}_n)$, let us prove that a sequence $(\tilde{f}_n) \in S_H(\mathcal{M},\mathcal{N})$ exists such that $\tilde{f}_n = f_n$ in the closed set $\mathcal{M} - D^\varphi_\varepsilon$ and $\mathbf{N}^+_{\mathcal{M}}\big((\tilde{f}_n)\big) \cap \overline{D^\varphi_\varepsilon} = \emptyset$ (in other words, we can eliminate all "upwards" wagons from $\mathcal{M}$ to $\mathcal{N}$, beginning in $\overline{D^\varphi_\varepsilon}$). We shall use $(\tilde{f}_n)$ as an intermediate step to define $(\hat{f}_n)$.

So, we set $\tilde{f}_n(P) = f_n(P)$ for $P \in \mathcal{M} - D^\varphi_\varepsilon$.

Let us define $\tilde{f}_n(P)$ in the case $P \in D^\varphi_\varepsilon$.

If $P \in D^\varphi_\varepsilon$ and $f_n(P)$ is not a critical point for ψ, consider the open set ℓ, defined as the connected component of $f_n(D^\varphi_\varepsilon) - K_\psi$ containing $f_n(P)$. Then call E_1 and E_2 the ends of the closure of the arc ℓ. By construction, each of them is either a point belonging to $f_n(\partial D^\varphi_\varepsilon)$ or a critical point for ψ belonging to $f_n(D^\varphi_\varepsilon)$ (see Figure 5).

It is easy to see that, since $D^\varphi_\varepsilon \cap K_\varphi = \emptyset$ and $\ell \cap K_\psi = \emptyset$, a real number λ_P in the open interval $(0,1)$ and a point $Q = Q(P) \in \ell$ exist (and are uniquely defined) such that

$$\begin{aligned} \varphi(P) &= \lambda_P \varphi(f_n^{-1}(E_1)) + (1 - \lambda_P)\varphi(f_n^{-1}(E_2)) \\ \psi(Q) &= \lambda_P \psi(E_1) + (1 - \lambda_P)\psi(E_2). \end{aligned}$$

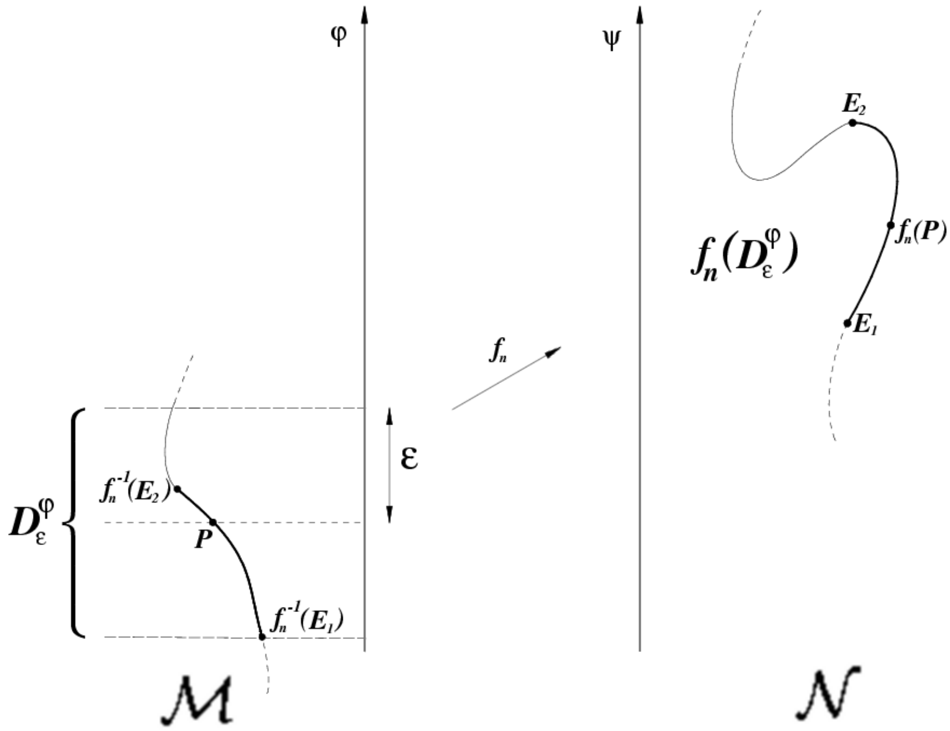

Fig. 5 In this example a connected component ℓ of the set $f_n(D^{\varphi}_{\varepsilon}) - K_{\psi}$, containing P, is represented on the right (thick line).

Hence, we shall set $\tilde{f}_n(P) = Q$.

If $P \in D^{\varphi}_{\varepsilon}$ but $f_n(P)$ is a critical point for ψ, we shall simply set $\tilde{f}_n(P) = f_n(P)$ (so the two homeomorphisms $\tilde{f}_n$ and f_n will even coincide in the set $\mathcal{M} - (D^{\varphi}_{\varepsilon} - f_n^{-1}(K_{\psi})))$.

In practice, we proceed this way in order to get the function $\psi \circ \tilde{f}_n \circ \varphi^{-1}$ to be linear on the interval $\varphi(f_n^{-1}(\ell))$ for each connected component ℓ of $f_n(D^{\varphi}_{\varepsilon}) - K_{\psi}$ (note that the restriction of φ to $f_n^{-1}(\ell)$ is invertible). This relaxing procedure is the tool we use to eliminate all "upwards" wagons from $\mathcal{M}$ to $\mathcal{N}$, beginning in $D^{\varphi}_{\varepsilon}$.

From the linearity of $\psi \circ \tilde{f}_n \circ \varphi^{-1}$ (and hence of $\psi \circ \tilde{f}_n \circ \varphi^{-1} - id$) on the interval $\varphi(f_n^{-1}(\ell))$, for each connected component ℓ in $f_n(D^{\varphi}_{\varepsilon}) - K_{\psi}$, it follows that the extrema of $\psi \circ \tilde{f}_n \circ \varphi^{-1} - id$ are taken at the two points on the boundary of $\varphi(f_n^{-1}(\ell))$, that is, two points $\varphi(A)$, $\varphi(B)$ with $A, B \in \partial(f_n^{-1}(\ell))$.

This means that, for every $P \in D^{\varphi}_{\varepsilon}$, two points A and B exist ($A = B = P$ if $f_n(P) \in K_{\psi}$) such that each of them is either a point belonging to $\partial D^{\varphi}_{\varepsilon}$ or a point of $D^{\varphi}_{\varepsilon}$ that is also the inverse image by f_n of a critical point for ψ, and

$$\begin{aligned} \psi(f_n(A)) - \varphi(A) &= \psi(\tilde{f}_n(A)) - \varphi(A) \\ &\leq \psi(\tilde{f}_n(P)) - \varphi(P) \\ &\leq \psi(\tilde{f}_n(B)) - \varphi(B) = \psi(f_n(B) - \varphi(B)). \end{aligned} \tag{4.0.1}$$

Hence $\Theta(\tilde{f}_n) \leq \Theta(f_n)$, and so the new sequence $(\tilde{f}_n)$ is a d-approximating sequence, too.

Since f_n and $\tilde{f}_n$ coincide outside $D^{\varphi}_{\varepsilon}$, and in particular outside the closed set $\overline{D^{\varphi}_{\varepsilon}}$, by construction, Lemma 4.1 guarantees that the equalities

$$\begin{aligned} \mathbf{N}^+_{\mathcal{M}}\big((\tilde{f}_n)\big) \cap (\mathcal{M} - \overline{D^{\varphi}_{\varepsilon}}) &= \mathbf{N}^+_{\mathcal{M}}\big((f_n)\big) \cap (\mathcal{M} - \overline{D^{\varphi}_{\varepsilon}}) \\ \mathbf{N}^-_{\mathcal{M}}\big((\tilde{f}_n)\big) \cap (\mathcal{M} - \overline{D^{\varphi}_{\varepsilon}}) &= \mathbf{N}^-_{\mathcal{M}}\big((f_n)\big) \cap (\mathcal{M} - \overline{D^{\varphi}_{\varepsilon}}) \end{aligned} \tag{4.0.2}$$

hold, i.e., every upwards (downwards) wagon for (f_n) arriving from $\mathcal{M} - \overline{D_\varepsilon^\varphi}$ at the set $\mathcal{N}$ is also an upwards (downwards) wagon for $(\tilde{f}_n)$ arriving from $\mathcal{M} - \overline{D_\varepsilon^\varphi}$ at $\mathcal{N}$, and vice versa.

Now we are going to prove the equality

$$\mathbf{N}^+_{\mathcal{M}}\big((\tilde{f}_n)\big) \cap \overline{D_\varepsilon^\varphi} \quad = \quad \emptyset, \tag{4.0.3}$$

saying that no upwards wagon from $\mathcal{M}$ to $\mathcal{N}$ beginning in $\overline{D_\varepsilon^\varphi}$ exists.

If a $N' \in \mathbf{N}^+_{\mathcal{M}}\big((\tilde{f}_n)\big) \cap \overline{D_\varepsilon^\varphi}$ existed, then a sequence (P_n) in D_ε^φ would exist such that, possibly by extracting a subsequence, $\psi(\tilde{f}_n(P_n)) - \varphi(P_n) \to d$. (The possibility of assuming $P_n \in D_\varepsilon^\varphi$ can easily be obtained in the case $N' \in D_\varepsilon^\varphi$; in the case $N' \in \partial D_\varepsilon^\varphi$ this is ensured by the coincidence of f_n and $\tilde{f}_n$ outside D_ε^φ and by hypothesis $ii)$, which states that $\partial D_\varepsilon^\varphi$ does not contain nodes belonging to $\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)$).

Hence, because of (4.0.1), a sequence (B_n) would also exist such that $\psi(f_n(B_n)) - \varphi(B_n) \to d$, where each B_n is either a point belonging to $\partial D_\varepsilon^\varphi$ or a point of D_ε^φ that is also the inverse image by f_n of a critical point for ψ.
Since we have seen that, for any large enough n, $\psi \circ f_n - \varphi \leq d - \eta$ in $\partial D_\varepsilon^\varphi$, then $B_n \notin \partial D_\varepsilon^\varphi$, and hence it is the inverse image by f_n of a critical point for ψ belonging to $f_n(D_\varepsilon^\varphi)$. Therefore, $B_n \in D_\varepsilon^\varphi$, and so $|\varphi(B_n) - \varphi(N)| < \varepsilon$, for any large enough n (where N is the node cited in the statement of our Lemma 4.2). Moreover,

$$|\psi(f_n(B_n)) - (\varphi(N) + d)| \leq |\psi(f_n(B_n)) - \varphi(B_n) - d| + |\varphi(B_n) - \varphi(N)|.$$

Therefore, for any large enough n the inequality

$$|\psi(f_n(B_n)) - (\varphi(N) + d)| < 2\varepsilon$$

holds, contradicting property $iii)$, since $f_n(B_n)$ is a critical point of ψ.

So we have proved that the new d-approximating sequence $(\tilde{f}_n)$ has no wagon going upwards from $\overline{D_\varepsilon^\varphi}$ to $\mathcal{N}$.

Unfortunately, we cannot ensure that $(\tilde{f}_n) \preceq (f_n)$, since the inclusion $\mathbf{N}^-_{\mathcal{M}}\big((\tilde{f}_n)\big) \subseteq \mathbf{N}^-_{\mathcal{M}}\big((f_n)\big)$ is not guaranteed (in fact, it could be $\mathbf{N}^-_{\mathcal{M}}\big((\tilde{f}_n)\big) \cap \overline{D_\varepsilon^\varphi} \not\subseteq \mathbf{N}^-_{\mathcal{M}}\big((f_n)\big) \cap \overline{D_\varepsilon^\varphi}$).

We can remedy this problem by applying Lemma 3.6 to $(\tilde{f}_n)$. In this way we get a new sequence $(\hat{f}_n)$ whose wagons are wagons for $(\tilde{f}_n)$, too. Therefore (4.0.2) and (4.0.3) imply that the equalities

$$\mathbf{N}^+_{\mathcal{M}}\big((\hat{f}_n)\big) \cap (\mathcal{M} - \overline{D_\varepsilon^\varphi}) \subseteq \mathbf{N}^+_{\mathcal{M}}\big((f_n)\big) \cap (\mathcal{M} - \overline{D_\varepsilon^\varphi})$$

$$\mathbf{N}^-_{\mathcal{M}}\big((\hat{f}_n)\big) \cap (\mathcal{M} - \overline{D_\varepsilon^\varphi}) \subseteq \mathbf{N}^-_{\mathcal{M}}\big((f_n)\big) \cap (\mathcal{M} - \overline{D_\varepsilon^\varphi})$$

$$\mathbf{N}^+_{\mathcal{M}}\big((\hat{f}_n)\big) \cap \overline{D_\varepsilon^\varphi} = \emptyset$$

hold.

Since all maximal trains for $(\hat{f}_n)$ begin and end at critical points of the measuring functions, and the set $\overline{D_\varepsilon^\varphi}$ does not contain critical points for φ (Property $i)$), the equality $\mathbf{N}^+_{\mathcal{M}}\big((\hat{f}_n)\big) \cap \overline{D_\varepsilon^\varphi} = \emptyset$ implies the equality $\mathbf{N}^-_{\mathcal{M}}\big((\hat{f}_n)\big) \cap \overline{D_\varepsilon^\varphi} = \emptyset$.

It follows that $(\hat{f}_n) \preceq (f_n)$.

Since $\overline{D_\varepsilon^\varphi}$ does not meet $\mathbf{N}^+_{\mathcal{M}}\big((\hat{f}_n)\big)$, but contains at least one node of $\mathbf{N}^+_{\mathcal{M}}\big((f_n)\big)$, it follows that $(\hat{f}_n) \prec (f_n)$, contradicting the hypothesis that (f_n) is a minimal sequence. □

Let us apply Lemma 4.2 to prove that the analytic folding number is never greater than 2 for the closed curves.

Theorem 4.3 *Assume that $\mathcal{M}$ and $\mathcal{N}$ are two homeomorphic smooth closed curves and that $\varphi : \mathcal{M} \to \mathbb{R}$ and $\psi : \mathcal{N} \to \mathbb{R}$ are two Morse functions. Then, if d denotes the natural pseudo-distance between the size pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$, at least one of the following properties holds:*

$a)$ *d equals the distance between a critical value of φ and a critical value of ψ;*

$b)$ *d equals half the distance between two critical values of φ.*

$c)$ *d equals half the distance between two critical values of ψ.*

Proof. If $d = 0$, then φ and ψ have the same global minimum μ. Hence, $d = |\mu - \mu|$ and the statement is trivial. So let us assume $d > 0$. Let (f_n) be a minimal sequence whose maximal trains begin and end at critical points of the measuring functions (Proposition 3.8(b)) and suppose $N' \in \mathcal{N}$ is the initial node of a maximal train (if no maximal train begins in $\mathcal{N}$, it is sufficient to exchange the roles of our curves). Therefore $\psi(N')$ is a critical value for the measuring function ψ. Let $N \in \mathcal{N}$ be the next node in the train. If $\varphi(N) \in \varphi(K_\varphi)$, then condition $a)$ holds. Otherwise, call $N'' \in \mathcal{N}$ the next node (N'' exists because $\varphi(N) \notin \varphi(K_\varphi)$, and so it is not the final node of the train). Lemma 4.2 ensures that $\psi(N'') = \varphi(N) + d \in \psi(K_\psi)$. Therefore $d = \frac{1}{2}|\psi(N'') - \psi(N')|$ and property $c)$ holds. □

Remark 4.4 Lemma 4.2 may be considered analogous to Lemma 5.2 proved in [8] for surfaces, but the techniques used in the proof are different, since here we can use linear maps in place of harmonic maps. Consequently, we succeed in getting more information about the position of the images of critical points. As a result, in both the 1-dimensional and the 2-dimensional case, we can prove that if we consider the heights of m consecutive nodes in a train of a minimal sequence, at least one of them is a critical value, but we have to suppose $m \geq 2$ for curves and $m \geq 3$ for surfaces, depending on the different techniques and dimensional constraints involved in our proofs. This difference explains why the statement of Theorem 5.7 in [8] (for surfaces) is weaker than the statement of Theorem 4.5 in this paper, concerning curves.

4.1 Weakening the hypotheses about the regularity of curves and measuring functions

Until now we have considered smooth closed curves and Morse measuring functions. By repeating the same proofs we have used in [6] to weaken our hypotheses about regularity, we can get our main result via an approximation procedure:

Theorem 4.5 *Assume that $\mathcal{M}$ and $\mathcal{N}$ are two homeomorphic closed curves of class C^1 and that $\varphi : \mathcal{M} \to \mathbb{R}$ and $\psi : \mathcal{N} \to \mathbb{R}$ are two functions of class C^1. Then, if d denotes the natural pseudo-distance between the size pairs $(\mathcal{M}, \varphi)$ and $(\mathcal{N}, \psi)$, at least one of the following properties holds:*

$a)$ *d equals the distance between a critical value of φ and a critical value of ψ;*

$b)$ *d equals half the distance between two critical values of φ.*

$c)$ *d equals half the distance between two critical values of ψ.*

Examples 2.6 and 2.7, showing two cases for which the analytic folding number equals 1 and 2 respectively, demonstrate that the result of Theorem 4.5 is the best possible result for closed curves.

5 Conclusions and further research

In this paper we have proved that for closed curves the relation between the natural pseudo-distance and the critical values of the measuring functions is stronger than the one we proved in [6] for the general dimensions and the one we proved in [8] for surfaces. In fact, Theorem 4.5 shows that the natural pseudo-distance between two closed curves is always either the distance or half the distance between two suitable critical values of the measuring functions. The examples we have given in this paper show that our result cannot be further improved. Therefore Theorem 4.5 completely solves the problem of determining the possible values for the natural pseudo-distance in the 1-dimensional case.

Unfortunately, both the linearization technique used in this paper and the one used in [8] do not seem to be suitable for studying the n-dimensional case. Hence the following question naturally arises. Do two n-manifolds associated with two regular measuring functions φ, ψ exist such that their pseudo-distance equals neither D, $D/2$ nor $D/3$, for D varying in the set of all distances between the critical values of φ and ψ?

In our further research we intend to study this problem and the availability of new techniques to consider the general n-dimensional case. However, it is interesting to note that we do not know of examples where the analytic folding number is strictly greater than 2, even in the bidimensional case.

The difficulty in finding examples where the analytic folding number equals 3, if any exist, deserves some further remarks. One technique that can be used for computing natural size pseudo-distances is based on *size functions* (cf. [7]). The computation of size functions is usually easy, and gives us a lower bound s for natural size pseudo-distances. Obviously, when we are able to show a sequence (f_i) of homeomorphisms for which $\lim_i \Theta(f_i) = s$, we can claim that the natural size pseudo-distance equals s. The key point is that the best lower bound s we can obtain is either the distance, or half the distance, between two suitable critical values of the measuring functions (cf. Theorem 2 in [7]). Therefore, even though an example where the analytic folding number equals 3 may exist, we are not able to find and recognize it using the previously described technique. Apparently, new techniques should be developed.

As regards the use of a linearization procedure in our study, this corresponds to the property that the deformation due to tension fields decreases both the energy and the maximum change of the measuring functions, provided that we are far from their critical points. In higher dimensions, the use of different kinds of deformation (e.g. curvature evolution of the level lines of the measuring functions) might be investigated. The main problem seems to be the possible birth of degeneracies.

As another line of research, it might be interesting to examine the possibility of moving from the study of trains of limit d-jumps to the study of relations obtained as limits of the d-approximating sequences of homeomorphisms, with respect to the Hausdorff (or another more suitable) topology.

In conclusion, various interesting questions remain open and deserve further study and research.

Acknowledgements This work was partially supported by MIUR (Italy), ARCES (Italy), INdAM-GNSAGA (Italy) and NATO. Investigation supported by the University of Bologna (Funds for selected research topics).

We wish to thank M. Ferri for our helpful discussions. Anyway, the authors are solely responsible for any mistakes.

This paper is dedicated to S. Endrigo.

References

[1] M. Allili, D. Corriveau, D. Ziou, *Morse Homology Descriptor for Shape Characterization*, Proceedings of the International Conference on Pattern Recognition, Cambridge (UK), (4) 2004, 27–30.

[2] D. Cohen-Steiner, H. Edelsbrunner and J. Harer, *Stability of Persistence Diagrams*, Proc. 21st Sympos. Comput. Geom. 2005, 263–271.

[3] M. d'Amico, P. Frosini and C. Landi, *Optimal matching between reduced size functions*, DISMI, Univ. di Modena e Reggio Emilia, Italy, Technical report no. **35** (2003).

[4] M. d'Amico, P. Frosini and C. Landi, *Using matching distance in Size Theory: a survey*, International Journal of Imaging Systems and Technology **16** (2006), no. 5, 154-161.

[5] F. Dibos, P. Frosini and D. Pasquignon, *The use of Size Functions for Comparison of Shapes through Differential Invariants*, Journal of Mathematical Imaging and Vision **21** (2004), no. 2, 107–118.

[6] P. Donatini and P. Frosini, *Natural pseudodistances between closed manifolds*, Forum Mathematicum **16** (2004), no. 5, 695–715.

[7] P. Donatini and P. Frosini, *Lower bounds for natural pseudodistances via size functions*, Archives of Inequalities and Applications **2** (2004), 1–12.

[8] P. Donatini and P. Frosini, *Natural pseudodistances between closed surfaces*, Journal of the European Mathematical Society **9** (2007), n. 2, 231-253.

[9] H. Edelsbrunner, D. Letscher and A. Zomorodian, *Topological persistence and simplification*, Discrete Comput. Geom. **28** (2002), no. 4, 511–533.

[10] P. Frosini, *A distance for similarity classes of submanifolds of a Euclidean space*, Bull. Austral. Math. Soc. **42** (1990), no. 3, 407–416.

[11] P. Frosini, *Connections between size functions and critical points*, Math. Meth. Appl. Sci. **19** (1996), 555–569.

[12] P. Frosini, *A note on the linearity of real-valued functions with respect to suitable metrics*, Geometriae Dedicata **108** (2004), n. 1, 105–110.

[13] P. Frosini and C. Landi, *Size functions and morphological transformations*, Acta Applicandae Mathematicae, **49**(1) (1997), 85–104.

[14] P. Frosini and C. Landi, *Intrinsic harmonicity of Morse functions*, Mathematika, **50** (2003), 167–170.

[15] P. Frosini and M. Mulazzani, *Size homotopy groups for computation of natural size distances*, Bull. Belg. Math. Soc. **6** (1999), 455–464.

[16] J. Jost and R. Schoen, *On the existence of harmonic diffeomorphisms between surfaces*, Invent. Math. **66** (1982), 353–359.

[17] T. Kaczynski, K. Mischaikow and M. Mrozek, Computational Homology, Applied Mathematical Sciences **157**, Springer-Verlag, New York, 2004.
[18] P.W. Michor and D. Mumford, *Riemannian geometries on spaces of plane curves*, Journal of the European Mathematical Society **8** (2006), n. 1, 1-48.
[19] P.W. Michor and D. Mumford, *An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach*, Appl. Comput. Harmon. Anal. **23** (2007), no. 1, 74–113.
[20] A. Verri, C. Uras, P. Frosini and M. Ferri, *On the use of size functions for shape analysis*, Biol. Cybern. **70** (1993), 99–107.
[21] A. Verri, C. Uras, *Computing size functions from edge maps*, Internat. J. Comput. Vision **23**(2) (1997), 169–183.
[22] A. Zomorodian, Topology for Computing, Cambridge Monographs on Applied and Computational Mathematics **16**, Cambridge University Press, 2005.