

ON THE DECOMPOSITION AND LOCAL DEGREE OF MULTIPLE SADDLES

SARA DERIVIÈRE, TOMASZ KACZYNSKI
AND PIERRE-OLIVIER VALLERAND-BEAUDRY

RÉSUMÉ. L'analyse topologique des images numériques nous motive vers l'analyse de la formule d'Euler-Maxwell en l'absence des conditions de non dégénérescence et d'isolation. Dans cet article, nous étudions le degré local des k -selles des champs de gradients et généralisons la formule aux points critiques dégénérés.

ABSTRACT. Topological analysis of digital images motivates exploring the Euler-Maxwell formula in the absence of non-degeneracy and isolation conditions. We study the local degree of gradient fields at a k -fold saddle, and provide a generalization of the formula for degenerate critical points.

1. Introduction

This paper is the first step towards the answer to questions posed in [1, 2], concerning the Euler-Maxwell formula in the context of topological analysis of digital images. In those papers, the object studied is a scalar function $f : \mathcal{X} \rightarrow \mathbb{R}$ on a discrete multi-dimensional data set \mathcal{X} . In the planar case studied in [1], f is geometrically interpreted as a height field. The features of interest are critical points of f , that is, peaks, pits, and saddles. Once the critical points are identified, various techniques are used to analyze relationships between them and to trace structures such as ridge lines, ravine lines and isolines. In the case of data of higher dimensions studied in [2], the geometric interpretation of critical points is more complex but those points play equally important role in further study, such as the construction of the level sets given by $f = c$. A good understanding of the nature of saddles is especially important because these are points where level sets intersect.

The smooth Morse theory [12] has inspired researchers in imaging science. However, in its rigorous applications such as [5], one spends a lot of effort on forcing, by local deformation of data, the main hypothesis of the Morse theory stating that critical points of f must be isolated and non-degenerate to hold. By this way, one adjusts the finite input to the theory, with the aim at validating practical implementations. There is a discrete Morse index theory due to Forman [6], but it also deforms the data and, moreover, its goals are different than those in the image analysis. In [1, 2], an effort is made to establish a discrete analogy of the Morse theory for a function f defined on

pixels (mathematically speaking, elementary cubes) while keeping the original geometry, that is, without forcing the isolation and non-degeneracy of critical points. The main results obtained there are the algorithms detecting and classifying critical regions, and constructing the so-called Morse connections graph, whose nodes are critical components and edges display the existence of trajectories connecting them. A computer experimentation is done in [1] on planar images.

Among questions addressed in [1, 2], one is related to extensions of the formula

$$(1) \quad \# \text{ pits} - \# \text{ passes} + \# \text{ peaks} = 2$$

for a height function defined on the surface of the globe, that is, the two-dimensional sphere S^2 . This formula is essentially due to Maxwell [13] but is often called *Euler formula* due to its similarity to the Euler characteristic of the sphere. In imaging science, it is used (often reinforced) as a criterion of correctness of programs extracting information on critical points from discrete data. A generalization of this formula to arbitrary dimensions and to Morse functions $f : M \rightarrow \mathbb{R}$ on compact smooth manifolds is the *Morse formula*

$$(2) \quad \sum_{i=1}^n (-1)^{\lambda(p_i)} = \chi(M),$$

where the p_i are the non-degenerate critical points of f and, for each i , $\lambda(p_i)$ is the Morse index of p_i defined as the sign of the Hessian of f at p_i . Finally, $\chi(M)$ is the Euler-Poincaré characteristic of M . The terminology related to Morse functions is recalled in Section 2.

In applications to digital $2D$ image analysis, the functions are neither defined on manifolds, nor on the sphere S^2 , but on some rectangular regions $D \subset \mathbb{R}^2$. One makes use of the formula (2) by assuming that D is “surrounded by a depression”, so that we may compactify the plane \mathbb{R}^2 to the sphere S^2 with a point at infinity where f assumes an absolute minimum. The same argument is used in an arbitrary dimension for a function on a bounded rectangular domain $\bar{D} \subset \mathbb{R}^d$. In mathematical terms, the assumption on a surrounding depression can be formulated by saying that f is decreasing through ∂D towards the exterior of D , or that ∇f points inward on ∂D . For the d -dimensional sphere S^d , we have $\chi(S^d) = 1 + (-1)^d$. Thus, by removing the added minimum point at infinity, we should obtain the formula

$$(3) \quad \sum_{i=1}^n (-1)^{\lambda(p_i)} = (-1)^d$$

for a function $f : \bar{D} \rightarrow \mathbb{R}$ whose gradient points inward on ∂D .

An observation which motivated the direction chosen in this paper is that the use of the passage through a theory of compact manifolds is somewhat artificial: the original function is defined on $\bar{D} \subset \mathbb{R}^d$ and we end up formulating the result for such functions. Thus, we want a more elementary and direct proof confined within the framework of functions on bounded domains in \mathbb{R}^d .

There is such a theory at our disposal. It is the Brouwer degree, also called topological degree theory of vector fields which we may apply here to the gradient of f .

Moreover, the degree theory remains valid for gradients of functions which have degenerate critical points, such as monkey saddles. Furthermore, it is valid for arbitrary continuous vector fields.

In the introductory Section 2, the basics of the degree theory are recalled and used to give an alternative proof of the formula (3) in the non-degenerate case.

In Section 3, we discuss a known model for the monkey saddle and use it to give a general definition of k -fold saddle using the terminology of stable and unstable manifolds from the theory of dynamical systems. We next give a combinatorial procedure for decomposing a k -fold saddle to k simple saddles.

Section 4 is concerned with any isolated, but possibly degenerated, critical points in \mathbb{R}^2 . We first state a version of Wilson and Yorke's isolating block, adapted to our context. We prove Theorem 4.7 stating that any isolated critical point is either a minimum, a maximum, or a k -fold saddle, whose local degree is 1 in the first two cases and $-k$ in the last one. We next use the additivity property of degree to provide a generalization of (3) in the presence of k -fold saddles.

As we said, the topological degree is valid for any continuous vector field on a domain in \mathbb{R}^d and any upper semi-continuous multivalued vector field with compact convex values (more generally, contractible or aspherical values) and it is additive with respect to unions of regions. Thus, it should suit better the applications to discrete data and critical regions in the context of [1, 2]. This is the project for the future work discussed in Section 5.

We finish this introduction with a little disclaimer. The study of dynamical systems in arbitrary dimensions is very extensive, and many statements presented in this paper can be derived from more general and abstract theorems formulated often in the language of algebraic topology. In particular:

(a) The most general and concise formulations of the local degree of a map are in terms of the homomorphism induced in homology or cohomology groups of spheres.

(b) The differential equation $\dot{z} = (k+1)\bar{z}^k$ related to the model (8) of k -fold saddle studied in Section 3 is well known, and it is a special case of differential equations studied in [15].

(c) Some conclusions in the proof of Theorem 4.7 can be derived from cohomological statements in [7].

However, the generality of a theory is often an obstruction to geometric visualization and accessibility to the applied mathematics, computer science, and engineering communities. We wish to emphasize that our goal is not to achieve the greatest possible generality but to give a presentation of the issue as elementary, as self confined, and as visual as possible within the framework of a mathematical paper. We hope that the understanding of geometric aspects of both analytical and homological tools will be helpful in designing adequate models in digital imaging.

2. Degree of a generic gradient field

The concept of topological degree goes back to Brouwer [3, 1912] but we use here a more recent analytic viewpoint on the degree due to Nagumo [14, 1951] which is a common choice in reference texts such as [10]. The degree is a tool for investigating the equation $F(x) = q$, where $F : \overline{D} \rightarrow \mathbb{R}^d$ is a continuous map of the closure of a bounded open subset D of \mathbb{R}^d and $q \in \mathbb{R}^d$. If F is admissible, that is, $F(x) \neq q$ for $x \in \partial D$, then one defines the *topological degree of F on D with respect to q* , which is an integer (denoted by $\deg(F, D, q)$) satisfying the additivity, homotopy and normalization axioms.

The analytic construction of the degree given, for example, in [10] goes in several steps of approximation. First, one assumes that F is *generic*, that is, it is of class C^1 and the *Jacobian* of F at p , defined by $J_F(p) = \det DF(p)$, is non-zero at any p such that $F(p) = q$. One proves that, in this case, the zeros of F are isolated and hence, since \overline{D} is compact, there are finitely many of them. Let $F^{-1}(q) = \{p_1, p_2, \dots, p_n\}$. Then the degree is defined by the formula

$$(4) \quad \deg(F, D, q) = \sum_{i=1}^n \operatorname{sgn} J_F(p_i).$$

In particular, if D is a bounded neighborhood of the origin of coordinates and id is the identity map, we instantly get the *normalization axiom* $\deg(\operatorname{id}, D, 0) = 1$.

One next proves that $\deg(F, D, q)$ is locally constant in the class of admissible generic C^1 maps with respect to the supremum norm. Finally, one proves that any admissible continuous map F is suitably approximated by an admissible generic map G . Hence, since the degree is locally constant, we may put $\deg(F, D, q) := \deg(G, D, q)$.

Let now $f : \overline{D} \rightarrow \mathbb{R}$ be a function of class C^2 . A point $p \in \overline{D}$ is called *critical* if the gradient $F = \nabla f$ vanishes at p and it is called *regular* otherwise. Hence, the critical points of f correspond to the zeros of F . The function f is called a *Morse function* if all of its critical points p are *non-degenerate*, that is, if the *Hessian* of f defined by $H_f = \det D^2 f$ does not vanish at p . Note that the Hessian of f is precisely the Jacobian of $F = \nabla f$. Thus, f is a Morse function if and only if its gradient F is generic for degree computation at $q = 0$. From now on, we assume that $q = 0$ and denote the degree of F on D with respect to 0 by $\deg(F, D)$ instead of $\deg(F, D, 0)$.

Given a Morse function f , the index of any critical point p , denoted by $\lambda(p)$, is the number of negative eigenvalues of $D^2 f(p)$. Thus $\operatorname{sgn} J_F(p) = (-1)^{\lambda(p)}$ and hence, the left-hand sides of the formulas (3) and (4) coincide. Here is a more visual, geometric way of introducing the Morse index. The Morse Lemma [12] says that there exist local C^2 coordinates originating at p such that, in those coordinates, f becomes a quadratic polynomial

$$(5) \quad f(x) = c + \sum_{i=1}^d \lambda_i x_i^2,$$

where $\lambda_i \in \{-1, 1\}$. Then $\lambda(p)$ is the number of indices i such that we have $\lambda_i = -1$. If $\lambda(p) = 0$, p is a local minimum and if $\lambda(p) = d$, p is a local maximum. The

intermediate values of $\lambda(p)$ correspond to simple (non-degenerate) saddles at p . The formula (4) instantly gives the following.

Proposition 2.1. *Let $F = \nabla f$ where f is the quadratic function in (5) and let $D = B^d$ be the unit ball in \mathbb{R}^d . Then*

$$\deg(F, D) = \lambda_1 \lambda_2 \cdots \lambda_d.$$

In particular, in \mathbb{R}^2 , $\deg(F, D) = 1$ when 0 is a local extremum (minimum or maximum), and $\deg(F, D) = -1$ when 0 is a saddle.

We may also deduce the formula (3) from the properties of degree in the case when $D = B^d$ is the open unit ball in \mathbb{R}^d . The condition that $F = \nabla f$ points inward on ∂D can be formulated in terms of the scalar product as $F(x) \cdot n(x) < 0$ for all $x \in \partial D$, where $n : \partial D \rightarrow \mathbb{R}^d$ is the outward normal vector field. When $D = B^d$, $\partial B^d = S^{d-1}$ is the unit sphere, $n(x) = x$ and we get the condition $F(x) \cdot x < 0$. Similarly, F points outward on ∂D if $F(x) \cdot n(x) > 0$ for all $x \in \partial D$, so if $D = B^d$, we get the condition $F(x) \cdot x > 0$. Thus, we want to prove the following result.

Theorem 2.2. *Let $f : \overline{B}^d \rightarrow \mathbb{R}$ be a Morse function satisfying the condition*

$$(6) \quad x \cdot \nabla f(x) < 0 \text{ for all } x \in S^{d-1}.$$

Let $\{p_1, p_2, \dots, p_n\}$ be the set of all critical points of f in B^d . Then the Euler-Maxwell-Morse formula (3) holds in \mathbb{R}^d .

Proof. We calculate the degree of $F = \nabla f$ on D with respect to $q = 0$. By (6), $F(x) \neq 0$ for $x \in S^{d-1}$ and hence, F is admissible. Since f is a Morse function, F is generic, so $\deg(F, D)$ is given by (4). Now, the degree of the linear map $-\text{id}$ given by $-\text{id}(x) = -x$ on B^d is $(-1)^d$. Hence, it remains to prove that

$$\deg(F, D) = \deg(-\text{id}, D).$$

For this, we will use the homotopy property of degree. Define $H : \overline{B}^d \times [0, 1] \rightarrow \mathbb{R}^d$ by

$$H(x, t) = (1 - t)F(x) - tx.$$

Then $H(x, 0) = F(x)$ and $H(x, 1) = -x$. It remains to show that H is admissible. Suppose, on the contrary, that there exists $t \in [0, 1]$ and $x \in S^{d-1}$ such that we have $H(x, t) = 0$. Since F and $-\text{id}$ are admissible, this is impossible that $t = 0, 1$ and we may assume that $0 < t < 1$. By (6), we get

$$0 = x \cdot H(x, t) = (1 - t)x \cdot F(x) - tx \cdot x < 0,$$

a contradiction. □

We wish to know if Proposition 2.1 remains true in original coordinates and if Theorem 2.2 can be extended to domains diffeomorphic to B^d . The Morse Lemma suggests that this is true but we need the following property of invariance of degree of ∇f under the change of coordinates in the domain of f . Its proof relies on lengthy but elementary vector calculus arguments. By a diffeomorphism between closed bounded regions of \mathbb{R}^d , we mean a homeomorphism extending to a diffeomorphism of their neighborhoods.

Lemma 2.3. *Let D be a bounded domain in \mathbb{R}^d with a C^2 boundary ∂D and let also $f : \overline{D} \rightarrow \mathbb{R}$ be a C^1 function. Suppose that there exists a C^2 diffeomorphism $\Phi : \overline{B}^d \rightarrow \overline{D}$ and put $g = f \circ \Phi$. Then we have the following:*

- (a) $F := \nabla f$ is admissible in D if and only if $G := \nabla g$ is admissible in B^d ,
- (b) If F is admissible, then $\deg(F, D) = \deg(G, B^d)$,
- (c) Moreover, F is inward (outward, respectively) at $\Phi(x) \in \partial D$ if and only if G is inward (outward, respectively) at $x \in S^{d-1}$.

Theorem 2.2 and Lemma 2.3 instantly imply the following.

Corollary 2.4. *Let \overline{D} be a region C^2 -diffeomorphic to a unit ball and $f : \overline{D} \rightarrow \mathbb{R}$ be a Morse function whose gradient is inward on ∂D . Let $\{p_1, p_2, \dots, p_n\}$ be the set of all critical points of f in B^d . Then the formula (3) holds in \mathbb{R}^d .*

The classical result of the Morse theory can now be deduced as an easy consequence of the previous statements.

Corollary 2.5. *Let $f : S^d \rightarrow \mathbb{R}$ be a Morse function and let $\{p_0, p_1, p_2, \dots, p_n\}$ be the set of all its critical points. Then the Euler-Maxwell-Morse formula*

$$(7) \quad \sum_{i=0}^n (-1)^{\lambda(p_i)} = 1 + (-1)^d$$

holds for S^d .

Proof. Since S^d is compact, f assumes its minimum at some point. Let it be p_0 . Let U be an open ball in S^d centered at p_0 , isolating it from other critical points, to which the Morse Lemma applies. The stereographic projection is a diffeomorphism of $S^d \setminus \{p_0\}$ onto \mathbb{R}^d which takes $S^d \setminus U$ to some closed ball $\overline{D} \subset \mathbb{R}^d$ centered at the origin. Since the Morse index of a minimum point is 1, it is enough to show that

$$\sum_{i=1}^n (-1)^{\lambda(p_i)} = (-1)^d.$$

This can be deduced from Corollary 2.4 applied for the composition of the inverse stereographic projection with the restriction of f to $S^d \setminus \{p_0\}$. \square

3. Local degree at a k -fold saddle

In this section, we study functions f in the plane \mathbb{R}^2 whose critical points are isolated but possibly degenerate.

3.1. A model of a k -fold saddle

The most commonly seen case of an isolated degenerate critical point is a monkey saddle. First, if

$$f_1(x, y) = x^2 - y^2,$$

then the origin of coordinates is a simple saddle of f . The two vectors $(1, 0)$ and $(-1, 0)$ define two *ascending directions* or, in terms of the topography of the surface

$h = f(x, y)$, two ridge lines emanating from the origin. Similarly, the vectors $(0, 1)$ and $(0, -1)$ define two descending directions or two ravine lines. The monkey saddle is, roughly speaking, a critical point which is the origin of three ridge lines separated by three ravine lines. By the Morse Lemma, this is of course impossible if $H_f(0, 0) \neq 0$. A simple model for the monkey saddle, illustrated by Figure 1, is given by

$$f_2(x, y) = x^3 - 3xy^2.$$

More generally, a k -fold saddle is a critical point originating $(k + 1)$ ridge lines separated by $(k + 1)$ ravine lines. A simple saddle is a 1-fold saddle and a monkey saddle is a 2-fold saddle.

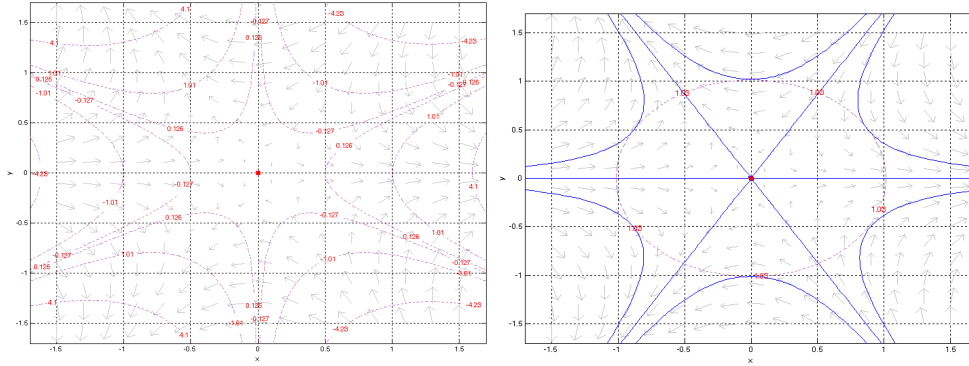


FIGURE 1. Left: Level lines and the gradient field for the monkey saddle. Right: trajectories, ridge lines and ravine lines. The displayed vector field permits tracing the winding of F as q moves on counter-clockwise on a circle described in Remark 3.2.

The most transparent formula for a function giving rise to a k -fold saddle is in terms of complex numbers. We identify \mathbb{R}^2 with the complex plane \mathbb{C} and we use the variable $z = x + iy = (x, y)^T$. Then,

$$z^2 = (x^2 - y^2) + 2ixy$$

and

$$z^3 = (x^3 - 3xy^2) + i(3x^2y - y^3),$$

that is, $f_1 = \Re e(z^2)$ and $f_2 = \Re e(z^3)$. Consider the function

$$(8) \quad f(z) = \Re e(z^{k+1}).$$

Note that f is positively homogeneous in the sense that $f(tx, ty) = t^{k+1}f(x, y)$ for all $t > 0$, so the ascending and descending directions are determined by the values of f on the circle S^1 given by $|z| = 1$. The maximum of f is 1, assumed at the roots of $z^{k+1} = 1$ and the minimum is -1 , assumed at the roots of $z^{k+1} = -1$. In polar coordinates, the ridge lines are the rays emanating from the origin at the angles $\theta_j = \frac{2\pi j}{k+1}$ and the ravine lines are rays at the angles $\theta_j + \frac{\pi}{k+1}$.

Theorem 3.1. *Let $F = \nabla f$, where f is given by (8). Then,*

$$\deg(F, B^2) = -k.$$

Proof. Consider the function $g : \mathbb{C} \rightarrow \mathbb{C}$, given by $g(z) = z^{k+1}$. Let $u(x, y)$ and $v(x, y)$ be the real and imaginary parts of g , respectively, so that $f(z) = u(x, y)$. Then $\nabla f = (u_x, u_y)^T$ where u_x and u_y denote the partial derivatives of u with respect to x and y , respectively. On one hand, using the Cauchy-Riemann equations, we get

$$\begin{aligned} g'(z) &= u_x(z) + iv_x(z) \\ &= u_x(z) - iu_y(z) \\ &\cong (u_x, -u_y)^T. \end{aligned}$$

On the other hand, $g'(z) = (k+1)z^k$, so $\overline{\nabla f} = (k+1)z^k$, where \bar{z} stands for the complex conjugate of z . It is known that the topological degree of a holomorphic function $h : \overline{B^2} \rightarrow \mathbb{C}$ which has no roots on S^1 is the number of roots of h in B^2 counting their multiplicity (see [10, Sec. 1.4]). In our case, $h = g'$ and this number is k .

Next, consider $G : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with $G(z) = (x, -y) = \bar{z}$. Then, $\overline{F} = G \circ F$. Since G is a linear isomorphism taking B^2 to itself, the multiplication Theorem [10, Sec. 2.3] implies that

$$\begin{aligned} \deg(\overline{F}, B^2) &= \deg(G \circ F, B^2, 0) \\ &= \deg(G, B^2, 0) \deg(F, B^2, 0). \end{aligned}$$

By definition,

$$\deg(G, B^2, 0) = \text{sgn det } \nabla G = -1,$$

so

$$k = \deg(\overline{F}, B^2) = -\deg(F, B^2, 0). \quad \square$$

Remark 3.2. Here is a geometric interpretation of the analytic proof provided above, based on the interpretation of degree as the winding number, and of ∇f as the vector pointing the direction of the steepest ascent of a height function f . We refer to Figure 1. We register the angle traced by the vector $F = \nabla f(q)$ attached to the origin 0, as the point $q = (x, y)$ moves counterclockwise on the unit circle. When $q = (1, 0)$ is on the first ridge line at $\theta_0 = 0$, F points in the same direction as $\vec{0}q$. When q moves counterclockwise towards the isoline at the angle $\frac{\pi}{2(k+1)}$, f decreases, so F rotates clockwise towards the left ridge line. When q reaches the first ravine line at the angle $\frac{\pi}{k+1}$, F points in the opposite direction of the angle $\frac{\pi}{k+1} - \pi$, and when the point is at the next ridge line $\theta_1 = \frac{2\pi}{k+1}$, F points again the same direction as $\vec{0}q$. Thus, the angle traced by F between the first two ridge lines is $\alpha = \frac{2\pi}{k+1} - 2\pi$. The same scenario repeats between any two consecutive ridge lines, so when q is back at $\theta = 2\pi$, the angle traced by F is $(k+1)\alpha = -2k\pi$. Thus, the winding number of F around 0 is $-k$.

3.2. Stable and unstable manifolds

In order to generalize Theorem 3.1, we need to give a more precise definition of a k -fold saddle of some function \tilde{f} . One possible way is to define it in similar terms as the Corollary 2.4, by requiring that there exists a diffeomorphism $\varphi : \overline{B^2} \rightarrow \overline{D}$ with $\varphi(0) = p$ and $\tilde{f} \circ \varphi = f$, where f is the model function given in (8). Then Lemma 2.3 can be used to conclude that the degree of \tilde{f} is $-k$. However, such a condition is hard to verify in practice.

In order to define the k -fold saddle for any C^2 function, we should first state what is meant by ridge lines and ravine lines in the discussion opening Section 3.1. This can be done in terms of the flow $\varphi(t, z)$ generated by the differential equation $\dot{z} = F(z)$, where $F = \nabla f$. Since $F(z)$ shows the direction of the fastest ascent, the ridge lines are formed by trajectories of φ “climbing up” from p as time increases, that is, converging to p as $t \rightarrow -\infty$. The points on those trajectories belong to the *unstable manifold* of p defined by

$$W^u(p) = \left\{ z \in M \mid \lim_{t \rightarrow -\infty} \varphi(t, z) = p \right\}.$$

The ravine lines are formed by trajectories of φ “sliding down” from p or, more precisely, converging to p as $t \rightarrow \infty$. The points on those trajectories belong to the *stable manifold* of p defined by

$$W^s(p) = \left\{ z \in M \mid \lim_{t \rightarrow \infty} \varphi(t, z) = p \right\}.$$

It is easy to check for the function in (8) that its unstable and stable manifolds are indeed the described rays θ_j and $\theta_j + \frac{\pi}{k+1}$, respectively.

Note that the terminology “manifold” for $W^u(p)$ and $W^s(p)$ is only justified if f is a Morse function. In this case, the dimensions of those manifolds are equal to the number of positive and negative eigenvalues of the hessian of f at p , respectively. Thus $\dim W^s(p) = \lambda(p)$ is the Morse index (In the literature, one often considers the reverse flow of the equation $\dot{x} = -\nabla f$, in order to make the potential of the gravitation field increasing along the trajectories as t increases. In this case, the roles of stable and unstable manifolds are reversed) of p . In a degenerate case, one may encounter, for example, W^u containing a cone of ridge lines ascending from p not separated by ravine lines. In order to handle such cases, we introduce the following sets. Let N be an *isolating neighborhood* of p , that is, a closed neighborhood of p which does not contain other critical points. We put

$$(9) \quad N_p = \{z \in N \mid f(z) > f(p)\},$$

$$(10) \quad N_n = \{z \in N \mid f(z) < f(p)\},$$

and

$$(11) \quad N_z = \{z \in N \setminus \{p\} \mid f(z) = f(p)\}.$$

In the case of an isolated minimum, $N_p = N \setminus \{p\}$ and $N_n = \emptyset$. For an isolated maximum, the roles of N_p and N_n are exchanged. From the isolation condition and the hypothesis that f is of class C^2 , it follows that $N_z \cup \{p\} = \overline{N}_p \cap \overline{N}_n$ and that it consists of isolines. For our model (8) of a k -saddle, the connected components of N_p are cones given by

$$|\theta - \theta_j| < \frac{\pi}{2(k+1)}$$

and those of N_n are given by

$$\left| \theta - \theta_j - \frac{\pi}{k+1} \right| < \frac{\pi}{2(k+1)}.$$

The set N_z is given by $z^{k+1} = \pm i$ and consists of rays at the angles $\theta_j \pm \frac{\pi}{2(k+1)}$.

Definition 3.3. A k -fold saddle of a C^2 function f is a critical point p of f whose unstable and stable manifolds contain $(k + 1)$ ridge lines S_0, S_1, \dots, S_k and $(k + 1)$ ravine lines V_0, V_1, \dots, V_k which satisfy the following conditions:

(i) The sets $\mathcal{S} = \{S_0, S_1, \dots, S_k\}$ and $\mathcal{V} = \{V_0, V_1, \dots, V_k\}$ are *interlaced* in the following sense. The set $N \setminus (\{p\} \cup \mathcal{S} \cup \mathcal{V})$ has $2(k + 1)$ connected components called *wedges*. Each wedge is bounded in $N \setminus \{p\}$ by one ridge line and one ravine line.

(ii) Each connected component of N_p contains one ridge line from \mathcal{S} and each connected component of N_n contains one ravine line from \mathcal{V} .

This definition allows ordering ridge lines and ravine lines as

$$(12) \quad (S_0, V_0, S_1, V_1, \dots, S_k, V_k)$$

in a circle around $\{p\}$, so that the two consecutive elements in this sequence (where S_0 follows V_k) bound a wedge.

3.3. Decomposition of a k -fold saddle

The degree theory assures that a degenerate critical point p can be replaced by a number of non-degenerate ones by a small perturbation of the vector field which does not change the global degree. More precisely, the vector field F is replaced by a shifted vector field $F(x) - q$ in a small region D around p . The measure theoretical arguments imply that there exist arbitrary small values of q for which the zeros of the perturbed field are non-degenerated. This is illustrated in Figure 2. It is however not always easy to explicitly determine q and analytically calculate the local degrees at the new critical points. The goal of this section is to establish a combinatorial graph-theoretical procedure for the decomposition of k -fold saddles into k -simple saddles, without relying on the smoothness and transversality assumptions. The main idea comes from Edelsbrunner *et al.*, see [5]. We show that the decomposition preserves the degree on D . The described procedure is useful for understanding and constructing the Morse connections graph described below.

We introduce first some terminology from [1] related to *Morse connections graph*. This is a graph whose nodes are critical points of the flow (minimum, maximum or k -fold saddle). Each node is connected to other nodes using oriented edges of the graph. To a pair of critical points (p, q) , we associate an edge called an *ascending direction* if there is a trajectory converging to p as $t \rightarrow -\infty$ and to q as $t \rightarrow \infty$, or equivalently, if $W^s(p) \cap W^u(q) \neq \emptyset$. It is called *descending direction* if there is a reverse trajectory. For example, if p is a minimum, then all the edges attached to p are ascending directions. Similarly, if p is a maximum, then all the edges attached to p are descending directions. However, if p is a k -fold saddle, there are $(k+1)$ ascending directions which correspond to ridge lines and $(k+1)$ descending directions which correspond to ravine lines.

In practical applications to imaging, one doesn't work in a compact manifold but in a bounded rectangular region of a plane, so ridge lines and ravine lines may escape the boundary picture. In this case, we assume that f is decreasing towards the boundary so that the escaping lines can be regarded as lines connecting a given critical point to the point compactifying the plane to the sphere, where f assumes its global minimum.

Now, we consider a k -fold saddle p , an isolating neighborhood N of p , and a portion of the Morse connections graph corresponding to ridge lines and ravine lines which

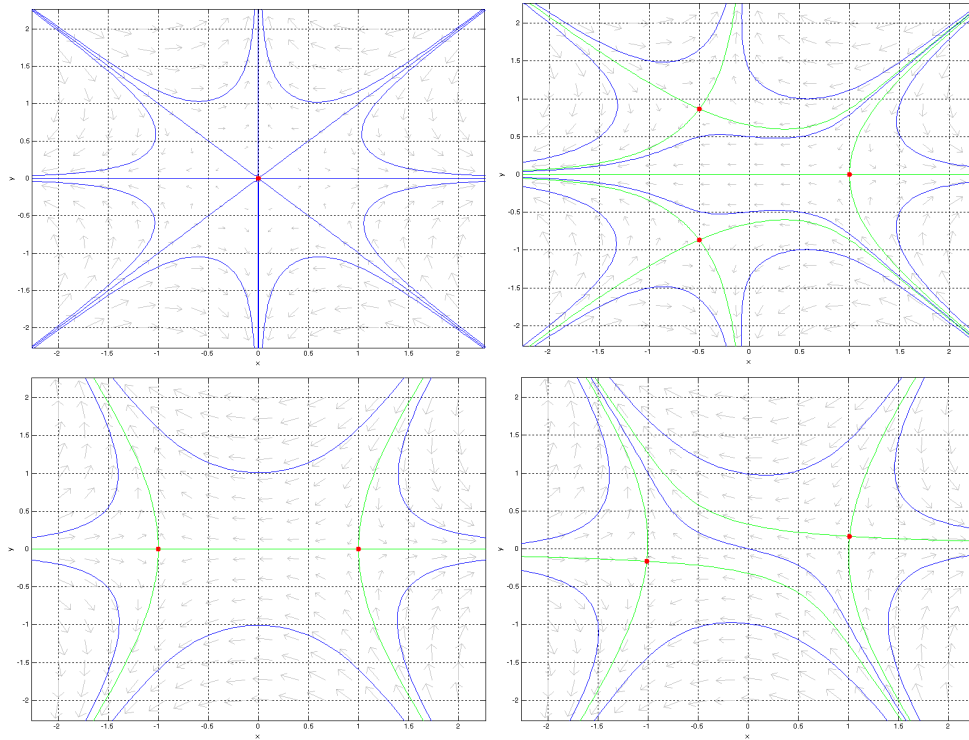


FIGURE 2. Above: the 3-fold saddle and its unfolding to two simple saddles by a shift of F in the x -direction. Below: two different decompositions of the monkey saddle. In the phase portrait on the left, a shift of F along the x -axis is applied. The ridge line of one saddle and the ravine line of another produce a connecting trajectory between the two, as described in Algorithm 3.5. On the right, a small shift in the y -direction makes those two lines separate and escape outside of the picture.

leave or enter N . Thus, this part of the graph consists of exactly $(k + 1)$ ascending directions or ridge lines $\mathcal{S} = \{S_0, S_1, \dots, S_k\}$, and $(k + 1)$ descending directions or ravine lines $\mathcal{V} = \{V_0, V_1, \dots, V_k\}$. We order ascending and descending directions in the abstract graph such as $(S_0, V_0, S_1, V_1, \dots, S_k, V_k)$ is the ordered set (12).

Definition 3.4. Let \mathcal{V}_i be a set of i descending directions and \mathcal{S}_j be a set of j ascending directions. Then \mathcal{V}_i and \mathcal{S}_j are said to be *interlaced*, see Figure 3, if we can alternate the elements of \mathcal{V}_i with those of \mathcal{S}_j such that the obtained sequence is a subsequence of (12) consisting to $i + 1$ consecutive elements, where we set $V_{k+1} = V_0$ and $S_{k+1} = S_0$. Note that, necessarily, $|i - j| \leq 1$.

We are now ready to present the procedure for the decomposition of a k -fold saddle p , see Figure 4, into two saddles p_i, p_j of multiplicity $1 \leq i, j \leq k$ with $i + j = k$.

Algorithm 3.5 (Decomposition procedure). Let p be a k -fold saddle and N be an isolating neighborhood of p .

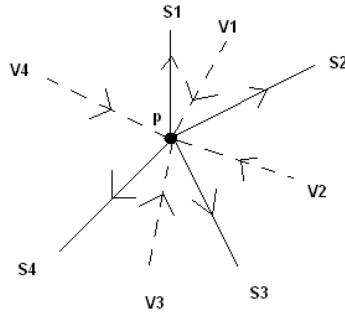


FIGURE 3. $\mathcal{V}_3 = \{V1, V2, V3\}$ and $\mathcal{S}_4 = \{S1, S2, S3, S4\}$ are interlaced.

(a) Choose arbitrarily a set \mathcal{S}_{i+1} of $i + 1$ ascending directions and a set \mathcal{V}_i of i descending directions originating at p such that \mathcal{V}_i and \mathcal{S}_{i+1} are interlaced. At the end of this step, we have the critical point p , \mathcal{S}_{i+1} and \mathcal{V}_i , see Figure 4(left part).

(b) As there are $i + 1$ ridge lines and i ravine lines originating at p , there exist two ridge lines bounding the same wedge. Modify the flow in $\text{Int } N$ by creating a ravine line inside this wedge, merging from p and ending at a new critical point $p_j \in \text{Int } N$. This new ravine line for p is a ridge line for p_j , see Figure 4(middle part).

(c) Attach at p_j the remaining $k - i = j$ ascending directions and the $k - i + 1$ ($= j + 1$) descending directions with the same ordering, see Figure 4(right part). At the end of this step, p is a i -fold saddle and p_j is a j -fold saddle.

(d) Repeat the step (a) for p_i and p_j and re-initialize k to, respectively, i and j .

At the end of this process, a k -fold saddle p is decomposed into k simple saddles.

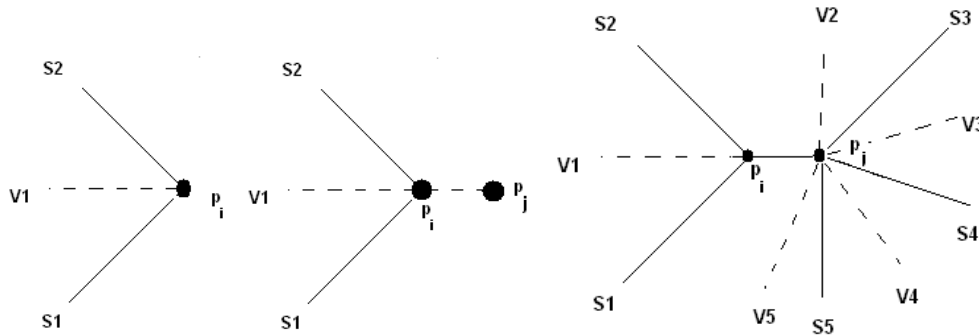


FIGURE 4. Left: isolating $i + 1 = 2$ ascending directions and i descending directions originated at p . Middle: creating a ridge line from p to p_j . Right: completing the graph; p is a 1-fold saddle (a simple saddle) and p_j is a 3-fold saddle.

The choices of edges to decompose in Algorithm 3.5 are not unique but they all lead to the same result on the sum of the local degrees.

Theorem 3.6. *Let p_1, p_2, \dots, p_k be k simple saddles in N produced from a k -fold saddle p by Algorithm 3.5. Let $N_i \subset N$, $i = 1, \dots, k$, be isolating neighborhoods for p_1, p_2, \dots, p_k , respectively, for the modified flow G . Then*

$$\deg(F', N) = \sum_{i=1}^k \deg(F', N_i).$$

Proof. By standard analytical arguments on smooth extensions of functions, it is possible to modify the surface $u = f(x, y)$ inside N so that the flow lines are modified as described in the algorithm. This is done without modifying it on ∂N . By the homotopy axiom,

$$\deg(F, N) = \deg(G, N).$$

There are no new critical points created in N other than p_1, p_2, \dots, p_k . Thus, the conclusion follows from the additivity axiom. \square

We would like to use Proposition 2.1 and Theorem 3.6 to conclude that

$$(13) \quad \deg(F, N) = \sum_{i=1}^k \deg(G, N_i) = (-1) + (-1) + \dots + (-1) = -k.$$

Unfortunately, a simple saddle may possibly have a null Hessian, so we are not ready yet to make use of Proposition 2.1. The conclusion on (13) could only be derived after the classification of degenerate critical points, which is the main goal of the next section.

Remark 3.7. The decomposition produced by Algorithm 3.5 creates an edge in the Morse connections graph corresponding to a connection between two new saddles. Those connections may not be desirable in the Morse theory. The two phase portraits in the bottom of Figure 2 show a possibility of modifying the algorithm to split that connection.

4. Classification of critical points

4.1. Isolating blocks

We recall here a definition of an isolating block from [17] adapted to the context of our paper. The general Wilson and Yorke's definition is given for an isolated invariant set of a flow in \mathbb{R}^n , but we restrict it to an isolated critical point p of a C^2 function f in \mathbb{R}^2 . The hypothesis that f is C^2 could be relaxed by assuming that it is C^1 and its gradient is locally Lipschitz, so the associated flow φ is well defined.

A *manifold with corners* in \mathbb{R}^2 is a closed bounded region N whose boundary is either smooth (i.e., of class C^2) or it consists of a finite number of smooth arcs connected at endpoints, called *corners*, where the smoothness fails.

If A is an open smooth arc on the boundary of N , then $n : A \rightarrow \mathbb{R}^2$ denotes the normal vector field on A pointing outward of N . We say that a C^1 vector field F is *strongly inward* (*strongly outward*, respectively) on A if $F \neq 0$ on A and there is a constant $\delta > 0$ such that $F/\|F\| \cdot n < \delta < 0$ ($F/\|F\| \cdot n > \delta > 0$, respectively).

Definition 4.1. Let p be an isolated zero of a vector field F . An isolating neighborhood N of p is called an *isolating block* if it is a manifold with corners homeomorphic to a closed unit disc in \mathbb{R}^2 and satisfy the following conditions:

(i) If A is a smooth arc of ∂N , then F is either strongly inward or strongly outward on A .

(ii) If $x \in \partial N$ is a corner point, then the orbit of the flow φ of F bounces off at x in the sense that

$$\varphi(\mathbb{R}, x) \cap N = \{x\}.$$

The closed union of the arcs at which F is outward is called the *exit set* of N and is denoted by N^- .

The purpose of using manifolds with corners rather than smooth manifolds for isolating blocks is that they are stable, in the sense that their inward and outward arcs are stable under small perturbations of the vector field F . Here is a standard example from the Conley index theory.

Example 4.2. Consider $f(x, y) = x^2 - y^2$. Its gradient field is given by the equation $F(x, y) = \nabla f(x, y) = 2(x, -y)$ and the flow trajectories are branches of hyperbolas $xy = c$. The square $N = [-1, 1]^2$ is an isolating block of F . The vector field is inward on the upper and lower open edges and outward on the closed left and right edges. The absolute value of the angle between F and each edge, counted at points of ∂N , takes the maximum $\pi/4$ at the vertices of the square. Since $\|F\| \geq 2$ on ∂N , for a sufficiently small perturbation G of F and any $q \in \partial N$, the angle between $G(x, y)$ and $F(x, y)$ is less than $\pi/4$ for all $(x, y) \in \partial N$. Hence, the conditions (i) and (ii) in Definition 4.1 remain valid for G .

Example 4.3. Consider the function given by (8) providing the model for a k -fold saddle. Let P be a closed convex equilateral polygon with $2(k+1)$ sides centered at the origin and whose vertices are on the rays $\theta = \frac{\pi}{2(k+1)} + \frac{\pi j}{k+1}$, $j = 0, \dots, 2k+1$. Then P is an isolating block of the origin.

As the above examples suggest, it is useful to state the following polyhedral version of Lemma 2.3. Its proof is analogous.

Lemma 4.4. Let N be a manifold with corners in \mathbb{R}^2 and let $f : N \rightarrow \mathbb{R}$ be a C^1 function. Suppose that there exists a C^2 diffeomorphism $\Phi : P \rightarrow N$, where P is a convex polyhedron and put $g = f \circ \Phi$. Then we have:

- (a) $F := \nabla f$ is admissible in $\text{Int } N$ if and only if $G := \nabla g$ is admissible in $\text{Int } P$,
- (b) If F is admissible, then $\deg(F, \text{Int } N) = \deg(G, \text{Int } P)$,

(c) Moreover, F is strongly inward (strongly outward, respectively) on smooth arcs of ∂D if and only if G is strongly inward (strongly outward, respectively) on the corresponding edges of P .

Lemma 4.5. Suppose that two C^1 fields F and G share an isolating block N and the same inward and outward arcs of ∂N . Then $\deg(F, \text{Int } N) = \deg(G, \text{Int } N)$.

Proof. One instantly verifies that the homotopy

$$H(x, t) = (1 - t) F(x) + t G(x)$$

satisfies the same strong inward and outward conditions as F and G . This implies that $H(x, t) \neq 0$ for all $t \in [0, 1]$ and all $x \in \partial N$. \square

Lemma 4.6. *Let $N \subset \mathbb{R}^2$ be an isolating block for p and $F = \nabla f$. If its exit set N^- is empty or if it is the whole ∂N , then $\deg(F, \text{Int } N) = 1$. Otherwise, N^- is disconnected. Let $k + 1$ be the number of its connected components. Then $\deg(F, \text{Int } N) = -k$.*

Proof. Since an isolating block of a critical point is homeomorphic to the disc \overline{B} , it must be either diffeomorphic to \overline{B} or to a closed convex polyhedron P . When either $N^- = \emptyset$ or $N^- = \partial N$, we get N which is diffeomorphic to \overline{B} .

In the first case, F is inward on ∂N . By Lemma 2.3, Lemma 4.5 and Proposition 2.1, we have

$$\deg(F, \text{Int } N) = \deg(-\text{id}, B) = 1.$$

By the same arguments, if $N^- = \partial N$, then F is outward on ∂N and we get

$$\deg(F, \text{Int } N) = \deg(\text{id}, B) = 1.$$

Suppose now that N^- is disconnected. The condition (ii) in Definition 4.1 and the continuity of the flow imply that if two smooth arcs of ∂N meet at a corner point, then F is strongly inward on one of them and strongly outward on the other. Therefore, the inward arcs are interlaced with outward arcs as the ridge lines and ravine lines in Definition 3.3. Since the arcs complete a circle, the number of inward arcs is the same as the number of outward arcs and is equal to $(k + 1)$. In particular, the convex polyhedron P to which N is homeomorphic has $2(k + 1)$ edges. By Lemma 4.4, Lemma 4.5, Theorem 3.1 and Example 4.3,

$$\deg(F, \text{Int } N) = \deg(\overline{z}^k, P) = -k. \quad \square$$

We note that Lemma 4.5 provides a link between the local degree at p and the Conley index [4] of the singleton $\{p\}$, which is the pointed homotopy type of the pair (N, N^-) . In the case where N^- is disconnected with $(k + 1)$ connected components, (N, N^-) has the homotopy type of the wedge of k circles.

4.2. Extension of the Euler-Maxwell formula

We are now ready to prove the main results of this section.

Theorem 4.7 (Classification of isolated critical points). *Let p be an isolated critical point of a C^2 function $f : \overline{D} \rightarrow \mathbb{R}$.*

- (i) *Then any isolating neighborhood of p contains an isolating block N of p ,*
- (ii) *Moreover, p is either a maximum point, a minimum point or a k -fold saddle,*
- (iii) *Finally, $\deg(\nabla f, \text{Int } N)$ is 1 in the first two cases and $-k$ in the last one.*

Proof. Since a gradient field has no periodic orbits, the singleton $\{p\}$ is an isolated invariant set in the sense of [17, Definition 1.1]. By [17, Theorem 2.5], any isolating neighborhood of p contains an isolating block in the sense of [17, Definition 1.2]. It follows from the proof of [17, Theorem 2.5] and from [16, Corollary 3.5] that one can construct a Wilson and Yorke's isolating block N which is deformable to $\{p\}$. It is known that a manifold homotopic to a disc is also homeomorphic to \overline{B} , hence it is an isolating block in the sense of Definition 4.1. One can also derive this conclusion from [7, Remark 3.1]. This proves (i). Then (iii) follows from Lemma 4.6.

We now prove (ii). If N^- is empty, then ∇f is inward on ∂N , so f must assume a maximum in N . Since there are no other critical points, that maximum is assumed at p . If $N^- = \partial N$, then ∇f is outward on ∂N so, by the same argument, f has a minimum on N at p .

Consider the remaining case when N^- is disconnected with $(k + 1)$ connected components. We already showed in the proof of Lemma 4.6 that the inward arcs are interlaced with outward arcs and their numbers are both equal to $(k + 1)$. Moreover, it follows again from [17, Theorem 2.5] and from [16, Corollary 3.5] that any outward arc deforms to its intersection with $W^u(p)$ and any inward arc deforms to its intersection with $W^s(p)$. In particular, those intersections are non-empty. This means that each outward arc contains at least one ridge line and each inward arc contains at least one ravine line. This conclusion can also be derived from cohomological description of isolating blocks in [7]. Thus, we proved that p is exactly the k -fold saddle according to Definition 3.3. \square

Remark 4.8. In spaces of higher dimensions, namely in \mathbb{R}^4 and \mathbb{R}^5 , the proof of the fact that a Wilson and Yorke's isolating block of an isolated critical point is homeomorphic to the unit ball relies on the famous Poincaré conjecture, proved just a few years ago.

Theorem 4.9 (Maxwell formula for degenerate critical points). *Let \overline{D} be a region in \mathbb{R}^2 , C^2 -diffeomorphic to the closed unit ball or to a closed convex polyhedron, and $f : \overline{D} \rightarrow \mathbb{R}$ a C^2 function whose gradient ∇f is inward on ∂D . Suppose that all critical points of f are isolated. Then there are finitely many of them and they are local minima, local maxima or extended k -fold saddles. Moreover, we have the formula*

$$\# \min - \Sigma(k \cdot \#(k - \text{saddles})) + \# \max = 1.$$

Proof. By the same arguments as those in the proof of Theorem 2.2 and Corollary 2.4,

$$\deg(\nabla f, D) = \deg(-\text{id}, B^2) = (-1)^2 = 1.$$

Since \overline{D} is compact and the critical points of f are isolated, there are finitely many of them. Let $\{p_i\}_{i=1,2,\dots,n}$ be the corresponding set of critical points. By Theorem 4.7 (i), each point p_i admits an isolating block N_i . By the additivity axiom,

$$1 = \deg(F, D) = \sum_{i=1}^n \deg(F, \text{Int } N_i).$$

The conclusion follows from Theorem 4.7 parts (ii) and (iii). \square

By the same arguments as in the proof of Corollary 2.5, we get the following.

Corollary 4.10. *Let $f : S^d \rightarrow \mathbb{R}$ be a C^2 function. Suppose that all critical points are isolated. Then there are finitely many of them and they are local minima, local maxima or extended k -fold saddles. Moreover, we have the formula*

$$\# \min - \Sigma(k \cdot \#(k - \text{saddles})) + \# \max = 2.$$

5. Conclusion

As we mentioned in the introduction, the main motivation for this paper is to improve existing models for analysis of digital images, where functions are not defined on points in \mathbb{R}^2 but pixels in a finite lattice. Our first numerical experiments showed that we need to relax the hypothesis that critical points are isolated. Note that a typical example from mathematical analysis is a critical point p which is a limit of a sequence of other critical points p_i . Such cases are not really of concern in the digital image analysis, because the sets of pixels are finite. However, flat critical regions are common in digital images. In analysis of a height function in topography, for example, one cares about flat regions such as bottoms of lakes, flat mountain tops, or volcano craters, which are extremum regions; and about long sand bars at sea shore, which are saddle regions. An algorithm detecting and classifying critical regions is produced in [1] but it requires improvements, especially with regard to the concept of topological boundary in the digital setting, and of identification of k -saddle regions. Understanding saddle regions is crucial for the construction of isolines, because these are places where smooth continuation techniques fail. Also the model of discrete multivalued dynamical system used for the Morse connections graph Algorithm needs to be rethought in terms of the degree theory for multivalued maps.

Another obvious direction for future studies is to provide an analogous analysis of critical points and regions for dimensions 3 and higher. An initial work on this topic is [2]. The analysis of saddle pixels and saddle regions is more difficult in high dimensions because the numbers of connected components of inward and outward portions of an isolating block is not sufficient to distinguish between a saddle and an extremum or between two different types of saddles. Thus, one has to search for more advanced topological tools.

Acknowledgements. S. Derivière was supported by a Tomlinson Visiting Scholarship of Bishop's University, T. Kaczynski was supported by a NSERC Discovery Grant and P.-O. Vallerand was supported by a NSERC USRA.

REFERENCES

- [1] M. Allili, D. Corriveau, S. Derivière, T. Kaczynski and A. Trahan, *Discrete dynamical system framework for construction of connections between critical regions in lattice height data*, J. Math. Imaging Vision **28**, 2007, no. 2, 99–111.
- [2] M. Allili, D. Corriveau, S. Derivière and T. Kaczynski, *Detecting critical regions in multi-dimensional data sets*, preprint.

- [3] L.E.J. Brouwer, *Über Abbildung von Mannigfaltigkeiten*, Math. Ann. **71**, 1912, no. 4, 97–115.
- [4] C. Conley, *Isolated invariant sets and the Morse index*, CBMS Regional Conference Series in Mathematics, **38**, AMS, Providence, 1978, iii+89 pp.
- [5] H. Edelsbrunner, J. Harer and A. Zomorodian, *Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds*, 17th ACM Symposium on Computational Geometry, Medford, MA, 2001.
- [6] R. Forman, *Morse theory for cell complexes*, Adv. Math. **134**, 1998, no. 1, 90–145.
- [7] A. Gierzkiewicz and K. Wójcik, *On the cohomology of an isolating block and its invariant part*, Topol. Methods in Nonlinear Anal. **32**, 2008, 313–326.
- [8] F.P. Greenleaf, *Introduction to complex variables*, W.B. Saunders Co., Philadelphia-London-Toronto, Ont., 1972, xii+588 pp.
- [9] J. Leray and J. Schauder, *Topologie et équations fonctionnelles*, Ann. Sci. Ecole Norm. Sup. **3**, 1934, no. 51, 45–78.
- [10] N. G. Lloyd, *Degree theory*, Cambridge Tracts in Mathematics, **73**, Cambridge University Press, Great Britain, 1978, vi+172 pp.
- [11] W.S. Massey, *A basic course in algebraic topology*, Graduate Texts in Mathematics, **127**, Springer-Verlag, New-York, 1991, xvi+428 pp.
- [12] Y. Matsumoto, *An introduction to Morse theory*, Translations of Mathematical Monographs, **208**, Iwanami Series in Modern Mathematics, AMS, Providence, 2002, xiv+219 pp.
- [13] J.C. Maxwell, *On hills and dales*, The London, Edinburg and Dublin Philosophical Magazine and Journal of Science, **40**, 1870, 421–425.
- [14] M. Nagumo, *A theory of degree of mappings based on infinitesimal analysis*, Amer. J. Math. **73**, 1951, 485–496.
- [15] R. Srzednicki, *On periodic solutions of planar polynomial differential equations with periodic coefficients*, J. Differential Equations **114**, 1994, no. 1, 77–100.
- [16] F.W. Wilson, *The structure of the level surfaces of a Lyapunov function*, J. Differential Equations **3**, 1967, 323–329.
- [17] F.W. Wilson and J.A. Yorke, *Lyapunov functions and isolating blocks*, J. Differential Equations **13**, 1973, 106–123.

S. DERIVIÈRE, DÉP. DE MATH., U. DE SHERBROOKE, 2500 BOUL. DE L'UNIVERSITÉ, SHERBROOKE, QC, CANADA, J1K 2R1.
s.deriviere@usherbrooke.ca

T. KACZYNSKI, DÉP. DE MATH., U. DE SHERBROOKE, 2500 BOUL. DE L'UNIVERSITÉ, SHERBROOKE, QC, CANADA, J1K 2R1.
t.kaczynski@usherbrooke.ca

P.-O. VALLERAND-BEAUDRY, DÉP. DE MATH., U. DE SHERBROOKE, 2500 BOUL. DE L'UNIVERSITÉ, SHERBROOKE, QC, CANADA, J1K 2R1.
p-o.vallerand@usherbrooke.ca