Molecular Shape Analysis based uponthe Morse-Smale Complexand the Connolly Function

Frédéric Cazals, F. Chazal, T. Lewiner

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Molecular Shape Analysis based upon the Morse-Smale Complex and the Connolly Function

F. Cazals — F. Chazal — T. Lewiner

N° 4673

Decembre 2002

THÈME 2

ISSN 0249-6399 ISRN INRIA/RR--4673--FR+ENG

INRIA
SOPHIA ANTIPOLIS

Molecular Shape Analysis based upon the Morse-Smale Complex and the Connolly Function

F. Cazals * , F. Chazal † , T. Lewiner ‡

Thème 2 — Génie logiciel
et calcul symbolique
Projets Prisme

Rapport de recherche n° 4673 — Decembre 2002 — 30 pages

Abstract: Docking is the process by which two or several molecules form a complex. Docking involves the geometry of the molecular surfaces, as well as chemical and energetical considerations. In the mid-eighties, Connolly proposed a docking algorithm matching surface *knobs* with surface *depressions*. Knobs and depressions refer to the extrema of the *Connolly* function, which is defined as follows. Given a surface $\mathcal{M}$ bounding a three-dimensional domain X, and a sphere S centered at a point p of $\mathcal{M}$, the Connolly function is equal to the solid angle of the portion of S contained within X.

We recast the notions of knob and depression of the Connolly function in the framework of Morse theory for functions defined over two-dimensional manifolds. First, we study the critical points of the Connolly function for smooth surfaces. Second, we provide an efficient algorithm for computing the Connolly function over a triangulated surface. Third, we introduce a Morse-Smale decomposition based on Forman's discrete Morse theory, and provide an $O(n \log n)$ algorithm to construct it. This decomposition induces a partition of the surface into regions of homogeneous flow, and provides an elegant way to relate local quantities to global ones —from critical points to Euler's characteristic of the surface. Fourth, we apply this Morse-Smale decomposition to the discrete gradient vector field induced by Connolly's function, and present experimental results for several mesh models.

Key-words: Morse Theory, Differential Geometry, Meshes, Shape Matching, Docking.

* INRIA Sophia-Antipolis, 2004 route des Lucioles, F-06902 Sophia-Antipolis; Frederic.Cazals@inria.fr
† Laboratoire de Topologie, UMR CNRS 5584 UFR des Sciences et Techniques, Université de Bourgogne; 9, avenue Alain Savary - B.P. 47 870 - F-21078 Dijon cedex; fchazal@u-bourgogne.fr
‡ INRIA Sophia-Antipolis, 2004 route des Lucioles, F-06902 Sophia-Antipolis; Thomas.Lewiner@inria.fr

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Sur une Analyse des Formes Moléculaires basée sur le Complexe de Morse-Smale et la Fonction de Connolly

Résumé : Le docking fait référence à l'ensemble des mécanismes intervenant lorsque deux ou plusieurs molécules forment un complexe moléculaire. Ces mécanismes font intervenir, outre la géométrie des molécules, leurs propriétés physico-chimiques. Dans le registre géométrique du docking, Connolly a proposé au milieu des années 80 un algorithme associant certains *creux* à certaines *bosses* des molécules impliquées. Par creux et bosses on entend plus précisément les extrema locaux de la fonction de Connolly qui se définit comme suit. Soit $\mathcal{M}$ une surface de $\mathbb{R}^3$ délimitant un volume X, et soit S une sphère centrée en un point p de S. La fonction de Connolly est définie comme l'angle solide associé à la portion de S contenue dans X.

Ce travail fait le lien entre l'algorithme de Connolly et la théorie de Morse, et plus précisément les fonction bivariées définies sur des surfaces de $\mathbb{R}^3$. Les contributions sont de quatre ordres. Premièrement, on s'intéresse à la nature des points critiques de la fonction de Connolly pour des surfaces lisses. Deuxièmement, on décrit un algorithme efficace de calcul de la fonction de Connolly pour les surfaces triangulées. Troisièmement, en s'appuyant sur la théorie de Morse discrète, on introduit la notion de diagramme discret de Morse-Smale. Étant donnée une fonction de Morse —ou de façon équivalente un champs de gradients, ce diagramme induit une partition de la surface en régions de flot homogène, et établit un lien entre certaines propriétés locales et globales —des points critiques à la caractéristique d'Euler. On décrit également un algorithme de calcul de ce complexe en $O(n \log n)$. Des résultats expérimentaux sur plusieurs modèles concluent le travail.

Mots-clés : Théorie de Morse, Géométrie Différentielle, Triangulations de surfaces, Analyse des formes, Docking.

Contents

≪Geometry is not everything, but it is the most fundamental thing ≫M. Connolly, [8, p1231]

1 Introduction

1.1 Docking

Docking is the process by which two or several molecules form a complex. These molecular complexes either involve maro-molecules only —DNA, RNA, proteins—, or macro-molecules together with small molecules —typically drugs. The computational study of docking has long been recognized as a major tool to understand biological reactions and to design drugs. Although (bio-)chemist may argue against Connolly that physical and chemical properties play a more prominent role than geometry, geometry is often used as the foundation for a unified approach including chemical and energetic considerations.

When one thinks of docking in terms of geometry, it is appealing to believe that shape complementarity materializes through the pairing of *knobs* and *depressions* of the molecular surfaces —see Figure 1. Knobs and depressions intuitively correspond to local maxima and minima on the surface. But talking of maxima and minima subsumes that (i)a function is defined on the surface (ii)the outputs of the function can be compared —typically a real-valued function. Extrema are then special cases of critical points of the function. Figure 1 features such a function which is a height function. Height functions, however, are obviously poor candidates for comparing extrema since these depend on the z direction chosen. Following the idea that critical points should be those of a function invariant under rigid motions, Connolly [7, 8] introduced a quantity encoding the portion of a probe sphere centered on the surface and contained within the volume bounded by the surface.

This paper is dedicated to a careful study of Connolly's function. The first contribution deals with smooth curves and surfaces, while the three others are concerned with triangulated surfaces. Before presenting them, we define precisely Connolly's function and review some relevant literature.

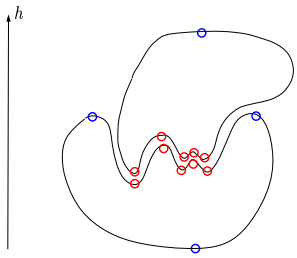


Figure 1: Shape complementarity and critical points

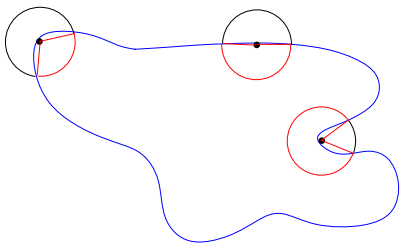


Figure 2: The Connolly function for a curve

1.2 Probing a surface with Connolly's function

We begin with the precise definition of Connolly's function [7, 8]. Let X be a domain of $\mathbb{R}^3$ whose boundary $\mathcal{M}$ is a surface. Let $S(p,r)$ be a sphere of radius r centered at a point p of the surface $\mathcal{M}$. Assume that r is small enough so that $\mathcal{M}$ cuts S into two two-dimensional topological disks D_i and D_o —the subscripts respectively stand for inside and outside X. The Connolly function is defined as the solid angle associated with D_i, that is:

$$C(p,r) = \frac{SurfaceArea(D_i)}{r^2}. \tag{1}$$

The Connolly function therefore maps a point of the surface to a real value within the interval $]0, 4\pi[$. Notice that a similar definition holds for curves using a circle instead of a sphere. The intuition behind Connolly's function is illustrated on Figure 2 with a curve for which three probing circles have been depicted. From left to right, they respectively correspond to a peak, a flat area, and a depression of the curve. Over these three configurations, the Connolly function takes an increasing value, and the initial incentive for defining such a function was "the sharper the peak, the smaller the value; the more depressed the crevice, the higher the value".

The previous definition does not assume anything about the geometry of $\mathcal{M}$. For example, domain X could be defined by a set of equations —e.g. a semi-algebraic set, or could have a more combinatorial description —e.g. a union of Van der Waals balls representing the atoms of a molecule. We shall, in this paper, carry out a general study of the Connolly function and in particular of its critical points for (i)smooth surfaces (ii)triangulated surfaces. Although the definition of Connolly's function stems from a focus on molecular modeling, our investigations are independent from the physical meaning of the surface studied. The last section however, deals with molecular solvent-accessible triangulated surfaces [5].

1.3 Previous work

Our review of previous work discusses in turn (i)contributions dealing with the definition of a function on a molecular surface (ii)Morse Theory (iii)contributions concerned with the global characterization of molecular features. We focus on the geometric aspects and do not discuss in depth the underlying implications for applications. See [30, 16] for recent references on docking.

Functions defined over molecular surfaces. The Connolly function was first introduced to study docking [7, 8]. The initial idea was to seek local extrema of the function and to match maxima with minima. For a real valued function of two variables, local maxima and minima form, together with saddle points, the so-called critical points. More precisely, any function of two variables whose Hessian is non degenerate at a critical point can be re-written, up to a diffeomorphism, as a local maximum, minimum, or saddle [3]. Connolly, however, points out that saddle points are difficult to deal with [8, p1236] $\ll$ saddle points

have not proven to be useful in docking because too many critical points over-complicate matters≫. Refinements of Connolly's algorithm [25, 24] do not use the information provided by saddle points neither.

An interesting docking algorithm implicitly using critical points is described in [4]. Star-shaped molecules can be modeled using spherical harmonics. The molecular surface is then defined as the graph of a function of two variables $r(\theta, \varphi)$. The contour map of this height function is used to define a fingerprint of the surface. The topology of the contour map naturally involves the critical points of the height function since contour lines are integral curves for the vector field $(\partial r/\partial\varphi, -\partial r/\partial\theta)$.

Morse-Theory and dynamical systems. The subject of critical points naturally calls for Morse Theory [23] and dynamical systems [26]. Critical points of a Morse function defined on a manifold can be used to partition the manifold into regions of homogeneous flow [33], the so-called stable and unstable manifolds. The intersection of stable and unstable manifolds defines the Morse-Smale complex. An algorithm computing an approximation of the Morse-Smale complex of a function defined on a piecewise linear surface by mimicking the smooth setting has been introduced in [12].

In addition to the *smooth* Morse theory just mentioned, a combinatorial Morse theory [17, 18] has been developed recently. Discrete gradient vector fields actually underly our Morse-Smale decomposition algorithm. A decomposition algorithm based on a flow diagram for the particular case of disks in the plane is presented in [14].

Global molecular properties. Connolly's contribution discussed so far is mainly concerned with local geometric features. More global analysis are concerned with the characterization of voids —that is regions of a macro-molecule not accessible from the outside, as well as cavities —that is excavations with a small opening on the outside. The reader is referred to [11] for algorithms reporting voids and cavities based upon weighted Voronoi diagrams. An interesting source of information for global characterization of molecular properties is also [28].

In a different application area, namely that computer vision, and for the particular issue of characterizing human faces, the reader is referred to [15]. This book indeed provides several methods for characterizing patterns of the face.

1.4 Contributions and paper overview

We are now ready to state more precisely the contributions of this paper:
—First, we investigate Connolly's function for smooth curves and surfaces. The incentive for such a study is to relate the critical points of the function to differential quantities of the surface and in particular extrema of principal curvatures along curvature lines —refer again to [15, 27].
—Second, we provide an efficient algorithm to compute Connolly's function over triangulated surfaces.

—Third, we introduce the notion of discrete Morse-Smale decomposition based upon a discrete gradient vector field [17, 18]. This decomposition bridges the gap between local quantities —the critical points— and the global geometry of the triangulated manifold. This decomposition is simpler than the one presented in [12] —in particular it does not follow the *simulation of differentiability* paradigm, and can be computed in $O(n \log n)$ time. —Fourth, we apply the previous Morse-Smale decomposition to the discrete gradient vector field induced by Connolly's function, and present experimental results for several molecules as well as standard computer graphics models.

2 The Connolly function for smooth curves and surfaces

Spherical caps on van der Waals balls or triangles contributing to a piecewise linear surface are not the most accurate primitives to model molecular surfaces. For example, spherical harmonics or electron density contours [6] provide smooth models for such surfaces. In this section, we investigate some properties of the Connolly function for smooth i.e. infinitely differentiable surfaces. As a preliminary step, we examine the case of curves. The conclusions we are heading over are twofold: for small values of the probe sphere radius, the critical points of Connolly's function are nearby extrema of curvature for curves, and nearby critical points of the mean curvature function for surfaces.

From a differential geometry perspective, a difficulty with Connolly's function is that it is not just local —as opposed to the curvature for example. For the case of curves, these global interactions can be managed through explicit differential equations —as opposed to mere Taylor expansions for surfaces.

2.1 The Connolly function for smooth curves

The Connolly function for curves. Let $\mathcal{C}$ be a smooth compact curve in $\mathbb{R}^2$ and let $\gamma(s)$ be a parameterization of $\mathcal{C}$ by arc length —see Figures 3 and 4. The Frenet frame and the curvature of $\mathcal{C}$ at s are respectively denoted $(T(s), N(s))$ and $\kappa(s)$. Let η be a positive real number so that the circle of radiusη centered at $\gamma(s)$ intersects the curve $\mathcal{C}$ in two points $\gamma(t^+(s))$ and $\gamma(t^-(s))$, with $t^-(s) < s < t^+(s)$. Let $u^+(s)$ and $u^-(s)$ be the unitary vectors defined by

$$u^+(s) = \frac{\gamma(t^+(s)) - \gamma(s)}{\eta}, \quad u^-(s) = \frac{\gamma(t^-(s)) - \gamma(s)}{\eta}.$$

We denote by $\theta^+(s)$ (resp. $\theta^-(s)$) the oriented angle between $\gamma'(s)$ and $u^+(s)$ (resp. $u^-(s)$ and $-\gamma'(s)$) :

$$\theta^+(s) = \widehat{\gamma'(s), u^+(s)} \quad \text{and} \quad \theta^-(s) = \widehat{u^-(s), -\gamma'(s)}.$$

At point $\gamma(s)$, the Connolly function $C(s)$ is defined by the length of the circle arc from $\gamma(t^-(s))$ to $\gamma(t^+(s))$ normalized by η, or equivalently by the angle

$$C(s) = \pi + \theta^+(s) + \theta^-(s). \tag{2}$$

The aim of this section is to study Connolly's function for fixed radius $\eta > 0$.

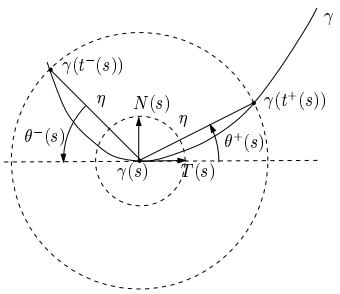


Figure 3: Connolly function for a curve: notations

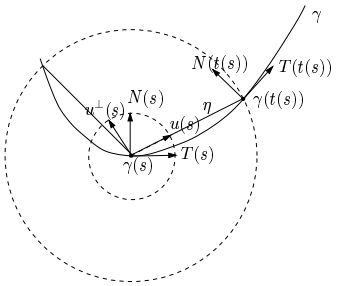


Figure 4: Differential Eq. for $\theta(s)$: notations —$\pm$ exponents omitted

Preliminary lemmas. We first seek equations satisfied by $t^\pm(s)$, $u^\pm(s)$ and $\theta^\pm(s)$. The equations (and proofs) involving quantities with + exponent are similar to the ones involving − exponents. So, to make notations more clear, we forget exponents in the following and we just use a symbol $\varepsilon \in \{+,-\}$ —that is $\varepsilon = +(-)$ corresponds to relations involving quantities with the + (−) exponent.

Lemma. 1 *Denote by $\langle .|. \rangle$ the scalar product in $\mathbb{R}^2$.*

1. *The function $t(s)$ satisfies the following differential equation*

$$\langle t^{'}(s)\gamma^{'}(t(s)) - \gamma^{'}(s) | u(s) \rangle = 0.$$

2. *Let $u^\perp$ the unitary vector orthogonal to u, and denote $\delta(s)$ the function*

$$\delta(s) = \frac{\langle t^{'}(s)\gamma^{'}(t(s)) - \gamma^{'}(s) | u^\perp(s) \rangle}{\eta}$$

If $\theta(s) \neq 0$ it satisfies

$$\theta^{'}(s) = \varepsilon(\delta(s) - \kappa(s)).$$

Proof. It follows from $\langle u(s)|u(s)\rangle = 1$ that there exists a function δ such that $u^{'}(s) = \delta(s)u^\perp(s)$. Differentiating the equality $\eta u(s) = \gamma(t(s)) - \gamma(s)$ we obtain

$$\eta u^{'}(s) = t^{'}(s)\gamma^{'}(t(s)) - \gamma^{'}(s) = \eta\delta(s)u^\perp(s).$$

Doing scalar product with $u^\perp(s)$ gives the desired formula for δ.
Now we have $\cos\theta(s) = \varepsilon\langle u(s)|\gamma^{'}(s)\rangle$. Hence

$$-\theta^{'}(s)\varepsilon\sin\theta(s) = \delta(s)\langle u^\perp|\gamma^{'}(s)\rangle + \kappa(s)\langle u(s)|(\gamma^{'}(s))^\perp\rangle.$$

Note that

$$\langle u^{\perp}|\gamma^{'}(s)\rangle = -\langle u(s)|(\gamma^{'}(s))^{\perp}\rangle = -\sin\theta(s),$$

so

$$\theta^{'}(s)\varepsilon\sin\theta(s) = (\delta(s)-\kappa(s))\sin\theta(s).$$

□

From Eq. (2), the derivative of Connolly's function is equal to $\delta^{+}(s)-\delta^{-}(s)$. Hence one deduces immediately the following corollary.

Corollary. 1 *Let $s_0 \in \mathbb{R}$ be such that for any s in a neighborhood of 0, $\gamma(s_0+s)$ $\gamma(s_0-s)$ are symmetric with respect to the line directed by $\gamma^{'}(s_0)^{\perp}$ passing through $\gamma(s_0)$. Then there exists η_0 such that $\gamma(s_0)$ is a singular point of Connolly functions with circle radius $\eta \leq \eta_0$.*

For example, the origin $(0,0)$ of $\mathbb{R}^2$ is a singular point of the Connolly function of the parabola $y = px^2$, $p \in \mathbb{R}\setminus\{0\}$.

On the critical points of Connolly's function. In lemma 1 we derived the expression of the derivative $\partial\theta(s)/\partial s$, which involves the function δ. Considering that s is fixed, we now work out an expansion of $\theta^{'}(s)$ as a function of the radius η.

Lemma. 2 *Using the previous notations and s being fixed, one has :*

$$t(s,\eta) = s + \varepsilon\eta + o(\eta^2), \tag{3}$$

$$\theta^{'}(s,\eta) = \frac{\partial\theta}{\partial s}(s,\eta) = \frac{1}{2}\frac{d\kappa}{ds}(s)\eta + o(\eta). \tag{4}$$

Proof. Without loss of generality, we may suppose that $s=0$ and we denote $t(\eta) = t(0,\eta)$. We first prove that t is a $\mathcal{C}^{\infty}$-function of η in a neighborhood of 0. The values $t^{+}(\eta)$ and $t^{-}(\eta)$ are solutions of equation

$$\langle\gamma(t)-\gamma(0)|\gamma(t)-\gamma(0)\rangle = \eta^2. \tag{5}$$

Hence if we consider only positive solution of this equation, t satisfies

$$\Phi(t,\eta) = t\sqrt{\langle\frac{\gamma(t)-\gamma(0)}{t}|\frac{\gamma(t)-\gamma(0)}{t}\rangle} - \eta = 0.$$

Function $g(t) = \sqrt{\langle\frac{\gamma(t)-\gamma(0)}{t}|\frac{\gamma(t)-\gamma(0)}{t}\rangle}$ is $\mathcal{C}^{\infty}$. Moreover, since γ is parameterized by arc-length, $g(0) = 1$ and $g^{'}(0) = 0$ —this last equality stemming from the fact that $\gamma^{'}(0)$ and γ"(0) are orthogonal. Applying the implicit function theorem to Φ at point $(0,0)$ ($\frac{\partial\Phi}{\partial t}(0,0) = g(0) \neq 0$), we deduce that $t(\eta)$ is a $\mathcal{C}^{\infty}$ function. Now, differentiating $tg(t)-\eta = 0$ two times with respect to η implies that $t(\eta) = \eta + o(\eta^2)$. Considering negative solution

of (5) leads in a similar way to $t(\eta) = -\eta + o(\eta^2)$. Putting everything together and for a general value of s yields Eq. (3).
Taking the derivative of the previous expression yields

$$\frac{\partial t}{\partial s}(0,\eta) = 1 + o(\eta^2).$$

Taking an order two Taylor expansion of $\gamma^{'}$ at $t(0,\eta) = \varepsilon\eta + o(\eta^2)$ yields

$$\gamma^{'}(t(0,\eta)) = \gamma^{'}(0) + \varepsilon\gamma"(0)\frac{\eta}{2} + o(\eta^2).$$

At last, since $u = (\gamma(t(0,\eta)) - \gamma(0))/\eta$, an order three Taylor expansion of γ at $t(0,\eta) = \varepsilon\eta + o(\eta^2)$ yields

$$u(0,\eta) = \varepsilon\gamma^{'}(0) + \gamma"(0)\eta + \gamma^{(3)}(0)\frac{\eta^2}{2} + o(\eta^2).$$

Injecting these equalities in the expression of δ gives

$$\delta(0,\eta) = \kappa(0) + \varepsilon\langle\gamma^{(3)}(0)|\gamma^{'}(0)^{\perp}\rangle\frac{\eta}{2} + o(\eta).$$

To conclude the proof, remark that

$$\gamma^{(3)}(0) = \frac{d}{ds}\gamma^{''}(s)(0) = \frac{d}{ds}(\kappa(s)\gamma^{'}(s)^{\perp})(0) = \kappa^{'}(0)\gamma^{'}(0)^{\perp} - \kappa^2(0)\gamma^{'}(0),$$

whence

$$\langle\gamma^{(3)}(0)|\gamma^{'}(0)^{\perp}\rangle = \kappa^{'}(0).$$

□

This last lemma helps to localize the singularities of Connolly's function. If s is such that the curvature of curve $\mathcal{C}$ at point $\gamma(s)$ is not singular, then for η sufficiently small, Connolly's function is regular at $\gamma(s)$. Hence for η small, singular points of Connolly function are located in a neighborhood of singular points of the curvature. This coincides with the intuition that the extrema of Connolly functions are located near the extrema of curvature of $\mathcal{C}$.

More precisely, let $Sing(\kappa) = \{\gamma(s) : \kappa^{'}(s) = 0\}$ be the singular set of the curvature function $\kappa(s)$ and let $Sing(C,\eta) = \{\gamma(s) : C^{'}(s,\eta) = 0\}$ be the singular set of Connolly function corresponding to radius η. If A and B are two subset of the plane, denote by

$$\mathcal{H}(A|B) = \sup_{m\in A} d(m,B)$$

the one sided Hausdorff distance between A and B, where $d(.,B)$ is the Euclidean distance to B. Be careful that $\mathcal{H}(.|.)$ is not a distance between sets : it is not symmetric and $\mathcal{H}(A|B) = 0$ does not imply $A = B$.

Theorem. 1 *With previous notations one has*

$$\lim_{\eta \to 0} \mathcal{H}(Sing(C,\eta)|Sing(\kappa)) = 0.$$

Proof. Computations done in the proof of the previous lemma depend continuously of s. From Eq. (4), there exists a continuous function $\phi(s,\eta)$ such that

$$\lim_{\eta \to 0} \phi(s,\eta) = 0,$$

and such that Connolly function satisfies

$$C^{'}(s,\eta) = \frac{\partial C}{\partial s}(s,\eta) = \frac{d\kappa}{ds}(s)\eta + \eta\phi(s,\eta).$$

Now, continuity of ϕ and compactness of $\mathcal{C}$ implies that for any $\varepsilon > 0$ there exists $\alpha > 0$ such that $(\eta < \alpha) \Longrightarrow (|\phi(s,\eta)| < \varepsilon$ for any $s)$. Hence if s is such that $|\kappa^{'}(s)|$ is greater than ε then $C^{'}(s,\eta) \neq 0$ whenever $0 < \eta < \alpha$. The critical points of Connolly's function are thus contained in the set $S_\varepsilon = \{s : |\kappa^{'}(s)| < \varepsilon\}$. The derivative of the curvature $\kappa^{'}(s)$ being continuous, one has

$$\lim_{\varepsilon \to 0} \mathcal{H}(S_\varepsilon|Sing(\kappa)) = 0$$

which conclude the proof. □

2.2 The Connolly function for smooth surfaces

On the critical points of Connolly's function. Consider now the Connolly function given by Eq. (1) assuming $\mathcal{M}$ is a $\mathcal{C}^\infty$ compact surface embedded in $\mathbb{R}^3$. Let p be a point of $\mathcal{M}$. Denote κ_1, κ_2 and H the principal and mean curvatures of $\mathcal{M}$ at p. Also, let $A(p,\eta)$ stand for the surface area of the sphere $S(p,\eta)$ contained within the volume bounded by $\mathcal{M}$. We first generalize lemma 2 to the case of surfaces.

Lemma. 3 *There exists a $\mathcal{C}^\infty$ function ε defined on $S \times \mathbb{R}_+$ such that*

$$C(p,\eta) = \frac{A(p,\eta)}{\eta^2} = 2\pi + \pi H(p)\eta + \eta^2\varepsilon(p,\eta) \tag{6}$$

with $\lim_{\eta \to 0} \varepsilon(p,\eta) = 0$.

Proof. Let p be a point of $\mathcal{M}$. We may assume that $p = (0,0,0)$ and $\mathcal{M}$ is locally the graph of a smooth function

$$z = f(x,y) = \frac{1}{2}(\kappa_1 x^2 + \kappa_2 y^2) + P_3(x,y) + \cdots$$

where $P_3(x,y) = (ax^3 + 3bx^2y + 3cxy^2 + dy^3)/6$ is a homogeneous polynomial of degree 3. Suppose that the compact domain bounded by $\mathcal{M}$ is (locally) $z \leq f(x,y)$. Denote by (r,θ,ϕ)

the spherical coordinates in $\mathbb{R}^3$ —see Figure 5. For a given angle θ, let $\mathcal{M}_\theta$ be the plane curve intersection between $\mathcal{M}$ and the half plane $\{r \geq 0,\ \theta\}$ —see Figure 6. At last, let $g_\theta(t)$ the expression of $\mathcal{M}_\theta$ as the graph of a smooth function near p, that is $\mathcal{M}_\theta = \{(t, g_\theta(t))\}$. Using the expression of the Jacobian of the parameterization of the sphere with spherical coordinates, one finds:

$$\begin{aligned}\frac{A(p,\eta)}{\eta^2} &= \int_{\theta=0}^{2\pi}\int_{\phi(\theta)}^{\pi} \sin\phi \; d\phi \; d\theta \\ &= \int_0^{2\pi} (1+\cos\phi(\theta))d\theta \\ &= 2\pi + \int_0^{2\pi} \cos\phi(\theta))d\theta\end{aligned}$$

where $\phi(\theta_0)$ is the ϕ-coordinate of the intersection point of $\mathcal{M}$ with the half plane $\{r \geq 0,\ \theta\}$. Applying Lemma 1 from [2] yields, for $\eta > 0$,

$$\frac{\pi}{2} - \phi(\theta) = \frac{g_\theta^{"}(0)}{2}\eta + \frac{g_\theta^{(3)}(0)}{6}\eta^2 + o(\eta^2).$$

But $g_\theta^{"}(0)$ is the curvature of $\mathcal{M}_\theta$ at the origin, that is the directional curvature of the normal section of $\mathcal{M}$. Expressing this directional curvature using Euler's relation yields

$$\frac{\pi}{2} - \phi(\theta) = \frac{1}{2}(\kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta)\eta + P_3(\cos\theta, \sin\theta)\eta^2 + o(\eta^2).$$

But

$$\int_0^{2\pi} P_3(\cos\theta, \sin\theta)d\theta = 0,$$

so that

$$\begin{aligned}\frac{A(p,\eta)}{\eta^2} &= 2\pi + \int_0^{2\pi} \sin\{\frac{1}{2}(\kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta)\eta + P_3(\cos\theta, \sin\theta)\eta^2 + o(\eta^2)\}d\theta \\ &= 2\pi + \eta\int_0^{2\pi} \frac{1}{2}(\kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta)d\theta + o(\eta^2) \\ &= 2\pi + \pi\frac{\kappa_1+\kappa_2}{2}\eta + o(\eta^2).\end{aligned}$$

As earlier noticed for curves, computation of $\phi(\theta)$ depends smoothly from parameter t in $\mathcal{M}_\theta = \{(t, g_\theta(t))\}$. Therefore, the previous computations also depends smoothly from point p. □

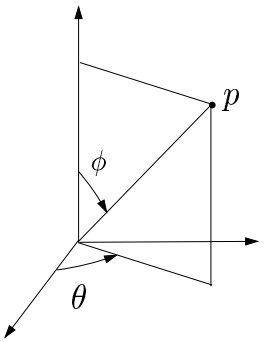


Figure 5: Spherical coordinates

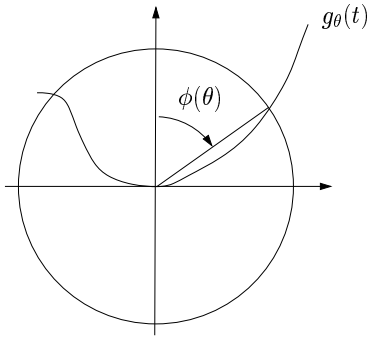


Figure 6: Normal section of $\mathcal{M}$ at angle θ

As in the case of curves, previous lemma allows to localize singularities of Connolly function using the one sided Hausdorff distance. Let $\nabla H(p)$ be the gradient of function H at point p and, η being fixed, let $\nabla C(p,\eta)$ be the gradient of Connolly function C at point p.

Theorem. 2 *Let* $Sing(H) = \{p \in \mathcal{M} : \nabla H(p) = 0\}$ *be the singular set of the mean curvature function* H *and let* $Sing(C,\eta) = \{p \in \mathcal{M} : \nabla C(p,\eta) = 0\}$ *be the singular set of the Connolly function corresponding to radius* η*. One has*

$$\lim_{\eta \to 0} \mathcal{H}(Sing(C,\eta) | Sing(H)) = 0.$$

Proof. Consider Eq. (6). Since $\mathcal{M}$ is compact the differential $d\varepsilon$ is bounded. There exists a constant $M > 0$ such that for any $\eta \in \mathbb{R}^+$ and any $m \in \mathcal{M}$, the norm of the gradient $\nabla \varepsilon_\eta(p)$ of ε is bounded by M. Radius η being fixed, we have

$$\nabla C(p,\eta) = \pi \nabla H(p) \eta + \eta^2 \nabla \varepsilon_\eta(p).$$

It follows that if H is not singular at point $p \in \mathcal{M}$, then for any $\eta < \frac{\pi \|\nabla H(p)\|}{M}$, the Connolly function $C(.,\eta)$ is regular at p. In other words, $Sing(C,\eta)$ is included in the set $S_\eta = \{p \in \mathcal{M} : \|\nabla H(p)\| < \frac{M\eta}{\pi}\}$. The gradient vector field ∇H being continuous, one has

$$\lim_{\eta \to 0} \mathcal{H}(S_\eta | Sing(H)) = 0$$

which concludes the proof. □

Relationship to ridges. Important curvature properties of surfaces are encoded by ridges, that is extrema of principal curvature along curvature lines [13, 27]. For example, points

located on elliptic ridges may be the contact points of maximal spheres contributing to the skeleton —but not the medial axis— of the complement $\mathbb{R}^3 \setminus \mathcal{M}$ of $\mathcal{M}$.

As pointed out in introduction —refer again to Figure 2, Connolly's function is obviously related to curvature properties. As just proved, the critical points of Connolly's function are located nearby critical points of the mean curvature function H. We therefore take a quick look at the exact relationship between critical points of H and ridges.

Consider the expression of surface $\mathcal{M}$ as a height function near point p. More precisely, express this height function in the coordinate system associated with the principal directions and whose z axis is the normal to $\mathcal{M}$ at p —this is the so-called Monge form of the surface. The following lemma is a straightforward application of [1, p.439], and shows that the critical points of H do not lie on ridges since required conditions there are $a = 0$ or $d = 0$:

Lemma. 4 *Assume the Monge form of the surface $\mathcal{M}$ at p is given by*

$$f(x,y) = \frac{1}{2}(\kappa_1 x^2 + \kappa_2 y^2) + \frac{1}{6}(ax^3 + 3bx^2y + 3cxy^2 + dy^3) + \cdots .$$

Then, the local expansion of the mean curvature H satisfies

$$2\ H(x,y) = \kappa_1 + \kappa_2 + (c+a)x + (b+d)y + \|(x,y)\|\delta(x,y) \quad \textit{with} \quad \lim_{(x,y)\to 0} \delta(x,y) = 0.$$

3 Computing the Connolly function for a mesh

This section presents the calculation of the Connolly function defined by Eq. (1) for a triangulated surface $\mathcal{M}$. We assume the surface is represented by a halfedge data structure [22]. The sphere of radius r centered at a vertex p of the mesh is denoted $S(p,r)$.

3.1 Outline

Consider the sphere $S(p,r)$. If r is small enough, the surface $\mathcal{M}$ cuts S into two topological disks, one of which entirely lies within the volume bounded by $\mathcal{M}$. The intersections between the triangles of the mesh and S consist of circle arcs. Such an arc is a great circle arc if the plane of its defining triangle contains the center of S, and a small circle arc otherwise. Let $C = \{C_1, \ldots, C_k\}$ be the set of circle arcs. As observed in [7], the local Gauss-Bonnet theorem can be invoked to compute the surface area of the patch we are interested in.

To see how, orientate the intersection curve C so as to leave the interior of the patch we want to compute the area of to the left. Also, let $\Delta\theta_i$ be the tangent turn between the endpoint of C_i and the starting point of C_{i+1}. Also denote k_g the geodesic curvature. The local Gauss-Bonnet theorem states that

$$\sum_i \Delta\theta_i + \sum_i \int_{C_i} k_g ds + \frac{1}{r^2} Area = 2\pi, \tag{7}$$

from which the surface area is easily derived if the $\Delta\theta_i$s and the geodesic curvatures are known.

3.2 The three steps

Computing the intersection curve C. To report the intersections between the triangles of the mesh and the sphere S, we proceed by first finding an edge crossing the boundary of S, and second by walking along C.

More precisely, let h_0 be a halfedge crossing S. Let T be a triangle incident to h_0 and denote $\{h_0, h_1, h_2\}$ its edges. The intersection between T and S is one of the two configurations depicted on Figure 8 —that is the second intersecting edge is either h_1 or h_2. Therefore, once a starting edge h_0 has been found, one just needs to find the second intersected edge of the triangle under consideration, move into the adjacent triangle, and iterate.

The key issue consists of finding h_0. Let p be the vertex where the probe sphere is centered. A greedy strategy is to (i) select the farthest neighbors v of p (ii)check whether the edge pv crosses S (iii)if not, iterate the *farthest neighbor of p* strategy until a crossing edge has been found. For big probes however, this strategy mail lead to a dead-end. On Figure 9 e.g., the search oscillates between p_1 and p_2. When such a situation is detected, an exhaustive search of the neighboring graph anchored at p can be performed.

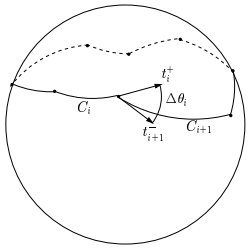


Figure 7: Gauss-Bonnet on a spherical cap

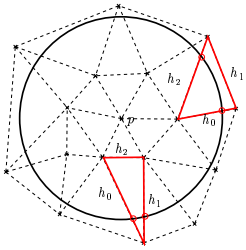


Figure 8: Walking around the intersection curve

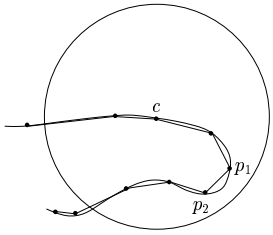


Figure 9: Greedy strategy fails for big probe spheres

Computing the angle turns $\Delta\theta_i$. Computing the angle turns $\Delta\theta_i$s requires the knowledge of the tangents at the endpoints of the circle arcs C_is. We therefore seek a parameterization of these circle arcs.

Consider a triangle T, two edges of which intersect S. Let p_1 and p_2 be the intersection points of these two edges with S, and let c be the center of the corresponding circle arc —see Figure 10. The line-segment p_1p_2 is parameterized by $X(t) = (1-t)p_1 + tp_2$, $t \in [0,1]$. Assuming no triangle is degenerate —i.e. has an angle equal to π, $X(t)$ can be uniquely scaled so as to lie on the sphere S. Let $L(t)$ be the parameterization sought. If r_c denotes the radius of the small circle, i.e. $r_c = \mid cp_1 \mid$, we have

$$cL(t) = r_c \frac{cX(t)}{\mid cX(t) \mid},$$

from which the tangent vector at any point of the circle arc is derived.

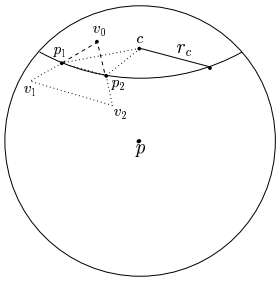


Figure 10: Parameterizing a small circle (1/2)

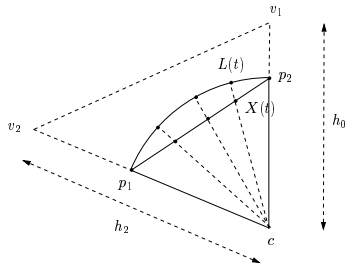


Figure 11: Parameterizing a small circle (2/2)

Computing the geodesic curvatures. The last step required to evaluate the surface area from the local Gauss-Bonnet theorem is to compute the geodesic curvatures k_g along the circle arcs. Denote d_c the distance between the sphere center and the small circle center, r the radius of S, and r_c the small circle radius. As noticed in [7], the absolute value of k_g is given by $k_g = d_c/(r\ r_c)$. However, this does not give the sign of k_g.

We therefore resort to the definition of the geodesic curvature based upon the covariant derivative [9]. Let γ be a curve parameterized by arc length on a smooth surface. Denote $\gamma^{'}(s) = t(s)$ the tangent vector and $\gamma^{''}(s) = dt/ds = kn$ be the curvature vector of γ. Also, let N stand for the normal of the surface —see Figure 12. The geodesic curvature is defined by the covariant derivative of the tangent vector t, i.e.

$$k_g = [\frac{d\gamma^{'}}{ds}] =< \frac{d\gamma^{'}}{ds}, N \wedge \frac{d\gamma^{'}}{ds} > =< kn, N \wedge t > .$$

For a small circle, the latter formula is especially convenient since kn points to the center c of the small circle, and has magnitude $1/r_c$.

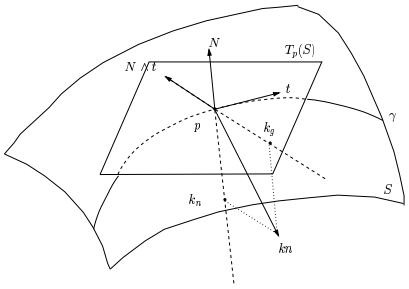


Figure 12: Geodesic curvature as covariant derivative

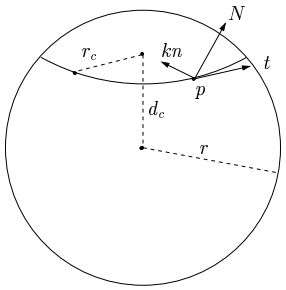


Figure 13: Geodesic curvature of a small circle

4 Constructing a discrete Morse–Smale decomposition

4.1 Introduction

We aim to analyze a scalar field f defined at the vertices of a mesh model. In particular, we want to decompose the domain of f into patches where the flow of f is uniform. In the smooth case, this decomposition has been introduced by Thom [33] and Smale [29] based on Morse theory [23]. Although an algorithm for computing quasi–Morse complexes already exists [12], the algorithm introduced here is simpler, faster, and does not follow the *simulation of differentiability* paradigm.

Is is based upon Forman's Morse theory [18], a discrete version of Morse theory. This theory extends the classical Morse theory to cell complexes, connecting the topology of such a structure to the study of scalar fields defined on it. The cell complex structure is particularly well suited for the study of meshes, and allows direct computations with a strong theoretical framework. Since Forman's Morse theory is combinatorial, it does not require any approximation to be implemented on meshes. The algorithm following our analysis is simple and accommodates a hierarchy, which can be used to detect topological noise.

We shall prove that Connolly's function is admissible for the discrete framework, and show how the hierarchy can be used to get rid of useless patches in particular when the radius of the Connolly ball is not adapted.

4.2 Pre-requisites

Smooth Morse–Smale decomposition. Given a smooth 2–manifold M and a smooth map $f : M \to \mathbb{R}$ defined on it, an *integral curve* is a curve everywhere tangent to the gradient of f. The *critical points* are the points where the gradient vanishes. The *stable manifold* of a critical point c is the union of all the integral curves converging to c. The *unstable manifold* of a critical point c is the union of all the integral curves originating from c.

A function is *Morse–Smale* if its critical points are isolated and if its integral curves cross transversally. It has been proven in [12, Lemma 1] that the intersections of the stable and unstable manifolds form a cell complex, where each 2–cell is a quadrangle with critical vertices of index 0,1,2,1 in this order around the region. This property characterizes a *quasi Morse complex*. The decomposition introduced here also verifies this property.

Forman's discrete Morse theory. The building block of Forman's theory is the *discrete gradient vector field*. It corresponds to the gradient of a smooth Morse function. The discrete gradient vector field is equivalent to discrete Morse functions [18].

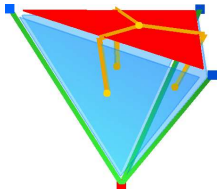


Figure 14: A discrete gradient vector field on a tetrahedron: 1 critical vertex, 1 critical face (in red).

A *combinatorial vector field* $\mathcal{V}$ defined on a manifold is a collection of disjoint pairs of incident cells: vertex/ edge, edge/face...

A $\mathcal{V}$*–path* is an alternating sequence α_0, β_0, α_1, β_1, ...of cells such that:

- all the cells α_i are of the same dimension p, and all the cells β_i are of the same dimension $p+1$.
- α_i and β_i are incident cells, and so are β_i and α_{i+1}.
- the pairs $\{\alpha_i, \beta_i\}$ are elements of the vector field $\mathcal{V}$.

A *discrete gradient vector field* is a combinatorial vector field with no non–trivial closed $\mathcal{V}$–path.

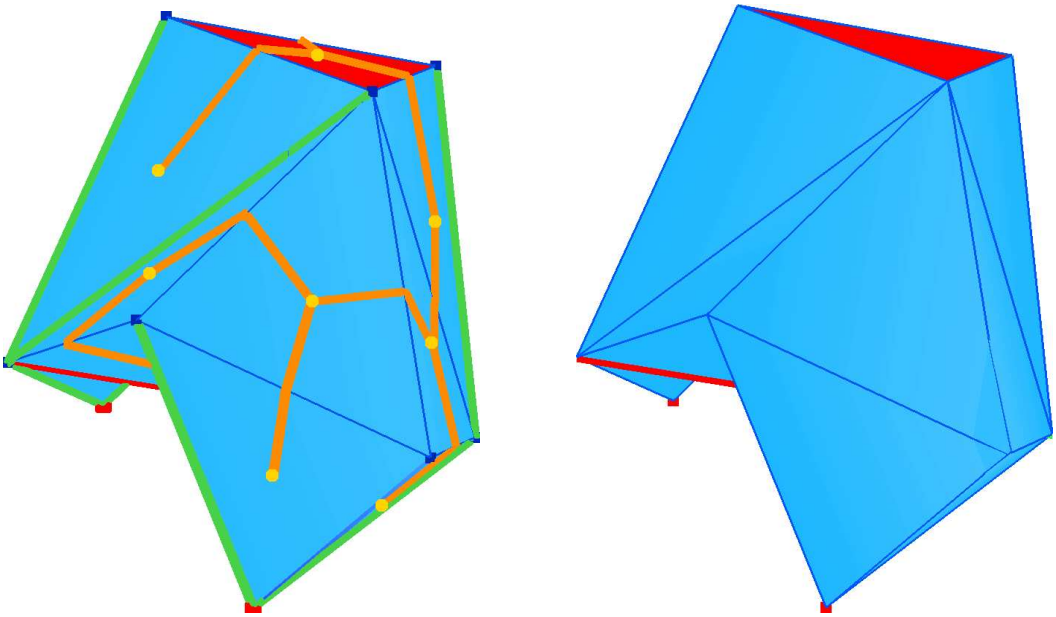
Morse proved that the topology of a manifold is related to the critical elements of a smooth function defined on it. Forman gave analogous results, with the following definition for the critical cells: a cell is *critical* if it is not paired with any other cell.

Figure 15: The primal and dual rooted trees of the discrete gradient vector field of figure 14.

Discrete gradient vector fields and rooted forests. It has been proven in [19, 20] that a discrete gradient vector field on a closed 2–manifold is in fact a pair of interlaced primal and dual rooted spanning forests. The dual graph D of a closed 2–manifold M represents each face of M by a node, and the edges of M by a link joining adjacent nodes. A dual spanning forest DF of M is a subgraph of D with no cycle, containing all the nodes of D.

A rooted forest has one node for each tree component distinguished as the root. A discrete gradient vector field on M is equivalent to a primal and a dual spanning forests PF and DF such that **(i)** all vertices of M are represented in PF, all faces in DF and **(ii)** an edge of M is not represented in both DF and PF — it is possible that an edge is represented neither in DF nor in PF. For example, figure 15 shows the primal and dual trees of the discrete gradient vector field of figure 14.

The critical vertices (minima) are the roots of the primal forest, the critical faces (maxima) are the roots of the dual forest and the critical edges (saddles) are the edges that belong neither to the primal nor to the dual forest. For example, figure 16 shows the interlaced forest and their critical points on a small sphere with saddle model.



(a) The interlaced primal (green) and dual (orange) forests

(b) : one maximum (top triangle), two minima (bottommost vertices) and one saddle (horizontal edge)

Figure 16: A z–coordinate Morse function on a topological sphere with a saddle and its critical elements (in red).

4.3 Discrete decomposition

Discrete stable and unstable regions. The discrete gradient vector field points from the roots to the leaves in the primal forest, and from the leaves to the roots in the dual forest. Therefore, an integral curve is of two kinds: a $\mathcal{V}$–path formed by vertices and edges, or a $\mathcal{V}$– path formed by faces and edges.

The stable region of a maximum (face) is its component in the dual forest. The unstable region of a minimum (vertex) is its component in the primal forest. The stable region of a minimum and the unstable region of a maximum are reduced to themselves (as in the smooth case). Notice that, as a difference with the smooth case, the stable and unstable regions are defined on different elements: the stable region is defined on faces, and the unstable one on vertices.

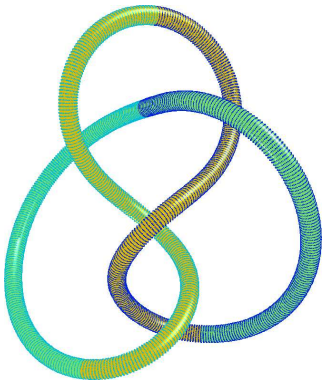


Figure 17: The discrete Morse–Smale decomposition of a figure–eight–knot model with the z–coordinate as a scalar field.

Discrete Morse–Smale decomposition. The Morse-Smale decomposition is usually computed by intersecting the stable and unstable regions, creating *patches*. As in our case they rely on separate cells, this cannot be done directly. A stable region can be extended to a manifold by including all its inner edges and vertices. The edges and vertices on the boundary of the region will be the boundary of the manifold. An unstable region can be extended to a 2–manifold by including all the triangles with all its vertices in the unstable region. The boundary is then a set of faces. It can be left empty, be attributed to both bounding regions (overlapping boundary), which could be suitable for docking applications. We can also attribute each face to the region that contains most of its vertices, starting from the saddle points (adapted boundary). Figure 17 shows the result on a figure–eight–knot model.

The property of quasi Morse complex remains for closed 2–manifolds: minima and maxima cannot be connected, as the primal and dual forests are disjoint. For the same reason, a

saddle cell cannot be connected to another saddle. Therefore, the boundary of every patch of the discrete Morse–Smale decomposition is composed of sequences of minimum, saddle, maximum and saddle. One patch corresponds to the intersection of the tree component of a maximum with the tree component of a minimum. Therefore, it cannot contain more than one minimum, and its boundary consists of only one sequence minimum, saddle, maximum, saddle.

A discrete gradient vector field from Connolly's function. Defining a smooth Morse function on a mesh based on a scalar field usually requires that the values of the scalar field on the vertices are all different. This can be obtained by a small perturbation, and ensures the uniqueness of the Morse–Smale decomposition. Then, the function is interpolated on the mesh's edges and faces to comply theoretically with the requirements of a Morse function.

The theoretical framework of Forman's theory gives a more straightforward construction. A discrete gradient vector field on a closed 2–manifold is equivalent to interlaced forests such that, on each component of the forest, the discrete gradient is monotone along the edges from the leaves to the root [19]. Any scalar field f given on the vertices of a mesh can also be decomposed into primal trees where f is monotone. The dual trees are obtained in a similar way by assigning to each face the mean value of its vertices. (To avoid having more maximal faces than the number of maxima of f, this argument is refined in the algorithm.) Therefore, any scalar field defined on the vertices can be completed to a discrete gradient vector field (which is equivalent to a discrete Morse function [18]). This completion is unique when the scalar field is injective.

Decomposition algorithm. Given a function f on every vertex, the algorithm proceeds in four steps. Each cell (vertex, edge, face) has an extra pointer to store the patch structure (the union/find structure [31]) for vertices and faces, and to mark edges as visited. The algorithm computes first the stable regions, and then the unstable ones.

1. It identifies the local minima and the local maxima, by comparing the value of f on a vertex to the value of f on each of its adjacent vertices.

2. The dual spanning forest is built, with the maxima of f as roots (remind that in the dual setting, values of f refer to values of f over the facets, and that edges actually connect two facets). The classical algorithm for minimal spanning tree [32] is used in the following manner. The edges are processed by increasing value (mean value of their incident vertices). When testing if an edge joins two different components, using the union/find data structure, it does not select an edge that would join two trees with different maxima. Otherwise, the edge is selected and marked as visited, and the union/find structure is updated. Therefore, we get a face–spanning forest, with one component per maxima. This maximum will be used as the root of the component.

3. It builds the primal spanning forest, with the minima of f as roots, in the same way as at the second step —processing the edges in a reverse order.

4. It computes the intersection of the stable and unstable regions, by marking the edges of each region with a different identifier. It uses a dictionary to store the pair unstable/stable identifier of each patch —with lexicographic key. The different options for the extension of the stable manifold, i.e. empty boundary, overlapping boundary and adapted boundary are given as an optional argument.

The algorithm's complexity is $O(n \log(n))$ for the edge sorting, and $O(n\alpha(n))$ for the forest creation, where n denotes the size of the mesh and $\alpha(n)$ is the inverse of Ackerman's function. At the end of the algorithm, each vertex and each face carries its patch number — as a union/find structure. For faces [resp. vertices], this denotes the stable [resp. unstable] region it belongs to. In the algorithm, we only care about the minima and the maxima. We know the Euler characteristic can be expressed as the alternated sum of the critical cells, or equivalently of all the cells. This guarantees we did not create any meaningless saddle critical cell.

Smale's order and Noise removal. Critical cells correspond to local maxima, minima and saddles. Therefore, a small perturbation of the scalar field would create a critical cell that would perturb the overall readability of the result. It would be convenient to rate each critical cell by its contribution to the overall scalar field, and to be able to cancel meaningless critical cells.

Forman's theory allows a very easy computation of a cancellation. If we want to cancel a minimum (resp. maximum) m with a saddle s incident to the tree component of m in the primal (resp. dual) forest, it is sufficient to add this saddle to the primal (resp. dual) forest. This is possible only if it does not create cycle, i.e. if the saddle s is also incident to a different tree component of the primal (resp. dual) forest. This is equivalent to computing the whole decomposition, removing m from the list of local minima (resp. maxima). From the computational side, it is just a union operation in the union/ find structure.

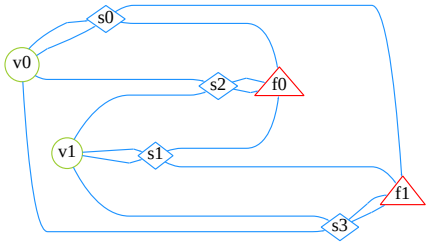


Figure 18: The Smale's order as a hierarchy of the decomposition of figure 17.

The hierarchy can be obtained by the Smale's order: a critical cell c_1 is inferior to a critical cell c_2 if the unstable region of c_1 intersects the stable region of c_2. This partial order can be represented as a graph. This graph is easily built by looking at which maxima and minima the saddles would join. We see on figure 18 that this graph points out the quasi–Morse complex property of the decomposition.

Consider a minimum–saddle cancellation C. For a possible cancellation C, denote $\delta(C)$ the difference of the scalar field at the two critical cells. C is possible if the saddle is connected with another minimum. Therefore, the saddle can be canceled in two ways, with both minima. The natural option is the cancellation with the smaller δ. It is now possible to order the possible cancellations by their δ, and to cancel each pair that is below a given threshold. This is actually equivalent to the hierarchy introduced in [12].

The possible cancellations of figure 18 are v0/s2, v1/s2, v0/s3, v1/s3, s0/f0, s0/f1, s1/f0, s1/f1. The figure is graduated with the value of scalar field f. Thus, we can read the most likely cancellations: v1/s2 and s1/f0.

5 Experimental results

5.1 Implementation issues

The algorithms presented in sections 3 and 4 have been implemented using CGAL. As already mentioned, the data structure used to represent surface triangulations is a halfedge data structure.

The algorithm computing the Connolly function requires a great deal of numerical calculations: computing intersection points between edges and the probe sphere, computing tangent vectors along small circles, evaluating angles and curvatures. All these calculations are performed using doubles and we did not observe any significant numerical instability.

The Morse–Smale decomposition algorithm proceeds directly on the halfedge data structure. It requires for each mesh element (vertex, halfedge, facet) a float for the scalar field and an additional void pointer for the union/find data structure. One halfedge is selected for representing the entire edge. The edges are quick–sorted in an independent array of pointers. The halfedge are marked as visited by an non–null void pointer. After the forest creation algorithm, each vertex' (resp. facet) void pointer points to the minimum (resp. maximum) of the unstable (resp. stable) region it belongs to. The region's intersection is computed as a STL set data structure, whose elements are pairs minimum/maximum, i.e. vertex/facet. Smale's order and cancellation hierarchy are computed by only testing the saddles. The graphs are exported using graphviz [21]. The cancellations are stored as pairs minimum/saddle or saddle/maximum. The cancellation hierarchy is stored as a STL priority queue. One cancellation can invalidate others. Therefore, the first element of the hierarchy are tested and eventually updated before any cancellation.

In [12], special kinds of saddles, namely monkey saddles, needs to be worked out separately. Those saddles are points whose upper and lower stars contain three or more wedges. With Forman's framework, those saddles are not distinguished from others: an edge is a critical cell whose end vertices belong to separate components of the primal forest, and whose incident faces belong to separate components of the dual forest.

5.2 Experimental results

Experimental setup. The outputs of the decomposition algorithm on several closed surfaces are reported below. In particular, the columns of table 1 read as follows: (i)model name (ii)number of vertices (iii)length of the model bounding box diagonal(iv)Connolly sphere radius (v,vi,vii)number of critical points (max,saddle,min) (viii)Euler characteristic i.e. $\mathcal{X} = \sum_{c\ crit.point} index(c)$ (ix)genus of the surface i.e. $g = (2 - \mathcal{X})/2$.

From a visual standpoint, a given picture displays the stable, unstable or mixed patches. For such a picture, the patches colors are chosen at random. Consider for example Figure 19: each colored patch is a stable manifold and corresponds to a maximum of the Connolly function, namely the red point located at to bottom of the *depression*. Similarly on Figure 20, one finds an unstable patch for each maximum of the Connolly function —the blue vertex at the tip of the *knob*.

The calculations presented thereafter required less than a minute on a PC at 1.8 GHz. The final version of the paper will report separate timings for the calculation of Connolly's function and the Morse-Smale decomposition algorithm.

Experiments on molecules. We tested the algorithms on three molecules: a small one —retinal, an enzyme —papain, and a DNA fragment. Retinal is a photo-pigment involved in the capture of photons during photo-synthesis. Papain is a proteolytic enzyme. Of particular interest is the DNA molecule due to the geometry of its two grooves. In each case, the solvent accessible molecular surfaces were produced from a `pdb` file using Michel Sanner's `msms` software [1], with a probe radius of 1.5 Angstrom.

The results displayed for the three molecules correspond to different radii of the Connolly sphere. As shown on pictures 19 to 21 with the retinal molecule, the maxima and minima of the Connolly function indeed correspond to *depressions* and *knobs*. Some topological noise materializing through tiny patches can be observed at the surface of papain. Getting rid of this noise calls for cancellations, but due to the the lack of space we omit results on hierarchical Morse complexes. At last, visual inspection of the DNA molecule is particularly pleasing since the grooves are paved with maxima of the Connolly function [2].

Experiments on computer graphics models. We performed additional experiments on standard computer graphics models —see Figures 26 and 27. The critical points enable one to recover the genus of the surfaces, an overkill process however! More interestingly, the Morse-Smale patches can certainly be useful for the segmentation of surfaces before entering a parameterization or a re-meshing process [10].

[1] available from `http://www.scripps.edu/pub/olson-web/people/sanner/html/msms_home.html`

[2] search `dna+groove` from `http://images.google.com` and compare!

	#vces	$\emptyset$	r	#max	#saddle	#min	$\mathcal{X}$	g
retinal	3643	22	1.5	55	97	44	2	0
papaine	17114	81	7	172	335	165	2	0
dna	31431	112	5	218	577	361	2	0
feline	49864	1.88	0.1	477	966	487	-2	2
torus3	4000	231	20	27	81	50	-4	3
fandisk	6475	1.45	0.2	66	122	58	2	0

Table 1: Morse-Smale decompositions: statistics

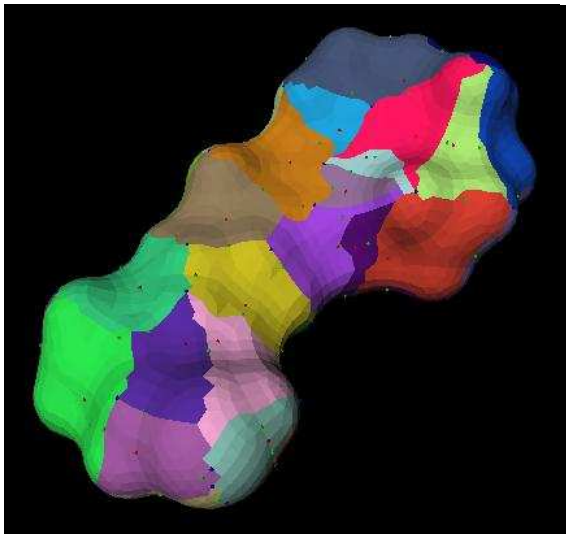


Figure 19: Retinal; stable

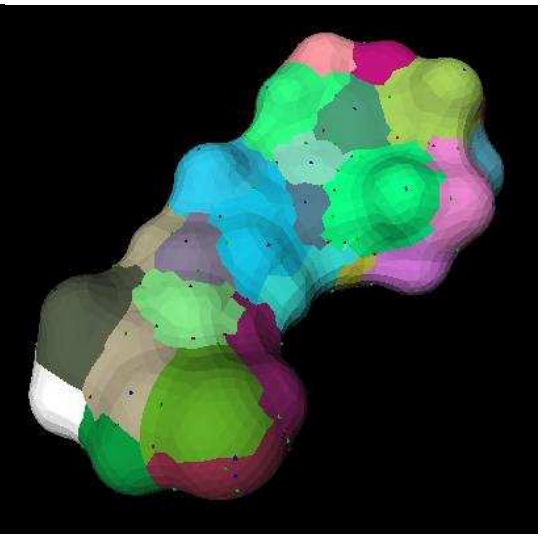


Figure 20: Retinal; unstable

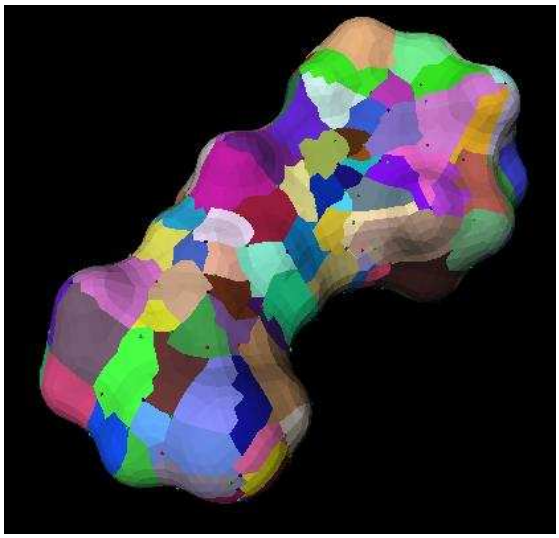


Figure 21: Retinal; mixed

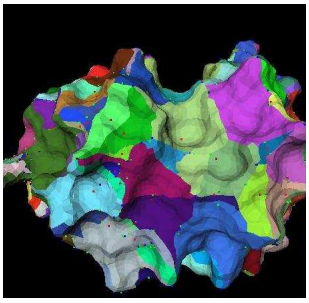


Figure 22: Papaine; stable

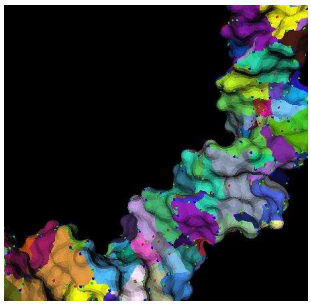


Figure 24: DNA; stable

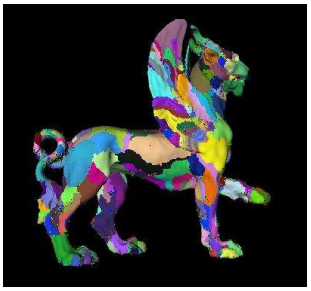


Figure 26: feline; unstable

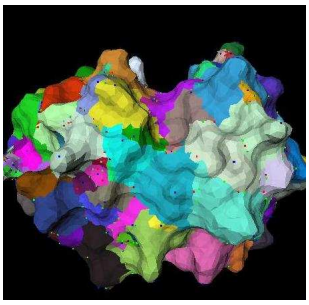


Figure 23: Papaine; unstable

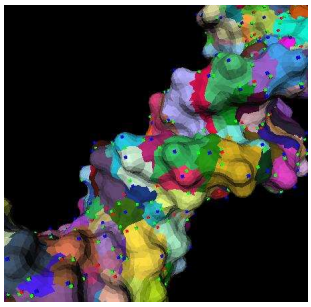


Figure 25: DNA; unstable

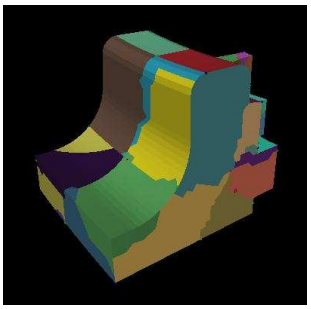


Figure 27: fandisk; mixed

6 Conclusion

In this paper we invigorated Connolly's function by recasting it into the framework of Morse theory for functions defined over two-dimensional manifolds. Contributions include a careful study of the function for smooth curves and surfaces, an algorithm to evaluate the function on a mesh, an algorithm based upon Forman's discrete Morse theory to construct a discrete Morse-Smale decomposition, as well as experiments on several mesh models. Future work will address the following directions.

On the application side, an urgent task is to investigate the relevance of stable and unstable patches for docking as well as other molecular algorithms. Further investigations will aim at bridging the gap between pure geometric filters and filters incorporating physical considerations.

From an algorithmic perspective, an interesting contribution would be to be able to perform Morse-Smale decompositions for molecular surface models more accurate than piecewise linear surfaces —e.g. spherical harmonics. The relevance of Morse-Smale diagrams for the

general problem of comparing shapes —a ubiquitous problem in computer vision— is also posed.

On a more mathematical perspective, several issues deserve investigation. The fact that Connolly's function is close to the mean curvature function is of interest due to the connexion between mean curvature, minimal surfaces and evolution equations. An interesting question is also to check whether or not the Connolly function is generically a Morse function.

References

[1] M. Berger and B. Gostiaux. *Géométrie différentielle : variétés, courbes et surfaces (2^{nd} edition)*. PUF, 1992.

[2] V. Borrelli, F. Cazals, and J.-M. Morvan. On the angular defect of triangulations and the pointwise approximation of curvatures. In *Curves and Surfaces*, St Malo, France, 2002. INRIA Research Report RR-4590.

[3] J. Bruce and P. Giblin. *Curves and singularities (2nd Ed.)*. Cambridge, 1992.

[4] W. Cai, M. Zhang, and B. Maigret. Protein-ligand recognition using spherical harmonic molecular surfaces: towards a fast and efficient filter for large virtual throughput screening. *J. Molecular Graphics and Modeling*, 53(01):1–16, 2001.

[5] M. Connolly. Molecular surfaces: A review. *Network Science*, ??, 1996.

[6] M. Connolly. Molecular surfaces: A review. *Network Science*, 14, 1996.

[7] M. L. Connolly. Measurement of protein surface shape by solid angles. *J. Mol. Graphics*, 4, 1986.

[8] M. L. Connolly. Shape complementarity at the hemoglobin a1b1 subunit interface. *Biopolymers*, 25, 1986.

[9] M. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.

[10] M. Eck, T. DeRose, T. Duchamp, H. Hoppe, T. Lounsbery, and W. Stuetzle. Multiresolution analysis of arbitrary meshes. In *ACM Siggraph*, 1995.

[11] H. Edelsbrunner, M. Facello, and J. Liang. On the definition and the construction of pockets in macromolecules. *Discrete Appl. Math.*, 88:83–102, 1998.

[12] H. Edelsbrunner, J. Harer, and A. Zomorodian. Hierarchical Morse complexes for piecewise linear 2-manifolds. In *Proc. 17th Annu. ACM Sympos. Comput. Geom.*, pages 70–79, 2001.

[13] A. Fomenko and T. Kunii. *Topological Modeling for visualization*. Springer, 1997.

[14] J. Giesen and M. John. A dynamical system for disks in the plane. *Preprint*, 2002.

[15] P. W. Hallinan, G. Gordon, A. Yuille, P. Giblin, and D. Mumford. *Two-and Three-Dimensional Patterns of the Face.* A.K.Peters, 1999.

[16] I. Halperin, B. Ma, H. Wolfson, and R. Nussinov. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins*, 47(4), 2002.

[17] R. Forman. A discrete Morse theory for cell complexes. In S. T. Yau, editor, *Geometry, Topology and Physics for Raoul Bott.* International Press, 1995.

[18] R. Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134:90–145, 1998.

[19] T. Lewiner, H. Lopes, and G. Tavares. Optimal discrete Morse functions for 2-manifolds. preprint, 2001.

[20] T. Lewiner, H. Lopes, and G. Tavares. Visualizing Forman's discrete vector field. In H.-C. Hege and K. Polthier, editors, *Mathematical Visualization III.* Springer, Berlin, 2002.

[21] E. Koutsofios and S. C. North. Drawing graphs with dot. Technical report, AT&T Bell Laboratories, Murray Hill, NJ, 1993.

[22] L. Kettner. Designing a data structure for polyhedral surfaces. In *Proc. 14th Annu. ACM Sympos. Comput. Geom.*, pages 146–154, 1998.

[23] J. W. Milnor. *Morse Theory.* Princeton University Press, Princeton, NJ, 1963.

[24] R. Norel, , H. Wolfson, and R. Nussinov. Small molecule recognition: Solid angles surface representation and molecular shape complementarity. *Combinatorial Chemistry & High Throughput Screening*, 2, 1999.

[25] R. Norel, S. L. Lin, H. Wolfson, and R. Nussinov. Shape complementarity at protein-protein interfaces. *Biopolymers*, 34, 1994.

[26] J. Palis and W. de Melo. *Geometric Theory of Dynamical Systems.* Springer, 1982.

[27] I. Porteous. *Geometric Differentiation (2nd Edition).* Cambridge University Press, 2001.

[28] F. M. Richards. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.*, 6:151–176, 1977.

[29] S. Smale. Morse inequalities for a dynamical system. *Bulletin of the AMS*, 66:43–49, 1960.

[30] G. Smith and M. Sternberg. Prediction of protein-protein interactions by docking methods. *Curr. Opin. Struct. Biol.*, 12(1), 2002.

[31] R. E. Tarjan. Efficiency of a good but not linear set union algorithm. *Journal of the ACM*, 22(2):215–225, 1975.

[32] R. E. Tarjan. *Data Structures and Network Algorithms.* Society for Industrial and Applied Mathematics, Philadelphia, 1983.

[33] R. Thom. Sur une partition en cellules associée à une fonction sur une variété. *Comptes Rendus de l'Académie de Sciences*, 228:973–975, 1949.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399