Détection de structure géométrique dans les nuages de points

Quentin Mérigot

UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS

ÉCOLE DOCTORALE STIC
SCIENCES ET TECHNOLOGIES
DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE

pour obtenir le titre de Docteur en Sciences
de l'Université de Nice Sophia-Antipolis
Mention: Informatique

présentée et soutenue publiquement le 10 décembre 2009 par

Quentin MÉRIGOT

Détection de structure géométrique dans les nuages de points

Rapporteurs :

M. Herbert Edelsbrunner	PR	IST Austria
M. Jean-Michel Morel	PR	École normale supérieure de Cachan

Composition du Jury :

M. Frédéric Chazal	DR	INRIA Saclay	*Directeur*
M. David Cohen-Steiner	CR	INRIA Sophia-Antipolis	*Directeur*
M. Jérôme Dedecker	MCF	Université Paris 6	*Examinateur*
M. Yann Ollivier	CR	École normale supérieure de Lyon	*Examinateur*
M. Jean-Michel Morel	PR	École normale supérieure de Cachan	*Rapporteur*
M. Ludovic Rifford	PR	Université de Nice Sophia-Antipolis	*Président*

Remerciements

Je souhaiterais commencer par remercier Frédéric Chazal et David Cohen-Steiner pour leur encadrement durant ces trois années. Le sujet sur lequel il m'ont aiguillé m'a permis de découvrir les belles mathématiques pures et appliquées qui interviennent en inférence géométrique J'ai également beaucoup apprécié la grande liberté qu'ils m'ont laissé, et qui m'a permis de mûrir tranquillement mes idées.

En second lieu, je voudrais exprimer ma reconnaissance à Herbert Edelsbrunner et à Jean-Michel Morel, qui ont bien voulu accorder un peu de leur emploi du temps que je sais très chargé à une lecture attentive de cette thèse. Un remerciement également à Jérôme Dedecker, Yann Ollivier et Ludovic Rifford qui m'ont fait l'honneur et le plaisir de participer au jury de soutenance.

Ensuite, je voudrais remercier tous les membres permanents de l'équipe Géométrica pour leur accueil — j'y ai trouvé un cadre très agréable et propice au travail. Un mot aussi pour les doctorants, stagiaires, post-doctorants, ingénieurs, et en particulier pour les participants plus ou moins assidus de la rituelle pause café de quatre heure: Pooran, Tom, Jane, Nader, Pedro, Stéphane, Marie, Camille, Sébastien. Je ne peux pas oublier Agnès et Caroline, dont l'aide a été plus que précieuse, et très appréciée.

J'ai eu la chance de pouvoir effectuer deux séjours longs à l'université Stanford au cours de ces trois ans, ce qui m'a permis (en plus de découvrir la Californie!) d'amorcer des projets avec Maks Ovsjanikov et Dmitriy Morozov. Ces deux séjours ont été rendus possible grâce à l'équipe associée Stanford/INRIA TGDA, créée par Steve Oudot et Leo Guibas en 2008: je les en remercie chaleureusement.

Je finirais par un salut à tous les amis et proches qui liraient ces remerciements, et qui se reconnaitront: matheux et "psychologues" de Nice, collègues de l'ENS Lyon, maintenant parisiens et grenoblois, et amis plus anciens dont ces six années à Lyon et à Nice m'ont parfois un peu éloignés. *Last but not least*, bravo à la famille — et encore plus à Laurine — d'avoir réussi à me supporter dans les derniers mois, un peu éprouvants il faut bien l'avouer, de rédaction !

Palo Alto, le 10 mars 2010

Sommaire

INTRODUCTION GÉNÉRALE

L'EXTRACTION d'information géométrique et topologogique à partir de donnée géométriques — comme par exemple les nuages de points 3D mesurés par un scanner laser — est une étape préliminaire à l'application de nombreux algorithmes de traitement géométrique et d'analyse de donnée. Le besoin d'invariants géométriques robustes s'est fait ressentir en traitement de la géométrie 3D (*geometry processing*) depuis un certain temps, et ces invariants ont trouvé des applications dans des domaines aussi variés que la comparaison de formes, la détection de symmétrie, mais aussi la reconstruction et le maillage. Plus récemment, il est devenu apparent que de telles quantités géométriques et topologiques pouvaient jouer un rôle dans l'analyse de jeux de données plus généraux, provenant de biologie structurale, de larges bases de données d'images, etc. Certaines techniques développées en topologie algorithmiques ont ainsi été utilisées pour estimer le nombre de composantes connexes de ce type de jeu de données, ce qui s'avère être une information indispensable à de nombreux algorithmes de *clustering*. La plupart des jeux de donnéees ayant un réel contenu géométrique sont concentrés près de sous-ensembles de très petite dimension intrinsèque. C'est le cas par exemple lorsqu'on considère l'ensemble des configurations réalisables d'un système mécanique, qui à cause des contraintes est souvent bien plus petit que l'espace des paramètres. De même, l'ensemble des images 2D n par n d'une même scène 3D prises à partir différents points de vue, bien que naturellement plongé dans $\mathbb{R}^{n\times n}$ a clairement une dimension intrinsèque bien plus petite (de l'ordre de grandeur de la dimension du groupe des isométries de l'espace euclidien). Le but des techniques de réduction de dimension non-linéaire est de proposer des paramétrisations de tels nuages de points, dépendant de peu de paramètre. De telles techniques pourraient bénéficier d'une estimation préalable de la dimension intrinsèque du nuage de points. Ces quelques applications motivent l'extension des questions d'estimation géométrique à des nuages de points plongés dans un espace de dimension plus grande que 3.

Une telle extension est rendue difficile par la présence de points aberrants, la "malédiction de la dimension", et la difficulté d'estimer visuellement la qualité des résultats.

De nombreux travaux se sont intéressés à l'estimation de quantités géométriques à partir d'un échantillonage sans bruit d'un objet physique 3D, et cette question est maintenant assez bien comprise. Si la surface de l'objet est lisse et compacte, il est non seulement possible d'estimer de telles quantités, mais il est aussi possible d'obtenir une *reconstruction* fidèle de la surface sous-jacente. On appelle reconstruction d'un nuage de points une surface implicite, ou un maillage, proche (au sens Hausdorff) du nuage. Sans hypothèse supplémentaire sur l'échantillonage et sur l'objet sous-jacent, le problème la reconstruction de surface est très mal posé. Si on sait que le nuage de points est densément échantilloné *sur* une surface peu courbée, une bonne reconstruction (au sens $\mathcal{C}^1$) peut être obtenue à partir d'un sous-complexe de la triangulation de Delaunay du nuage de points. Le maillage obtenu est isotope à la surface sous-jacente et la déviation des normales est contrôlée. Des travaux ultérieurs ont montré que ce type de reconstruction avec de bonnes normales peut être utilisé pour estimer correctement la courbure, le volume, etc. de la surface sous-jacente. La reconstruction d'objets (même 2D) plus complexes, échantillonés avec bruit ou plongés dans un espace euclidien de dimension plus grande est par contre une question encore largement ouverte: en réalité, même pour une surface lisse par morceaux dans $\mathbb{R}^3$ échantillonée densément avec un bruit modéré, il n'y a pas d'algorithme de reconstruction connu qui garantisse une topologie correcte et une faible déviation des normales.

Le but de l'inférence géométrique est dans un certain sens beaucoup plus modeste que la reconstruction complète d'un hypothétique objet sous-jacent. La question principale est: étant donnée une approximation discrète (potentiellement bruitée) d'un objet de l'espace euclidien, sous quelles conditions est-il possible d'obtenir de manière fiable des informations topologiques ou géométriques sur l'objet sous-jacent en utilisant seulement l'approximation ? Le point de vue adopté dans cette thèse sur la question de l'inférence géométrique est que ce problème peut (et devrait) être considéré comme un problème symétrique: si un nuage de points est une approximation d'un compact donné, alors ce compact est aussi une approximation du nuage de points. Autrement dit, nous cherchons à ne faire aucune différence dans les hypothèses faites sur le compact sous-jacent et celui qui l'approche. Ceci implique en particulier de ne considérer que les quantités géométriques définies pour *tous* les compacts de l'espace euclidien, ou alors pour une classe suffisamment large qui contienne à la fois les objets continus et leur approximation discrète. Ce point de vue est assez inhabituel; une des raisons vient de ce que les définitions classiques de quantités géométriques comme la courbure font des hypothèses très fortes de régularité.

On ne peut espérer pouvoir inférer une quantité $\mathcal{F}(K)$ dépendant d'un compact K de l'espace Euclidien que si elle est stable par perturbation Hausdorff de K. Cela signifie si K et K' sont proches au sens de la distance de

Hausdorff, alors les quantités correspondantes $\mathcal{F}(K)$ doivent également être proches. Notons que cette stabilité Hausdorff impose déjà des restrictions sur l'espace dans lequel $\mathcal{F}(K)$ prend ses valeurs: une notion de dimension Hausdorff-stable ne peut pas prendre (uniquement) des valeurs entières. Ensuite, pour déterminer le le sens géométrique d'une telle quantité $\mathcal{F}(K)$, il faut pouvoir la relier à une notion existante, par exemple dans le cas où K est un convexe, une surface lisse, un polyèdre, etc. Finalement, et parce que l'inférence géométrique est un problème pratique, il doit être possible de calculer ou au moins d'approcher $\mathcal{F}(K)$ lorsque K est un objet discrétisé, comme par exemple un nuage de points ou un maillage.

Un certain nombre de quantités de nature géométrique ou topologique Hausdorff-stables peuvent être définies en utilisant la notion de fonction distance. Dans le paragraphe suivant, nous donnons un aperçu de quelques résultats d'inférence géométrique et topologique qui reposent sur l'utilisation de la fonction distance. Avant de commencer, il faut souligner que ces méthodes ne sont robuste qu'au «bruit Hausdorff». C'est évidemment bien plus satisfaisant que de demander un échantillonage sans bruit, mais cela interdit la présence d'*outliers*[1] aussi bien en théorie qu'en pratique.

Inférence géométrique utilisant la fonction distance

Un nuage de points ne contient aucune information géom'etrique ou topologique *en lui même*. À grande échelle, il est impossible de le distinguer d'un unique point; et à une trés petite échelle, ce nuage de points ne contient pas plus d'information que son cardinal. Ce probléme du choix de l'échelle n'est pas seulement un artefact de la discrétisation de la géométrie. Alors qu'à échelle humaine, la surface d'une table en bois paraît parfaitement lisse, cela est probablement beaucoup moins vrai du point de vue d'une fourmi. Et cette même table considérée à à l'échelle atomique ne sera qu'une union de boules éloignées les unes des autres. La notion de *fonction distance* à un objet K permet d'encoder toutes ces échelles géométriques en un objet unique. Dans ce cadre, on peut considérer que l'information concernant l'objet K vu à l'échelle R est contenue dans le sous-niveau R de la fonction distance à K. Ce sous-niveau correspond à l'union des boules de rayon R centrées en un point K, et sera appelé R-voisinage tubulaire de K.

Il est bien connu (et c'est très élémentaire) que la fonction distance à un compact dépend de manière continue de celui-ci. Ainsi on peut espérer que si un nuage de points C est suffisamment proche au sens de Hausdorff d'un objet K les R-voisinages tubulaires de C et R encodent une géométrie et une topologie similaire. Il y a beaucoup de choses à préciser dans cette idée pour la rendre utilisable: il faut définir les quantités qu'on peut espérer retrouver dans ces voisinages tubulaires, déterminer les conditions d'échantillonnage (ou de proximité Hausdorff) qui permettent de retrouver cette information, et finalement proposer des algorithmes permettant d'effectuer des calculs

[1] c'est-à-dire des erreurs de mesures, des points aberrants situés loin de l'objet échantillonné

utilisant voisinages tubulaires de nuages de points. Nous mentionnons maintenant certains de ces résultats.

Il a été démontré que le type d'homotopie d'une variété peut être retrouvé en considérant certains voisinages tubulaires d'un nuage de points proche de la variété, sous des hypothèse d'échantillonage semblables de celles utilisées en reconstruction de surface [NSW06]. Cette condition est satisfaite si la densité des échantillons est plus grande qu'une certaine constante multipliée par la *portée*[2] de K. Comme la portée devient nulle dés que l'objet comporte des singularités (arêtes vives), même un simple polyhèdre ne possède pas déchantillon fini vérifiant cette condition! Dans [CCSL09, CCSL08], les auteurs proposent une condition d'échantillonage moins restrictive sous laquelle la topologie de K peut être déduite de la topologie d'un voisinage tubulaire du nuage de points. De manière très vague, une surface admettra un échantillon fini vérifiant cette condition si les angles dihèdres de ses singularités sont minorés.

Au cours de la dernière décennie la théorie de la persistence topologique a été developpée afin d'estimer de manière robuste la topologie, ou plus précisément le type d'homologie, des sous-niveaux d'une fonction (cf [EH08]). Cette théorie associe à tout compact de $\mathbb{R}^d$ et toute dimension k un sous-ensemble du plan appelé diagramme de persistence. Ce diagramme encode l'évolution des nombres de Betti de dimension k du α-sous niveau de la fonction distance à K, lorsque α décrit l'ensemble des réels positifs. Une propriét'e cruciale des diagrammes de persistence est leur stabilité par perturbation Hausdorff de K [CSEH07]. Ce résultat de stabilité ne fait aucune hypothèse sur la nature des compacts (que ce soit l'ensemble sous-jacent ou son approximation discrète), et constitue un des premiers exemples d'inférence géométrique «symétrique» au sens défini un peu plus haut.

Il est possible algorithmiquement de construire un complexe simplicial possédant le même type d'homotopie que le α sous-niveau de la fonction distance. Ce complexe, appelé *α-complex* [Ede95] est obtenu comme sous-complexe de la triangulation de Delaunay, et est un deformation-retract du voisinage tubulaire P^α. Comme la taille de la triangulation de Delaunay dépend de manière exponentielle de la dimension ambiente, le calcul de cet α-complexe n'est pas réalisable en grande dimension. Mais même dans ce cas, on peut tout de même obtenir une approximation des nombres de Betti du voisinage tubulaire *via* le complexe de Vietoris-Rips du nuage de points [CO08].

Contributions

Comme nous l'avons vu, il existe une littérature abondante pour l'inférence de la topologie d'un objet échantilloné, même lorsqu'il est non lisse ou plongé dans un espace Euclidien de grande dimension. En revanche, on ne sait que très peu de choses en ce qui concerne l'inférence de la géométrie — que ce soit la courbure, la dimension intrinsèque ou le lieu de certaines singularités.

[2] notion plus connue sous les noms anglais de *reach* ou *minimum local feature size*

Les chapitre I–III, sont consacrés à l'estimation de propriétés différentielles en utilisant la fonction distance, tandis que dans le chapitre IV on cherche à étendre les méthodes utilisant la fonction distance au cas où le jeu de données peut contenir des erreurs de mesures (*outliers*). Avant de donner un bref aperçu des résultats principaux de cette thèse, je souhaiterais mettre en avant ce que je pense être les contributions de ce travail au domaine de l'inférence géometrique, dans un sens assez large.

L'exploitation de la *semi-concavité* de la fonction distance à un compact est un des fils directeurs de cette thése. Cette propriété de régularité traduit le fait que la fonction distance est presque $\mathcal{V}^2$, et que ses singularités sont celles d'une fonction concave. La semi-concavité de la fonction distance a été remarquée et largement utilisée en géométrie et en analyse, mais ce n'est que récemment qu'elle a commencé à être utilisée (implicitement) en inférence géométrique. Nous avons apporté notre pierre à cet édifice. Dans le premier chapitre, on exploite la semi-concavité de la fonction distance pour obtenir des bornes assez fines sur le volume d'une filtration de l'axe médian d'un compact. On s'en sert également pour démontrer un théorème de stabilité L^1 pour des dérivées de premier ordre de la fonction distance, comme la fonction de projection sur un compact — qui envoie un point sur son plus proche voisin dans ce compact. Finalement, nous montrons comment certains résultats existants sur la topologie de certains voisinages tubulaires peuvent être obtenus pour les sous-niveaux de certaines fonctions dont le carré est 1-semiconcave (Chapitre IV).

La plupart des méthodes existantes pour l'estimation de quantités différentielles comme les normales ou la courbure cherchent à obtenir des estimations ponctuelles. Par exemple, on pense souvent à un champs de normal sur une surface de $\mathbb{R}^3$ comme une fonction de cette surface dans l'espace des droites. Cette approche pose de nombreux problèmes. Tout d'abord, lorsque l'objet sous-jacent n'est pas lisse, il peut y avoir non pas une unique normale en chaque point mais tout un cône convexe — souvent appelé cône normal. Cela refléte une concentration infinie de normales, qu'une simple fonction ne permet pas de traduire. En second lieu, même ces cônes normaux ne sont pas ponctuellement Hausdorff-stables — ne serait-ce qu'à cause des translations de l'objet. Finalement, la possibilité d'ajouter du bruit de haute fréquence rend nécessaire le choix d'une échelle à laquelle on regarde notre objet. Une des contributions de ce travail est l'introduction de mesures (au sens de Lebesgue) qui encodent l'information contenue dans la distribution des cônes normaux à une certaine échelle (Chapitres II–III). Tout théorème de stabilité par approximation Hausdorff de ces mesures nécessite de prendre en compte ces trois phénomènes. Dans ce contexte, l'utilisation des notions de distances entre mesures définies par le tranport optimal (distances de Wasserstein) s'est révélé extrêmement utiles.

Les méthodes d'inférence utilisant la fonction distance, comme celles des chapitres II–III, ne fonctionnent pas en présence d'*outliers* dans le jeu de donnée. Même un seul *outlier* peut perturber suffisamment la fonction distance pour rendre inopérant tout résultat d'inférence qui l'utilise. La

difficulté principale est qu'en acceptant l'existance de telles erreurs de mesure, on quitte le cadre de la géométrie pure (cadre Hausdorff) pour un cadre où la quantité, et non plus seulement la position, des points compte. Notre proposition pour prendre en compte ce fait est de ne plus considérer nos objets comme des compacts, mais de les considérer comme des distributions de masse (mesures de probabilité), auxquelles on associe une géométrie. L'objet sous-jacent ne sera plus (par exemple) une surface, mais la mesure uniforme sur cette surface; notre nuage de points sera remplacé lui aussi par la mesure uniforme qu'il supporte; etc. On remplace également la distance de Hausdorff par la distance de Wasserstein. Dans le chapitre IV nous montrons comment étendre la plupart des résultats «classiques» d'inférence géometrique à ce nouveau cadre, via l'introduction d'une notion de fonction distance à une mesure de probabilité.

CHAPITRE I

La notion d'axe médian d'un compact est couramment utilisée en reconnaissance de formes, reconstruction, segmentation, etc. Dans ce chapitres, nous suvolons les principaux résultats de régularité et de dimension; cela complète, dans une autre direction, le *survey* de Attali, Boissonnat et Edelsbrunner. Dans une deuxième partie, nous donnons des bornes optimales sur le $(d-1)$-volume (et mieux, sur les nombres de recouvrement de Lebesgue) de certaines parties de l'axe médian. Ce résultat sera utilisé plus tard dans le texte.

CHAPITRE II

Dans ce chapitre, nous exploitons l'information géométrique contenue dans la croissance du volume des voisinages tubulaires K^r d'un compact K. Les formules du tube montrent que si ce compact est convexe, lisse, ou de portée positive, ce volume est un polynôme de degré d en r, pour r suffisamment petit. Les coefficients de ce polynôme permettent de retrouver certaines courbures intégrales de K, sa dimension intrinsèque, et même caractéristique d'Euler-Poincaré (théorème de Gauss-Bonnet généralisé). Les *mesures de courbure* de K, introduites par Federer, détaillent la contribution de chaque partie de K à la croissance de ses voisinages tubulaires et donc de localiser les informations données par la formule du tube.

Notre principal résultat est un théorème de stabilité des mesures de courbure d'un compact K lorsque celui-ci est remplacé par une approximation Hausdorff K. Un aspect intéressant de ce théorème est de ne faire aucune hypothèse géométrique sur C, ce qui le distingue de résultats similaires en géométrie Riemannienne qui nécessitent une borne supérieure sur la courbure et une borne inférieure sur le rayon d'injectivité. Finalement, un algorithme Monte-Carlo très simple à implanter permet de calculer les mesures de courbure (approchées) et mesures de bord d'un nuage de points.

En appendice, nous demandons si l'existence d'une formule du tube (approchée) peut-être utilisée pour estimer la portée d'un compact. Cette quantité, qui est en particulier une borne inférieure sur les rayons de courbures

dans le cas d'une variété, est utilisée par de nombreux algorithmes de reconstruction. Dans cette direction, nous obtenons de manière plus élémentaire un résultat de Hug, Heveling and Last.

CHAPITRE III

La détection (robuste) d'arètes vives dans un maillage est maintenant assez bien comprise; pour ce qui est des nuages de points en revanche, les méthodes existantes sont très heuristiques et expérimentales. En particulier, l'analyse de la convergence (ou des conditions d'échantillonage) n'est jamais donnée, et très souvent une hypothès d'échantillonage uniforme est implicitement faite. Dans ce chapitre nous introduisons et étudions une version anisotrope des mesures de bord que nous appelons mesure de covariance de Voronoi (VCM), qui peut être utilisee pour la détection d'arêtes vives. Sous une hypothèse similaire á celle de portée positive, la VCM d'un polyèdre permet de retrouver ses arêtes vives. En utilisant la stabilité Hausdorff de cette mesure qui se déduit des mêmes arguments que ceux utilisés dans le chapitre précédent, on obtient une méthode prouvée de détection d'arête vive á partir d'une approximation Hausdorff.

Nous expliquons comment la VCM peut être calculée de manière efficace en 3D, et expérimentons notre méthode sur plusieurs nuages de points, sous différentes conditions de bruit et d'échantillonage.

CHAPITRE IV

Nous introduisons une notion de fonction distance à une mesure de probabilité qui est stable par approximation Wasserstein. Cette fonction distance permet d'associer de manière naturelle une géométrie et une topologie à la mesure, via ses sous-niveaux. Nous montrons que cette fonction distance à une mesure possède toutes les propriétés de la fonction distance usuelle utilisées par un certain nombres de résultats d'inférence géométrique. Ceci permet d'étendre, parfois sans changer les preuves, ces résultats au nouveau cadre que nous proposons. Finalement, le calcul de la topologie et de la géométrie de certains sous-niveaux de ces fonctions distances soulèves plusieurs questions algorithmiques, que nous discuterons.

Publications

— *Size of the medial axis and Federer's curvature measures.*, Q. Mérigot. (soumis)

— *Boundary measures for geometric inference.*, F. Chazal, D. Cohen-Steiner, et Q. Mérigot. Found Comput Math **10**, 221-240 (2010)

— *Robust Voronoi-based Curvature and Feature Estimation*, Q. Mérigot, M. Ovsjanikov, et L. Guibas. Proc. ACM Symp. Solid Phys. Modeling, 1–12 (2009). (Prix du meilleur article)

— *Geometric inference for measures*, F. Chazal, D. Cohen-Steiner et Q. Mérigot. (soumis)

General introduction

EXTRACTING geometric and topological information from geometric data, such as 3D point clouds obtained from laser scanners, is a requirement for many geometry processing and data analysis algorithms. The need for robust estimation of geometric invariants have been recognized long times ago in geometry processing, and such invariants have found applications in fields as different as shape matching, registration, symmetry detection in 3D models or more generally structure discovery, reconstruction, meshing to name just a few. More recently, it has been apparent that such geometric and topological quantities could also be used to analyze more general data sets coming from computational structural biology, large image databases, etc. For instance, methods from computational topology have been used to estimate the number of connected components of such sets which is a useful input for many clustering algorithms. Data sets that have a real geometric content are often concentrated around manifolds (or more general sets) of very low dimension, sometimes several order of magnitude smaller than the ambient dimension. Examples of this situation arise when considering configurations of a mechanical system: because of constraints, the dimension of the space of transformations is usually much larger than the space of realizable configurations. Another example: the set of n-by-n 2D pictures of a given 3D scene, taken from different viewpoints. This data set is naturally embedded into $\mathbb{R}^{n\times n}$; however, its real underlying dimension is probably much lower (close to the dimension of the group of isometries of the Euclidean 3-space). The goal of non-linear dimensionality reduction techniques, also known as "manifold learning" is to provide automatic parametrizations of such point clouds with few parameters; again, they could benefit from a preliminary reliable estimation of the intrinsic dimension. These applications motivates the extension of geometric estimation to point clouds embedded in spaces of dimension larger than 3. The lack of visual feedback, the presence of outliers, and what is referred to as the "curse of

dimensionality" are among the difficulties that are raised by this extension.

A lot of work has been done to estimate geometric quantities from a noiseless sampling of a physical object in 3D, and this question is quite well understood. If the surface of the object is a smooth compact manifold, not only is it possible to estimate geometric quantities such as those mentioned above, but also to obtain a full *reconstruction* of the underlying surface. Reconstructing a surface from a point cloud consists in building a mesh or an implicit surface that lie close to a given point cloud. Without any further assumption, surface reconstruction is not a very well-posed problem. If the surface is smooth and well behaved, and the point cloud is sampled densely enough *on* S, a reconstruction can be obtained through a subcomplex of its Delaunay triangulation. The obtained mesh then is isotopic to the underlying surface, and the deviation of the normals can be kept low. Subsequent work shows that this reconstruction with good normals can be used to estimate the curvature, volume, etc. of the underlying surface correctly. Reconstructing more complex objects, sampled with noise, or in higher-dimensional ambient space is far from being well understood. As a matter of fact, even for a piecewise smooth surface in $\mathbb{R}^3$ sampled densely with moderate noise, there is no known reconstruction algorithm that guarantees the correct topology and low normal deviation.

The goal of geometric inference is in some sense more modest than obtaining a full-fledged reconstruction of an unknown underlying shape. The main question is: we are given a (perhaps noisy) discrete approximation of an object embedded in the Euclidean space ; under what conditions is it possible to reliably recover geometric and topological informations about the underlying object knowing only the approximation ? Our personal view of the subject is that geometric inference is a "symmetric" problem: if a point cloud is an approximation of a compact set, then the compact set is also certainly an approximation of the point cloud! Said otherwise, we prefer to make no difference between the input data and the unknown underlying set. This motivates us to only consider those geometric quantities that are defined for *every* compact subsets of the Euclidean d-space. This viewpoint is quite unusual, mostly because almost no "classical" geometric quantity such as curvature or dimension are defined regardless of any regularity assumption.

The most important requirement to be able to infer a quantity $\mathcal{F}(K)$ that depends on a compact subset K of an Euclidean space is its *Hausdorff-stability*. This means that if K and K' are close in the Hausdorff sense, then the quantities $\mathcal{F}(K)$ and $\mathcal{F}(K')$ should also be close, in a sense that depends on their nature. Notice that this requirement already puts a restriction on the space in which $\mathcal{F}$ takes its value: a Hausdorff-stable notion of dimension *cannot* be integer-valued. In order to understand the geometric meaning of $\mathcal{F}(K)$, one can then study it for some class of subsets of $\mathbb{R}^d$ to see if it coincides or approximates an already known notion: smooth compact surfaces, polyhedron, convex sets, etc. Last but not least, because geometric inference is a practical problem, it should be possible *in practice* to compute or at least approximate $\mathcal{F}(K)$ when K is a discrete object, such as a point cloud or a mesh.

A possible way to define geometric quantities as above that are Hausdorff-stable is through the use of distance functions. In the next paragraph we give a brief overview of distance-based topological and geometric inference. Before starting, let us stress that these distance-based approach allow "Hausdorff-noise" *only*. This is already an improvement over geometric estimation methods that require noiseless sampling, but this discards the possibility of outliers in the point cloud (both in theory and practice).

Distance-based inference

By itself, a point cloud does not carry any geometric or topological information. At a very large scale, there is no way to tell the difference between the cloud and a single point; and at a very fine scale, on the other hand, it is nothing but a bunch of unrelated points. The importance of the scale parameter is not only for point clouds: at our scale, the surface of a wooden table is smooth; from the point of view of an ant, it is probably much less clear — not to speak of the atomic level. The *distance function* to an object K in the Euclidean space is a very convenient way to encode these scales in a single way. With this point of view, the object seen *at scale* R is simply the R-sublevel set of the distance function, or equivalently the union of balls of radius R centered at the points of K. We call this sublevel set the R-offset of K.

It is very well-known (and quite elementary, too) that the distance function depends continuously on the underlying compact set. As a consequence, if we are given a point cloud C sampled close to a compact set K, the R-offsets of K and C are also geometrically and topologically close to each other — provided that the Hausdorff distance between K and C is slightly smaller than the offset parameter. All of this needs of course to be made precise: this is the object of distance-based geometric and topological inference. There are two almost orthogonal questions. The first one is to determine what information can be stably extracted from offsets, with (preferably) quantitative Hausdorff closeness assumptions. The second one is about the design of efficient algorithms to deal with offsets of point clouds.

Distance functions and offsets have been used extensively in topological inference. Under sampling conditions similar to the one used in 3-dimensional surface reconstruction, it has been shown that appropriate offsets of a point cloud sampled on a smooth submanifold S of $\mathbb{R}^d$ had the same homotopy type as the underlying submanifold [NSW06]. The condition is satisfied if the density of the sample points is greater than some constant multiplied by the inverse of the *reach* of S. Since the reach vanishes on a sharp feature, even a mere polyhedron does not admit a finite sample satisfying this condition! In [CCSL09, CCSL08], the authors introduced a less restrictive sampling condition under which the topology of the underlying set can be estimated from the topology of the offsets of the sample points. Very loosely speaking, a compact set K will admit a finite sample satisfying this condition if the dihedral angle of its sharp features is bounded from below.

In the last decade, the theory of persistent homology has been developped

to estimate the topology of sublevel sets of functions in a robust way (see [EH08] for a survey). This theory can be used to associate to any compact subset K of $\mathbb{R}^d$ and any dimension k a finite subset of the plane, called a persistence diagram. This diagram encodes the evolution of the Betti numbers of dimension k of the α-sublevel sets of the distance function to K as α goes from zero to infinity. The crucial property is that persistence diagrams are stable under Hausdorff approximation [CSEH07]. This stability result makes no assumption on the nature of the compact sets (the underlying set and its discrete approximation), and is one of the first interesting example of "symmetric geometric inference" as we defined a few paragraphs earlier.

Algorithmically, the homotopy type of the α-offset of a point cloud P can be obtained through the α-complex of P. The α-shape is subcomplex of the Delaunay triangulation of P, and is proven to be a deformation-retract of the offset P^α [Ede95]. Computing the alpha-shape of a data set P requires computing the whole Delaunay triangulation of P, and is impractical in high dimension. On the other hand, approximate persistent homology computations can be done using the so-called Vietoris-Rips complex [CO08], thus avoiding the exponential growth of α-shape or Čech complexes with the ambient dimension.

Contributions

As described above, substantial progress has been made for inferring topological invariants of possibly non-smooth sampled objects embedded in arbitrary dimensional spaces. However, very little is known on how to infer more geometric invariants, such as curvature, dimension or singularities, for such objects. Chapters I–III are devoted to distance-based geometric inference to estimate differential properties, while Chapter IV is devoted to the extension of distance-based geometric inference to handle point clouds that have been corrupted by outliers. Before we give a brief overview of the organization of the thesis, let us highlight what we believe are the three main contributions of this work in a very broad sense.

The *semiconcavity* of the distance function to a compact set can very well be considered as a guiding thread in our thesis. Semiconcavity is a smoothness property that ensures that the distance function is not only practically $\mathcal{C}^1$, as implied by the Lipschitz property, but also practically $\mathcal{C}^2$. Semiconcavity of the distance function has been remarked and used for a long time in some area of analysis and geometry[3], but it is only quite recently that it started being (implicitely) used in geometric inference. The identification of its central role in existing distance-based geometric inference results probably counts as one of our contribution to the topic. We use it to obtain fine bounds on the $(d-1)$-volume of some refinements the medial axis of K in Chapter I. It is also used to prove L^1 stability results for first-order derivatives of the distance function such as the *projection function* on the compact set K, which maps a point to its closest neighbor in K (Chapter II–III). Finally, the 1-semiconcavity is the

[3] e.g. in Hamilton-Jacobi theory, in the study of Alexandrov spaces, in optimal transportation

main requirement for a *distance-like* function, a notion that we introduce and use in Chapter IV to extend "classical" distance-based inference to a more general setting.

The goal of most existing work on estimating differential quantities such as normals and curvature in geometry processing is to obtain pointwise estimates. For instance, a normal field is often thought of as a function from the surface to the space of directions. There are several problems with this approach. First, when the underlying shape S is not smooth, there is not necessarily a single normal to a point, but rather a convex cone of normals which is reffered to as the *normal cone*. This correspond to an "infinite concentration" of normals, which isn't allowed by a mere function. Second, even normal cones are not pointwise stable under Hausdorff approximation — if only because the domain of the function can change. Finally the possibility to add high frequency noise even when the approximation is restricted to manifolds makes it necessary to introduce a notion of scale. A contribution of this work is the introduction of *measures* in the sense of Lebesgue which encode the information provided by the distribution of normal cones at a certain scale (Chapters II–III). The stability theorems under Hausdorff approximation need to take into account these three phenomena. In this context, the use of distances on the space of measures related to optimal transportation — such as the bounded-Lipschitz distance or Wasserstein distances — proved very helpful.

All geometric inference results based on distance functions (such as those of Chapter II–III) do not work at all if the input data cloud is corrupted with outliers. The reason is that the addition of even a single outlier changes the distance function dramatically. Outliers are not easy to deal with in practice; they are also not easy to define from a theoretical viewpoint. The main reason is that by accepting outliers, one leaves the world of purely geometric approximation for some other place where the amount of sample points is of importance. Indeed, it is possible to deal with outliers only if there are much less of them than actual data points. Our proposition to deal with this, and define what "much less" means, is to consider the input point cloud no more as a geometric object, but as a probability measure, to which one attaches geometry. The underlying object is no longer a surface (or a more complex compact set), but *eg.* the uniform measure on this surface. The Hausdorff condition is then replaced by a meaningful distance on probability measures: the (exponent 2) Wasserstein distance. Under this metric, the uniform measure μ on a point cloud C, and the uniform measure $\tilde{\mu}$ on the same point cloud corrupted with a few outliers are close. In Chapter IV we show how this "classical" distance-based geometric inference can be carried out in this new setting.

CHAPTER I

The *medial axis* of a compact set K is a very common object in shape recognition, reconstruction, shape segmentation. In this chapter, we survey results on its local regularity and dimension, which complements the survey

on the stability and computations of medial axes by Attali, Boissonnat and Edelsbrunner. These results are obtained by using the convexity – or more precisely semiconcavity – properties of the distance function. In a second part of the chapter, we give an optimal bound on the $(d-1)$-volume of a refinement of the medial axes, that will be used later in the thesis.

CHAPTER II

In this chapter, we exploit the geometric information contained in the growth of the volume of the offsets of a compact set K in the d-dimensional Euclidean space in order to retrieve information about the geometry of K. The tube formula states that for positive reach compact sets (a class that includes convex subsets and smooth compact submanifolds of $\mathbb{R}^d$), this volume is a degree d polynomial in r provided that r is small enough. The coefficients of this polynomial encode important geometric information about K, such as dimension, curvatures, or even Euler characteristic. Federer introduced the notion of *curvature measures* of K, which loosely speaking details the contribution of each part of K to the volume of offsets, and hence gives local information about curvatures, dimensions, and sharp features angles.

Our main result is a stability result for curvature measures of K under Hausdorff approximation by C. This result does not make any assumption on the smoothness of C, which is in contrast with most Hausdorff stability results under upper curvature bounds and lower bounds on the injectivity radius — including Federer's own stability theorem for curvature measures. Finally, boundary measures and curvature measure can be computed in practice using a very simple Monte-Carlo algorithm, which makes them good candidate for geometric inference.

In an appendix, we ask whether the existence of an (approximate) tube formula can be used for estimating the *reach* of a compact set, a quantity that measures the "extrinsic smoothness" of the compact set, and is widely used in provable reconstruction algorithms. In particular, we obtain a simpler proof of a result by Hug, Heveling and Last.

CHAPTER III

Robustly detecting sharp features in meshes is quite well understood; for point clouds on the other hand, existing methods are very heuristic and experimental – in particular the analysis of the convergence (*ie.* the sampling conditions) is never given, and very often uniform sampling is an implicit requirement. This chapter is devoted to an anisotropic version of boundary measures (called Voronoi covariance measure, or VCM) that can be used for detecting sharp features and estimate sharp feature directions. Under a condition similar to positive reach, the VCM of a polyhedron P can be used to detect sharp edges in P. Thanks to the Hausdorff stability of VCM, that follows from the same arguments developped in Chapter II, the VCM of a point cloud which is also Hausdorff-close to P will retain the information on sharp edges. We explain how to implement approximate computations of the

VCM in an efficient way in 3D, and study the methods experimentally by applying them on various point cloud models under noise and strong sampling bias.

CHAPTER IV

We introduce a notion of distance function to a probability measure, that is stable under Wasserstein approximation. This distance-to-measure function naturally associates a geometry and a topology to every probability measures on $\mathbb{R}^d$, through the sublevel sets of the function. We show that the distance-to-measure functions shares all features of the usual distance functions (such as semiconcavity) which are used in the proof of many geometric inference results. This allows to extend, sometimes almost verbatim, these results to the new setting. Finally, this new setting raises some interesting computational issues, which we discuss.

Publications

— *Size of the medial axis and Federer's curvature measures.*, Q. Mérigot. (submitted)

— *Boundary measures for geometric inference.*, F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Found Comput Math **10**, 221-240 (2010)

— *Robust Voronoi-based Curvature and Feature Estimation*, Q. Mérigot, M. Ovsjanikov, and L. Guibas. Proc. ACM Symp. Solid Phys. Modeling, 1–12 (2009). (Best paper award)

— *Geometric inference for measures*, F. Chazal, D. Cohen-Steiner and Q. Mérigot. (submitted)

Chapter I

REGULARITY AND SIZE OF THE MEDIAL AXIS

Abstract

In this short chapter, we introduce some of the objects that will be used throughout the thesis: the *distance functions*, *projection functions* and *medial axes*. We survey some known result concerning the local regularity of medial axis (eg. Hausdorff dimension, rectifiability), a few of which come from the semi-concavity of the distance function. The compilation of these results complements the survey on the stability and computation of medial axes by Attali, Boissonnat and Edelsbrunner.

In the second part of the chapter, we study the $(d-1)$-volume and the covering numbers of the medial axis of a compact set. In general, this volume is infinite; however, the $(d-1)$-volume and covering numbers of a filtered medial axis (the μ-medial axis) that is at distance greater than ε from the compact set will be explicitely bounded. The behaviour of the bound we obtain with respect to μ, ε and the covering numbers of K is optimal. Let us stress that bounding the covering numbers is stronger than bounding the $(d-1)$-volume, since we cannot disregard the lower-dimensional parts of the medial axis. This result will be used in the next chapter to give a first stability statement for projection functions.

Contents

General notations. The canonical scalar product between two vectors $v, w \in \mathbb{R}^d$ is denoted by $\langle v|w\rangle = \sum_{i=1}^{d} v_i w_i$; and the Euclidean norm of a vector $v \in \mathbb{R}^d$ is $\|v\| := \sqrt{\langle v|v\rangle}$. The open ball of radius r centered at x is denoted by $B(x,r)$ while the closed ball is denoted by $\overline{B}(x,r)$. Finally, the Lebesgue measure of a Borel set $B \subseteq \mathbb{R}^d$ is denoted by $\mathcal{H}^d(B)$.

I.1 REGULARITY OF DISTANCE FUNCTIONS AND MEDIAL AXES

In this §, we review several well-known (and less well-known) properties of the distance function to a compact set K, as well as some results on the dimension and regularity of the medial axes. We also take this as an opportunity to introduce mathematical tools and notions that will be used throughout the thesis.

Estimating medial axes in a stable way has many applications. In image analysis and shape recognition, the skeleton of a shape is often used as an idealized version of the shape [SHB07], that is known to have the same homotopy type as the original shape [Lie04]. In the reconstruction of curves and surfaces from point cloud approximations, the distance to the medial axis provides an estimation of the size of the local features that can be used to give sampling conditions for provably correct reconstruction [AB99]. The flow associated with the distance function d_K to a compact set K, that flows away from K toward local maxima of d_K (that lie in the medial axis of K) can be used for shape segmentation [DGG03, CGOS09]. The reader that is interested by the computation and stability of the medial axis, with some of these applications in mind can refer to the survey [ABE07].

Distance function, offsets, Hausdorff distance. The set of compact subsets of a metric space X is denoted by $\mathcal{K}(X)$. If X is equipped with a metric d, one can associate to any compact subset K of X a function $d_K : X \to \mathbb{R}^+$, called the *distance function* to K; $d_K(x)$ is the minimum distance between x and any point in K, *ie.* $d_K(x) = \min_{y\in K} d(x,y)$. The distance function is then used to define a metric on $\mathcal{K}(X)$ called the the Hausdorff distance:

$$d_H(K,K') := \|d_K - d_{K'}\|_\infty = \sup_{x\in X} \left|d_K(x) - d'_K(x)\right|$$

Equivalently, the Hausdorff distance can be defined in term of offsets. The r-offset of K is the set of points at distance at most r of K; we denote it by $K^r = d_K^{-1}(r)$. With this notation, $d_H(K,K')$ is the minimum non-negative ε such that both $K \subseteq K'^{\varepsilon}$ and $K' \subseteq K^{\varepsilon}$. A useful fact about the Hausdorff topology is *Blaschke selection theorem*: if a metric space X is locally compact, then so is the set of compact subsets of X endowed with the Hausdorff metric.

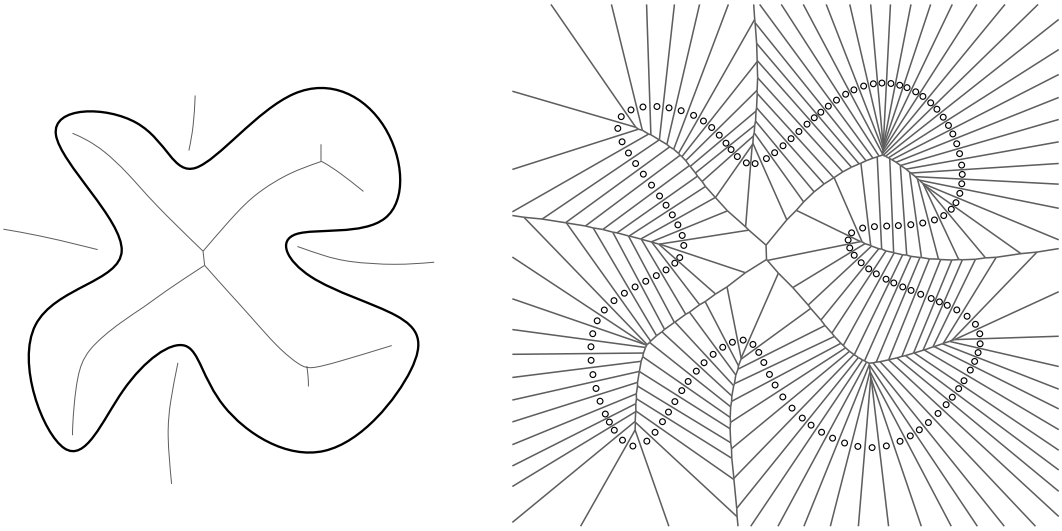


Figure I.1 – Medial axis of a curve C in the plane, and Voronoi diagram of a point cloud P sampled on the curve. Notice that while P and C are Hausdorff close, their respective medial axes are very different.

Projection, Voronoi diagram. Throughout this work, K will always denote a compact set in the Euclidean d-space $\mathbb{R}^d$, with no additional regularity assumption unless specified otherwise. A point p of K that realizes the minimum in the definition of $d_K(x)$ is called an *orthogonal projection* of x on K, or a *closest (or nearest) neighbor* of x in K. The set of orthogonal projections of x on K is denoted by $\mathrm{proj}_K(x)$. The locus of the points $x \in \mathbb{R}^d$ which have more than one projection on K is called the *medial axis* of K. Denote this set by $\mathrm{Med}(K)$. For every point x of $\mathbb{R}^d$ not lying in the medial axis of K, we let $p_K(x)$ be the unique orthogonal projection of x on K. This defines a map $p_K : \mathbb{R}^d \setminus \mathrm{Med}(K) \to K$, which we will refer to as the *projection function* on the compact set K.

EXAMPLE (Voronoi diagram). The above definitions are related to the notion of Voronoi diagram of a finite set of points $C = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d$. The algorithmic construction and complexity analysis of Voronoi diagrams and the dual Delaunay triangulations is a standard topic in computational geometry. From a geometric point of view, the Voronoi diagram of C is a decomposition of the space in n cells, one for each input point, which will be denoted $\mathrm{Vor}_C(p_1), \ldots, \mathrm{Vor}_C(p_n)$. These cells are defined by the property

$$\mathrm{Vor}_C(p_i) = \{p \in \mathbb{R}^d \,;\, \forall j \in \{1, \ldots, n\},\ d(x, p_i) \leqslant d(x, p_j)\}$$

The medial axis of the point cloud C is the union of boundaries of Voronoi cells: $\mathrm{Med}(C) = \cup_{p \in C} \partial\, \mathrm{Vor}_C(p)$. The projection function is defined on the union of the interior of Voronoi cells, and maps a point x to the center of its Voronoi cell.

Medial axis, cut locus, nerve. The medial axis defined above is also known as *ambiguous locus* in Riemannian geometry and geometry of Banach spaces [Ste63, Zam04]. A related notion is the *cut locus* of Riemannian geometry: a point $x \in M$ belongs to the cut locus of a compact set $K \subseteq M$ iff there exists a distance-minimizing geodesic from K to x that lose the minimizing property when extended beyond x. This is the case, for instance, if there exists two distinct distance-minimizing geodesics from x to K *even if they have the same endpoints*.

In the Euclidean space, the cut locus of a compact set (which has been also been studied under the name *nerve* of K, *cf.* [Riv01]) is the set of centers of open balls contained in the complement of K that are maximal for the inclusion. One has the sequence of inclusions:

$$\mathrm{Med}(K) \subseteq \mathrm{Cut}(K) \subseteq \overline{\mathrm{Med}}(K) \tag{I.1}$$

where $\overline{\mathrm{Med}}(K)$ is the closure of the medial axis. All of these inclusions can be strict even in the Euclidean space.

I.1.1 — Negligibility of the medial axis.

As we will see, the medial axis is *always* negligible both according to Baire and Lebesgue. Despite this, the medial axis of a generic compact set is dense in $\mathbb{R}^d$.

Negligibility and Genericity. Given a compact or complete metric space X, a subset $A \subseteq X$ is called *Baire-negligible* if it is a countable union of closed sets with empty interior, and *Baire-generic* if its complement is Baire-negligible. The Baire Theorem asserts that any Baire-negligible set has empty interior, while any Baire-generic subset is dense in X.

PROPOSITION I.1. *Let E be a Banach space and K a compact subset of E. Then* $\mathrm{Med}(K)$ *is Baire-negligible.*

This proposition also holds for any closed subset provided that the space E is complete and uniformly convex (this result is known as Stečkin's Theorem [Ste63]). The proof in the case we consider is simplified due to compactness, and by the use of the λ-medial axis:

Proof. Let $z \in \mathrm{Med}(K)$ and x be any projection of z on K. Then, for any point z_t in the interior of the segment $[z, x]$, the closed ball $\overline{B}(z_t, d(z_t, K)) \setminus \{x\}$ is included in the open ball $B(z, d(z, K))$ and hence cannot intersect K. This proves that z_t is not in the medial axis of K. Hence, z cannot be an interior point of $\mathrm{Med}(K)$.

Let $\mathrm{Med}^\lambda(K)$ be the set of points in $x \in \mathrm{Med}(K)$ such that $\mathrm{diam}(\mathrm{proj}_K(x)) \geqslant \lambda$. This set has obviously empty interior. Let (x_i) be a sequence of points of $\mathrm{Med}^\lambda(K)$ converging to $x \in E$. By definition of $\mathrm{Med}^\lambda(K)$, for every i there exists two points p_i^1 and p_i^2 in $\mathrm{proj}_K(x_i)$ with $\|p_i^1 - p_i^2\| \geqslant \lambda$. By compactness of K we can make the assumption that the sequences (p_i^1) and (p_i^2) both converge

to p^1 and p^2. Then, p^1 and p^2 both belong to $\mathrm{proj}_K(x)$, and $\|p^1 - p^2\| \geqslant \lambda$. This means that x belongs to the λ-medial axis $\mathrm{Med}^\lambda(K)$, which is then closed.

One concludes the proof by remarking that $\mathrm{Med}(K)$ is the countable union of the $1/n$-medial axes $\mathrm{Med}^{1/n}(K)$, with n ranging from one to infinity. □

Following the same kind of ideas, one can prove that the medial axis of a compact set K is generically dense. A similar statement for compact sets in Alexandrov spaces has been obtained in [Zam04]. The proof in the Euclidean space is very simple, and follows from a weak stability theorem for medial axes.

PROPOSITION I.2. *Let $\mathcal{K}(F)$ denote the set of compact subsets of $\mathbb{R}^d$ that are contained in a closed set F. Then, generically, the medial axis $\mathrm{Med}(K)$ of a compact set $K \in \mathcal{K}(F)$ is dense in $\mathbb{R}^d \setminus F$.*

A particular consequence of this proposition is that the compact subsets with dense medial axis are dense within the compact subsets of $\mathbb{R}^d$ endowed with the Hausdorff metric. In order to prove this statement, we need the following lemma, which can be deduced from *eg.* [CL05, Theorem 3]:

LEMMA I.3. *For any point x in $\mathrm{Med}(K)$ and $\varepsilon > 0$, there exists an $\eta > 0$ such that the medial axis of any compact set K' with $d_H(K, K') \leqslant \eta$ intersects the ball $B(x, \varepsilon)$.*

Proof of Proposition I.2. Let Ω be an open set in $\mathbb{R}^d$, and demote by $A(\Omega)$ the set of compact subsets of $\mathbb{R}^d$ whose medial axis intersect Ω. Using Lemma I.3, one can prove that $A(\Omega)$ is open in $\mathcal{K}(\mathbb{R}^d)$.

Let us now prove that $A(\Omega)$ is dense in $\mathcal{K}(\mathbb{R}^d)$. Let K be any compact set in $\mathbb{R}^d$; we will distinguish two cases. First, if K contains Ω, we denote by K_ε the set $K \setminus B(x, \varepsilon)$ ($x \in \Omega$). This set is compact for ε small enough by assumption, and x belongs to its medial axis, which proves that K is adherent to $A(\Omega)$. If however K doesn't contain Ω, there exists a point $x \in \Omega$ with $d_K(x) > 0$. For any $\varepsilon > 0$, let y_ε be a point on the sphere $S(x, d_K(x))$, at distance at most ε of K. Then $K_\varepsilon = K \cup \{y_\varepsilon\}$ is Hausdorff-close to K and x belongs to its medial axis.

Now, let $(\Omega_n)_{n\in\mathbb{N}}$ be a countable family of open sets generating the topology of $\mathbb{R}^d$. The set of compact subsets whose medial axis is dense is then exactly $\cup_n A(\Omega_n)$. This proves that the property of having a dense medial axis is generic. □

Despite being generically dense, the medial axis is always Lebesgue-negligible, *ie.* $\mathcal{H}^d(\mathrm{Med}(K)) = 0$ (this is a stronger statement than Proposition I.1). This fact can be deduced from the following proposition:

PROPOSITION I.4. *If p is an orthogonal projection of $x \in \mathbb{R}^d$ on K, then*

$$d_K^2(x + h) \leqslant d_K^2(x) + 2\langle x - p | h\rangle + \|h\|^2$$

In particular, if the distance function d_K is differentiable at x, then $\nabla_x d_K^2 = 2p$. In fact, one has:

$$\mathrm{Med}(K) = \{x \in \mathbb{R}^d \setminus K\,;\ d_K \textit{ is not differentiable at } x\}$$

Proof. Let us first remark that for any point p in K,

$$d_K^2(x+h) \leqslant \|x+h-p\|^2 = \|x-p\|^2 + 2\langle x-p|h\rangle + \|h\|^2 . \tag{I.2}$$

If moreover p be an orthogonal projection of x on K, then $\|x-p\| = d_K(x)$, thus proving the inequality. If d_K is differentiable at x, then so is d_K^2. Hence, there exists a vector $v \in \mathbb{R}^d$ s.t

$$d_K^2(x+h) = d_K^2(x) + \langle v|h\rangle + o(h) \leqslant d_K^2(x) + \langle 2p|h\rangle + o(h) \tag{I.3}$$

This inequality is possible if and only if $v = 2p$. The lemma follows. □

THEOREM (Rademacher differentiability theorem). *Any Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ is differentiable $\mathcal{H}^d$-almost everywhere.*

COROLLARY I.5. *The medial axis of any compact set $K \subseteq \mathbb{R}^d$ has zero Lebesgue measure.*

Proof. Rademacher's differentiability theorem applied to the 1-Lipschitz function d_K shows that the set of non-differentiability points of the distance function has zero Lebesgue measure. One concludes using the lemma. □

I.1.2 — Hausdorff dimension and rectifiability

In the next paragraph, we recall the definition and main properties of the Hausdorff dimension and Lebesgue covering numbers. The reader already acquainted with these two notions can safely skip it.

Hausdorff measure and Hausdorff dimension. If M is an n-dimensional smooth submanifold of $\mathbb{R}^d$, the n-volume of M can be defined either by Riemannian geometry, or more simply by tesselating M in domains that can be parametrized by open sets in $\mathbb{R}^d$ and using the change-of-variable formula to compute the volume of each of these domains.

The Hausdorff measures allow to give a more synthetic definition to the α-volume of a compact set K, that also has two additional advantages: it is defined for any (even non-integer) positive dimension α, and it does not rely on any smooth structure on K.

There are some subtleties in the definition of the Hausdorff measure of a set. The notion we consider is sometimes called *spherical Hausdorff measure*. However, for our purpose — measuring the volumes of rectifiable sets (see Def. I.5) — all definitions agree. A detailed study of the properties of Hausdorff measures can be found in most textbooks on geometric measure theory, such as [Mat95], [Mor88] or [Fed69].

DEFINITION I.1. Let A be a subset of a metric space X. For any $\varepsilon > 0$, we define the premeasure $\mathcal{H}^\alpha_\varepsilon(K) = \inf_{\mathcal{C}} \sum_{\overline{B}(x,r)\in\mathcal{C}} (2r)^\alpha$, where the infimum is taken over all coverings of the set A by closed balls of diameter at most ε.

The *α-Hausdorff measure of* A, denoted by $\mathcal{H}^\alpha(A)$ is then by definition the limit of $\mathcal{H}^\alpha_\varepsilon(A)$ as the parameter ε goes to zero. With this definition, the

d-dimensional Hausdorff measure of a (Borel) subset of $\mathbb{R}^d$ coincides with its Lebesgue measure.

DEFINITION I.2. Let A be a subset of a metric space X. For any $\varepsilon > 0$, the *ε-Lebesgue covering number* of A, denoted by $\mathcal{N}(A, \varepsilon)$ is the minimum number of closed balls of radius ε needed to cover A.

There is of course a relation between the covering number and the Hausdorff measure of a given set A: $\mathcal{H}^\alpha_\varepsilon(K)$ is always smaller than $(2\varepsilon)^\alpha \mathcal{N}(A, \varepsilon)$, which means in particular that

$$\mathcal{H}^\alpha(A) \leqslant \liminf_{\varepsilon \to 0} \mathcal{N}(A, \varepsilon)(2\varepsilon)^\alpha \tag{I.4}$$

One should not expect any reverse inequality to hold in general. Indeed if A is a compact subset of $\mathbb{R}^d$ with finite α-Hausdorff measure ($\alpha > 0$), and $X = (x_i)$ is a countable, dense family of points in $\mathbb{R}^d$, the covering number $\mathcal{N}(A \cup X, \varepsilon)$ is infinite for any positive ε, while $\mathcal{H}^\alpha(A \cup X) = \mathcal{H}^\alpha(A) < +\infty$.

DEFINITION I.3. Given a subset A of X, the function that maps $\alpha \in [0, +\infty[$ to $\mathcal{H}^\alpha(A)$ is decreasing; in fact, there exists a unique number α_0 such that for $\alpha < \alpha_0$, the Hausdorff measure $\mathcal{H}^\alpha(A)$ is $+\infty$, and for all $\alpha > \alpha_0$, the Hausdorff measure $\mathcal{H}^\alpha(A)$ vanishes. This number is called the Hausdorff dimension of A, and denoted by $\dim_H A$.

The previous remark shows that if the ε-Lebesgue covering number of a set A behaves as $O(\varepsilon^{-\alpha})$, then the Hausdorff dimension of A is at most α.

Behaviour under Lipschitz mapping and Hausdorff approximation. The Hausdorff measure and Lebesgue covering number behave well under Lipschitz mappings. Recall that a map $f : X \to Y$ between two metric spaces (X, d) and (Y, d') is called L-Lipschitz if for any pair $(x, x') \in X^2$, $d'(f(x), f(x')) \leqslant L d(x, x')$.

PROPOSITION I.6. *For any compact subset K of a metric space X, and $f : X \to Y$ any L-Lipschitz function, one can bound the Hausdorff measures and Lebesgue covering numbers of $f(K)$ from those of K:*

$$\forall \varepsilon > 0,\ \mathcal{N}(f(K), \varepsilon) \leqslant \mathcal{N}(K, \varepsilon/L)$$
$$\forall \alpha > 0,\ \mathcal{H}^\alpha(f(K)) \leqslant L^\alpha \mathcal{H}^\alpha(K)$$

The following very simple result shows that the covering numbers of a set A can be estimated from the covering numbers of a Hausdorff-close set B.

LEMMA I.7. *Let A, B be two subsets of a metric space X. Suppose moreover that their Hausdorff distance $d_H(A, B)$ is bounded from above by a positive ε. Then, for any $r > \varepsilon$,*

$$\mathcal{N}(A, r + \varepsilon) \leqslant \mathcal{N}(B, r) \leqslant \mathcal{N}(A, r - \varepsilon)$$

There is no such result for Hausdorff measures. In fact, even there is not even a lower or upper semicontinuity for Hausdorff measures under

Hausdorff approximation in general. Indeed, if $K = [0,1] \times \{0\} \subseteq \mathbb{R}^d$ and $K_\varepsilon = [0,1] \times [-\varepsilon, \varepsilon]$. Then, $\mathcal{H}^1(K_\varepsilon)$ is infinite for any positive ε, while $\mathcal{H}^1(K) = 1$. Conversely, let $K_n = \{k/n\,;\, 0 \leqslant k \leqslant n\} \subseteq \mathbb{R}$ and $K = [0,1]$. Then K_n Hausdorff converges to K, while the Hausdorff measures $\mathcal{H}^1(K_n) = 0$ does not converge to $\mathcal{H}^1(K) = 1$.

I.1.3 — Dimension and smoothness of the medial axis

Hausdorff dimension of the medial axis. The following proposition is due to P. Erdös (*in* [Erd46]). A consequence of this proposition is that the Hausdorff dimension on $\mathrm{Med}(K)$ is at most $d-1$, which is already a big improvement over Proposition I.5.

THEOREM I.8. *For a given compact set $K \subseteq \mathbb{R}^d$, let $\mathrm{Med}^k(K)$ denote the set of points of $\mathbb{R}^d$ such that the affine space spanned by $\mathrm{proj}_B(x, d_K(x))$ has dimension at least $k+1$. Then, $\mathrm{Med}^k(K)$ has σ-finite $(d-k)$-Hausdorff measure, ie. it is a countable union of subsets with finite $(d-k)$-Hausdorff measure.*

Another consequence of this result is that $\mathcal{H}^{d-1}$-almost every point of the medial axis has only two projections on K.

Note that thanks to Proposition I.2, the Hausdorff dimension of the closure of the medial axis $\overline{\mathrm{Med}}(K)$ is d for most compact sets. In [Riv01], Rivière produced a family of counter-examples concerning the dimension of the set of points of the cut-locus that are not in the medial axis:

THEOREM I.9 (Rivière). *For any $d \geqslant 2$ and $s \in [0,d]$, there exists an open bounded convex set $\Omega \subseteq \mathbb{R}^d$ such that the Hausdorff dimension of $\mathrm{Cut}(\partial\Omega) \setminus \mathrm{Med}(\partial\Omega) \cap \Omega$ is s.*

Semiconcavity of the distance functions. In this paragraph, we review the semiconcavity properties of the distance function and the squared distance function, which will play an important role in Chapter II and IV. Then, we recall some consequences of the semiconcavity of d_K on the local regularity of the medial axis, that can be obtained as particular cases of theorems on the set non-differentiability points of a convex function.

DEFINITION I.4. Let Ω be a connected open subset of $\mathbb{R}^d$. Recall that a function $\varphi : \Omega \to \mathbb{R}$ is called *convex* if for any segment $[a,b]$ included in Ω, the restriction of φ to $[a,b]$ is convex in the usual sense, *ie.* $\varphi(ta + (1-t)b) \leqslant t\varphi(a) + (1-t)\varphi(b)$. The function φ is called λ-convex if the function $x \mapsto \varphi(x) + \lambda\|x\|^2$ is convex on Ω.

As expected, a function $\psi : \Omega \to \mathbb{R}$ will be called λ-concave if $-\varphi$ is λ-convex.

The first remark to be done about λ-convexity is that if the function is $\mathcal{C}^2$, it is λ-convex iff the second derivative is bounded from below: $D^2\varphi \geqslant -\lambda \mathrm{id}$. Another interesting characterization for λ-convexity is the following:

PROPOSITION I.10. *A function $\varphi : \Omega \to \mathbb{R}^d$ is λ-convex iff for any point x in Ω, there exists a vector v such that:*

$$\varphi(x+h) \leqslant \varphi(x) + \langle h|v\rangle - \lambda \|h\|^2$$

General references on semi-concavity are [CS04] and [Pet07] for semi-concave functions on Alexandrov spaces. Proposition I.4 of the previous section can be rephrased:

PROPOSITION I.11. *The squared distance function to a compact set $K \subseteq \mathbb{R}^d$ is 1-semiconcave (or, said otherwise, the function $v_K : x \mapsto \|x\|^2 - d_K^2(x)$ is convex).*

This proposition seems to have been discovered many times; it is attributed to P.-L. Lions in J.H.G. Fu's paper on tubular neighborhood [Fu85]. It is a recurrent example in [CS04] and [Pet07], and is underlying the techniques of [Lie04]. As a consequence of Proposition I.11, one also has:

PROPOSITION I.12. *For any $\rho > 0$ and any compact set $K \subseteq \mathbb{R}^d$, the distance function d_K is $1/\rho$ semiconcave on the complement of the ρ-offset of K, $\mathbb{R}^d \setminus K^\rho$.*

Rectifiability of the medial axis. In order to give more precise results on the regularity of the medial axis, we need to introduce the notion of rectifiability:

DEFINITION I.5. A subset S of $\mathbb{R}^d$ is called k-rectifiable if there exists a countable family of Lipschitz maps $f_i : \mathbb{R}^k \to \mathbb{R}^d$ such that the union of the $f_i(\mathbb{R}^k)$ cover S up to a $\mathcal{H}^k$-negligible set, *ie.* $\mathcal{H}^k(S \setminus \cup_i f_i(\mathbb{R}^k)) = 0$.

Although it might seem surprising at first glance, it is equivalent to require that the set S is included in a countable union of k-dimensional $\mathcal{C}^1$ submanifolds of $\mathbb{R}^d$, up to a $\mathcal{H}^k$-negligible set. A set is called $\mathcal{C}^m$ k-rectifiable if it can be written as a countable union of k-dimensional $\mathcal{C}^m$ submanifolds of $\mathbb{R}^d$ up to a $\mathcal{H}^k$-negligible set.

DEFINITION I.6. Given a function $\varphi : \mathbb{R}^d \to \mathbb{R}$, the subdifferential at a point $x \in \mathbb{R}^d$, denoted by $\partial\varphi(x)$, is the set of vectors v such that $\varphi(x+h) \geqslant \varphi(x) + \langle v|h\rangle$ for every $h \in \mathbb{R}^d$. Notice that $\partial\varphi(x)$ is a convex set; its dimension is by definition the dimension of the affine subspace that it spans.

When φ is a convex function, we will denote by $\Sigma^k(\varphi)$ the set of points $x \in \mathbb{R}^d$ where the dimension of the subdifferential $\partial\varphi(x)$ is at least k. In particular, the set of non-differentiability point of φ is $\Sigma^1(\varphi)$.

The following theorem is due to Alberti, Ambrosio and Cannarsa, *cf.* [AAC92, Alb94]:

THEOREM I.13. *If $\varphi : \mathbb{R}^d \to \mathbb{R}$ is a λ-convex function, then the singular set $\Sigma^k(\varphi)$ is $(d-k)$-rectifiable of class $\mathcal{C}^2$.*

The medial axis of a compact set $K \subseteq \mathbb{R}^d$ is equal to the set of non-differentiability points of the function v_K, *ie.* $\Sigma^1(v_K)$, while the set of points with three or more closest neighbors is equal to $\Sigma^2(v_K)$. As a consequence of Theorem I.13, one gets the following improvement over Theorem I.8:

COROLLARY I.14. *For any compact subset* K *of* $\mathbb{R}^d$*:*

(i) the medial axis $\mathrm{Med}(K)$ K *is* $\mathcal{C}^2$ $(d-1)$*-rectifiable;*

(ii) the set of point of $\mathrm{Med}(K)$ *with three or more closest points in* K *is* $\mathcal{C}^2$ $(d-2)$*-rectifiable.*

I.2 SIZE AND VOLUME OF THE μ-MEDIAL AXIS

The results that we recalled in the two previous sections show that the dimension of the medial axis is $(d-1)$ both in the Hausdorff sense and in the sense of rectifiability. Our goal in this section is to obtain informations on the *size* of the medial axis, as measured by either its $(d-1)$-Hausdorff measure, or better, by its covering numbers. Recall that the medial axis $\mathrm{Med}(K)$ of a generic compact set K is dense in $\mathbb{R}^d$; hence, for every such compact set the covering number $\mathcal{N}(\mathrm{Med}(K), \varepsilon)$ is infinite. In order to obtain finite bounds without assumptions on K, we have to consider suitable subsets of the medial axis $\mathrm{Med}(K)$. In this section, we obtain bounds on the covering numbers of the part of the μ–medial axis [CCSL09], which lie at distance at least ε from K. In Chapter II, we will use this use this result to give quantitative continuity results for the map $p : K \in \mathcal{K}(\mathbb{R}^d) \mapsto p_K|_E \in L^1(E)$. Before turning to the definition of the μ-medial axis, let us briefly recap some known facts on the gradient of distance functions.

Generalized gradients for the distance function. Thanks to its semiconcavity one is able to define a notion of generalized gradient for the distance function d_K, that is defined even at points where d_K isn't differentiable. Given a compact set $K \subseteq \mathbb{R}^d$, the subdifferential of the distance function to K at a point $x \in \mathbb{R}^d$ is by definition the set of vectors $v \in \mathbb{R}^d$ such that

$$d_K^2(x+h) \leqslant d_K^2(x) + \langle h|v\rangle - \|h\|^2 \tag{I.5}$$

for all $h \in \mathbb{R}^d$. The subdifferential of d_K at a point x is denoted by $\partial_x d_K$, it is the convex hull of the set $\{(p-x)/\|p-x\| \ ; \ p \in \mathrm{proj}_K(x)\}$.

The gradient $\nabla_x d_K$ of the distance function d_K at a point $x \in \mathbb{R}^d$ is defined as the vector of $\partial_x d_K$ whose Euclidean norm is the smallest, or equivalently as the projection of the origin on $\partial_x d_K$ (see [Pet07] or [Lie04]). Given a point $x \in \mathbb{R}^d$, denote by $\gamma_K(x)$ the center and $r_K(x)$ the radius of the smallest ball enclosing the set of orthogonal projections of x on K. Then,

$$\begin{aligned} \nabla_x d_K &= \frac{x-\gamma_K(x)}{d_K(x)} \\ \|\nabla_x d_K\| &= \left(1 - \frac{r_K^2(x)}{d_K^2(x)}\right)^{1/2} = \cos(\theta) \end{aligned} \tag{I.6}$$

where θ is the (half) angle of the cone joining x to $B(\gamma_K(x), r_K(x))$

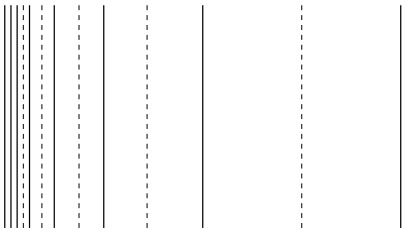


Figure I.2 – The "comb" and a part of its medial axis (dotted)

I.2.1 — μ-Medial axis of a compact set

The notion of μ-medial axes and μ-critical point of the distance function to a compact subset K of $\mathbb{R}^d$ were introduced by Chazal, Cohen-Steiner and Lieutier in [CCSL09]. We recall the definitions and properties we will need later.

A point x of $\mathbb{R}^d$ will be called a μ-critical point for the distance function to K (with $\mu \geqslant 0$), or simply a μ-critical point of K if for every $h \in \mathbb{R}^d$,

$$d_K^2(x+h) \leqslant d_K^2(x) + \mu \|h\| d_K(x) + \|h\|^2 .$$

By the definition of the subdifferential of d_K (Eq. (I.5)), the point x is μ-critical iff the norm of the gradient $\|\nabla_x d_K\|$ is at most μ. The μ-*medial axis* $\mathrm{Med}_\mu(K)$ of a compact set $K \subseteq \mathbb{R}^d$ is the set of μ-critical points of the distance function. Is is easily seen that the medial axis is the union of all μ-medial axes, with $0 \leqslant \mu < 1$:

$$\mathrm{Med}(K) = \bigcup_{0 \leqslant \mu < 1} \mathrm{Med}_\mu(K).$$

Moreover, from the lower semicontinuity of the map $x \mapsto \|\nabla_x d_K\|$, one obtains that for every $\mu < 1$, the μ-medial axis $\mathrm{Med}_\mu(K)$ of K is a compact subset of $\mathbb{R}^d$.

Because of its compactness, one could expect that the μ-medial axis of a well-behaved compact set will have finite $\mathcal{H}^{d-1}$-measure. This is not the case in general: if one considers a "comb", *ie.* an infinite union of parallel segments of fixed length in $\mathbb{R}^2$, such as $\mathcal{C} = \cup_{i \in \mathbb{N}^*} [0,1] \times \{2^{-i}\} \subseteq \mathbb{R}^2$ (see Figure I.2), the set of critical points of the distance fonction to $\mathcal{C}$ contains an imbricate comb. Hence $\mathcal{H}^{d-1}(\mathrm{Med}_\mu(\mathcal{C}))$ is infinite for any $\mu > 0$.

However, for any positive ε, the set of points of the μ-medial axis of $\mathcal{C}$ that are ε-away from $\mathcal{C}$ (that is $\mathrm{Med}_\mu(\mathcal{C}) \cap \mathbb{R}^d \setminus \mathcal{C}^\varepsilon$) only contains a finite union of segments, and has finite $\mathcal{H}^{d-1}$-measure. The goal of this section is to prove (quantitatively) that this remains true for any compact set. Precisely, we have:

THEOREM I.15. *For any compact set* $K \subseteq \mathbb{R}^d$, $\varepsilon \leqslant \mathrm{diam}(K)$, *and* η *small enough,*

$$\mathcal{N}\left(\mathrm{Med}_\mu(K) \cap (\mathbb{R}^d \setminus K^\varepsilon), \eta\right) \leqslant \mathcal{N}(\partial K, \varepsilon/2)\, O\left(\left\lceil \frac{\mathrm{diam}(K)}{\eta\sqrt{1-\mu}} \right\rceil^{d-1}\right)$$

In particular, one can bound the $(d-1)$*-volume of the* μ*-medial axis*

$$\mathcal{H}^{d-1}\left(\mathbf{Med}_\mu(K)\cap(\mathbb{R}^d\setminus K^\varepsilon)\right)\leqslant \mathcal{N}(\partial K,\varepsilon/2)\,O\left(\left[\frac{\operatorname{diam}(K)}{\sqrt{1-\mu}}\right]^{d-1}\right)$$

REMARK (Sharpness of the bound). Let x,y be two points at distance D in $\mathbb{R}^d$ and $K=\{x,y\}$. Then, $\mathrm{Med}(K)$ is simply the medial hyperplane between x and y. A point m in $\mathrm{Med}(K)$ belongs to $\mathrm{Med}_\mu(K)$ iff the cosine of the angle $\theta=\frac{1}{2}\angle(x-m,y-m)$ is at most μ.

$$\cos^2(\theta)=1-\frac{\|x-y\|^2}{4d_K^2(m)}=1-\frac{\operatorname{diam}(K)^2}{4d_K^2(m)}$$

Hence, $\cos(\theta)\geqslant\mu$ iff $d_K(m)\leqslant\frac{1}{2}\operatorname{diam}(K)/\sqrt{1-\mu^2}$. Let z denote the mid-point between x and y; then $d_K(m)^2=\|z-m\|^2+\operatorname{diam}(K)^2/4$. Then, $\mathrm{Med}_\mu(K)$ is simply the intersection of the ball centered at z and of radius $\frac{1}{2}\operatorname{diam}(K)\sqrt{\mu^2/(1-\mu^2)}$ with the medial hyperplane. Hence,

$$\mathcal{H}^{d-1}(\mathbf{Med}_\mu(K))=\Omega\left(\left[\frac{\operatorname{diam}(K)\mu^2}{\sqrt{1-\mu}}\right]^{d-1}\right)$$

This shows that the behaviour in $\operatorname{diam}(K)$ and μ of the theorem is sharp as μ converges to one.

Outline of the proof. In order to obtain the bound on the covering numbers of the 2ε-away μ-medial axis $\mathrm{Med}_\mu(K)\cap(\mathbb{R}^d\setminus K^{2\varepsilon})$ given in Theorem I.15, we prove that this set can be written as the image of a part of the level set ∂K^ε under the so-called *normal projection on the medial axis* $\ell:\mathbb{R}^d\setminus K\to\overline{\mathbf{Med}}(K)$.

There are two difficulties. First, one need to obtain a bound on the $(d-1)$-volume (or Lebesgue covering numbers) of the hypersurface ∂K^ε; this is achieved in §I.2.2. The second and main ingredient is a Lipschitz regularity result for the normal distance to the medial axis τ_K. When K is a compact submanifold of class $\mathcal{C}^{2,1}$, this function is globally Lipschitz on any r-level set of the distance function to K, when the radius r is small enough [IT01, LN05, CR09]. When K is the analytic boundary of a bounded domain Ω of $\mathbb{R}^2$, the normal distance to the medial axis of $\partial\Omega$ is $2/3$-Hölder on Ω [CCG07].

However, without strong regularity assumption on the compact set K, it is hopeless to obtain a global Lipschitz regularity result for τ_K on a parallel set of K. Indeed, such a result would imply the finiteness of $(d-1)$-Hausdorff measure of the medial axis, which is known to be false (recall for instance that the medial axis of a generic compact set is dense). We show however, that the normal distance to the medial axis is Lipschitz on a suitable subset of a parallel set. This will be enough to conclude, thanks to Proposition I.6.

I.2.2 — Volume of the boundary of the offsets of a compact set

The next proposition gives a bound for the measure of the r-level set of a compact set depending only on its covering number. A similar result with a different bound has been previously obtained in [OP85]. As always, K^r is the set of points of $\mathbb{R}^d$ at distance less than r of K, and ∂K^r is the boundary of this set, *ie.* the r-level set of the distance function d_K.

Volume of spheres and balls. We will denote by ω_{d-1} (resp. $\omega_{d-1}(r)$) the $(d-1)$-volume the $(d-1)$-dimensional unit (resp. radius r) *sphere* in $\mathbb{R}^d$. Similarly, β_d and $\beta_d(r)$ will denote the d-volume of d-balls.

PROPOSITION I.16. *If K is a compact set in $\mathbb{R}^d$, for every positive r, ∂K^r is rectifiable and*

$$\mathcal{H}^{d-1}(\partial K^r) \leqslant \mathcal{N}(\partial K, r) \times \omega_{d-1}(2r) \tag{I.7}$$

$$\mathcal{N}(\partial K^r, \varepsilon) \leqslant \mathcal{N}(\partial K, r)\mathcal{N}(\mathcal{S}^{d-1}, \varepsilon/2r) \tag{I.8}$$

This proposition will also be used in the next chapter, §II.1.2. Before proving it, let us make a few comments on this result:

— The bound in the theorem is tight, as one can check by considering K to be the ball $B(0,r)$ (*ie.* $K^r = B(0,2r)$).

— There exists a constant $C(d)$ such that the Lebesgue number $\mathcal{N}(B(0,1),r)$ of the unit ball in $\mathbb{R}^d$ is bounded by $1+C(d)r^{-d}$. From this and Proposition I.16 it follows that

$$\begin{aligned}\mathcal{H}^{d-1}(\partial K^r) &\leqslant \mathcal{N}(\partial K, r) \times \omega_{d-1}(2r)\\ &\leqslant \mathcal{N}(B(0, \operatorname{diam}(K)/2), r) \times \omega_{d-1}(2r)\\ &\leqslant (1 + C(d) \times (\operatorname{diam}(K)/r)^d)\omega_{d-1}(2r)\\ &\leqslant C'(d) \times \left(1 + \frac{\operatorname{diam}(K)^d}{r}\right)\end{aligned}$$

for some universal constant $C'(d)$ depending only on the ambient dimension d. This last inequality was the one actually proved in [OP85].

— A corollary of Proposition I.16 is that if K is a compact subset of $\mathbb{R}^d$ of metric codimension k, *ie.* $\mathcal{N}(K,\varepsilon) \leqslant C\varepsilon^{k-d}$, then $\mathcal{H}^{d-1}(\partial K^\varepsilon) \leqslant 2^{d-1}\omega_{d-1} \times C\varepsilon^{d-1}$. Of course, no such inequality can be expected if we only assume a bound on the Hausdorff dimension of K.

We first prove Proposition I.16 in the special case of r-*flowers*. A r-flower F is the boundary of the r-offset of a compact set contained in a ball $B(x,r)$, *ie.* $F = \partial K^r$ where K is contained in some ball $B(x,r)$. The difference with the general case is that when K is contained in a ball $B(x,r)$, the offset K^r is star-shaped with respect to x. This allows us to define a ray-shooting application $s_K : \mathcal{S}^{d-1} \to \partial K^r$ which maps any $v \in \mathcal{S}^{d-1}$ to the intersection of the ray emanating from x with direction v with the hypersurface ∂K^r.

LEMMA I.17. *If* K *is a compact set contained in a ball* $B(x,r)$*, the ray-shooting application* s_K *defined above is* $2r$*-Lipschitz, so that* $\mathcal{H}^{d-1}(\partial K^r) \leqslant \omega_{d-1}(2r)$.

Proof. Since $\partial K^r = s_K(B(0,1))$, assuming that s_K is $2r$-Lipschitz, we will indeed have: $\mathcal{H}^{d-1}(K^r) \leqslant (2r)^{d-1}\mathcal{H}^{d-1}(B(0,1)) = \omega_{d-1}(2r)$. Let us now compute the Lipschitz constant of the ray-shooting map s_K.

If we let $t_K(v)$ be the distance between x and $s_K(v)$, we have $t_K(v) = \sup_{e\in K} t_e(v)$. Since t_K is the supremum of all the t_e, in order to prove that s_K is $2r$-Lipschitz, we only need to prove that each s_e is $2r$-Lipschitz. Solving the equation $\|x + tv - e\| = r$ with $t \geqslant 0$ gives

$$t_e(v) = \sqrt{\langle x-e|v\rangle^2 + r^2 - \|x-e\|^2} - \langle x-e|v\rangle$$

$$d_v t_e(w) = \frac{\langle x-e|v\rangle\langle x-e|w\rangle}{\sqrt{\langle x-e|v\rangle^2 + r^2 - \|x-e\|^2}} - \langle x-e|w\rangle$$

Since $s_e(v) = x + t_e(v)v$,

$$\begin{aligned} d_v s_e(w) &= \|s_e(v)-x\|\, w + d_v t_e(w) v \\ &= \langle s_e(v)-x|v\rangle w + \left(\frac{\langle x-e|v\rangle\langle x-e|w\rangle}{\langle s_e(v)-e|v\rangle} - \langle x-e|w\rangle\right) v \\ &= \langle s_e(v)-x|v\rangle \frac{\langle s_e(v)-e|v\rangle w - \langle x-e|w\rangle v}{\langle s_e(v)-e|v\rangle} \end{aligned}$$

If w is orthogonal to $x-e$, then $\|d_v s_e(w)\| \leqslant \|s_e(v)-x\|\,\|w\| \leqslant 2r\,\|w\|$ and we are done. We now suppose that w is contained in the plane spanned by $x-e$ and v. Since w is tangent to the sphere at v, it is also orthogonal v. Hence, $\langle s_e(v)-x|w\rangle = 0$, and $\langle x-e|w\rangle = \langle s_e(v)-e|w\rangle$.

$$d_v s_e(w) = \|s_e(v)-x\|^2 \frac{\langle s_e(v)-e|v\rangle w - \langle s_e(v)-e|w\rangle v}{\langle s_e(v)-e|s_e(v)-x\rangle}$$

If we suppose that both v and w are unit vectors, $\langle s_e(v)-e|v\rangle w - \langle s_e(v)-e|w\rangle v$ is the rotation of $s_e(v)-e$ by an angle of $\pi/4$. By linearity we get $\|\langle s_e(v)-e|v\rangle w - \langle s_e(v)-e|w\rangle v\| = \|w\|\,\|s_e(v)-e\|$ (we still have $\|v\| = 1$). Now let us remark that

$$\begin{aligned} \|s_e(v)-e\|\,\|s_e(v)-x\|\,|\cos(\theta)| &= |\langle s_e(v)-e|s_e(v)-x\rangle| \\ &= \frac{1}{2}(\|x-s_e(v)\|^2 + \|s_e(v)-e\|^2 - \|x-e\|^2) \\ &\geqslant \frac{1}{2}\|x-s_e(v)\|^2 \end{aligned}$$

Using this we deduce:

$$\|d_v s_e(w)\| \leqslant \|s_e(v)-x\|^2 \frac{\|w\|\,\|s_e(v)-e\|}{\frac{1}{2}\|x-s_e(v)\|^2} = 2\,\|s_e(v)-e\|\,\|w\| \leqslant 2r\,\|w\|$$

We have just proved that for any w tangent to v, $\|d_v s_e(w)\| \leqslant 2r\,\|w\|$, from

which we can conclude that s_e is $2r$-Lipschitz. □

Proof of Proposition I.16. By definition of the covering number, there exist a finite family of points $x_1, \dots, x_n$, with $n = \mathcal{N}(\partial K, r)$, such that union of the open balls $B(x_i)$ covers ∂K. If one denotes by K_i the intersection of ∂K with $\overline{B}(x_i, r)$, the boundary ∂K^r is contained in the union $\cup_i \partial K_i^r$. Hence its Hausdorff measure does not exceed the sum $\sum_i \mathcal{H}^{d-1}(\partial K_i^r)$. Since for each i, ∂K_i^r is a flower, one concludes by applying the preceding lemma.

The second inequality is proven likewise, using Proposition I.6 to bound the Lebesgue covering number of the image of a metric space by a Lipschitz map. □

I.2.3 — Covering numbers of the μ-medial axis

We now proceed to the proof of Theorem I.15.

DEFINITION I.7. For any point $x \in \mathbb{R}^d$, we define the *normal distance of* x *to the medial axis* as $\tau_K(x) := \inf\{t \geqslant 0\,;\, x + t\nabla_x d_K \in \mathrm{Med}(K)\}$. We will set $\tau_K(x)$ to zero at any point in K or in the medial axis $\mathrm{Med}(K)$.

For any time t smaller than $\tau(x)$, we denote by $\Psi_K^t(x)$ the point $\Psi_K^t(x) = x + t\nabla_x d_K$. Finally, for any $x \notin K$, we let $\ell_K(x)$ be the first intersection of the half-ray starting at x with direction $\nabla_x d_K$ with the medial axis. More precisely, we define $\ell_K(x) = \Psi_K^{\tau(x)}(x) \in \overline{\mathrm{Med}}(K)$.

LEMMA I.18. *Let* m *be a point of the medial axis* $\mathrm{Med}(K)$ *with* $d(x, K) > \varepsilon$*, and* x *be a projection of* m *on* ∂K^ε*. Then* $\ell(x) = m$.

Proof. By definition of K^ε, $d(m, K) = d(m, K^\varepsilon) + \varepsilon$, so that the projection p of x on K must also be a projection of m on K. Hence, m, x and p must be aligned. Since the open ball $B(m, d(m, p))$ does not intersect K, for any point $y \in]p, m[$ the ball $B(y, d(y, p))$ intersects K only at p. In particular, by definition of the gradient, $\nabla_x d_K$ must be the unit vector directing $]p, m[$, *ie.* $\nabla_x d_K = (m - x)/d(m, x)$. Moreover, since $[x, p[$ is contained in the complement of the medial axis, $\tau(x)$ must be equal to $d(x, m)$. Finally one gets $\Psi^{\tau(x)}(x) = x + d(x, m)\nabla_x d_K = m$. □

This statement means in particular that ε-away medial axis, that is the intersection $\mathrm{Med}(K) \cap (\mathbb{R}^d \setminus K^\varepsilon)$, is contained in the image of the piece of hypersurface $\{x \in \partial K^\varepsilon\,;\, \tau_K(x) \geqslant \varepsilon\}$ by the map ℓ.

Recall that the radius of a set $K \subseteq \mathbb{R}^d$ is the radius of the smallest ball enclosing K, while the diameter of K is the maximum distance between two points in K. The following inequality between the radius and the diameter is known as Jung's theorem [Jun10]: $\mathrm{radius}(K)\sqrt{2(1 + 1/d)} \leqslant \mathrm{diam}(K)$.

LEMMA I.19. *For any point* m *in the* μ*-medial axis* $\mathrm{Med}_\mu(K)$*, there exist two projections* $x, y \in \mathrm{proj}_K(m)$ *of* m *on* K *such that the cosine of the angle* $\frac{1}{2}\angle(x - m, y - m)$ *is smaller than* $\left(\frac{1+\mu^2}{2}\right)^{1/2}$.

Proof. We use the characterization of the gradient of the distance function given in equation (I.6). If $B(\gamma_K(m), r_K(m))$ denotes the smallest ball enclosing $\mathrm{proj}_K(m)$, then $\mu^2 \geqslant 1 - r_K^2(m)/d_K^2(m)$. Using Jung's theorem and the definition of the diameter, there must exists two points x, y in $\mathrm{proj}_K(m)$ whose distance r' is larger than $\sqrt{2}r_K(m)$. The following bound on the cosine of the angle $\theta = \frac{1}{2}\angle(x-m, y-m)$ concludes the proof:

$$\cos^2(\theta) = 1 - \frac{(r'/2)^2}{d_K^2(m)} \leqslant 1 - \frac{1}{2}\frac{r_K^2(m)}{d_K^2(m)} \leqslant (1+\mu^2)/2 \tag{I.9}$$

□

LEMMA I.20. *The maximum distance from a point in* $\mathrm{Med}_\mu(K)$ *to* K *is bounded by* $\frac{1}{\sqrt{2}}\,\mathrm{diam}(K)/\left(1-\mu^2\right)^{1/2}$

Proof. Let x, y be two orthogonal projections of $m \in \mathrm{Med}_\mu(K)$ on K as given by the previous lemma. Then, using equation (I.9), one obtains

$$1 - \frac{\|x-y\|^2/4}{d_K^2(m)} \leqslant (1+\mu^2)/2.$$

Hence, $d_K^2(m) \leqslant \frac{1}{2}(1-\mu^2)^{-1}\|x-y\|^2$, which proves the result. □

Let us denote by S_μ^ε the set of points x of the hypersurface ∂K^ε that satisfies the three conditions below:

(i) the normal distance to the medial axis is bounded below: $\tau(x) \geqslant \varepsilon$;

(ii) the image of x by ℓ is in the μ-medial axis of K: $\ell(x) \in \mathrm{Med}_\mu(K)$;

(iii) there exists another projection y of $m = \ell(x)$ on ∂K^ε with

$$\cos\left(\frac{1}{2}\angle(p-m, q-m)\right) \leqslant \sqrt{\frac{1+\mu^2}{2}}$$

A reformulation of Lemmas I.19 and I.18 is the following corollary:

COROLLARY I.21. *The image of* S_μ^ε *by the map* ℓ *covers the whole* 2ε*-away* μ*-medial axis:* $\ell(S_\mu^\varepsilon) = \mathrm{Med}_\mu(K) \cap (\mathbb{R}^d \setminus K^{2\varepsilon})$

Lipschitz estimations for the map ℓ. In this paragraph, we bound the Lipschitz constants of the restriction of the maps ∇d_K, τ and (finally) ℓ to the subset $S_\mu^\varepsilon \subseteq \partial K^\varepsilon$.

First, let $\partial K^{\varepsilon,t}$ be the set of points x in ∂K^ε where the distance function is differentiable, and such that $\tau(x)$ is bounded from below by t. In particular, notice that S_μ^ε is contained in $\partial K^{\varepsilon,\varepsilon}$. The following Lemma proves that the functions Ψ^t and $\nabla_x d_K$ are Lipschitz on $\partial K^{\varepsilon,t}$:

LEMMA I.22. *(i) The restriction of* Ψ^t *to* $\partial K^{\varepsilon,t}$ *is* $(1+t/\varepsilon)$*-Lipschitz.*

(ii) The gradient of the distance function, $x \mapsto \nabla_x d_K$*, is* $3/\varepsilon$*-Lipschitz on* $\partial K^{\varepsilon,\varepsilon}$.

Proof. (i) Let x and x' be two points of ∂K^ε with $\tau(x), \tau(x') > t$, p and p' their projections on K and y and y' their image by Ψ^t. We let $u = 1 + t/\varepsilon$ be the scale factor between $x - p$ and $y - p$, *ie.* :

$$(*) \quad y' - y = u(x' - x) + (1 - u)(p' - p)$$

Using the fact that y projects to p, and the definition of u, we have:

$$\|y - p\|^2 \leqslant \|y - p'\|^2 = \|y - p\|^2 + \|p - p'\|^2 + 2\langle y - p | p - p'\rangle$$

$$\textit{ie. } 0 \leqslant \|p - p'\|^2 + 2u\langle x - p | p - p'\rangle$$
$$\textit{ie. } \langle p - x | p - p'\rangle \leqslant \frac{1}{2}u^{-1}\|p - p'\|^2$$

Summing this last inequality, the same inequality with primes and the equality $\langle p' - p | p - p'\rangle = -\|p' - p\|^2$ gives

$$(**) \quad \langle x' - x | p' - p\rangle \geqslant (1 - u^{-1})\|p' - p\|^2$$

Using $(*)$ and $(**)$ we get the desired Lipschitz inequality

$$\begin{aligned}\|y - y'\|^2 &= u^2\|x - x'\|^2 + (1 - u)^2\|p' - p\|^2 + 2u(1 - u)\langle x' - x | p' - p\rangle \\ &\leqslant u^2\|x - x'\|^2 - (1 - u)^2\|p' - p\|^2 \leqslant (1 + t/\varepsilon)^2\|x - x'\|^2\end{aligned}$$

(ii) If x belongs to $\partial K^{\varepsilon,\varepsilon}$, then $\nabla_x d_K = \frac{1}{\varepsilon}(\Psi^\varepsilon(x) - x)$. The result follows from the Lipschitz estimation of (i). □

The second step is to prove that the restriction of τ to the set S^ε_μ is also Lipschitz. The technical core of the proof is contained in the following geometric lemma:

LEMMA I.23. *Let $t(x, v)$ denote the intersection time of the ray $x + tv$ with the medial hyperplane $H_{x,y}$ between x and another point y, and $t(x', v')$ the intersection time between the ray $x' + tv'$ and $H_{x'y}$. Then, assuming:*

$$\alpha\|x - y\| \leqslant \langle v | x - y\rangle, \tag{I.10}$$
$$\|x' - y\| \leqslant D, \tag{I.11}$$
$$\|v' - v\| \leqslant \lambda\|x' - x\|, \tag{I.12}$$
$$\varepsilon \leqslant t(x, v) \tag{I.13}$$

one obtains the following bound:

$$t(x', v') \leqslant t(x, v) + \frac{6}{\alpha^2}(1 + \lambda D)\|x' - x\|$$

as soon as $\|x' - x\|$ is small enough (namely, smaller than $\varepsilon\alpha^2(1 + 3\lambda D)^{-1}$).

Proof. We search the time t such that $\|x' + tv' - x'\|^2 = \|x' + tv' - y\|^2$, *ie.*

$$t^2\|v'\|^2 = \|x' - y\|^2 + 2t\langle x' - y | v'\rangle + t^2\|v'\|^2$$

Hence, the intersection time is $t(x', v') = \|x' - y\|^2 / 2\langle y - x' | v' \rangle$. The lower bound on $t(x, y)$ translates as

$$\varepsilon \leqslant \frac{1}{2} \frac{\|x - y\|^2}{\langle x - y | v \rangle} \leqslant \frac{1}{2\alpha} \|x - y\|$$

If $\nabla_{x'} t$ and $\nabla_{v'} t$ denote the gradients of this function in the direction of v' and x', one has:

$$\nabla_{v'} t(x', v') = \frac{1}{2} \frac{\|x' - y\|^2 (x' - y)}{\langle y - x' | v' \rangle^2}$$
$$\nabla_{x'} t(x', v') = \frac{1}{2} \frac{\|x' - y\|^2 v' + 2\langle y - x' | v' \rangle (x' - y)}{\langle y - x' | v' \rangle^2}$$

Now, we bound the denominator of this expression:

$$\begin{aligned} \langle x' - y | v' \rangle &= \langle x' - y | v' - v \rangle + \langle x' - x | v \rangle + \langle x - y | v \rangle \\ &\geqslant \alpha \|x - y\| - (1 + \lambda \|x' - y\|) \|x' - x\| \\ &\geqslant \alpha \|x' - y\| - (2 + \lambda D) \|x' - x\| \end{aligned}$$

The scalar product $\langle x' - y | v' \rangle$ will be larger than (say) $\frac{\alpha}{2} \|x' - y\|$ provided that

$$(2 + \lambda D) \|x' - x\| \leqslant \frac{\alpha}{2} \|x' - y\|$$

or, bounding from below $\|x' - y\|$ by $\|x - y\| - \|x - x'\| \geqslant 2\alpha\varepsilon - \|x - x'\|$, provided that:

$$(3 + \lambda D) \|x' - x\| \leqslant \alpha^2 \varepsilon$$

This is the case in particular if $\|x' - x\| \leqslant \alpha^2 \varepsilon (3 + \lambda D)^{-1}$. Under that assumption, we have the following bound on the norm of the gradient, from which the Lipschitz inequality follows:

$$\|\nabla_{x'} t(x', v')\| \leqslant 6/\alpha^2 \quad \text{and} \quad \|\nabla_{v'} t(x', v')\| \leqslant 4D/\alpha^2$$

□

Using this Lemma, we are able to show that the function ℓ is locally Lipschitz on the subset $S_\mu^\varepsilon \subseteq \partial K^\varepsilon$:

PROPOSITION I.24. *The restriction of τ to S_μ^ε is locally L-Lipschitz, in the sense that if $(x, y) \in S_\mu^\varepsilon$ are such that $\|x - y\| \leqslant \delta_0$, then $\|\ell(x) - \ell(y)\| \leqslant L \|x - y\|$ with*

$$L = O\left(\frac{1 + \operatorname{diam}(K)/\varepsilon}{(1 - \mu)^{1/2}} \right) \text{ and } \delta_0 = O(\varepsilon / L)$$

In order to simplify the proof of this Proposition, we will make use of the following notation, where f is any function from $X \subseteq \mathbb{R}^d$ to $\mathbb{R}$ or $\mathbb{R}^d$:

$$\operatorname{Lip}_\delta f|_X := \sup\{ \|f(x) - f(y)\| / \|x - y\| \, ; (x, y) \in X^2 \text{ and } \|x - y\| \leqslant \delta \}.$$

Proof. We start the proof by evaluating the Lipschitz constant of the restriction of τ to S^ε_μ, using Lemma I.23 (Step 1), and then deduce the Lipschitz estimate for the function ℓ (Step 2).

STEP 1 Thanks to Lemma I.19, for any x in S^ε_μ, there exists another projection y of $m = \ell(x)$ on ∂K^ε such that the cosine of the angle $\theta = \frac{1}{2}\angle(x-m, y-m)$ is at most $\sqrt{(1+\mu^2)/2}$. Let us denote by $v = \nabla_x d_K$ the unit vector from x to m. The angle between $\overrightarrow{xy}$ and v is $\pi/2 - \theta$. Then,

$$\cos(\pi/2-\theta) = \sin(\theta) = \sqrt{1-\cos^2(\theta)} \geqslant \alpha := \left(\frac{1-\mu^2}{2}\right)^{1/2}$$

As a consequence, with the α introduced above, one has $\alpha\|x-y\| \leqslant |\langle v|x-y\rangle|$. Moreover, $\|x-y\|$ is smaller than $D = \operatorname{diam}(K^\varepsilon) \leqslant \operatorname{diam}(K) + \varepsilon$. For any other point x' in S^ε_μ, and $v' = \nabla_{x'} d_K$, one has $\|v - v'\| \leqslant \lambda\|x-x'\|$ with $\lambda = 3/\varepsilon$ (thanks to Lemma I.22).

These remarks allow us to apply Lemma I.23. Using the notations of this lemma, one sees that $t(x,v)$ is simply $\tau(x)$ while $t(x',v')$ is an upper bound for $\tau(x')$. This gives us:

$$\begin{aligned}\tau(x') &\leqslant \tau(x) + \frac{6}{\alpha^2}(1+\lambda D)\left\|x-x'\right\| \\ &\leqslant \tau(x) + M\left\|x-x'\right\| \\ \text{where } M &= O\left(\frac{1+\operatorname{diam}(K)/\varepsilon}{\sqrt{1-\mu^2}}\right)\end{aligned}$$

as soon as x' is close enough to x. From the statement of Lemma I.23, one sees that $\|x-x'\| \leqslant \delta_0$ with $\delta_0 = O(\varepsilon/M)$ is enough. Exchanging the role of x and x', one proves that $|\tau(x) - \tau(x')| \leqslant M\|x-x'\|$, provided that $\|x-x'\| \leqslant \delta_0$. As a conclusion,

$$\operatorname{Lip}_{\delta_0}\left[\tau|_{S^\varepsilon_\mu}\right] = O\left(\frac{1+\operatorname{diam}(K)/\varepsilon}{\sqrt{1-\mu^2}}\right) \tag{I.14}$$

STEP 2 We can use the following decomposition of the difference $\ell(x) - \ell(x')$:

$$\ell(x) - \ell(x') = (x-x') + (\tau(x)-\tau(x'))\nabla_x d_K + \tau(x')(\nabla_x d_K - \nabla_{x'} d_K) \tag{I.15}$$

in order to bound the (local) Lipschitz constant of the restriction of ℓ to S^ε_μ from those computed earlier. One deduces from this equation that

$$\operatorname{Lip}_{\delta_0}\left[\ell|_{S^\varepsilon_\mu}\right] \leqslant 1 + \operatorname{Lip}_{\delta_0}\left[\tau|_{S^\varepsilon_\mu}\right] + \|\tau\|_\infty \operatorname{Lip}_{\delta_0}\left[\nabla d_K|_{S^\varepsilon_\mu}\right] \tag{I.16}$$

Thanks to Lemma I.20, one has $|\tau(x)| = O(\mathrm{diam}(K)/(1-\mu)^{1/2})$; combining this with the estimate from Lemma I.22 that $\mathrm{Lip}\, \nabla d_K|_{S^\varepsilon_\mu} \leqslant 3/\varepsilon$, this gives

$$\|\tau\|_\infty \,\mathrm{Lip}_{\delta_0}\left[\nabla d_K|_{S^\varepsilon_\mu}\right] = O(\mathrm{diam}(K)/[\varepsilon(1-\mu)^{1/2}]) \tag{I.17}$$

Putting the estimates (I.14) and (I.17) into (I.16) concludes the proof. □

Proof of Theorem I.15. Applying Proposition I.24, we get the existence of

$$L = \mathrm{Lip}_{\delta_0}\left[\ell|_{S^{\varepsilon/2}_\mu}\right] = O(\mathrm{diam}(K)/(\varepsilon\sqrt{1-\mu})) \text{ and } \delta_0 = O(\varepsilon/L)$$

such that ℓ is locally L-Lipschitz. In particular, for any η smaller than δ_0,

$$\begin{aligned}\mathcal{N}\left(\mathbf{Med}_\mu(K)\cap(\mathbb{R}^d\setminus K^\varepsilon),\eta\right) &= \mathcal{N}\left(\ell(S^{\varepsilon/2}_\mu),\eta\right)\\ &\leqslant \mathcal{N}\left(S^{\varepsilon/2}_\mu,\eta/L\right)\\ &\leqslant \mathcal{N}(\partial K^{\varepsilon/2},\eta/L).\end{aligned} \tag{I.18}$$

The bound on the covering number of the boundary of tubular neighborhoods (Proposition I.16) gives:

$$\mathcal{N}(\partial K^{\varepsilon/2},\eta/L) \leqslant \mathcal{N}(\partial K,\varepsilon/2)\mathcal{N}\left(\mathcal{S}^{d-1},\frac{\eta}{L\varepsilon}\right). \tag{I.19}$$

Equations (I.18) and (I.19), and the estimation $\mathcal{N}(\mathcal{S}^{d-1},\rho)\sim\omega_{d-1}\rho^{d-1}$ yield

$$\mathcal{N}\left(\mathbf{Med}_\mu(K)\cap(\mathbb{R}^d\setminus K^\varepsilon),\eta\right) = \mathcal{N}(\partial K,\varepsilon/2)\,O\left(\left[\frac{\eta}{L\varepsilon}\right]^{d-1}\right).$$

It suffices to replace L by its value from Proposition I.24 to finish the proof. □

Chapter II

BOUNDARY AND CURVATURE MEASURES

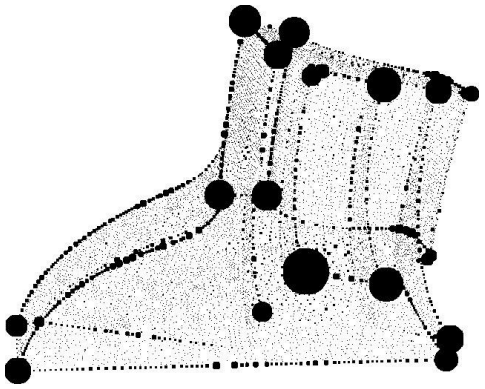


Figure II.1 – Boundary measure for a sampled mechanical part. As expected, the relevant features of the shape are visually highlighted.

Abstract

In this chapter, we introduce and study the *boundary measures* of compact subsets of the d-dimensional Euclidean space, which are closely related to Federer's curvature measures. Both boundary and curvature measures can be computed efficiently for point clouds, through a Monte-Carlo algorithm.

The main contribution of this work is the proof of a quantitative stability theorem for boundary measures, under Hausdorff approximation. Hausdorff stability makes boundary measures a useful tool for geometric inference. As a corollary we obtain a stability result for Federer's *curvature measures* of a compact set, showing that they can be reliably estimated from point-cloud approximations.

These stability results follow from bounds for the L^1 norm between the projection functions p_K and $p_{K'}$ on a bounded open set E as a function of the Hausdorff distance between K and K'. A first bound is obtained using the

results of the previous chapter on the μ-medial axis. A better bound, with optimal exponents, is obtained using tools from convex analysis and geometric measure theory.

In the appendix, we discuss how the existence of approximate curvature measure might be used for estimating the reach of a compact set, a quantity that plays an important role in most reconstruction approaches.

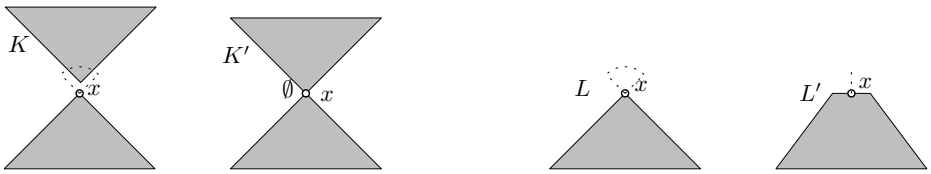
Contents

INTRODUCTION

This chapter (as well as the next chapter) deals with approximate extrinsic geometry, and more precisely, with the geometric information given by normal cones. We are especially interested in the "size" of the normal cones, which give precious informations concerning the smoothness of the considered shape.

Loosely speaking, the normal cone to a compact set $K \subseteq \mathbb{R}^d$ at a point p is the set of normals to K at p. The definition can be made precise in terme of the projection function: a vector $\overrightarrow{n} \in \mathbb{S}^{d-1}$ is normal to K at p if there is a point $x \in \mathbb{R}^d$ whose projection in K is p, and such that the vector $\overrightarrow{px}$ is (positively) collinear to $\overrightarrow{n}$. Of course, the normal cone at a point x in K is a Hausdorff-unstable piece of information, for at least two reasons. First, because there is no notion of scale attached to a normal cone (Fig. II.2.(a)), and second because a stable notion of this kind cannot be pointwise (Fig. II.2.(b))

In order to account for these two phenomena, we define the *boundary measure* of K at a scale r as a mass distribution (rather than as a function)



(a) Normal cones don't have scale.

(b) Normal cones are pointwise unstable.

Figure II.2 – Unstability of normal cones. In Fig. (a), the normal cone to K at x is a circle arc while the normal cone to K' at x is empty. In Fig. (b), the normal cone to L at x is the same circle arc, while the normal to L' at the same point x is a singleton (because L' is smooth at x).

μ_{K,K^r} concentrated on K. The amount of mass contained in a region $B \subset K$, denoted by $\mu_{K,K^r}(B)$, measures the volume of the set of points in the offset K^r whose projection on K lies in B. When K is convex or smooth (or more generally has *positive reach*) the dependence of the boundary measure on the scale parameter r is related to generalized notions of curvature, as proven by the tube formulas from various author (Steiner, Minkowski, Weyl and Federer). In particular, these generalized curvatures can be retrieved knowing only the boundary measures at several scales.

Stability and computability. The usability of boundary measures for geometric inference depends on two questions. First, is this notion Hausdorff stable? *ie.* if C is a good approximation of K (dense enough and without too much noise), does the boundary measure μ_{C,C^r} carry approximately the same geometric information as μ_{K,K^r}? Second, is it practically feasible to compute, or approximate the boundary measure of a point cloud $C \subseteq \mathbb{R}^d$?

The answer to the second question is given in section 5 in the form of a very simple to implement Monte-Carlo algorithm allowing to compute the boundary measure of a point cloud C embedded in the space $\mathbb{R}^d$. Standard arguments show that if C has n points, a ε-approximation of μ_{C,C^r} can be obtained with high probability (*eg.* 99%) in time $O(dn^2 \ln(1/\varepsilon)/\varepsilon^2)$ without using any sophisticated data structure. A more careful analysis shows that the n^2 behavior can be replaced by n times an appropriate covering number of C, which indicates that the cost is linear both in n and d for low entropy point clouds. Hence this algorithm is practical at least for moderate size point clouds in high dimension.

The study of the Hausdorff-stability of boundary and curvature measures is the main contribution of this chapter. We prove a stability theorem for boundary measures, showing that the dependence on K is $1/2$-Hölder — this exponent being optimal. This theorem can be stated as follows, endowing the set of compact subsets of $\mathbb{R}^d$ with the Hausdorff distance d_H, and the set of mass distributions on $\mathbb{R}^d$ with the so-called bounded-Lipschitz distance d_{bL} (see below for definitions):

THEOREM. *For every compact set* $K \subseteq \mathbb{R}^d$*, there exists some constant* $C(K)$ *such that*

$$d_{bL}(\mu_{K,K^r}, \mu_{C,C^r}) \leqslant C(K) d_H(C,K)^{1/2}$$

In the sequel we will make this statement more precise by giving explicit constants. A similar stability result for a generalization of curvature measures is deduced from this theorem. At the heart of these two stability results is a L^1 stability property for (closest point) projections on compact sets. The proof of the projection stability theorem involves a apparently new inequality in convex analysis, which may also be interesting in its own:

THEOREM. *Let* E *be an open subset of* $\mathbb{R}^d$ *with* $(d-1)$*–rectifiable boundary, and* f, g *be two convex functions such that* $\operatorname{diam}(\nabla f(E) \cup \nabla g(E)) \leqslant k$*. Then there exists a constant* $C(d,E,k)$ *such that for* $\|f-g\|_\infty$ *small enough,*

$$\|\nabla f - \nabla g\|_{L^1(E)} \leqslant C(d,E,k)\, \|f-g\|_\infty^{1/2}$$

Outline. In §II.1, we recall the geometric background on tube formulas and curvature measures. We especially insist on Federer's contribution to the theory, namely the definition of sets with positive reach, and the introduction of curvature measures for these sets (an alternative proof of existence of which is given in Appendix §II.A.1). In §II.2, we state and prove our Hausdorff-stability results for boundary and curvature measures. In §II.2.2, we give a first stability results that uses the bound on the covering numbers of the μ-medial axis of the previous chapter, while in §II.2.3, we prove the bound with optimal exponent stated above. The Hausdorff-stability of curvature measures is deduced in §II.2.5. In §II.3, we introduce and study a Monte-Carlo scheme for approximating boundary measures. We also discuss some of the possibilities for extracting information (*eg.*on the location of features) from these measures.

Appendix §II.A is dedicated to the relation between local Steiner (*ie.* tube) formula and reach. We give an alternative proof of existence of tube formula for compact set with positive reach, as well as a partial reverse statement (barely stronger than the one given in [HHL04]), using some of the tools introduced in the previous chapter. We then discuss the question of reach estimation.

II.1 CURVATURE MEASURES AND REACH OF A COMPACT SET

In this section, we review some aspects of the theory of tube formulas. We start with the global tube formulas of Steiner, Minkowski and Weyl, before turning to the more quantitative and localized tube formulas by Federer, that he used to define the curvature measures of a compact set with positive reach in the Euclidean space.

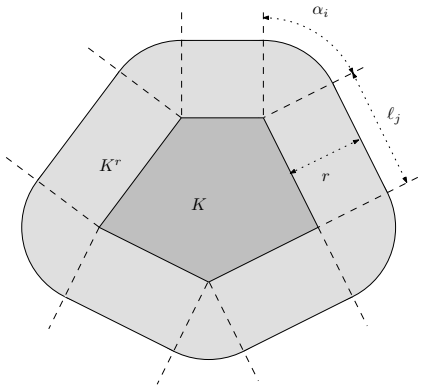


Figure II.3 – Offset of a polygon in the Euclidean plane

II.1.1 — Tube formulas: from Steiner to Weyl

Steiner-Minkowski formula. A tube formula for a compact set K in Euclidean space is a formula for the (Lebesgue) volume of the tubular neighborhoods K^r as a function of r. The offset $\mathcal{H}^d(K^r)$ grows as $O(r^d)$ as r goes to infinity. Steiner was the first to remark that, when K is a convex polygon in the Euclidean plane, the function $r \mapsto \mathcal{H}^d(K^r)$ is in fact a polynomial of degree two. Namely,

$$\mathcal{H}^d(K^r) = \mathcal{H}^2(K) + r\mathcal{H}^1(\partial K) + \pi r^2$$

The proof of this fact is (almost) contained in Figure II.3: every vertex with exterior angle α_i contributes a volume of $\alpha_i r^2$ to K^r, while every segment contributes $r \times \ell_j$. Summing these up on every segment and vertex yields the 2D Steiner formula.

Minkowski proved a similar polynomial behaviour for the volume of the offsets of any convex compact set in $\mathbb{R}^d$. Actually, he even proved that if K and L are two compact convex sets, and $\lambda K \oplus \mu L$ denotes the Minkowski sum $\{\lambda p + \mu q\,;\, p \in K,\ q \in L\}$, the volume $\mathcal{H}^d(\lambda K \oplus \mu L)$ is a homogeneous polynomial of degree d in the two variables $(\lambda, \mu) \in \mathbb{R}^+ \times \mathbb{R}^+$. Notice that if L is the unit ball, and $\lambda = 1$, we get the desired tube formula.

Weyl tube formula. Weyl [Wey39] proved that the polynomial behavior for $r \mapsto \mathcal{H}^d(K^r)$ shown above is also true for small values of r when K is a compact and smooth submanifold of $\mathbb{R}^d$. Moreover he also proved that the coefficients of this polynomial can be computed from the second fundamental form of K. The following proposition is an example of such a result in the case of an hypersurface bounding a domain:

PROPOSITION II.1. *Let K be a bounded d-dimensional domain with smooth*

boundary M. *Then, for sufficiently small values of* $r > 0$,

$$\mathcal{H}^d(K^r) = \mathcal{H}^d(K) + \sum_{k=1}^{d-1} \text{const}(d,k) r^{k+1} \int_M \left[\sum_{i_1 < \cdots < i_k} \kappa_{i_1} \dots \kappa_{i_k}(p) \right] dp \quad \text{(II.1)}$$

where $\kappa_1(p), \dots, \kappa_d(p)$ *are the principal curvatures at point* p *of* $\partial K = M$.

Proof. Let $\mathbf{n}$ be an oriented normal field on M. The map $\Phi : M \times \mathbb{R} \to \mathbb{R}^d$, $(p,t) \mapsto p + t\mathbf{n}$ is locally injective; by compactness of M, it is also injective on $M \times [0,r]$ for r small enough. One has $d_{(p,t)}\Phi = \text{id}_{T_pM} + t d_p\mathbf{n} + \mathbf{n}$, *ie.* $\left|\det(d_{(p,t)}\Phi)\right| = |\det(\text{id} + t d_p\mathbf{n})|$. For $t = 0$, $\det(d_{(p,t)}\Phi) = 1 > 0$ at any point $p \in M$; as a consequence, and by compactness of M again, this determinant remains positive for small enough values of t.

All of this allows us to apply the following change-of-variable formula for small valus of r:

$$\begin{aligned} \mathcal{H}^d(K^r) &= \mathcal{H}^d(K) + \int_{K^r \setminus K} 1 dx \\ &= \mathcal{H}^d(K) + \int_M \int_0^r \det(\text{id} + t d_p\mathbf{n}) dt d\mathbf{n} \end{aligned} \quad \text{(II.2)}$$

The eigenvalues of the map $d_p\mathbf{n}$ are the d principal curvatures of M at p, which means:

$$\det(\text{id} + t d_p\mathbf{n}) = \prod_{i=1}^{d-1} (1 + t\kappa_i(p)) = \sum_{k=1}^{d} t^k \left[\sum_{i_1 < \cdots < i_k} \kappa_{i_1}(p) \dots \kappa_{i_k}(p) \right] \quad \text{(II.3)}$$

We conclude the proof by putting Equation II.3 in Equation II.2. □

II.1.2 — Federer curvature measures

The contribution of Federer in [Fed59] to this theory is twofolds. First, he defines the notion of *reach* of a compact set, denoted by $\text{reach}(K)$, and proves that $r \mapsto \mathcal{H}^r(K)$ is polynomial in r for r smaller than the reach of K. The class of compact with positive reach includes both convex sets and smooth submanifolds. More importantly, he gives a *local* version of the tube formula, which takes into account the origin of every point of the offset (*ie.* where it projects on K). Using this formula, he is able to associate to any compact set K with positive reach $(d+1)$ curvature measures $\Phi_0(K), \dots, \Phi_d(K)$.

The reach of a compact set. The *reach* of a compact set K, denoted by $\text{reach}(K)$ is the minimum distance between K and its medial axis, see Figure II.4. Similar notion has been introduced and used under a couple of other names.

The *local feature size* of a compact K at a point $x \in K$, denoted by $\text{lfs}_K(x)$, was defined by Amenta and Bern [AB99]. It is the distance between x and the

medial axis of K. With this definition, the reach of K is equal to the *minimum local feature size* $\min_{x\in K} \mathrm{lfs}_K(x)$.

Sets with positive reach in $\mathbb{R}^d$ or in a Riemannian manifold are also said to have the *unique footpoint property* (UFP) [Kle80, Ban82] or to be *proximally smooth* [CS04, p. 75].

EXAMPLES. — It is well known that the projection to a closed convex set $K \subseteq \mathbb{R}^d$ is defined on the whole space, from what one deduces that the $\mathrm{reach}(K) = +\infty$. The reciprocal of this theorem (if $\mathrm{reach}(K) = +\infty$, then K is convex) is known as Motzkin theorem.

— From the proof of Proposition II.1, one sees that the reach of a smooth hypersurface is always positive. In fact, a smooth compact submanifold of $\mathbb{R}^d$ always has positive reach (this follows, for instance, from the tubular neighborhood theorem).

— The reach of submanifold M is always upper-bounded by the minimum radius of curvature of M; there is however no lower bound depending on this minimum radius of curvature. Indeed, consider two spheres of radius R at distance ε; then the reach of the union of those two sphere is $\frac{\varepsilon}{2}$ while the minimum curvature radius remains constant and equal to R. Similar examples can be constructed for connected manifolds.

Measures. Recall that a (nonnegative) *measure* μ associates to any (Borel) subset B of $\mathbb{R}^d$ a nonnegative number $\mu(B)$. It should also enjoy the following additivity property: if (B_i) is a countable family of disjoint subsets, then $\mu(\cup B_i) = \sum_i \mu(B_i)$. A measure on $\mathbb{R}^d$ can be thought of as a mass distribution on $\mathbb{R}^d$, which one can probe using Borel sets: $\mu(B)$ is the mass of μ contained in B.

The *restriction* of a measure μ to a subset $C \in \mathbb{R}^d$ is the measure $\mu_{|C}$ defined by $\mu_{|C}(B) = \mu(B \cap C)$. The simplest examples of measures are Dirac masses which, given a point $x \in \mathbb{R}^d$, are defined by $\delta_x(B) = 1$ if and only if x is in B. In what follows, we will also use the k-dimensional *Hausdorff measures* $\mathcal{H}^k$ whose definition has been recalled in chapter 1: if $S \subset \mathbb{R}^d$ is a k-dimensional submanifold, $\mathcal{H}^k_{|S}$ models a mass distribution uniformly distributed on S.

Curvature measures. The second contribution of Federer is to consider *local Steiner formulas* (or *local tube Formula*) instead of global ones. The global tube formula considers the volume of the offset K^r; for the local one, a Borel set B in K is fixed, and one considers the volume of the set of points in the offset K^r whose projection on K lie in B (*ie.* $p_K^{-1}(B) \cap K^r$). This defines a measure on K:

DEFINITION II.1. If E is a subset of $\mathbb{R}^d$ one introduces the measure $\mu_{K,E}$ defined as follows: for any subset $B \subseteq \mathbb{R}^d$, $\mu_{K,E}(B)$ is the d-volume of the set of points of E whose projection on K is in B, *ie.* $\mu_{K,E}(B) = \mathcal{H}^d(p_K^{-1}(B) \cap E)$. In other words, $\mu_{K,E}$ is the pushforward of the restriction of the Lebesgue

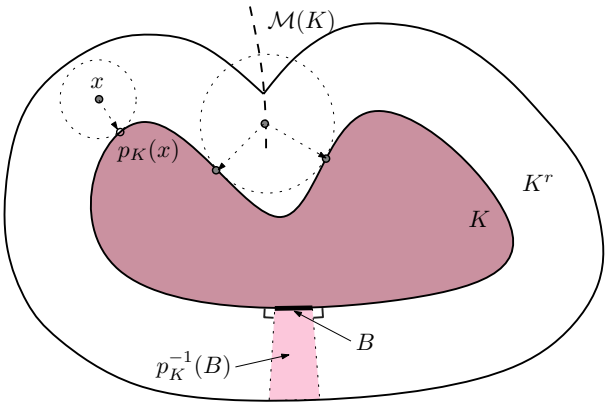


Figure II.4 – Boundary measure of $K \subset \mathbb{R}^d$. The medial axis $\mathrm{Med}(K)$ of K is the dashed line. Remark that the boundary of the offset ∂K^r is smooth everywhere but at its point of intersection with the medial axis.

measure to E by the projection p_K, which can be written more concisely as $\mu_{K,E} = p_{K\#}\, \mathcal{H}^d\big|_E$.

We will be particulary interested in the case where E is an offset of K, $E = K^r$ (Figure II.4). While the above definition makes sense for any compact $K \subseteq \mathbb{R}^d$, boundary measures have been mostly studied in the convex case and in the smooth case. Let us first give two examples in the convex case:

EXAMPLES. — Let S be a unit-length segment in the plane with endpoints a and b. The set S^r is the union of a rectangle of dimension $1 \times 2r$ whose points project on the segment and two half-disks of radius r whose points are projected on a and b. It follows that

$$\mu_{S,S^r} = 2r\, \mathcal{H}^1\big|_S + \frac{\pi}{2} r^2 \delta_a + \frac{\pi}{2} r^2 \delta_b$$

— If P is a convex solid polyhedron of $\mathbb{R}^3$, F its faces, E its edges and V its vertices, then one can see that:

$$\mu_{P,P^r} = \mathcal{H}^3\big|_P + r \sum_{f \in F} \mathcal{H}^2\big|_f + r^2 \sum_{e \in E} K(e)\, \mathcal{H}^1\big|_e + r^3 \sum_{v \in V} K(v) \delta_v$$

where $K(e)$ the angle between the normals of the faces adjacent to the edge e and $K(v)$ the solid angle formed by the normals of the faces adjacent to v. As shown by Steiner and Minkowski, for general convex polyhedra the measure μ_{K,K^r} can be written as a sum of weighted Hausdorff measures supported on the i-skeleton of K, and whose local density is the local external dihedral angle.

Federer showed that this polynomial behaviour for μ_{K,K^r} holds for any compact set $K \subset \mathbb{R}^d$ with reach at least R:

THEOREM II.2. *For any compact set $K \subseteq \mathbb{R}^d$ with reach greater than R, there exists $d+1$ (signed) measures $\Phi_K^0, \ldots, \Phi_K^d$ such that for any $r \leqslant R$, $\mu_{K,K^r} =$*

$\sum_{i=0}^{d} \omega_{d-i}\Phi_{K,i}r^i$ *(as usual,* ω_k *is the volume of the* k *unit sphere). These measures are uniquely defined and are called the curvature measures of* K.

The name *curvature measures* is motivated by the construction of Federer. Federer first proves that if $\mathrm{reach}(K) \geqslant R$, then for any $r \in]0, R[$ the boundary of the offset K^r is a $\mathcal{C}^{1,1}$ hypersurface — that is a $\mathcal{C}^1$ hypersurface with a Lipschitz normal field. He then extends Proposition II.1 to this less smooth setting; this allows him to define the curvature measures $\Phi_{K^r}(i)$ of K^r for any $r \in]0, R[$. Finally, the existence of curvature measures for K, as well as the polynomial behaviour for the volume of the offsets is obtained by approximation, by letting r go to zero.

In §II.A.1 we give a direct proof of existence of curvature measures without using such an approximation argument (this prove has the disadvantage of hiding why curvature measures are related to curvature, though).

THEOREM. *Given any compact set* K *with positive reach,* $\Phi_K^0(K)$ *is equal to the Euler-Poincaré characteristic of* K.

II.2 STABILITY OF BOUNDARY MEASURES

In this section, we suppose that E is a fixed open set with rectifiable boundary, and we obtain a quantitative stability theorem for the map $K \mapsto \mu_{K,E}$. What we mean by *stable* is that if two compact sets K and K' are close, then the measures $\mu_{K,E}$ and $\mu_{K',E}$ are also close. In order to be able to formulate a precise statement we need to choose a notion of distance on the space of compact subsets of $\mathbb{R}^d$ and on the set of measures on $\mathbb{R}^d$.

To measure the distance between two compact subsets K and K' of $\mathbb{R}^d$, we will use the Hausdorff distance: $d_H(K, K')$ is by definition the smallest positive constant η such that both $K' \subseteq K^\eta$ and $K \subseteq K'^\eta$. It is also the uniform distance between the two distance functions d_K and $d_{K'}$: $d_H(K, K') = \sup_{x\in\mathbb{R}^d} |d_K(x) - d_{K'}(x)|$. The next paragraph describes the distance we use to compare measures.

Wasserstein distance. The Wasserstein distance (of exponent 1) between two measures μ and ν on $\mathbb{R}^d$ having the same total mass $\mu(\mathbb{R}^d) = \nu(\mathbb{R}^d)$ is a nonnegative number which quantifies the cost of the optimal transportation from the mass distribution defined by μ to the mass distribution defined by ν (*cf.* [Vil03]). It is denoted by $W_1(\mu, \nu)$. More precisely, it is defined as

$$W_1(\mu, \nu) = \inf_{X,Y} \mathbb{E}[d(X, Y)]$$

where the infimum is taken on all pairs of $\mathbb{R}^d$-valued random variables X and Y whose law are μ and ν respectively. This distance is also known as the earth-mover distance, and has been used in vision [PWR89] and image retrieval [RTG00]. One of the interesting properties of the Wasserstein distance is the Kantorovich-Rubinstein duality theorem. Recall that a function $f : \mathbb{R}^d \to \mathbb{R}$ is 1-Lipschitz if for every choice of x and y, $|f(x) - f(y)| \leqslant \|x - y\|$.

THEOREM (Kantorovich-Rubinstein). *If μ and ν are two probability measures on $\mathbb{R}^d$, then*

$$W_1(\mu,\nu) = \sup_f \left| \int f d\mu - \int f d\nu \right|$$

where the supremum is taken on all 1-Lipschitz function in $\mathbb{R}^d$.

The Lipschitz function f in the theorem can be thought as a way of probing the measure μ. For example, if f is a tent function, *e.g.* $f(y) = (1 - \|x - y\|)_+$, then $\int f d\mu$ gives an information about the local density of μ near x. The Kantorovich-Rubinstein theorem asserts that if two measures μ and ν are Wasserstein-close, then one can control the probing error between μ and ν by Lipschitz functions.

NOTATION. If E is a (measurable) subset of $\mathbb{R}^d$, we will denote by $\|f\|_{L^1(E)}$ the L^1 norm of the restriction of any integrable function $f : \mathbb{R}^d \to \mathbb{R}$ to E. Likewise, $\|g\|_{\infty,E}$ will denote the supremum of $|g(x)|$, where x ranges in the set E.

The following proposition reduces the stability result of the map $K \mapsto \mu_{K,E}$ with respect to the Wasserstein distance to a stability result for the map $K \mapsto p_K|_E \in L^1(E)$.

PROPOSITION II.3. *If E is a subset of $\mathbb{R}^d$, and K and K' two compact sets, then*

$$W_1(\mu_{K,E}, \mu_{K',E}) \leqslant \|p_K(x) - p_{K'}(x)\|_{L^1(E)}$$

Proof. Let Z be a random variable whose law is $\mathcal{H}^d|_E$. Then, $X = p_K \circ Z$ and $Y = p_{K'} \circ Z$ have law $\mu_{K,E}$ and $\mu_{K',E}$ respectively. Hence by definition, $W_1(\mu_{K,E}, \mu_{K',E}) \leqslant \mathbb{E}(d(p_K \circ Z, p_K \circ Z))$, which is the desired bound. □

II.2.1 — Federer's stability theorem

The question of whether projection maps are stable already drawn attention in the past. In particular, Federer proved the following result in [Fed59]:

THEOREM. *Let R be a positive number, $K_n \subseteq \mathbb{R}^d$ be a sequence of compact sets whose reach is greater than R, which Hausdorff-converges to some compact $K \subseteq \mathbb{R}^d$ with* reach$(K) > R$. *Then p_{K_n} converges to p_K uniformly on K^R as n grows to infinity.*

A drawback of this theorem is that it does not give any information about the speed of convergence. More importantly, the very strong assumptions that all compact sets in the sequence have their reach bounded from below makes it completely unusable in the setting of geometric inference:

LEMMA II.4. *If a sequence of point clouds $(K_n)_{n\in\mathbb{N}}$ Hausdorff-converges to an infinite compact set $K \subseteq \mathbb{R}^d$, then the reach of K_n converges to zero as n goes to infinity.*

Proof. The reach of a point cloud is the minimum distance between any two of its points. Since K is infinite, it has an adherence point p; this means that

there exists a sequence p_i of points of K converging to p. For any $\varepsilon > 0$, there exists an i such that $\|p_i - p\| \leqslant \varepsilon$; then for n large enough, there are two points x_n, y_n in K_n such that $\|x_n - p\| \leqslant \varepsilon/3$ and $\|y_n - p_i\| \leqslant \varepsilon < 3$. Hence, $\mathrm{reach}(K_n) \leqslant \varepsilon/3$, which concludes the proof. □

The conclusion of the lemma still hold if the point clouds are replaced by embedded simplicial complexes, and K is *eg.* a smooth submanifold of $\mathbb{R}^d$.

In fact, if K is the union of two distinct points x and y, and $K_n = \{x + \frac{1}{n}(x - y), y\}$, p_{K_n} does not converge uniformly to p_K near the medial hyperplane of x and y. Hence, it is hopeless to expect uniform convergence of p_{K_n} to p_K on an arbitrary open set E without assumption on both K and the K_n.

II.2.2 — A first attempt at a quantitative result

Intuitively, one expects that the projections $p_K(x)$ and $p_{K'}(x)$ of a point x on two Hausdorff-close compact subsets can differ dramatically only if x lies close to the medial axis of one of the compact sets. This makes it reasonable to expect a L^1 convergence property of the projections. However, since the medial axis of a compact subset of $\mathbb{R}^d$ is generically dense (Proposition I.2), translating the above intuition into a proof isn't completely straightforward. We will make use of the critical point stability theorem from [CCSL09] (a generalization of which is given in Proposition IV.9):

THEOREM II.5. *(Critical point stability) Let K, K' be two compact sets with $d_H(K, K') \leqslant \varepsilon$. For any point x in the μ-medial axis of K, there exists a point y in the μ'-medial axis of K' with $\mu' = \mu + 2\sqrt{\varepsilon/d_K(x)}$ and $\|x - y\| \leqslant 2\sqrt{\varepsilon d_K(x)}$.*

For any $L > 0$, and two compact sets K and K', we will denote $\Delta_L(K, K')$ the set of points x of $\mathbb{R}^d \setminus (K \cup K')$ whose projections on K and K' are at least at distance L, i.e. $\|p_K(x) - p_{K'}(x)\| \geqslant L$. For technical reasons, we remove all points of the medial axes of K and K' from $\Delta_L(K, K')$. Since the Lebesgue measure both medial axes vanishes, this does not affect the measure of $\Delta_L(K, K')$.

A consequence of the critical point stability theorem is that $\Delta_L(K, K')$ lie close to the μ-medial axis of K for a certain value of μ (this Lemma is similar to [CCSL08, Theorem 3.1]):

LEMMA II.6. *Let $L > 0$ and K, K' be two compact sets and $\delta \leqslant L/2$ denote their Hausdorff distance. Then for any positive radius R, one has*

$$\Delta_L(K, K') \cap K^R \subseteq \mathrm{Med}_\mu(K)^{2\sqrt{R\delta}}$$

with

$$\mu = \left(1 + \left[\frac{L - \delta}{4R}\right]^2\right)^{-1/2} + 4\sqrt{\frac{\delta}{L}}$$

Proof. Let x be a point in $\Delta_L(K, K')$ with $d_K(x) \leqslant R$, and denote by p and p' its projections on K and K' respectively. By assumption, $\|p - p'\|$ is at least L. We

let q be the projection of p' on the sphere $\mathcal{S}(x, d_K(x))$, and let K_0 be the union of K and q. By hypothesis on the Hausdorff distance between K and K', there exists a point p'' in K such that $\|p'' - p'\| \leqslant \delta$. By definition of the distance to K, $\|x - p''\| \geqslant d_K(x)$: this means that $\|x - p'\| \geqslant d_K(x) - \delta$. Thus, because q is the projection of p' on the sphere $\mathcal{S}(x, d_K(x))$, the distance between p' and q is at most δ. Hence, $d_H(K, K_0)$ is at most 2δ.

By construction, the point x has two projections on K_0, and must belong to the μ_0-medial axis of K_0 for some value of μ. Letting m be the midpoint of the segment $[p, q]$, we are able to upper bound the value of μ_0:

$$\mu_0^2 \leqslant \|\nabla_x d_{K_0}\|^2 \leqslant \cos\left(\frac{1}{2}\angle(p - x, q - x)\right)^2 = \|x - m\|^2 / \|x - p\|^2$$

Since p, q belong to the sphere $B(x, d_K(x))$, one has $(p - q) \perp (m - x)$ and $\|x - p\|^2 = \|x - m\|^2 + \frac{1}{4}\|p - q\|^2$. This gives

$$\mu_0 \leqslant \left(1 + \frac{1}{4}\frac{\|p - q\|^2}{\|x - m\|^2}\right)^{-1/2} \leqslant \left[1 + \left(\frac{L - \delta}{2R}\right)^2\right]^{-1/2}$$

To get the second inequality we used $\|x - m\| \leqslant R$ and $\|p - q\| \geqslant L - \delta$.

In order to conclude, one only need to apply the critical point stability theorem (Theorem II.5) to the compact sets K and K_0 with $d_H(K, K_0) \leqslant 2\delta$. Since x is in the μ_0-medial axis of K_0, there should exist a point y in $\mathrm{Med}_\mu(K)$ with $\|x - y\| \leqslant 2\sqrt{R\delta}$ and $\mu = \mu_0 + 4\sqrt{\delta/L}$. □

Using this Lemma, one is able to prove the following non-quantitative convergence result for projections:

PROPOSITION II.7. *If (K_n) Hausdorff converges to a compact $K \subseteq \mathbb{R}^d$, then for any bounded open set E, $\lim_{n\to+\infty} \|p_{K_n} - p_K\|_{L^1(E)} = 0$.*

Proof. Fix $L > 0$, and suppose K and K' are given. One can decompose the set E between the set of points where the projections differ by at least L (*ie.* $\Delta_L(K, K') \cap E$) and the remaining points. This gives the bound:

$$\|p_{K'} - p_K\|_{L^1(E)} \leqslant L\mathcal{H}^d(E) + \mathcal{H}^d(\Delta_L(K, K') \cap E)\,\mathrm{diam}(K \cup K')$$

Now, take $R = \sup_E \|d_K\|$, so that E is contained in the tubular neighborhood K^R, and fix $L = \varepsilon/\mathcal{H}^d(E)$. Then, for $\delta = d_H(K, K')$ small enough (e.g. less than some δ_0), the value of μ given in Lemma II.6 is smaller than one. Denote by μ_0 the value given by the lemma for δ_0. Then

$$\|p_{K'} - p_K\|_{L^1(E)} \leqslant \varepsilon + \mathcal{H}^d(\mathrm{Med}_{\mu_0}(K)^{2\sqrt{R\delta}})\,\mathrm{diam}(K \cup K') \tag{II.4}$$

Being compact, $\mathrm{Med}_{\mu_0}(K)$ is the intersection of its tubular neighborhoods. Combining this with the outer-regularity of the Lebesgue measure gives:

$$\lim_{\delta\to 0} \mathcal{H}^d(\mathrm{Med}_{\mu_0}(K)^{2\sqrt{R\delta}}) = \mathcal{H}^d(\mathrm{Med}_{\mu_0}(K)) = 0.$$

Putting this limit in equation (II.4) concludes the proof. □

Using Theorem I.15, one can give a more quantitative result. Notice that the meaning of *locally* in the next statement could also be made quantitative using the same proof.

PROPOSITION II.8. *The map* $K \mapsto p_K \in L^1(E)$ *is locally* h*-Hölder for any exponent* $h < \frac{1}{2(2d-1)}$.

Proof. As in the previous proof, we will let $R = \|d_K\|_{E,\infty}$, so that E is contained in the tubular neighborhood K^R. Remark first that if a point x is such that $d_K(x) \leqslant \frac{1}{2}L - d_H(K, K')$, then by definition of the Hausdorff distance, $d_{K'}(x) \leqslant \frac{1}{2}L$. In particular, the orthogonal projections of x on K and K' are at distance at most L. Said otherwise, the set $\Delta_L(K, K')$ is contained in the complementary of the $\frac{L}{2} - \delta$ tubular neighborhood of K, with $\delta := d_H(K, K')$. Using this fact and the result of Lemma II.6, we have:

$$\Delta_L(K, K') \cap K^R \subseteq \mathrm{Med}_\mu(K)^{2\sqrt{R\delta}} \cap (\mathbb{R}^d \setminus K^{\frac{L}{2}-\delta}) \tag{II.5}$$

$$\subseteq \left(\mathrm{Med}_\mu(K) \cap \left(\mathbb{R}^d \setminus K^{\frac{L}{2}-\delta-2\sqrt{R\delta}}\right)\right)^{2\sqrt{R\delta}} \tag{II.6}$$

$$\text{where } \mu = \left(1 + \left[\frac{L-\delta}{4R}\right]^2\right)^{-1/2} + 4\sqrt{\frac{\delta}{L}} \tag{II.7}$$

We now choose L to be δ^h, where $h > 0$, and see for which values of h we are able to get a converging bound. For $h < 1/2$, the radius $\frac{1}{2}(L-\delta) - 2\sqrt{R\delta}$ will be greater than $L/3$ as soon as as soon as δ is small enough. For these values,

$$\Delta_L(K, K') \cap K^R \subseteq \left(\mathrm{Med}_\mu(K) \cap (\mathbb{R}^d \setminus K^{L/3})\right)^{2\sqrt{R\delta}} \tag{II.8}$$

The μ above, given by Lemma II.6 can then be bounded as follows. Note that the constants in the "big O" will always be positive in the remaining of the proof. From Eq. (II.7), one deduces:

$$\mu = 1 + O(-\delta^{2h} + \delta^{1/2-h/2})$$

This term will be asymptotically smaller than 1 provided that $2h < 1/2 - h/2$ *ie.* $h < 1/5$, in which case $\mu = 1 - O(\delta^{2h})$. By definition of the covering number, one has:

$$\begin{aligned}\mathcal{H}^d(\Delta_L(K, K') \cap K^R) &\leqslant \mathcal{H}^d\left[\left(\mathrm{Med}_\mu(K) \cap \left(\mathbb{R}^d \setminus K^{L/3}\right)\right)^{2\sqrt{R\delta}}\right] \\ &\leqslant \mathcal{N}\left(\mathrm{Med}_\mu(K) \cap \left(\mathbb{R}^d \setminus K^{L/3}\right), 2\sqrt{R\delta}\right) \times O(\delta^{d/2})\end{aligned} \tag{II.9}$$

The covering numbers of the intersection $\mathrm{Med}_\mu(K) \cap \left(\mathbb{R}^d \setminus K^{L/3}\right)$ can be

bounded using Theorem I.15:

$$\begin{aligned}\mathcal{N}\left(\mathbf{Med}_\mu(K)\cap\left(\mathbb{R}^d\setminus K^{L/2}\right),2\sqrt{R\delta}\right)&\\ =\mathcal{N}(\partial K,L/4)\,O\left(\left[\frac{\operatorname{diam}(K)/\sqrt{R\delta}}{\sqrt{1-\mu^2}}\right]^{d-1}\right)&\\ =\mathcal{N}(\partial K,L/4)\,O\left(\delta^{-(h+\frac12)(d-1)}\right)&\end{aligned}\tag{II.10}$$

Combining equations (II.9) and (II.10), and using the (crude) estimation $\mathcal{N}(\partial K, L/4) = O(1/L^d) = O(\delta^{-hd})$,

$$\begin{aligned}\mathcal{H}^d(\Delta_L(K,K')\cap K^R)&\leqslant\mathcal{N}(\partial K,L/4)\,O(\delta^{-h(d-1)-\frac12(d-1)+\frac12 d})\\ &\leqslant O\left(\delta^{\frac12-h(2d-1)}\right)\end{aligned}$$

Hence, following the proof of Proposition II.7,

$$\begin{aligned}\|p_{K'}-p_K\|_{L^1(E)}&\leqslant L\mathcal{H}^d(E)+\mathcal{H}^d(\Delta_L(K,K')\cap E)\operatorname{diam}(K\cup K')\\ &=O(\delta^h+\delta^{1/2-h(2d-1)})\end{aligned}$$

The second term converges to zero as $\delta = d_H(K,K')$ does if $h < \frac{1}{2(2d-1)}$. This concludes the proof. □

II.2.3 — An optimal projection stability theorem

In this paragraph we give a second upper bound for the L^1-distance between the restriction of two projection functions p_K and p'_K to a subset $E\subset\mathbb{R}^d$, with an optimal Hölder exponent (which is also much better than the one of Proposition II.8). The actual proof of the projection stability theorem relies on a new theorem in convex analysis, and is postponed to the next section.

THEOREM II.9 (Projection Stability). *Let E be an open set in $\mathbb{R}^d$ with rectifiable boundary, K and K' be two compact subsets of $\mathbb{R}^d$ and $R_K = \|d_K\|_{\infty,E}$. Then, there is a constant $C_1(d)$ such that*

$$\|p_K-p_{K'}\|_{L^1(E)}\leqslant C_1(d)[\mathcal{H}^d(E)+\operatorname{diam}(K)\mathcal{H}^{d-1}(\partial E)]\times\sqrt{R_K d_H(K,K')}$$

assuming $d_H(K,K')\leqslant\min(R_K,\operatorname{diam}(K),\operatorname{diam}(K)^2/R_K)$.

Optimality of the projection stability theorem. Let us now comment on the optimality of this projection stability theorem. First, the speed of convergence of $\mu_{K',E}$ to $\mu_{K,E}$ cannot be (in general) faster than $O(d_H(K,K')^{1/2})$. Indeed, if D is the closed unit disk in the plane, and P_ℓ is a regular polygon inscribed in D with sidelength ℓ (see Fig. II.5), then $d_H(D,P_\ell)\simeq\ell^2$. Now we let E be the disk of radius $1+R$. Then, a constant fraction of the mass of E will be projected onto the vertices of P_ℓ by the projection p_{P_ℓ} (lightly shaded area in Figure II.5). The cost of spreading out the mass concentrated on these vertices

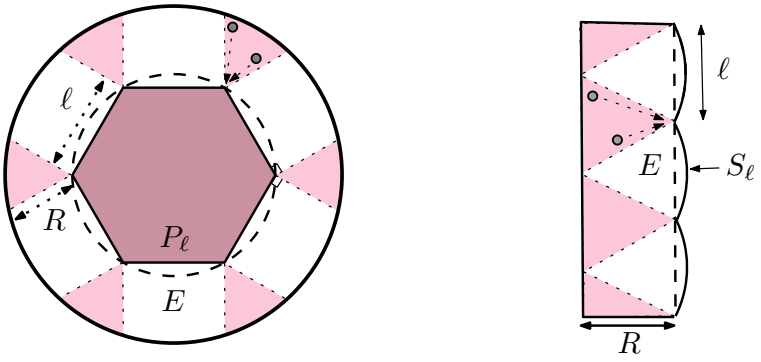


Figure II.5 – Optimality of the stability theorem. The compact set K — the dashed disk in the first figure and the dashed segment on the second one — is approximatex by a sequence of compact sets (resp. S_ℓ and P_ℓ) with sharp vertices. A constant fraction of the mass in E projects onto those vertices.

to get a uniform measure on the circle is proportional to the distance between consecutive vertices, so $W_1(\mu_{D,E}, \mu_{P_\ell,E}) = \Omega(\ell)$. Hence, by Proposition II.3, $\|p_D - p_{P_\ell}\|_{L^1(E)} = \Omega(\ell) = \Omega(d_H(D, P_\ell)^{1/2})$. Note that this $O(d_H(K, K')^{1/2})$ behavior does not come from the curvature of the disk, since one can also find an example of a family of compacts S_ℓ made of small circle arcs converging to the unit segment S such that $\|p_S - p_{S_\ell}\|_{L^1(E)} = \Omega(d_H(S, S_\ell)^{1/2})$ (see Fig. II.5).

The second remark concerning the optimality of the theorem is that the second term of the bound involving $\mathcal{H}^{d-1}(\partial E)$ cannot be avoided. Indeed, let us suppose that a bound $\|p_K - p_{K'}\|_{L^1(E)} \leqslant C(K)\mathcal{H}^d(E)\sqrt{\varepsilon}$ were true around K, where ε is the Hausdorff distance between K and K'. Now let K be the union of two parallel hyperplanes at distance R intersected with a big sphere centered at a point x of their medial hyperplane M. Let E_ε be a ball of radius $\varepsilon/2$ tangent to M at x and K_ε be the translate by ε of K along the common normal of the hyperplanes such that the ball E_ε lies in the slab bounded by the medial hyperplanes of K and K_ε. Then, $\|p_K - p_{K'}\|_{L^1(E_\varepsilon)} \simeq R \times \mathcal{H}^d(E_\varepsilon)$, which exceeds the assumed bound for a small enough ε.

Proof of the projection stability theorem. The projection stability theorem will follow from a more general theorem on the L^1 norm of the difference between the gradients of two convex functions defined on some open set E with rectifiable boundary. The connection between projections and convex analysis is that the projection function p_K derives from a convex potential ν_K:

LEMMA II.10. *The function $\nu_K : \mathbb{R}^d \to \mathbb{R}$ defined by $\nu_K(x) = \|x\|^2 - d_K(x)^2$ is convex with gradient $\nabla\nu_K = 2p_K$ almost everywhere.*

Proof. By definition, $\nu_K(x) = \sup_{y\in K} \|x\|^2 - \|x-y\|^2 = \sup_{y\in K} \nu_{K,y}(x)$ with $\nu_{K,y}(x) = 2\langle x|y\rangle - \|y\|^2$. Hence ν_K is convex as a supremum of affine functions, and is differentiable almost everywhere. Then, $\nabla\nu_K(x) = 2d_K(x)\nabla_x d_K - 2x$. The equality $\nabla d_K(x) = (p_K(x) - x)/d_K(x)$, valid when x is not in the medial axis, concludes the proof. □

Hence if K, K' are two compact subsets of $\mathbb{R}^d$, the L^1 distance between the two projections can be written in term of the L^1 distance between $\nabla\nu_K$ and $\nabla\nu_{K'}$: $\|p_K - p_{K'}\|_{L^1(E)} = 1/2\,\|\nabla\nu_K - \nabla\nu_{K'}\|_{L^1(E)}$. Moreover, denoting by $R_K = \|d_K\|_{\infty,E}$, the two following properties for the functions ν_K and $\nu_{K'}$ can be deduced from a simple calculation:

LEMMA II.11. *If* $d_H(K,K') \leqslant \min(R_K, \operatorname{diam}(K))$*, then*

$$\left\|d_K^2 - d_{K'}^2\right\|_{\infty,E} \leqslant 3 d_H(K,K') R_K$$
$$\operatorname{diam}(\nabla\nu_K(E) \cup \nabla\nu_{K'}(E)) \leqslant 3 \operatorname{diam}(K)$$

Let us now state the stability result for gradients of convex functions:

THEOREM II.12. *Let* E *be an open subset of* $\mathbb{R}^d$ *with rectifiable boundary, and* f, g *be two locally convex functions on* E *such that* $\operatorname{diam}(\nabla f(E) \cup \nabla g(E)) \leqslant k$*. Then,*

$$\|\nabla f - \nabla g\|_{L^1(E)} \leqslant C_2(n) \times \left(\mathcal{H}^d(E) + (k + \|f-g\|_{\infty,E}^{1/2})\mathcal{H}^{d-1}(\partial E)\right) \|f-g\|_{\infty,E}^{1/2}$$

From this theorem and the two previous lemmas one easily deduces the projection stability theorem.

II.2.4 — A L^1 stability theorem for gradient of convex functions

We now turn to the proof of the Theorem II.12, which is organized as follows. In Proposition II.13, we give a proof of the 1-dimensional case. The general case will follow using an argument of *integral geometry – ie.* we will integrate the 1-dimensional inequality over the set of lines of $\mathbb{R}^d$ and use the Cauchy-Crofton formulas to get the d-dimensional inequality.

Cauchy-Crofton formulas. The following formulas, usually called Cauchy-Crofton formulas, gives a way to compute the volume of a set E in terms of the expectation of the length of $E \cap \ell$ where ℓ is a random line in $\mathbb{R}^d$. More precisely, if one denotes by $\mathcal{L}^d$ the set of oriented lines of $\mathbb{R}^d$, and by $d\mathcal{L}^d$ the properly normalized rigid motion invariant measure on $\mathcal{L}^d$, we have

$$\mathcal{H}^d(E) = \frac{1}{\omega_{d-1}} \int_{\ell \in \mathcal{L}^d} \operatorname{length}(\ell \cap E)\, d\mathcal{L}^d \tag{II.11}$$

where ω_{d-1} is the $(d-1)$-volume of the unit sphere in $\mathbb{R}^d$. If S is a rectifiable hypersurface of $\mathbb{R}^d$, then

$$\mathcal{H}^{d-1}(S) = \frac{1}{2\beta_{d-1}} \int_{\ell \in \mathcal{L}^d} \#(\ell \cap S)\, d\mathcal{L}^d \tag{II.12}$$

where β_{d-1} is the $(d-1)$-volume of the unit ball in $\mathbb{R}^{d-1}$.

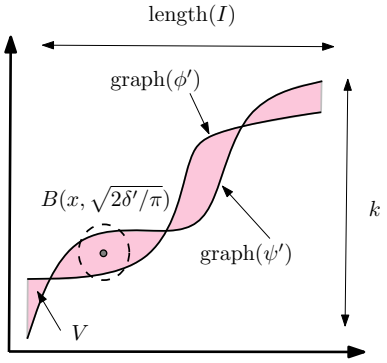


Figure II.6 – Proof of the 1-dimensional inequality.

PROPOSITION II.13. *If* I *is an interval, and* $\varphi : I \to \mathbb{R}$ *and* $\psi : I \to \mathbb{R}$ *are two convex functions such that* $\operatorname{diam}(\varphi'(I) \cup \psi'(I)) \leqslant k$*, then letting* $\delta = \|\varphi - \psi\|_{L^\infty(I)}$*,*

$$\int_I \left|\varphi' - \psi'\right| \leqslant 8\pi(\text{length}(I) + k + \delta^{1/2})\delta^{1/2}$$

Proof. Since φ and ψ are convex, their derivatives are non-decreasing. Let V be the closure of the set of points (x, y) in $I \times \mathbb{R}$ such that y is in the segment $[\varphi'(x), \psi'(x)]$ (or $[\psi'(x), \varphi'(x)]$ if $\varphi'(x) \geqslant \psi'(x)$). By definition of V, $\int_I |\varphi' - \psi'| = \mathcal{H}^2(V)$.

If D is a disk included in V and $[x_0, x_1] \subset I$ is the projection of D on the x-axis , then the sign of the derivative of the difference $\kappa = \varphi - \psi$ does not change on $[x_0, x_1]$. Assuming w.l.o.g. that κ is non decreasing on $[x_0, x_1]$, we have $|\kappa(x_0) - \kappa(x_1)| = \int_{x_0}^{x_1} |\kappa'| \geqslant \mathcal{H}^2(D)$

But since $\|\kappa\|_\infty = \delta$, the area of D cannot be greater than δ. Thus, if p is any point of V, for any $\delta' > \delta$ the disk $B(p, \sqrt{2\delta'/\pi})$ necessarily intersects the boundary ∂V. This proves that V is contained in the offset $(\partial V)^{\sqrt{2\delta/\pi}}$. It follows that:

$$\int_I \left|\varphi' - \psi'\right| \leqslant \mathcal{H}^2\left((\partial V)^{\sqrt{2\delta/\pi}}\right) \tag{II.13}$$

Now, ∂V can be written as the union of two xy-monotone curves Φ and Ψ joining the lower left corner of V and the upper right corner of V so that $(\partial V)^r \subseteq \Phi^r \cup \Psi^r$.

We now find a bound for $\mathcal{H}^2(\Phi^r)$ (the same bound will of course apply to Ψ). Since the curve Φ is xy-monotone, we have $\text{length}(\Phi) \leqslant \text{length}(I) + k$. Thus, for any $r > 0$ there exists a subset $X \subseteq \Phi$ of $N = \lceil (\text{length}(I) + k)/r \rceil$ points such that $\Phi \subseteq X^r$, implying

$$\mathcal{H}^2(\Phi^r) \leqslant \mathcal{H}^2(X^{2r}) \leqslant 4\pi r^2 N \leqslant 4\pi r(\text{length}(I) + k + r) \tag{II.14}$$

Using equations (II.13) and (II.14), and $\sqrt{2\delta/\pi} \leqslant \sqrt{\delta}$, one finally obtains:

$$\int_I \left|\varphi' - \psi'\right| \leqslant \mathcal{H}^2\left(\Phi^{\sqrt{2\delta/\pi}} \cup \Psi^{\sqrt{2\delta/\pi}}\right) \leqslant 8\pi(\text{length}(I) + k + \sqrt{\delta})\sqrt{\delta}$$

□

Proof of Theorem II.12. The 1-dimensional case follows directly from proposition II.13: in this case, E is a countable union of intervals on which f and g satisfy the hypothesis of the proposition. Summing the inequalities gives the result with $C_2(1) = 8\pi$.

We now turn to the general case. Given any L^1 vector-field X one has

$$\int_E \|X\| \, dx = \frac{d}{2\omega_{d-2}} \int_{\ell \in \mathcal{L}^d} \int_{y \in \ell \cap E} |\langle X(y) | u(\ell) \rangle| \, dy d\ell$$

where $u(\ell)$ is a unit directing vector for ℓ (see Lemma III.4 in [CCSM07] for a proof of this formula). Letting $X = \nabla f - \nabla g$, $f_\ell = f|_{\ell \cap E}$ and $g_\ell = g|_{\ell \cap E}$, one gets, with $D(d) = d/(2\omega_{d-2})$,

$$\begin{aligned} \int_E \|\nabla f - \nabla g\| &= D(d) \int_{\ell \in \mathcal{L}^d} \int_{y \in \ell \cap E} |\langle \nabla f - \nabla g | u(\ell) \rangle| \, dy d\ell \\ &= D(d) \int_{\ell \in \mathcal{L}^d} \int_{y \in \ell \cap E} \left| f'_\ell - g'_\ell \right| dy d\ell \end{aligned}$$

The functions f_ℓ and g_ℓ satisfy the hypothesis of the one-dimensional case, so that for each choice of ℓ, and with $\delta = \|f - g\|_{L^\infty(E)}$,

$$\int_{y \in \ell \cap E} \left| f'_\ell - g'_\ell \right| dy \leqslant 8\pi\delta^{1/2} \times (\mathcal{H}^1(E \cap \ell) + (k + \delta^{1/2})\mathcal{H}^0(\partial E \cap \ell))$$

It follows by integration on $\mathcal{L}^d$ that

$$\begin{aligned} &\int_E \|\nabla f - \nabla g\| \leqslant 8\pi D(d)\delta^{1/2} \\ &\quad \times \left(\int_{\mathcal{L}^d} \mathcal{H}^1(E \cap \ell) d\mathcal{L}^d + (k + \delta^{1/2}) \int_{\mathcal{L}^d} \mathcal{H}^0(\partial E \cap \ell) d\mathcal{L}^d \right) \end{aligned}$$

The formula (II.11) and (II.12) show that the first integral in the second term is equal (up to a constant) to the volume of E and the second one to the $(d-1)$-measure of ∂E. This proves the theorem with $C_2(d) = 8\pi D(d)(\omega_{d-1} + 2\beta_{d-1})$. □

II.2.5 — Stability of curvature measures

The definition of Wasserstein distance assumes that both measures are positive and have the same mass. While this is true for $\mu_{K,E}$ and $\mu_{K',E}$ (whose mass is the volume of E), this is not the case anymore when considering μ_{K,K^r} and μ_{K',K'^r} whose mass are respectively $\mathcal{H}^d(K^r)$ and $\mathcal{H}^d(K'^r)$. We thus need to introduce another distance on the space of (signed) measures.

DEFINITION II.2. The Kantorovich-Rubinstein theorem makes it natural to introduce the *bounded-Lipschitz distance* between two measures μ and ν as

follows:

$$d_{bL}(\mu,\nu) = \sup_{f\in BL_1(\mathbb{R}^d)} \left| \int f d\mu - \int f d\nu \right|$$

the supremum being taken on the space of 1-Lipschitz functions f on $\mathbb{R}^d$ such that $\sup_{\mathbb{R}^d} |f| \leqslant 1$.

PROPOSITION II.14. *If K, K' are compact subsets of $\mathbb{R}^d$,*

$$d_{bL}(\mu_{K,K^r}, \mu_{K',K'^r}) \leqslant \int_{K^r\cap K'^r} \|p_K(x) - p_{K'}(x)\|\, dx + \mathcal{H}^d(K^r \Delta K'^r)$$

Proof. Let φ is a 1-Lipschitz function on $\mathbb{R}^d$ bounded by 1. Using the change-of-variable formula one has:

$$\begin{aligned}
&\left| \int \varphi(x) d\mu_{K,K^r} - \int \varphi(x) d\mu_{K',K'^r} \right| \\
&\qquad = \left| \int_{K^r} \varphi\circ p_K(x) dx - \int_{K'^r} \varphi\circ p_{K'}(x) dx \right| \\
&\qquad \leqslant \int_{K'^r\cap K^r} |\varphi\circ p_K(x) - \varphi\circ p_{K'}(x)|\, dx + \int_{K\Delta K'} |\varphi(x)|\, dx
\end{aligned}$$

By the Lipschitz condition, $|\varphi\circ p_K(x) - \varphi\circ p_{K'}(x)| \leqslant \left|p_K(x) - p'_K(x)\right|$, thus giving the desired inequality. □

The new term appearing in this proposition involves the volume of the symmetric difference $K^r\Delta K'^r$. In order to get a result similar to the projection stability theorem but for the map $K \mapsto \mu_{K,K^r}$, we need to study how fast this symmetric difference vanishes as K' converges to K. It is not hard to see that if $d_H(K,K')$ is smaller than ε, then $K^r\Delta K'^r$ is contained in $K^{r+\varepsilon} \setminus K^{r-\varepsilon}$. Assuming that $d_H(K,K') < \varepsilon$, using the coarea formula (see [Mor88]), we can bound the volume of this annulus around K as follows:

$$\mathcal{H}^d(K^r\Delta K'^r) \leqslant \int_{r-\varepsilon}^{r+\varepsilon} \mathcal{H}^{d-1}(\partial K^s) ds$$

From the bound on the $\mathcal{H}^{d-1}$-volume of ∂K^r proved in the previous chapter (proposition I.16) we easily get:

COROLLARY II.15. *For any compact sets $K, K' \subseteq \mathbb{R}^d$, with $d_H(K,K') \leqslant r/2$,*

$$\mathcal{H}^d(K^r\Delta K'^r) \leqslant 2\mathcal{N}(K, r/2)\omega_{d-1}(3r) \times d_H(K,K') = O(d_H(K,K'))$$

Stability of approximate curvature measures. Combining Proposition II.14, Corollary II.15 and the projection stability theorem, one easily obtains the following stability result for boundary measures μ_{K,K^r}:

THEOREM II.16. *If K and K' are two compact sets of $\mathbb{R}^d$,*

$$d_{bL}(\mu_{K,K^r}, \mu_{K',K'^r}) \leqslant C_3(d)\mathcal{N}(K, r/2) r^d [r + \operatorname{diam}(K)] \sqrt{\frac{d_H(K,K')}{r}}$$

provided that $d_H(K, K') \leqslant \min(\operatorname{diam} K, r/2, r^2/\operatorname{diam} K)$.

To define the approximate curvature measures, let us fix a sequence (r_i) of $d+1$ distinct numbers $0 < r_0 < \ldots < r_d$. For any compact set K and Borel subset $B \subset K$, we let $\left[\Phi^{(r)}_{K,i}(B)\right]_i$ be the solutions of the linear system

$$\forall i \text{ s.t } 0 \leqslant i \leqslant d, \ \sum_{j=0}^{d} \omega_{d-j} \Phi^{(r)}_{K,j}(B) r_i^{d-j} = \mu_{K,K^{r_i}}(B)$$

We call $\Phi^{(r)}_{K,j}$ the (r)-approximate curvature measure. Since this is a linear system, the functions $\Phi^{(r)}_{K,i}$ also are additive. Hence the (r)-approximate curvature measures $\Phi^{(r)}_{K,i}$ are signed measures on $\mathbb{R}^d$. We also note that if K has a reach greater than r_d, then the measures $\Phi^{(r)}_{K,i}$ coincide with Federer's curvature measures of K, as introduced in §II.1.2. Thanks to these remarks and to theorem II.16, we have:

COROLLARY II.17. *Suppose* $(r_i)_{i\in\{0,\ldots,d\}}$ *is given as above. For each compact set* K *whose reach is greater than* r_d*, there exist a constant* $C(K,(r),d)$ *depending on* K*,* (r_i) *and* d *such that for any* $K' \subseteq \mathbb{R}^d$ *close enough to* K*,*

$$d_{bL}\left(\Phi^{(r)}_{K',i}, \Phi^{i}_{K}\right) \leqslant C(K,(r),d) d_H(K,K')^{1/2}$$

This corollary gives a way to approximate the curvature measures of a compact set K with positive reach from the (r)-approximate curvature measures of any point cloud close to K.

Stability of the pushforward of a measure by a projection. The boundary measure defined above are a special case of *pushforward* of a measure by a function. The pushforward of a measure μ on $\mathbb{R}^d$ by the projection p_K is another measure, denoted by $p_{K\#}\mu$, concentrated on K and defined by the formula: $p_{K\#}\mu(B) = \mu(p_K^{-1}(B))$.

The stability results for the boundary measures $K \mapsto \mu_{K,E}$ can be generalized to prove the stability of the map $K \mapsto p_{K\#}\mu$ where μ has a density $u : \mathbb{R}^d \to \mathbb{R}^+$, which means that $\mu(B) = \int_B u(x)dx$. We need the measure μ to be finite, which is the same as asking that the function u belongs to the space $L^1(\mathbb{R}^d)$ of integrable functions.

We also need the function $u \in L^1(\mathbb{R}^d)$ to have *bounded variation*. We recall the following basic facts of the theory of functions with bounded variation, taken from [AFP00]. If u is an integrable function on $\mathbb{R}^d$, the *total variation* of u is

$$\operatorname{var}(u) = \sup\left\{\int_{\mathbb{R}^d} u \operatorname{div}\varphi ; \varphi \in \mathcal{C}^1_c(\mathbb{R}^d), \|\varphi\|_\infty \leqslant 1\right\}$$

A function $u \in L^1(\mathbb{R}^d)$ has *bounded variation* if $\operatorname{var}(u) < +\infty$. The set of functions of bounded variation on Ω is denoted by $BV(\mathbb{R}^d)$. We also mention that if u is Lipschitz, then $\operatorname{var}(u) = \|\nabla u\|_{L^1(\mathbb{R}^d)}$.

THEOREM II.18. *Let μ be a measure with density $u \in BV(\mathbb{R}^d)$ with respect to the Lebesgue measure, and K be a compact subset of $\mathbb{R}^d$. We suppose that the support of u is contained in the offset K^R. Then, if $d_H(K, K')$ is small enough,*

$$d_{bL}(p_{K\#}\mu, p_{K'\#}\mu) \leqslant C(d)\left(\|u\|_{L^1(K^R)} + \operatorname{diam}(K)\operatorname{var}(u)\right) \times \sqrt{R} d_H(K, K')^{1/2}$$

Proof. We begin with the additional assumption that u has class $\mathcal{C}^\infty$. The function u can be written as an integral over $t \in \mathbb{R}$ of the characteristic functions of its superlevel sets $E_t = \{u > t\}$, *ie.* $u(x) = \int_0^\infty \chi_{E_t}(x)dt$. Fubini's theorem then ensures that for any 1-Lipschitz function f defined on $\mathbb{R}^d$ with $\|f\|_\infty \leqslant 1$,

$$\begin{aligned} p_{K'\#}\mu(f) &= \int_{\mathbb{R}^d} f \circ p_{K'}(x) u(x) dx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} f \circ p_{K'}(x) \chi_{\{u \geqslant t\}}(x) dx dt \end{aligned}$$

By Sard's theorem, for almost any t, $\partial E_t = u^{-1}(t)$ is a $(n-1)$-rectifiable subset of $\mathbb{R}^d$. Thus, for those t the previous corollary implies, for $\varepsilon = d_H(K, K') \leqslant \varepsilon_0 = \min(R, \operatorname{diam}(K), \operatorname{diam}(K)^2/R_K)$,

$$\begin{aligned} \int_{E_t} |f \circ p_K(x) - f \circ p_{K'}(x)|\, dx &\leqslant \|p_K - p_{K'}\|_{L^1(E_t)} \\ &\leqslant C(d)[\mathcal{H}^d(E_t) + \operatorname{diam}(K)\mathcal{H}^{d-1}(\partial E_t)]\sqrt{R\varepsilon} \end{aligned}$$

Putting this inequality into the last equality gives

$$|p_{K\#}\mu(f) - p_{K'\#}\mu(f)| \leqslant C(d)\left(\int_{\mathbb{R}} \mathcal{H}^d(E_t) + \operatorname{diam}(K)\mathcal{H}^{d-1}(\partial E_t) dt\right)\sqrt{R\varepsilon}$$

Using Fubini's theorem again and the coarea formula one finally gets that

$$|p_{K\#}\mu(f) - p_{K'\#}\mu(f)| \leqslant C(d)\left(\|u\|_{L^1(K^R)} + \operatorname{diam}(K)\operatorname{var}(u)\right)\sqrt{R\varepsilon}.$$

This proves the theorem in the case of a $\mathcal{C}^\infty$ function u. To conclude the proof in the general case, one has to approximate the bounded variation function u by a sequence of $\mathcal{C}^\infty$ functions (u_n) such that both $\|u - u_n\|_{L^1(K^R)}$ and $|\operatorname{var}(u) - \operatorname{var}(u_n)|$ converge to zero, which is possible thanks to Theorem 3.9 in [AFP00]. □

REMARK. Taking $u = \chi_E$ where E is a suitable open set shows that conversely, Theorem II.9 can also be recovered from Theorem II.18.

II.3 COMPUTING (WITH) BOUNDARY MEASURES

If $C = \{p_i; 1 \leqslant i \leqslant n\}$ is a point cloud, that is a finite set of points of $\mathbb{R}^d$, then μ_{C,C^r} is a sum of weighted Dirac measures: letting $\operatorname{Vor}_C(p_i)$ denote the

Voronoi cell of p_i, we have:

$$\mu_{C,C^r} = \sum_{i=1}^{n} \mathcal{H}^d(\mathrm{Vor}_C(p_i) \cap C^r)\delta_{p_i}$$

Hence, computing boundary measures amounts to find the volume of intersections of Voronoi cells with balls. This method is practical in dimension 3 (and we will detail it in Chap. III) but in higher dimensions it becomes prohibitive due to the exponential cost of Voronoi diagrams computations. We instead compute approximations of boundary measures using a Monte-Carlo method. Let us first recall some standard facts about this family of methods.

II.3.1 — Monte-Carlo approximation of boundary measures

Convergence of empirical measures. If μ is a probability measure on $\mathbb{R}^d$, one can define another measure as follows: let $X_1, \dots, X_N$ be a family of independent random vectors in $\mathbb{R}^d$ whose law is μ, and let $\overline{\mu}_N$ be the sum of Dirac $\frac{1}{N}\sum_i \delta_{X_i}$. Convergence results of the *empirical measure* $\overline{\mu}_N$ to μ are known as *uniform law of large numbers*. Using standard arguments based on Hoeffding's inequality, covering numbers of spaces of Lipschitz functions and the union bound, it can be shown that if μ supported in $K \subseteq \mathbb{R}^d$, then the following estimate on the bounded-Lipschitz distance between the empirical and the real measure holds:

PROPOSITION II.19. *Let μ be measure on $\mathbb{R}^d$ whose support is contained in a compact set K. Then μ_N converges to μ with high probability:*

$$\mathbb{P}\left[d_{bL}(\mu_N, \mu) \geqslant \varepsilon\right] \leqslant 2\exp\left(\ln(16/\varepsilon)\mathcal{N}(K, \varepsilon/16) - N\varepsilon^2/2\right)$$

In order to prove this proposition, we make use of the following inequality known as Hoeffding's inequality:

THEOREM II.20. *If (Y_i) is a sequence of independent $[0,1]$-valued random variables whose common law ν has a mean $m \in \mathbb{R}$, and $\overline{Y}_N = (1/N)\sum_{i\leqslant N} Y_i$ then*

$$\mathbb{P}(\left|\overline{Y}_N - m\right| \geqslant \varepsilon) \leqslant 2\exp(-2N\varepsilon^2)$$

Proof of Proposition II.19. Let us consider a family (X_i) of independent random variables distributed according to the law μ. Then, for any 1-Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$ with $\|f\|_\infty \leqslant 1$, one can apply Hoeffding's inequality to the family of random variables $Y_i = f(X_i)$:

$$\mathbb{P}\left[\left|\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \int f d\mu\right| \geqslant \varepsilon\right] \leqslant 2\exp(-2N\varepsilon^2)$$

We now let $\mathrm{BL}^1(K)$ be the set of Lipschitz functions f on K whose Lipschitz constant $\mathrm{Lip}\, f$ is at most 1 and $\|f\|_\infty \leqslant 1$. We let $\mathcal{N}(\mathrm{BL}^1(C), \|.\|_\infty, \varepsilon)$ be covering number of $\mathrm{BL}^1(C)$ with respect to the norm of uniform convergence.

Lemma II.21 below gives a bound for this number. It follows from the definition of the bounded-Lipschitz distance (see Def. II.2) and from the union bound that

$$\mathbb{P}\left[d_{bL}\left(p_{C\#}\mu_N, p_{C\#}\mu\right) \geqslant \varepsilon\right] \leqslant 2\mathcal{N}(BL^1(C), \|.\|_\infty, \varepsilon/4)\exp(-N\varepsilon^2/2)$$

□

LEMMA II.21. *For any compact metric space* K*,*

$$\mathcal{N}(BL^1(K), \|.\|_\infty, \varepsilon) \leqslant \left(\frac{4}{\varepsilon}\right)^{\mathcal{N}(K,\varepsilon/4)}$$

Proof. Let $X = \{x_i\}$ be an $\varepsilon/4$-dense family of points of K with $\#X = \mathcal{N}(K, \varepsilon/4)$. It is easily seen that for every 1-Lipschitz functions f, g on K, $\|f - g\|_\infty \leqslant \|(f-g)\|_{\infty,X} + \varepsilon/2$. Then, one concludes using that $\mathcal{N}(BL^1(X), \|.\|_\infty, \varepsilon/2) \leqslant (4/\varepsilon)^{\#X}$. □

In particular, if μ is supported on a point cloud C, with $\#C = n$, then $\mathcal{N}(C, \varepsilon/16) \leqslant n$. This shows that computing an ε-approximation of the measure μ with high probability (*eg.* 99%) can always be done with $N = O(n\ln(1/\varepsilon)/\varepsilon^2)$. However, if C is sampled near a k-dimensional object, then for ε in an appropriate range we have $\mathcal{N}(C, \varepsilon/16) \leqslant \text{const}(C)\varepsilon^{-k}$, in which case N is of the order of $-\ln(\varepsilon)\varepsilon^{k+2}$.

Application to boundary measures. Let $C = \{p_1, \ldots, p_n\}$ be a point cloud. Applying the ideas of the previous paragraph to the probability measure $\beta_{C,C^r} = \frac{\mu_{C,C^r}}{\mathcal{H}^d(C^r)}$, we get a Monte-Carlo approximation (Algorithm 1).

To simulate the uniform measure on C^r in step **I** of the algorithm one cannot simply generate points in a bounding box of C^r, keeping only those that are actually in C^r since the probability of finding a point in C^r decreases exponentially with the ambient dimension.

Luckily there is a simple way to generate points according to this law which relies on picking a random point x_i in the cloud C and then a point X in $B(x_i, r)$ — taking into account the overlap of the balls $B(x, r)$ where $x \in C$ (Algorithm 2).

Instead of completely rejecting a point if it lies in k balls with probability $1/k$, one can instead modify Algorithm 1 to attribute a weight $1/k$ to the Dirac mass added at this point. Step **III.** requires an estimate of $\mathcal{H}^d(C^r)$. Using the same empirical measure convergence argument, one can prove that if TL is the total number of times the loop of Algorithm 2 was run, and TN is the total number of points generated, then $TN/TL \times n\mathcal{H}^d(B(0,r))$ converges to $\mathcal{H}^d(C^r)$.

II.3.2 — Approximation of curvature measures

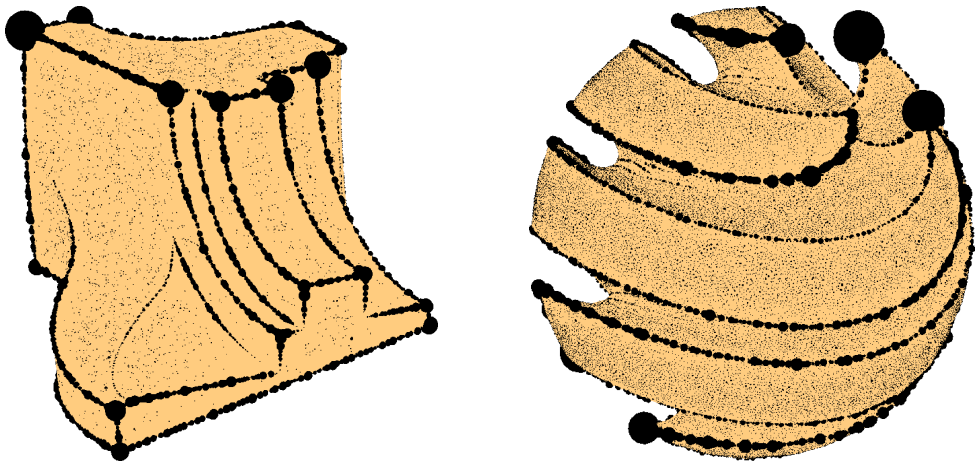
In the case the underlying compact set has positive reach, we can also try to estimate its curvature measures from a point cloud sampling. The straight-

Algorithm 1 Monte-Carlo algorithm to approximate μ_{C,C^r}

Input: a point cloud C, a scalar r, a number N
Output: an approximation of μ_{C,C^r} in the form $\frac{1}{N}\sum n(p_i)\delta_{p_i}$
while $k \leqslant N$ **do**
  [**I.**] Choose a random point X with probability distribution $\frac{1}{\mathcal{H}^d(C^r)}\,\mathcal{H}^d\big|_{C^r}$
  [**II.**] Finds its closest point p_i in the cloud C, add 1 to $n(p_i)$
end while
[**III.**] Multiply each $n(p_i)$ by $\mathcal{H}^d(C^r)$.

Algorithm 2 Simulating the uniform measure in C^r

Input: a point cloud $C = \{p_i\}$, a scalar r
Output: a random point in C^r whose law is $\mathcal{H}^d\big|_{K^r}$
repeat
  Pick a random point p_i in the point cloud C
  Pick a random point X in the ball $B(p_i, r)$
  Count the number k of points $p_j \in C$ at distance at most r from X
  Pick a random integer d between 1 and k
until $d = 1$
return X.



(a) Fandisk $R = 0.05\,\mathrm{diam}\,C$, $r = 0.02\,\mathrm{diam}\,C$

(b) Sharp sphere, $R = 0.1\,\mathrm{diam}\,C$, $r = 0.03\,\mathrm{diam}\,C$

Figure II.7 – Convolved boundary measures of 100k point clouds sampled from the Fandisk and Sharp sphere models. 50k points are drawn uniformy, while the 50k remaining points are (over) sampled on randomly chosen edge path on the two models.

forward way is to estimate the boundary measures for several choice of parameters, and then perform polynomial fitting. The difficulty with this approach is that one needs to compute the real boundary measures μ_{C,C^r} and not merely the normalized versions β_{C,C^r}.

There is, however, a modification of the previous Monte-Carlo algorithm that allows to compute the curvature measures in a single pass. For each point $p \in C$, one keeps track of the points X that are selected in Algorithm 1, and lie in the Voronoi cell of p; denote them by $x_1(p), \dots, x_{k(p)}(p)$. Denote by $\nu_p(s)$ ($s \in [0, r]$) the number of points at distance at most s of p among the $x_i(p)$:

$$\nu_p : s \in [0, r] \mapsto \frac{\#\{i \leqslant k(p); \|x_i(p) - p\| \leqslant s\}}{N}$$

$\nu_p(s)$ approximates (up to a constant factor) the volume of the intersection $\mathrm{Vor}_C(p) \cap B(p, s)$. Performing polynomial fitting on this function allows to determine (up to a constant factor again) the weights of the curvature measures at p.

II.3.3 — Exploiting boundary measures

Visualizing boundary measures. There is no obvious convenient way to visualize a sum of Dirac masses. The most straightforward idea is to display a sphere at each point whose radius is related to the weight at that point (eg. the d-th root of this weight is the point cloud is embedded in $\mathbb{R}^d$). But since the radii of these sphere do not add up, a highly concentrated area can be "invisible" if the mass is split among a lot of points.

Convolved measure. A consequence of Kantorovich-Rubinstein theorem is that if two probability measures μ and ν are close in the Wasserstein sense, then given 1-Lipschitz function $\chi : \mathbb{R}^d \to \mathbb{R}$, the convolution of μ and ν by χ are uniformly close. Precisely, $\|\mu * \chi - \nu * \chi\|_\infty \leqslant \mathrm{Lip}(\chi)\, W_1(\mu, \nu)$.

Thus, a way to display significant (*ie.* Hausdorff-stable) information from the boundary measure of a point cloud C is to chose a convolution kernel χ, and consider the function $\mu_{C,C^r} * \chi$. In the Fig. II.7, we display the convolved boundary measures of point clouds extracted from two models of the AIM@SHAPE repository. The offset radius R is given as a fraction of the diameter of the point cloud, and the convolution function is $\chi_r : x \mapsto \max(1 - \|x\|/r, 0)$. We display the function $\mu_{C,C^r} * \chi_r$ by putting a black sphere of radius $\mu_{C,C^r} * \chi_r(p)$ for each point p in the point cloud.

Feature extraction and support estimation. Extracting the set of sharp features of a compact sets known through a finite point cloud sampling is of interest in many geometry processing applications. Fig. II.7 (and intuition) suggests that the sharp corners carry more mass than the points on the sharp edges, which again carry more mass than the smooth points. We will give more precise statements on this topic in the Chapter III.

As for now, we can take it as a definition: a feature point is a point that carries more mass than a smooth point. Or, more precisely (for the same

reasons as in the previous paragraph), where the values of the function $\mu_{K,K^R} * \chi_r$ are higher than some threshold T. If the point x is on a flat area, the set of points that projects onto a r-ball around x is a cylinder of radius r and height $2R$. Hence one can choose T to be a constant T_0 times $2R \times \pi r^2$ if the point clouds is embedded in $\mathbb{R}^3$ (the constant is to allow some amount of curvature). An example of the result of such a thresholding is given in Fig. II.8

In the next chapter, we introduce an anisotropic version of the boundary measures, that allow to classify feature point based on the estimated feature angle, and not on such an arbitrary threshold.

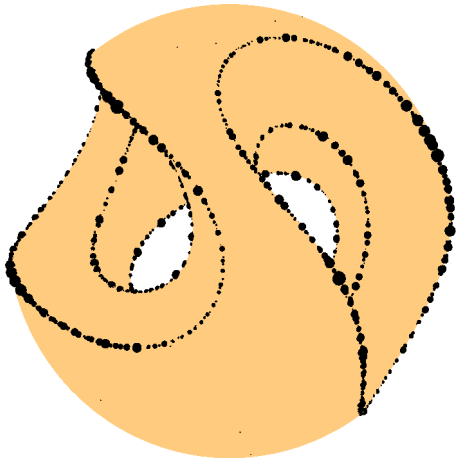


Figure II.8 – Feature points extracted from a point cloud sampling of a CSG model by thresholding low values of the convolved boundary measure ($T_0 = 3$).

II.A STEINER FORMULA AND REACH ESTIMATION

II.A.1 — Existence of Steiner tuber formula revisited

In this section, we give a proof of the existence and uniqueness of curvature measures for a compact set with positive reach, which is shorter than Federer's original proof [Fed59] based on approximation arguments. We make use of the notations of Definition I.7; τ denotes the normal distance to the medial axis of the considered compact set.

We begin with a few classical definitions from geometric measure theory (see [Mat95] or [Fed69] for more details). Actually, we are using approximate differentiability only through propositions and lemmas of those books: the definition below can be safely skipped and considered as a "black box".

DEFINITION II.3. If $x \in \partial K^r$ is such that $\tau(x) \geqslant \tau_0$, then for any $0 < t < \tau_0$, the distance function to K is differentiable at $\Psi^t(x)$, which means that the hypersurface ∂K^{r+t} has a tangent hyperplane $T_{\Psi^t(x)}\partial K^{r+t}$ at $\Psi^t(x)$.

For $B \subseteq \partial K^r$, with $\inf_{x\in B} \tau(x) \geqslant \tau_0$ and $t < \tau$, we say that the map $\Psi^t : B \to K^r$ is approximately differentiable at $x \in B$ if there exists a linear

map $\varphi : \mathbb{R}^d \to T_{\Psi^t(x)}\partial K^{r+t}$ that vanishes on the normal space $\{T_x\partial K^r\}^\perp$, and such that for any positive ε the set

$$\{y \in B\,;\, \left\|\Psi^t(y) - \Psi^t(x) - \varphi(y-x)\right\| \geqslant \varepsilon \|x-y\|\}$$

has zero $(d-1)$-Hausdorff density at x. The restriction of φ to the tangent space $T_x\partial K^r$ is called the approximate differential of Ψ^t at x, and denoted by $\operatorname{ap} d_x\Psi^t$. The determinant of $\operatorname{ap} d_x\Psi^t$ is called the approximate Jacobian of Ψ^t at x.

LEMMA II.22. *Let B be a Borel set in ∂K^r such that $\inf_B \tau_K > \tau_0$. Then Ψ^t is approximately differentiable at $\mathcal{H}^{d-1}$–almost every point of B in the sense of Definition II.3.*

Proof. The set B is $(d-1)$-rectifiable, as a subset of the rectifiable set ∂K^r. The function $\Psi^t : B \to \mathbb{R}^d$ is Lipschitz on B, by Lemma I.22. Hence, Corollary 3.2.20 in [Fed69] (which is analogous to the Rademacher theorem) proves that Ψ^t is approximately differentiable at $\mathcal{H}^{d-1}$–almost every point of B. □

PROPOSITION II.23. *Let K be a compact set, and B a Borel subset of ∂K^{t_0} such that for any point x in B, $\tau_K(x) \geqslant \tau_0$. Then the map*

$$t \mapsto \mathcal{H}^{d-1}(\{x \in \partial K^{t_0+t}\,;\, p_{K^{t_0}}(x) \in B\})$$

is a polynomial in t on $[0, \tau_0]$ whose degree is at most d.

In order to prove this proposition, we need the following adaptation of Lemma 5.1 of [Fed59]:

LEMMA II.24. *Let $r > 0$ be such that ∂K^r contains $\mathcal{H}^{d-1}$-almost no point of the medial axis, let B be a Borel set contained in ∂K^r, such that $\inf_B \tau_K \geqslant \tau_0 > 0$, and let $0 < t < \tau_0$. Then, for $\mathcal{H}^{d-1}$–almost any point x of B, Ψ^t is differentiable at x and $\ker \operatorname{ap} d_x\Psi^t = \{0\}$. In particular $J_x\Psi^t \neq 0$.*

Proof. Lemmas I.22 and II.22 show that Ψ^t is differentiable at $\mathcal{H}^{d-1}$–almost every $x \in B$. Let us prove the second assertion by contradiction. Suppose there exist a unit tangent vector $u \in T_xB = T_x\partial K^r$ such that $\operatorname{ap} d_x\Psi^t(u) = 0$. Let (x_n) be a sequence of points of B such that $(x_n - x)/\|x_n - x\|$ converges to u. For any positive ε, there exist N such that for $n \geqslant N$,

$$\frac{\left\|\Psi^t(x_n) - \Psi^t(x)\right\|}{\|x_n - x\|} \leqslant \varepsilon$$

In order to obtain a contradiction, we will prove that the the reciprocal map $(\Psi^t)^{-1} : \Psi^t(B) \to B$ is Lipschitz. Indeed, since for any $y \in B$, $\tau_K(\Psi^t(y))$ is greater than $\tau_0 - t$, we can apply the Lemma I.22 to $\Psi^t(B) : \Psi^{\frac{\tau_0-t}{2}}$ is Lipschitz on $\Psi^t(B)$. Hence, remarking that for every y in B,

$$(\Psi^t)^{-1}(y) = y - 2t\frac{\Psi^{\frac{\tau_0-t}{2}}(y) - y}{\tau_0 - t}$$

we see that $(\Psi^t)^{-1}$ is Lipschitz. This yields the desired contradiction, thus proving our statement. □

Proof of Proposition II.23. Applying the area formula of Corollary 3.2.20 in [Fed69] and the one-to-one property of the mapping Ψ^t on B (for $t < \tau_0$), we get

$$\mathcal{H}^{d-1}(\Psi^t(B)) = \int_B \left|\operatorname{ap} J_x\Psi^t\right| d\mathcal{H}^{d-1}(x)$$

For $t = 0$, Ψ^t coincides with the identity and $J_x\Psi^0 = 1$. The map $\operatorname{ap} J_x\Psi^t = \operatorname{ap} J_x(\mathrm{id} + t \operatorname{ap} \nabla_x d_K)$ is a polynomial in the variable t; hence $t \mapsto \left|J_x\Psi^t\right|$ is continuous and never vanishes (by Lemma II.24). Hence, the sign of $\operatorname{ap} J_x\Psi^t$ must be constant and positive. Thus we get

$$\mathcal{H}^{d-1}(\Psi^t(B)) = \int_B \operatorname{ap} J_x\Psi^t d\mathcal{H}^{d-1}(x)$$

One concludes as in Proposition II.1. □

COROLLARY II.25. *If the reach of K is greater than R, then K admits a local Steiner formula on $[0, R]$.*

Proof. Let B be a Borel set, and $\varepsilon > 0$. Applying the previous theorem to $B^\varepsilon = p_K^{-1}(B) \cap \partial K^\varepsilon$, shows that

$$\begin{aligned}\mathcal{H}^d(\{x \in K^{\varepsilon+t}\,;\, p_{K^\varepsilon}(x) \in B^\varepsilon) &= \mathcal{H}^d(\{x \in K^{\varepsilon+t} \setminus K^\varepsilon\,;\, p_K(x) \in B)\\ &= \mu_{K,K^{\varepsilon+t}}(B) - \mathcal{H}^d(K^\varepsilon)\end{aligned}$$

is polynomial in t on $[0, R - \varepsilon]$, of degree at most d. Since it is true for any positive ε, this proves that $\mu_{K,t}(B)$ is also a degree d polynomial on $[0, R]$. □

II.A.2 — Does a local Steiner formula imply positive reach?

In a recent paper [HHL04], Heveling, Hug and Last asked whether polynomial parallel volume should imply convexity. They answered positively in the case of the plane but exhibited a counter-example in higher dimension. Using the notion of support measure [HLW04], they were able to prove that if a compact set admits a strong version of a local Steiner formula on $[0, r]$ then its reach should be at least r.

DEFINITION II.4. (i) A compact subset K admits a *strong local Steiner formula on* $[0, R]$ if for any measurable, compactly supported and bounded function $f : \mathbb{R}^d \times S^{d-1} \to \mathbb{R}$ there exists constants $c_0(f), \dots, c_d(f)$ such that

$$\forall r \in [0, R],\ \int_{K^r \setminus K} f\left(p_K(x), \frac{x - p_K(x)}{d(x, K)}\right) = \sum_{i=0}^{d} c_i(f) r^i \qquad \text{(II.15)}$$

(ii) Following the definition of §II.1.2, we say that K admits a *local Steiner formula on* $[0, R]$ if for any Borel set $B \subseteq \mathbb{R}^d$, there exists constants

$c_0(B), \dots, c_d(B)$ such that

$$\forall r \in [0, R],\ \mathcal{H}^d(\{x \in K^r\,;\, p_K(x) \in B\}) = \sum_{i=0}^{d} c_i(B) r^i \qquad \text{(II.16)}$$

It is equivalent to require that there exists d signed Borel measures $\Phi_0, \dots, \Phi_d$ such that $\mu_{K,K^r} = \sum_{i=1}^{d} \Phi_i r^i$ for $r \in [0, R]$.

THEOREM II.26 (*cf.* [HHL04]). *If* K *admits a strong local Steiner formula on* $[0, R]$*, then the the reach of* K *is at least* R.

A strong local Steiner formula for K is, as its name suggests, stronger than the usual local Steiner formula for K. In fact it implies that all the offsets K^r of K (with $r < R$) admit a local Steiner formula:

PROPOSITION II.27. *Let* $K \subseteq \mathbb{R}^d$ *be a compact set that admits a strong local Steiner formula on* $[0, R]$ *(cf. eq. II.15). Then, for any* $r < R$*, the offset* K^r *admit a (usual) local Steiner formula on* $[0, R - r]$.

Proof. Indeed, if B is a Borel subset of K^r, let $f_B : \mathbb{R}^d \times \mathbb{S}^{d-1} \to \mathbb{R}$ be defined by $f_B(p, u) = 1$ iff there is an x in B such that $(p_K(x), \nabla_x d_K) = (p, u)$. Then,

$$\int_{K^{r+t} \setminus K} f(p_K(x), \nabla_x d_K) dx = \text{cst} + \mu_{K^r, K^{r+t}}(B)$$

By hypothesis, the first side of the equation is a degree d polynomial in $t \in [0, R - r]$. This shows K^r admits a local Steiner formula on $[0, R - r]$. □

The following statement is barely stronger than Hevelin, Hug and Last's result; however, the proof is completely elementary and doesn't rely on on the notion of support measure that was used in their original proof.

THEOREM II.28. *A compact set* $K \subseteq \mathbb{R}^d$ *for which one of the following property is true has a reach greater than* R *:*

(i) for every $r < R$*, the offset* K^r *admits a local Steiner formula on* $[0, R - r]$*;*

(ii) K *has a strong local Steiner formula on* $[0, R]$.

Before proving the theorem, we need to prove the two following lemmas. The normal distance of $x \in \mathbb{R}^d \setminus K$ to the medial axis of K as introduced in Definition I.7, will be denoted by $\tau_K(x)$.

LEMMA II.29. *If a Borel set* $B \subseteq \partial K^r$ *is such that* $\forall x \in B, \tau_K(x) \leqslant \tau_{\max}$*, then* $t \mapsto \mu_{K^r, K^{r+t}}(B)$ *is constant for* $t \geqslant \tau_{\max}$.

Proof. Let y be a point of $\mathbb{R}^d$ that is not in the medial axis of K, and $x = p_K(y)$. Then, by Lemma I.18, $\|x - y\| \leqslant \tau_K(x)$. The statement follows. □

LEMMA II.30. *Let* K *be a compact set,* $r \geqslant 0$ *such that for almost any* $x \in K^r$*,* $\tau_K(x) \geqslant r - d_K(x)$*. Then* $\text{reach}(K) \geqslant r$.

Proof. Let $x \in K^r \setminus K$, p a projection of x on K and u the unit vector from p to x. We want to show that p is the only projection of x on K, *ie.* the ball $B(x, d_K(x))$ intersects K only at x.

By hypothesis, there exist a sequence of points $x_i \in K^r \setminus K$ converging to x, such that $\tau_K(x_i) \geqslant r - d_K(x_i)$ and $p_i = p_K(x_i)$ converges to p. Now by hypothesis on the x_i, for any $r' < r$, the interior of the ball $B(p_i + r'\nabla_{x_i} d_K, r')$ does not contain any point of K. Hence, the interior of the ball $B(p + r'u, r')$ does not contain any point of K either. For $r' > d_K(x)$, this also shows that the (closed) ball $B(x, d_K(x))$ can intersect K only at p. □

Proof of Theorem II.28. By Proposition II.27, we can suppose that condition (i) is realized. By Lemma II.30 we only have to prove that for almost any point x in K^R, $\tau_K(x) \geqslant r - d_K(x)$. In order to to that, it suffices to show that the set B^ε of points of K^R such that $\tau_K(x) \leqslant R - d_K(x) - \varepsilon$ is negligible for any positive ε. Since $\mathrm{Med}(K)$ has zero $\mathcal{H}^d$–volume, we can remove the points of the medial axis from B^ε.

For any $r > 0$, by the condition (i), there exists constants $c_0, \ldots, c_n$ such that

$$\mu_{K^r,t}(B^\varepsilon \cap \partial K^r) = \sum_{i=0}^{d} c_0 t^i \tag{II.17}$$

Now lemma II.29 and the assumption on B^ε tells us that this polynomial is constant between $R - r - \varepsilon$ and $R - r$. Since its value at $t = 0$ is zero, $\mu_{K^r,t}(B^\varepsilon \cap \partial K^r)$ is identically null. Since

$$\{x \in \mathbb{R}^d \setminus \mathbf{Med}(K)\,;\, p_{K^r}(x) \in B^\varepsilon \cap \partial K^r\} = B^\varepsilon \setminus K^r$$

Eq. II.17 gives us $\mathcal{H}^d(B^\varepsilon) = \sup_{r>0} \mathcal{H}^d(B^\varepsilon \setminus K^r) = 0$ thus proving the theorem. □

II.A.3 — Estimating the reach of a compact set

Estimating the reach of a compact set is an interesting theoretical question, but it would also be a very useful piece of information in applications such as surface reconstruction. There are a few heuristics for estimating it (such as the distance to the nearest pole, see §III.1.3), but they often break badly in the presence of noise.

Even the precise meaning of estimating the reach is not so clear. There are two possible and natural questions that one would want to address:

— Given a compact subset $K \subseteq \mathbb{R}^d$, and a parameter $\varepsilon > 0$. What is the maximum reach of a compact subset at Hausdorff distance less than ε of K ?

— Given K and a parameter $R > 0$, what is the minimum Hausdorff distance between K and a compact set with reach at least R ? If one denotes by $\mathrm{Reach}_R(\mathbb{R}^d)$ the compact subsets of $\mathbb{R}^d$ with reach bounded from below by R, this amounts to computing $d(K, \mathrm{Reach}_R(\mathbb{R}^d))$.

The goal is of course to be able to come up with an algorithm that computes, or approximates these two quantities given a point cloud $C \subseteq \mathbb{R}^d$. It is not so hard to come up with one-sided bounds, at least theoretically:

— If there exists a segment of length R_0 joining two points in the cloud, and whose circumsphere does not enclose any other points of C (such a segment is called a Gabriel edge), then for any compact set K close enough to C, $\mathrm{reach}(K) \leqslant \frac{1}{2}R_0 - d_H(K, C)$.
Notice that this works because the midpoint of the Gabriel edge is a critical point for the distance function. Similar bounds based on other kind of critical points can also be given.

— If C is close to a compact set K with reach at least R, then the measure-valued function $r \in [0, R] \mapsto \mu_{K,K^r}$ is close to a degree d polynomial $r[0, R] \mapsto \sum_{i=1}^{d} \varphi_i r^i$ where φ_i are finite signed Borel measures supported on K.
If one denotes by $\mathrm{Poly}_R^d(\mathbb{R}^d)$ the set of such measure-valued polynomials, then the stability result of the previous section proves that

$$\inf\{ \sup_{r\in[0,R]} d_{bL}(\mu(r), \mu_{C,C^r}) \,;\, \mu \in \mathrm{Poly}_R^d(\mathbb{R}^d)\} \leqslant C(C)d(C, \mathrm{Reach}_R(\mathbb{R}^d))^{1/2} \tag{II.18}$$

Notice that original (unmatched) goal of the previous section was to show that the reach of a compact set K is at least R iff the measure-valued function $r \mapsto \mu_{K,K^r}$ is polynomial. A natural extension of this question would be the following:

QUESTION. Is there a reverse inequality to Ineq. (II.18) ? Or, said otherwise: If the function $r \mapsto \mu_{K,K^r}$ associated to a compact subset K of $\mathbb{R}^d$ is close[1] to being polynomial of degree d on $[0, R]$, is this compact subset close to a compact set with reach bounded from below by $\Omega(R)$?

[1] *close* in a meaning to be precisely defined, *eg.* in the uniform sense, as in Eq. (II.18)

Chapter III

Voronoi-based Feature Estimation

Abstract

Many algorithms for shape analysis and shape processing rely on accurate estimates of differential information such as normals and curvature. In most settings, however, care must be taken around non-smooth areas of the shape where these quantities are not easily defined. This problem is particularly prominent with point-cloud data, which are discontinuous everywhere. In this chapter we present an efficient and robust method for extracting normal directions sharp features and curvature information of a piecewise smooth surface from a point cloud sampling. Our method is integral in nature and uses convolved covariance matrices of Voronoi cells of the point cloud which makes it provably robust in the presence of noise. Using the same analysis we provide theoretical guarantees for a modification of a previously proposed normal estimation technique. We illustrate experimentally the correctness of both principal curvature estimation and feature extraction in the presence of varying levels of noise and sampling density on a variety of models.

This whole chapter derives from an article written with Maks Ovsjanikov

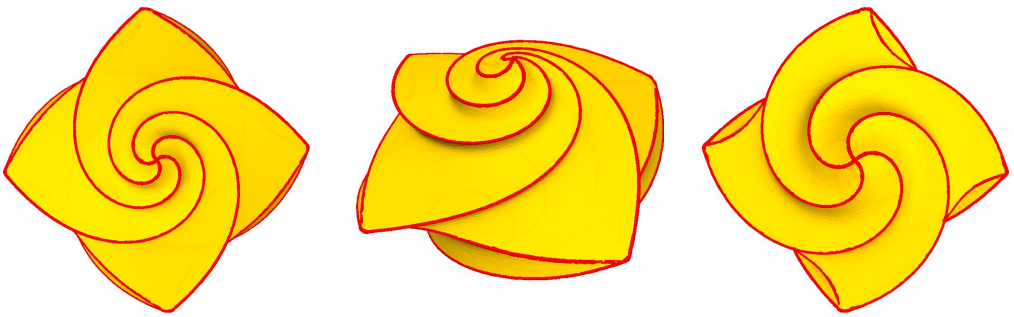


Figure III.1 – Features (in red) computed by our algorithm from a point cloud sampling of the surface in yellow.

and Leonidas Guibas, accepted for publication in the Proceedings of the 2009 SIAM/ACM Joint Conference on Geometric and Physical Modeling [MOG09].

Contents

INTRODUCTION

Estimating surface normals, principal curvatures and sharp edges from a noisy point cloud sampling, has many applications in computer graphics, geometry processing and reverse engineering. Principal curvatures are rotation-invariant local descriptors, which together with principal curvature directions have proven useful in detecting structural regularity [PMW+08], global matching [AMCO08], modeling and rendering of point-based surfaces [GTE+06], and anisotropic smoothing [LP05] to name just a few. In these applications, finding exact curvature informations (such as principal curvature directions) is not as important as to find robust local descriptors that encode second order variations of the surface. The location of sharp edges and highly curved areas of the surface is a precious piece of information in settings that include feature-aware reconstruction [JWS08], non photorealistic rendering [PKG03], and industrial metrology.

In practice, it is often interested to recover these information when the input is an unstructured collection of point coordinates, obtained by a range scanner, before any reconstruction step (besides registration). These point clouds can be noisy, and can exhibit strong sampling bias. The ability to reliably estimate surface normals, principal curvatures, and curvature directions as well as sharp features directly on such point clouds can be used in both geometry processing algorithms and in surface reconstruction to improve the quality of the resulting mesh.

Devising robust local descriptors, which can handle both non-uniform noise, sampling bias and sharp edges is a challenging task. This is mainly

because we are trying to estimate differential quantities which by nature are very sensitive to local perturbations. The lack of a natural parametrization of point cloud data introduces another challenge by making it difficult to estimate angles and areas on the surface. Finally, devising a method with theoretical guarantees on the approximation quality is not easy in the absence on a unified framework that would incorporate both point clouds and piecewise-smooth surfaces.

In this chapter, we address some of these challenges by presenting a method for robustly estimating principal curvatures and curvature directions, as well as sharp edges of the underlying surface from point cloud data. We provide theoretical guarantees on the robustness of the results, by deriving a bound on the quality of the estimation based on the Hausdorff distance between the point cloud and the underlying surface. We also address a certain class of outliers.

Prior work on curvature and feature estimation

Estimating curvature information of the underlying surface from a discrete approximation has been studied extensively over the past several decades (see e.g. [MD02] for a survey dating 2002 and [MSR07] for an extensive comparison of methods to estimate Gaussian and mean curvatures). Nearly all existing methods for reliably estimating principal curvatures and curvature directions, however, rely on meshes. These methods are difficult to extend to the point cloud setting, because the mesh defines a discrete parametrization of the surface making angle and area computations easier.

Curvature estimation using point with normals. Only recently several methods have been proposed for computing curvature information directly on point clouds. All of the following algorithms start by estimating normals to the surface, or assume that they are given. Any error made in this estimation is aggravated in the computation of principal curvatures.

Yang and Qian [YQ07] derive analytic expressions for computing principal curvatures based on the implicit definition of the *moving least squares* (MLS) surface by Amenta and Kil [AK04]. Another class of methods work by sampling curves on the surface passing by a given point p_0, and use their normal curvature to estimate the normal curvatures of the surface (by Meusnier theorem). Using several of these curves, they are then able to estimate the principal curvature directions: [Tan05, LP05, HM02, TT05]. A third kind of mehods rely on computing covariance matrices. Given a point p_0 on the surface, one projects all the normals of neighbouring points on the estimated tangent plane to p_0. The covariance of this set of vectors is used to estimate the shape operator [BC94].

Computer graphics approaches to feature estimation. Although extracting sharp edges and corners is closely related to curvature estimation, research in these areas has been relatively independent. Fleishman et al.

[FCOS05] detect sharp edges and corners by segmenting neighborhoods of points into regions corresponding to the same part of the surface. They achieve robustness by using a forward search technique which finds reference planes, corresponding to each segment. This work is extended by Daniels et al. [IOHS08] to extract feature-curves, by locally projecting onto feature points, and growing smooth polylines through the projected cloud. Lipman and colleagues [LCOL07] extract sharp edges within the MLS projection framework by defining a Singularity Indicator Field (SIF) based on the error of the MLS surface approximation. Jenke et al. [JWS08] detect sharp features by robustly fitting local surface patches and computing intersections of nearby patches with dissimilar normals. In a similar spirit, Oztireli et al. [OGG09] define a feature-preserving MLS projection operator by noting that angles between normal vectors, rather than point coordinates can be used to discard points from unrelated parts of a piecewise-smooth surface.

Provably correct curvature and feature estimation. Very few of the curvature and feature estimation methods come with theoretical guarantees on the convergence to the real principal curvature directions as the point cloud converges (in the Hausdorff or distribution sense) to the surface. Even fewer come with quantitative convergence results. The most notable results include:

Local fitting of polynomials. The most commonly used methods for computing curvatures on point clouds in practice, rely on least-square polynomial fitting (see *eg.* Cazals and Pouget [CP05] and references therein). Let p_0 be a point of the surface, $p_1, \dots, p_k$ some of its neighbors, and $\mathbf{n}_0$ a (not necessarily good) estimation of the normal at p_0. Denote by $p'_1, \dots, p'_k$ the projections of the points on the affine plane $P = p_0 + \{\mathbf{n}_0\}^\perp$, and $h_1, \dots, h_k$ the heights $h_i = \langle p_i - p'_i | \mathbf{n}_0 \rangle$. The height function of the surface is then estimated by the two-variable polynomial $h : P \to \mathbb{R}$ that best fits the couples (p'_i, h_i) in the least-square sense, and the curvature tensor of the unknown surface is estimated by the curvature tensor of the surface $(p', h(p'))$ at p_0.

Least-square fitting is very fast in practive, and its convergence properties are well understood. Given a function $f : D \subseteq \mathbb{R}^d \to \mathbb{R}$, let us denote by P_0 the solution of the polynomial least-square fitting for a domain D in $\mathbb{R}^d$, and P_X the solution for a finite subset $X \subseteq D$:

$$P_0 = \arg\min \left\{ \|P - f\|_{L^2(D)} \, ; P \in \mathbb{R}^d[X], \deg(P) \leqslant k \right\}$$

$$P_X = \arg\min \left\{ \sum_{x \in X} \|P(x) - f(x)\|^2 \, ; P \in \mathbb{R}^d[X], \deg(P) \leqslant k \right\}$$

Given a sequence of point clouds $X_n \subseteq D$, the polynomial P_{X_n} converges to P_0 provided that uniform measure on the point cloud X_n converges to the uniform measure on the domain Ω in the L^2-sense.

This means that the requirement for convergence is not geometric (Hausdorff) but measure-theoretic. In particular, curvature estimation methods

based on least-square fitting are sensitive to sampling. This can be a problem for some point clouds such as laser scans which often exhibit oversampled clusters of points along horizontal lines.

Normal cycle and second fundamental measure. Cohen-Steiner and Morvan [CSM03] use the normal cycle from geometric measure in order to define a notion of second fundamental measure of a *geometric set*. They prove that if K approximates S for some notion of closeness stronger than Hausdorff, then the normal cycles (and hence the second fundamental measure) of K and S are close. This theorem applies to a Delaunay mesh reconstructed from an ε-sampling of S, for instance, allowing to compute very reliable curvature informations on this kind of meshes.

Second fundamental measures can be applied in a meshless setting by considering the *offsets* of the compact set [CCSLT09]. If a point cloud C is close enough to an unknown compact set K with positive μ-reach, the second fundamental measures of C^r and K^r are close, when the Hausdorff distance $d_H(C, K)$ is close enough. To the best of our knowledge, there is no known way to pull-back these curvature measures to the original point cloud.

Integral methods (Connolly-like). These methods estimate the principal curvature directions and sharp features of a surface S bounding a domain D, by considering the volume or covariance matrices of intersection of small balls $B(x, r)$ ($x \in S$) with the domain D. A nice feature of these methods is that the estimation error made by replacing a domain D by another domain D' only depends only on the volume of the symmetric difference $D \Delta D'$, and not of any higher order approximation properties between $\partial D'$ and ∂D.

The drawback, however, is the need for an approximation of not only the surface S (or of a point-cloud approximation of it), but also of the interior domain D. For the original applications in computational structural biology [Con86, CCL03], this requirement is perfectly fine, since the domain is simply a union of balls. But for applications in geometry processing, this means that D has to be meshed, and the method doesn't apply to point clouds. Since these integral methods are close in spirit to our approach, we give more detail about them in §III.1.2.

Contributions

The main contribution presented in this chapter is a framework for estimating curvature and feature information of the underlying surface from a *point cloud* C, based on integral quantities. We rely on covariance matrices of Voronoi cells (as in the method of Alliez and colleagues [ACSTD07]); but instead of intersecting them with a large bounding box of the point cloud, we intersect them with an offset C^R. Intersecting with an offset gives us more *local* information about the variation in shape and size of the Voronoi cell, which is crucial for curvature estimation. We present two algorithms for computing covariance matrices of this intersection: a Monte Carlo method, and a method based on tessellating the intersection with tetrahedra.

The theoretical results are twofold: first, for any compact set K, we define

its Voronoi covariance measure (VCM), through the *projection function* on K. This allows us to study the discrete case of point clouds and the underlying continuous piecewise-smooth surfaces in a single framework. With this notion at hand, we prove that if K is a piecewise-smooth submanifold of $\mathbb{R}^d$, then the eigenvalues and eigenvectors of its convolved VCM provide information on the normal directions, principal curvatures and directions, directions of sharp features, and dihedral angles between its smooth parts. In the second part, we prove that if the point cloud C is a good Hausdorff approximation of K, then convolved versions of the VCM of C and K are uniformly close.

Outline. In §III.1, we start with a review some background material on normal, curvature and feature estimation (§§III.1.2,III.1.3). We then give the formal definition of the Voronoi-covariance measure (VCM). In §III.2, we study the theoretical properties of the VCM for feature (§III.2.2) estimation, and discuss the possible relation with curvature estimation (§III.2.1), before giving a stability result for VCM (§III.2.3). We finish by an experimental part (§III.3): after giving a possible algorithmic implementation in §III.3.1, we study the results obtained in practice under varying sampling, noise levels and for piecewise smooth surfaces with small dihedral angles between smooth parts.

III.1 VORONOI COVARIANCE MEASURE

Before introducing the main character in this chapter, the *Voronoi covariance measure* of a compact set in $\mathbb{R}^d$, we will review some geometric estimation approaches, which share some features with this work:

— integral-based approaches for estimating curvature and sharp features on a mesh, using the volume or covariance matrices of the intersection of this mesh with small balls.

— normal estimation techniques, from principal component analysis normal estimation to integral Voronoi-based normal estimation.

III.1.1 — Covariance matrices, stability of their eigenspaces

Throughought this chapter, we will make use of the following definition and results on covariance matrices of a domain $E \subseteq \mathbb{R}^d$.

DEFINITION III.1 (Covariance matrix). If $E \subseteq \mathbb{R}^d$ has finite volume, its covariance matrix is a 2-tensor whose eigenvectors capture the principal axes of E with respect to a base point p:

$$\mathrm{cov}(E, p) = \int_E (x - p) \otimes (x - p) dx$$

where $v \otimes w$ denotes the $n \times n$ matrix defined by $[v \otimes w]_{i,j} = w_i v_j$.

A nice feature of the eigenanalysis of covariance matrices (and more generally symmetric matrices) is that eigenvalues and eigenspaces are stable under

perturbation, in a sense to be made precise. As a very simple application, consider two domains E and E' that are close in the sense that the symmetric difference $E' \Delta E$ has small volume. Then, it is easy to see that the covariance matrices $\mathrm{cov}(E, p)$ and $\mathrm{cov}(E', p)$ ($p \in \mathbb{R}^d$) are also close:

$$\left\|\mathrm{cov}(E, p) - \mathrm{cov}(E', p)\right\|_{\mathrm{op}} \leqslant \mathcal{H}^d(E \Delta E') \sup_{x \in E \cup E'} \|x - p\|^2$$

The stability result mentioned above can be use to quantify the difference between the eigenvalues and eigenspaces of the two covariance matrices. These stability results also explain why the standard principal-component analysis techniques work.

All the following results are standard stability results from matrix perturbation theory. A good reference for this theory can be found in [SS90]. The set of eigenvalues of M is denoted by $\mathrm{Spec}(M) \subseteq \mathbb{R}$. An eigenpair of M is a pair (λ, v) made of an eigenvalue of M and a corresponding unit eigenvector v, *ie.* $Mv = \lambda v$ and $\|v\| = 1$. The norm we use is the operator norm:

$$\|M\|_{\mathrm{op}} = \sup_{v \in \mathbb{S}^{d-1}} \|Mv\| .$$

A first very classical stability result for eigenvalues is due to Weyl:

THEOREM III.1 (Weyl). *Let M and M' be two symmetric matrices and $\lambda_1 \geqslant \ldots \geqslant \lambda_n$ and $\lambda'_1 \geqslant \ldots \geqslant \lambda'_n$ be their ordered eigenvalues. Let m^- and m^+ denote the smallest and largest eigenvalues of $M' - M$. Then,*

$$\forall i,\ \lambda_i + m^- \leqslant \lambda'_i \leqslant \lambda_i + m^+$$

As a weaker consequence, one obtains that the Hausdorff distance between the spectra of M and M' is at most $\|M - M'\|_{\mathrm{op}}$.

Recall that the (global) eigengap of a diagonalizable matrix is the minimum distance between two eigenvalues — in particular, it vanishes if some eigenvalue of the matrix has multiplicity greater than one. The local eigengap at $\lambda \in \mathrm{Spec}(M)$, denoted by $\delta_\lambda(M)$ is the minimum distance between λ and any *distinct* eigenvalue of M. The following theorem is a simplified version of Davis-Kahan $\sin(\Theta)$ theorem.

THEOREM III.2 (Davis-Kahan). *Let M and M' be two symmetric matrices, λ an eigenvalue of M and $\delta = \delta_\lambda(M)$ be the eigengap of M at λ. Then for every eigenpair (λ, v) of M, there exists an eigenpair (λ', v') of M' such that*

$$\left|\lambda - \lambda'\right| \leqslant \sqrt{2} \left\|M - M'\right\|_{\mathrm{op}} \text{ and } \left\|v - v'\right\| \leqslant \frac{2 \left\|M - M'\right\|_{\mathrm{op}}}{\delta}$$

provided that $\|M - M'\|_{\mathrm{op}} < \delta\sqrt{2}$.

III.1.2 — Integral-based curvature and feature estimation

Connolly function and curvature. A prominent line of work to estimate some curvature properties of a surface S which bounds a domain D make use of the intersection a Euclidean ball $B(x,r)$ (or sphere $\mathcal{S}(x,r)$) with D, where $x \in S$. The function $r \mapsto \mathcal{H}^2(\mathcal{S}(x,r) \cap D)$ was introduces by Connolly in [Con86] in the context of molecular shape analysis. Its precise relation with the *mean curvature* $H(x)$ of the surface S at x is as follows (*cf.* for instance [CCL03] or [HT03]),

$$\begin{aligned} \mathcal{H}^3(B(x,r) \cap D) &= \frac{2\pi}{3} r^3 - \frac{\pi H(x)}{4} r^4 + O(r^5) \\ \mathcal{H}^2(\mathcal{S}(x,r) \cap D) &= 2\pi r^2 - \pi H(x) r^3 + O(r^4) \end{aligned} \tag{III.1}$$

Application to feature detection. In [CRT04], the authors use similar integral techniques for detecting sharp edges of a piecewise smooth manifold. They consider the centroid $M_r^0(x)$ of the intersection $B(x,r) \cap D$, *ie.*

$$M_r^0(x) = \frac{1}{\mathcal{H}^d(B(x,r) \cap D)} \int_{B(x,r) \cap D} y \, dy$$

They prove that the distance from x to the $M_r^0(x)$ is quadratic (*ie.* $O(r^2)$) if x is a smooth point of the surface:

PROPOSITION III.3. *If $S = \partial D$ is locally smooth around a point $x \in S$, then*

$$M_r^0(x) = x + C(d) r^2 H(x) \mathbf{n}_S(x) + o(r^2)$$

where $\mathbf{n}_S(x)$ is the (outward-pointing) normal to S at x, $H(x)$ is the mean curvature, and $C(d)$ is a constant depending on the ambient dimension.

On the contrary, $\left\| x - M_r^0(x) \right\|$ is linear (*ie.* $\Omega(r)$) in r if x belongs to a sharp edge. In this case, the eigenanalysis of the covariance matrix of $B(x,r) \cap D$ (with respect to the centroid) can be used to get the direction of the sharp edge. Following this work, Pottmann and coauthors [PWHY09] used principal component analysis of the intersections $B(x,r) \cap D$ to estimate principal curvature directions of ∂D at a smooth point x.

An important drawback of this kind of asymptotic analysis is that *eg.* "locally smooth" is not a quantitative statement. If the point x in the proposition is very close to a feature point (let's say at distance $r_0 \ll 1$), the behaviour of the function $r \mapsto \left\| M_r^0(x) - x \right\|$ will be linear as long as $r \geqslant \Omega(r_0)$, at which point only it will start being quadratic. This is of course unavoidable, but it seems nonetheless important to be able to quantify this, through notions such as reach, local feature size, etc.

Hausdorff-stability. Given a scale r and a point x in $\mathbb{R}^d$, most of the quantities defined above (at least the Connolly function function and $M_r^0(x)$) are Hausdorff-stable — though this question has not really been considered.

The reason of the stability is easy to grasp. Computing the volume (or any other integral quantity) of the intersection $B(x,r) \cap D$ only depends on the

measure $\mathcal{H}^3\big|_D$, and not at all on the smoothness of the boundary. Hence, if one replaces D with another domain D' such that the volume of the symmetric difference $D'\Delta D$ is small, the volumes $\mathcal{H}^3(D\cap B(x,r))$ and $\mathcal{H}^3(D'\cap B(x,r))$ will also be close. This is true, regardless of any higher order closeness properties for ∂D and $\partial D'$.

III.1.3 — Normal estimation: from PCA to Voronoi-PCA

In all the methods presented below, the assumption is that the point cloud C is drawn from a compact smooth surface S in $\mathbb{R}^3$ (often with assumptions on the reach of S). Only unoriented normals are estimated.

Principal component analysis estimation. A commonly used method for estimating normals is based on local principal component analysis. Given a point cloud C and $p \in C$, the normal at p is estimated by determining the plane that best approximates the k nearest neighbors to p in the least-square sense (*cf.* Algorithm 3). A commonly cited reference for this method is [HdRT+92].

Algorithm 3 local PCA-based normal estimation

1. select the k closest neighbors $p_1, \dots, p_k$ to p in C;
2. compute the affine hyperplane H_p that solves the least-square problem
$$H_p := \arg\min_H \sum_{i=1}^{k} d(p_i, H)^2$$
3. choose the unit normal in the orthogonal $\{H_p\}^\perp$.

There seems to have been no formal study of the convergence properties of this algorithm; however there are at least two requirements:

— The sampling should be rather uniform. When the sampling is not perfectly uniform, the k nearest neighbor can be replaced by all the neighbors in a ball of given radius r. However, this method cannot (easily) be made resilient to stronger sampling bias under structural noise, because the least-square fitting step will be biased toward oversampled areas.

— The intuitive geometric condition for the algorithm to work is that given a point $p \in S$, the patch of surface $B(p,r) \cap S$ should be close to the affine tangent plane $p + T_pS$. A necessary condition is that the (sectional) curvature is low. But this isn's sufficient: the ball $B(p,r)$ also shouldn't contain points from farther parts of the surface. This will be true, provided that the reach of S is high.

Since the algorithm tries to approximate the tangent space rather than the normal space to S, it does not take full advantage of having "large" normal cones. This explains why the method is not very robust to noise — unless the parameter r is chosen big enough, which leads to oversmoothing. Moreover,

this normal estimation methods behaves very poorly if the underlying object has sharp edges.

Voronoi-based normal estimation. These techniques have been pioneered by Amenta and Bern [AB99]; they rely on the fact that given a densely sampled point cloud C on a surface $S \subseteq \mathbb{R}^3$, the Voronoi cell of any point $p \in C$ is elongated in the direction of the normal to S at p. Such Voronoi cells are often called *pencil*-shaped in the litterature. Amenta and Bern calls the farthest Voronoi vertex to p in $\mathrm{Vor}_S(p)$ the positive *pole* v^+ of $\mathrm{Vor}_C(p)$. The normal $\mathbf{n}_C(p)$ is estimated by the direction of the vector $\mathbf{v}^+\mathbf{p}$.

The geometric intuition behind this method lies in the following simple fact. If p is a point on the hypersurface S, and B is an open ball whose closure contains p, and that does not contain any other point in S — we call such a ball an *empty ball* at p —, then the vector joining the center v of B and p must belong to the normal space to S at p.

Amenta and Bern's result quantify the (potential) deviation of $\mathbf{p}^+\mathbf{v}$ to the normal $\mathbf{n}_S(p)$ under approximation. In order to give a precise statement, we need to introduce the notion of ε-sample:

DEFINITION III.2. The *local feature size* at a point $p \in S$ is the distance from x to the medial axis: $\mathrm{lfs}_S(p) = d(p, \mathrm{Med}(S))$. A (finite) subset $C \subseteq S$ is called an ε-sample iff every ball $B(p, \mathrm{lfs}_S(p))$ ($p \in S$) contains a point of C.

The following proposition proves that pole-based normal estimation is consistent ([AB99]):

PROPOSITION III.4. *Let C be an ε-sample on a smooth closed hypersurface S; then:*

1. *For any point $p \in C$, there exists a point v in $\mathrm{Vor}_C(p)$ such that $\|p - v\| \geqslant \mathrm{lfs}_S(p)$.*
2. *If $v \in \mathrm{Vor}_C(p)$ is such that $\|v - p\| \geqslant v\,\mathrm{lfs}_S(p)$, then the angle between the vector $\mathbf{pv}$ and the normal $\mathbf{n}_S(p)$ oriented in the same direction is at most*

$$\arcsin(\varepsilon/v(1-\varepsilon)) + \arcsin(\varepsilon/(1-\varepsilon)) \sim_{\varepsilon\to 0} (1 + 1/v)\varepsilon$$

In the noisy case, the Voronoi cells are deformed, which can affect the orientation of the normal. More importantly, the Voronoi cells can be clamped, and become isotropic (see Fig. III.2) – this phenomenom completely breaks Amenta and Bern's method. A modification of this method by Dey and Sun [DS05] tries to recover from this situation this by looking for a large cell among the nearby Voronoi cells. The sampling conditions are a bit more intricate; however, it should be noticed that the deviation bounds for the estimated normals proven in Dey and Sun's paper degrades from $O(\varepsilon)$ to $O(\sqrt{\varepsilon})$.

Alliez and colleagues [ACSTD07] have a different approach. They compute the covariance matrix of every Voronoi cell; in the pencil case, the eigenvector corresponding to the largest eigenvalue will be aligned along the normal. In

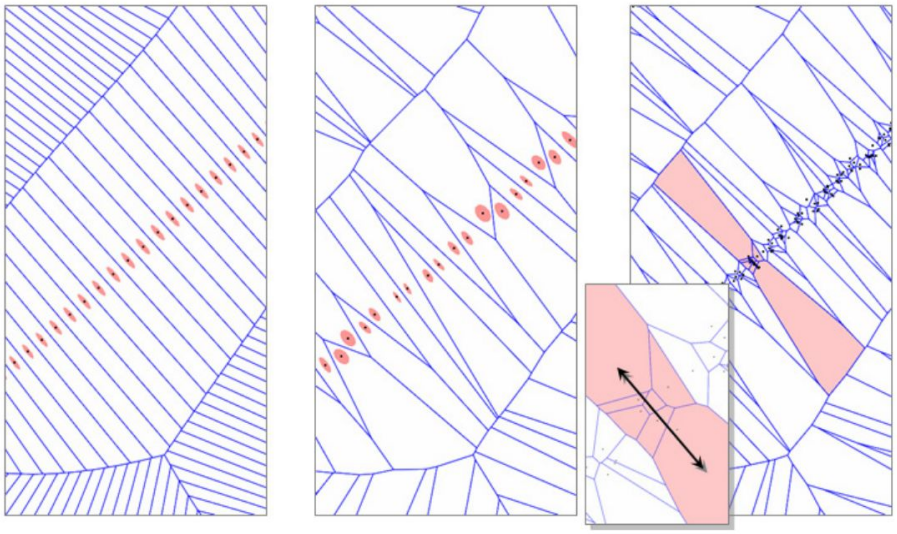


Figure III.2 – Voronoi-based normal estimation. In the second picture, one can see the angle defect in the normals estimated unsing the pole method, even if the points have been only moderately perturbed. The third picture shows clamped Voronoi cells under stronger noise. *(courtesy of P. Alliez et al.)*

order to accomodate noise, given a point $p \in C$, they first compute the average of the covariance matrices of the Voronoi cells of the k nearest neighbor. The largest eigenvector of this averaged covariance matrix is then considered as an approximation of the normal at p (see Fig. III.2).

As we will see, these covariance matrices appear naturally as convolved anisotropic boundary measures. They provide information one the normal cones, even in the non-smooth setting, which makes them useful for feature detection.

III.1.4 — Definition of the Voronoi covariance measure

Infinitesimal Voronoi cells. The projection function maps a point in $\mathbb{R}^d \setminus \mathrm{Med}(K)$ to its only closest point $p_K(x)$ in $K \subseteq \mathbb{R}^d$. Recall that thanks to Corollary I.5, this function is defined almost everywhere; being only interested in integral quantities defined from p_K, we will do as though p_K was defined on the whole space $\mathbb{R}^d$.

If $K = \{p_1, \ldots, p_n\}$ is a point cloud, and $\mathrm{Vor}_K(p_i)$ denote the Voronoi cell of p_i, the projection on K simply maps a point $x \in \mathbb{R}^d$ to the center of its Voronoi cell. By analogy to this case, we will refer to the set $p_K^{-1}(B)$ as the Voronoi cell of the set $B \subseteq K$, and to $p_K^{-1}(\{x\})$ as the infinitesimal Voronoi cell of the point $x \in K$.

DEFINITION III.3. The *Voronoi covariance measure* (or VCM) is defined for *any* compact set K of $\mathbb{R}^d$. This can be a finite point cloud, a (piecewise) smooth manifold, or an even wilder object. We also need a scale parameter R, which will be used to define an offset K^R of K.

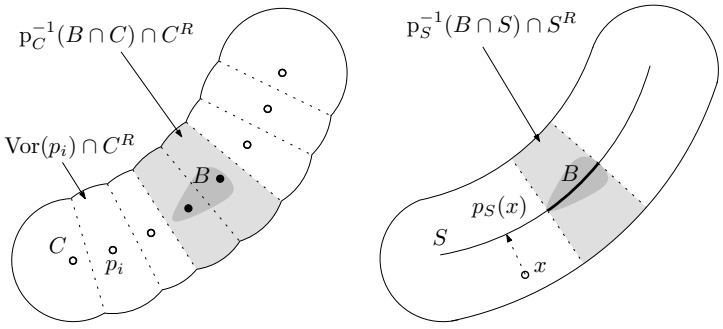


Figure III.3 – Voronoi covariance measures $\mathcal{V}_{C,R}(B)$ of a 2D point cloud C and of the underlying curve S, with respect to the same probing set B.

The Voronoi covariance measure of K with respect to K^R is a tensor-valued measure denoted by $\mathcal{V}_{K,R}$. Being tensor-valued means that unlike a usual measure μ in the sense of Lebesgue, which maps every (Borel) subset B of $\mathbb{R}^d$ to a non-negative number $\mu(B)$, the Voronoi covariance measure maps every such $B \subseteq \mathbb{R}^d$ to a non-negative definite covariance matrix $\mathcal{V}_{K,R}(B)$. This covariance matrix is defined as follows:

$$\mathcal{V}_{K,R}(B) = \int_{K^R \cap \mathrm{p}_K^{-1}(B \cap K)} (x - \mathrm{p}_K(x)) \otimes (x - \mathrm{p}_K(x)) dx$$

Figure III.3 illutrates the domain of integration in the definition of $\mathcal{V}_{K,R}(B)$ when K is a point cloud or a curve. Intuitively, the VCM $\mathcal{V}_{K,R}(B)$ is the covariance matrix of the "Voronoi cell" $\mathrm{p}_K^{-1}(B \cap K) \cap K^R$, but with a varying base point: one can think of it as the integral over all $p \in B \cap K$ of the covariance matrices of the infinitesimal Voronoi cell $\mathrm{p}_K^{-1}(\{p\}) \cap E$, with base point p.

DEFINITION III.4. Given a compact set $K \subseteq \mathbb{R}^d$, the *normal cone* $\mathcal{N}_p K$ at a point $p \in S$ is the positive cone generated by all the vectors $x - p$, where x is any point that projects onto S. That is: $\mathcal{N}_p K = \{\lambda(x - p)\,;\, \lambda > 0 \text{ and } \mathrm{p}_S(x) = p\}$.

We will call *normal set to* K *at* p *at scale* R, and denote by $\mathcal{N}_{p,R} K$ the intersection of the infinitesimal Voronoi cell $\mathrm{p}_K^{-1}(p)$ translated at the origin with the ball $B(0, R)$: $\mathcal{N}_{p,R} K = \{x - p;\ \mathrm{p}_K(x) = p \text{ and } x \in K^R\}$. Notice that for any positive R, the normal cone is the positive cone generated by $\mathcal{N}_{p,R}$.

The normals in the cone $\mathcal{N}_p K$ are often called *proximal normals* to K at p ([CS04, p. 75]); the normal cone $\mathcal{N}_p K$ also coincides with the definition of Clarke ([Cla83, p. 10]). The normal set $\mathcal{N}_{p,R} K$ (which is not a cone!) bears some resemblance with the approximate normal cones defined in [CCSL08]. Note that while $\mathcal{N}_p K$ is always convex, this is not necessarily the case for the set $\mathcal{N}_{p,R}$ ($R > 0$).

EXAMPLE. If K is smooth and $R < \mathrm{reach}(K)$ then $\mathcal{N}_{p,R} K$ is the segment $[-R\mathbf{n}(p), R\mathbf{n}(p)]$ (Figure III.3). On the other hand, if $K \subseteq \mathbb{R}^3$ is a convex polyhedron, $\mathcal{N}_{p,R} K$ corresponds to the is the intersection of the normal cone

at p with the ball $B(p, R)$. It is a 2-dimensional subset of $\mathbb{R}^3$ when p lies on an edge and 3-dimensional when p is a vertex (*cf.* Figure III.4).

If B is a small neighborhood of a point $p \in K$, the covariance matrix $\mathcal{V}_{K,R}(B)$ captures the variation of the normal cone of K around p, which is related to the curvature at p when K is locally smooth. This intuition is made more precise in §III.2.1.

DEFINITION III.5. The tensor-valued measure $\mathcal{V}_{K,R}$ can be convolved by any continuous and integrable function $\chi : \mathbb{R}^d \to \mathbb{R}$. This turns it into a (tensor-valued) density function $\mathcal{V}_{K,R} * \chi : \mathbb{R}^d \to \mathrm{Sym}(\mathbb{R}^d)$, defined by

$$\mathcal{V}_{K,R} * \chi(p) := \int_{K^R} (x - p_K(x)) \otimes (x - p_K(x)) \chi(p_K(x) - p) dx$$

Let $\chi_r : \mathbb{R}^d \to \mathbb{R}$ be the indicator function of the ball $B(0, r)$, defined by $\chi_r(p) = 1$ if $x \in B(0, r)$ and $\chi_r(p) = 0$ otherwise. Note that in this case the convolved VCM has a particularly simple expression:

$$\mathcal{V}_{K,R} * \chi_r(p) = \mathcal{V}_{K,R}(B(p, r)) \tag{III.2}$$

In this work, the convolution kernel is always chosen to be a Lipschitz approximation of such an indicator function, like the "hat function" $\chi(p) = \max(0, r - \|p\|^2)$, for which we can prove that $\mathcal{V}_{K',R} * \chi$ converges to $\mathcal{V}_{K,R} * \chi$ as K' converges to K (see §III.2.3). We also often use the indicator function itself for which we get good results even though the theoretical guarantees of convergence do not apply directly.

VCM of a point cloud. In this paragraph, we derive a simple formula for the VCM of a point cloud $C = \{p_1, \dots, p_n\}$ in $\mathbb{R}^d$. Given a set $B \subseteq \mathbb{R}^d$, the inverse image $p_C^{-1}(B)$ is the union of Voronoi cells whose center lie in B; hence:

$$\begin{aligned} \mathcal{V}_{C,R}(B) &= \int_{C^R \cap p_C^{-1}(B)} (x - p_C(x)) \otimes (x - p_C(x)) \\ &= \sum_{p_i \in B} \int_{C^R \cap \mathrm{Vor}_C(p_i)} (x - p_i) \otimes (x - p_i) \\ &= \sum_{p_i \in B} \mathrm{cov}\left(\mathrm{Vor}(p_i) \cap B(p_i, R), p_i\right) \end{aligned} \tag{III.3}$$

The last equality holds because the intersection $\mathrm{Vor}_C(p_i) \cap C^R$ is equal to the intersection$\mathrm{Vor}_C(p_i) \cap B(p_i, R)$. Notice the resemblance with the covariance matrices used in [ACSTD07]: the main differences is that in their definition, the Voronoi cells are intersected with an enlarged bounding box of the point cloud instead of the balls $B(p_i, R)$.

Convolved VCM of a point cloud. Using equation III.2, and choosing the convolution kernel to be the indicator function of a ball with radius r, we

obtain the following expression:

$$\mathcal{V}_{C,R} * \chi_r(p) = \sum_{p_i \in B(p,r) \cap C} \mathbf{cov}\left(\mathbf{Vor}(p_i) \cap \mathbf{B}(p_i, R), p_i\right)$$

We describe practical algorithms to compute the convolved VCM of a point cloud in §III.3.1.

III.2 THEORETICAL PROPERTIES OF THE VCM

The analysis of the geometric inference using the Voronoi covariance measure is organized as follows:

— First, we study the "asymptotic case", *ie.* we try to understand what information is contained in $\mathcal{V}_{S,R}$ when S is the underlying object. In §III.2.1, we study the VCM of a smooth surface S, and its relation with normals and curvature. In §III.2.2, we show that the Voronoi covariance measure of a polyhedron P in $\mathbb{R}^3$ contains information about its sharp edges — under a lower bound on the 1-sided reach (a condition that forbids concave corners).

— In a second step, we study the stability of the VCM under Hausdorff approximation (in §III.2.3). We also show that the VCM is resilient to a certain class of outliers.

These stability results and the study of the asymptotic case can be combined to show that VCM of a point cloud sampled close to a smooth or piecewise-smooth surface can be used to infer geometric properties such as normals or sharp feature directions of that surface.

III.2.1 — Voronoi covariance of a smooth hypersurface

Let S be a compact smooth hypersurface embedded in $\mathbb{R}^d$, $\mathbf{n}$ an oriented normal vector field on S and R_S the reach of S. As before, S^R denotes the R-offset of S.

Let us recall a few facts about the curvature of embedded hypersurfaces. The map $x \in S \mapsto \mathbf{n}(x)$ is called the *Gauss map* of S, while its derivative $d\mathbf{n}$ is called the *shape operator* of S. If $\kappa_1(x), \ldots, \kappa_{d-1}(x)$ denote the $(d-1)$ principal curvatures at x and $P_1(x), \ldots, P_{d-1}(x)$ is a set of vectors in the tangent plane spanning the principal curvature directions, with $\|P_i(x)\| = 1$, one has: $\forall v \in \mathbf{T}_x S,\ d\mathbf{n}(x)(v) = \sum_{i=1}^{d-1} \kappa_i(x) \langle P_i(x) | v \rangle P_i(x)$.

In the following, we show that the Voronoi Covariance Measure $\mathcal{V}_{S,R}$ evaluated on a subset $B \subseteq S$ can be described in terms of the covariance matrix of the normal vectors $\mathbf{n}(x),\ x \in B$:

THEOREM III.5. *If* $R < R_S$, *the reach of* S, *then for every* $B \subseteq S$ *one has*

$$\mathcal{V}_{S,R}(B) = \sum_{k=0}^{\lfloor (d-1)/2 \rfloor} \frac{R^{2k+3}}{k + \frac{3}{2}} \int_{p \in B \cap S} \sigma_{2k}(p) [\mathbf{n}(p) \otimes \mathbf{n}(p)]\, dp$$

Proof. This computation is almost the same than in the proof of Prop. II.1. If R is smaller than the reach of S, the map $\varphi : S \times [-R, R] \to S^R, (p, t) \mapsto p + t\mathbf{n}(p)$, where S^R is the R-neighborhood of S, is a diffeomorphism. Thus, the change-of-variable formula yields:

$$\begin{aligned}\mathcal{V}_{S,R}(B) &= \int_{S^R \cap p_S^{-1}(B \cap S)} (x - p_S(x)) \otimes (x - p_S(x)) dx \\ &= \int_{p \in B \cap S} \int_{-R}^{R} \mathbf{n}(p) \otimes \mathbf{n}(p)\ t^2\ J\varphi(p,t)\ dt\ dp\end{aligned}$$

where $|J\varphi(p,t)|$ is the Jacobian determinant of φ at $(p,t) \in S \times [-R, R]$. In the local frame of the tangent space given by the $d-1$ principal curvature directions $P_1(p), \ldots, P_{d-1}(p)$, the derivative of φ is the diagonal matrix with components $1 + t\kappa_1(p), \ldots, 1 + t\kappa_{d-1}(p), 1$. From this, we are able to compute the Jacobian determinant:

$$|J\varphi(p,t)| = \prod_{i=1}^{d-1} (1 + t\kappa_i(p)) = \sum_{i=0}^{d-1} t^i \sigma_i(p)$$

In this expression, $\sigma_i(p) = \sum_{j_1 < \ldots < j_i} \kappa_{j_1}(p) \ldots \kappa_{j_i}(p)$ is the i^{th} symmetric polynomial of the principal curvatures at p. Since all terms with odd powers of t vanish when we integrate from $-R$ to R, we get the desired polynomial expansion for the VCM. □

As a consequence, the eigenvector corresponding to the largest eigenvalue of $\mathcal{V}_{S,R}(B(p_0, r))$ converges to the normal to S at p_0 as r goes to zero:

COROLLARY III.6. *If $R < R_S$, the reach of S, then for every $p_0 \in S$ and $r > 0$ small enough, one has*

$$\mathcal{V}_{S,R}(B(p_0, r)) = \text{const}(S, R, p_0)\ (\mathbf{n}(p_0) \otimes \mathbf{n}(p_0) + O(r/R_S))$$

where

$$\text{const}(S, R, p_0) = \sum_{k=0}^{\lfloor (d-1)/2 \rfloor} \left[\frac{R^{2k+3}}{k + \frac{3}{2}} \int_{p \in B(p_0, r) \cap S} \sigma_{2k}(p)\ dp \right]$$

Proof. The map $p \in S \mapsto \mathbf{n}(p)$ is locally $1/R_S$ Lipschitz for the induced metric on S (because the principal curvatures are at most $1/R_S$). One concludes using $d_S(p_0, p) \geqslant \|p - p_0\|$. □

Voronoi-covariance and curvature. Analysing the relationship between Voronoi covariance measure of S and the curvature of the surface in a way that allows to give precise sampling conditions that allow to recover (approximately) the principal curvature directions from a point cloud sampling is quite difficult.

Moreover, in practice, the radii R and r will have to be chosen big enough in order to compensate for noise, and such conditions would seldom be met. As

we mentioned in the introduction, the most important feature of a local signature like $\mathcal{V}_{S,R}(B(p_0,r))$ is the stability of the VCM with respect to Hausdorff approximation; which is proven for VCM in §III.2.3.

Let us still mention the following lemma, which together with Theorem III.5 indicate a strong link between $\mathcal{V}_{S,R}(B(p_0,r))$ and the local second-order behaviour of S. In the experimental section, we will see that under good sampling conditions, the two eigenvectors corresponding to the two smallest eigenvalues of $\mathcal{V}_{S,R}(B(p_0,r))$ indeed align with the principal curvature directions.

LEMMA III.7. *Let $p_0 \in S$, and note $\mathcal{M}_S(p_0,r)$ be the matrix defined by*

$$\int_{B(0,r)} \mathbf{n}(\exp_{p_0}(v)) \otimes \mathbf{n}(\exp_{p_0}(v)) dv$$

Then,

$$\mathcal{M}_S(p_0,r) = \omega_{d-1} r^{d-1} \mathbf{n}(p_0) \otimes \mathbf{n}(p_0) + O(r^{d+1})$$

Moreover, the restriction of $\mathcal{M}_S(p_0,r)$ to the tangent plane $T_{p_0}S$ satisfies:

$$\begin{aligned}\mathcal{M}_S(p_0,r)|_{T_{p_0}S} &= \omega_{d-1} r^{d-1} \mathbf{n}(p_0) \otimes \mathbf{n}(p_0) \\ &\quad + \text{const}(d) r^{d+1} \sum_{i=1}^{d-1} \kappa_i^2(p_0)[P_i(p_0) \otimes P_i(p_0)] + O(r^{d+2})\end{aligned}$$

Proof. First, write the Taylor expansion formula for normals:

$$\mathbf{n}(\exp_{p_0}(v)) = \mathbf{n}(p_0) + d_{p_0}\mathbf{n}(v) + O(\|v\|^2)$$

Now, since $\int_{B(0,r)} d_{p_0}\mathbf{n}(v) = 0$, the first order terms (in $\|v\|$) in $\mathbf{n}(\exp_{p_0}(v)) \otimes \mathbf{n}(\exp_{p_0}(v))$ vanish when integrated on the ball $B(0,r)$. This gives the first formula. Moreover, when w is in the tangent plane to S at p_0, the scalar product $\langle \mathbf{n}(p_0) | w \rangle$ vanishes. This means that the term $O(\|v\|^2) \otimes \mathbf{n}(p_0)$ (and the symmetric term) don't contribute.

Let $P_i(p_0)$ be the ith principal curvature direction. Then,

$$\begin{aligned}&\int_{B(0,r)} [d_{p_0}\mathbf{n}(v) \otimes d_{p_0}\mathbf{n}(v)](P_i(p_0), P_i(p_0)) dv \\ &\qquad = \int_{B(0,r)} \kappa_i^2(p_0) \langle v | P_i(p_0) \rangle^2 dv \\ &\qquad = 2\kappa_i^2(p_0) \int_0^r s^2 \beta_{d-2}\left(\sqrt{r^2 - s^2}\right) ds\end{aligned}$$

Using a proper change of variable, this last integral is equal to

$$\beta_{d-2} r^{d+1} \int_0^1 u^2 \left(1 - u^2\right)^{(d-1)/2} ds = \text{const}(d) r^{d+1}$$

The second formula follows. □

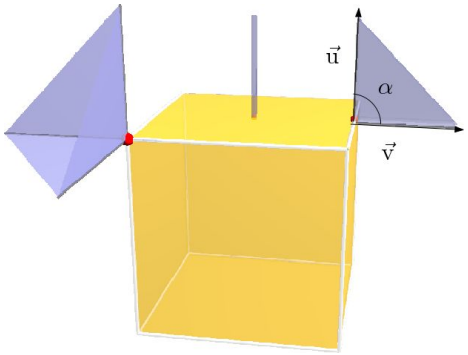


Figure III.4 – The infinitesimal Voronoi cell $p_C^{-1}(x)$ of a point x on a cube is pencil, triangle or cone-shaped depending on the dimension of the normal cone.

III.2.2 — Voronoi covariance and sharp features

The dimension of the Voronoi cell of a point, and more precisely the number of "small" eigenvalues s of its covariance matrices, can help to distinguish between a smooth point ($s = 2$), an edge point ($s = 1$) and a vertex ($s = 0$) on a polyhedron (*cf.* Figure III.4). In this paragraph, we use this intuition to make a precise link between between the eigenanalysis of the convolved VCM of a polyhedron of $\mathbb{R}^3$ with positive 1-reach (this notion is defined below) and the location and angle of its sharp edges.

DEFINITION III.6 (Edge, vertex, dihedral angle). Let S be a (not necessarily smoothly) embedded compact close surface in $\mathbb{R}^3$. We distinguish three type of points depending on the dimension of the normal cone:

— If the normal cone $\mathcal{N}_p S$ at a point $p \subseteq S$ contains two opposite vectors, then S admit a tangent plane at p — because the surface is stuck between two balls tangent to each other at p. Such a point is a *smooth* point of S.

— The points where the dimension of $\mathcal{N}_p S$ is either 3 or 0 are called *vertices* of S and denoted by $\mathrm{Vert}(S)$. Vertices with $\dim \mathcal{N}_p S = 3$ are *concave*, while those with $\dim \mathcal{N}_p S = 0$ are *convex*.

— The remaining points, *ie.* those where $\dim(\mathcal{N}_{p_0} S) = 2$ are called *edge points*; their locus is denoted by $\mathrm{Edg}(S)$. If p_0 is a point in $\mathrm{Edg}(S)$, the normal cone $\mathcal{N}_{p_0} S$ is a two-dimensional convex cone. We will denote by $\mathbf{e}(p_0)$ the *feature direction* at p_0, *ie.* a unit vector orthogonal to this normal cone. Also denote by $\mathbf{u}^{\pm}(p_0)$ the leftmost and rightmost unit vectors in the circle arc corresponding to the intersection of the normal cone with the unit sphere. The positive angle between these two vectors is denoted by $\alpha(p_0)$ and called *dihedral angle*.

DEFINITION III.7 (One-sided local feature size). In general, if the local feature size of S at a point p is greater than R, then the infinitesimal Voronoi cell can be simply written in term of normal cone: $p_K^{-1}(\{p\}) \cap B(p, R) = p + (\mathcal{N}_p S \cap B(0, R))$. In those cases, the VCM $\mathcal{V}_{S,S^R}$ at that point is fully described by the normal cone.

Because the lfs vanishes at any non-smooth point, we have to introduce a weaker notion of feature size. The 1-*sided local feature size*, or $\mathrm{lfs}^1_S(p)$, of a point p is defined the supremum of all radii $R \geqslant 0$ such that for all unit normal vector $\mathbf{n}$ in $\mathcal{N}_p S$ the projection on S of either $p + R\mathbf{n}$ or $p - R\mathbf{n}$ is at p. The 1-reach of S, denoted by $\mathrm{reach}_1(S)$ is the infimum of $\mathrm{lfs}^1_S(p)$ among all $p \in S$. (Although the notation can be a bit misleading, this notion of reach has nothing to do with the μ-reach defined in [CCSL09] with $\mu = 1$.)

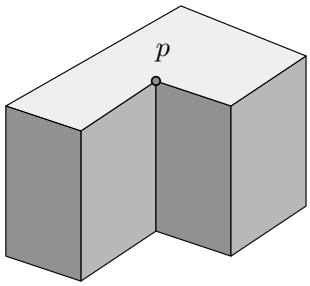


Figure III.5 – The corner at the point p is *concave*, meaning that the normal cone at p only contains the zero vector.

REMARKS. — For any edge point p_0, let $Q(\mathbf{u}^{\pm}(p_0))$ be the cone at the origin positively generated by the vectors $\mathbf{u}^{\pm}(p_0)$, *ie.* $Q(\mathbf{u}^{\pm}(p_0)) = \{\lambda\mathbf{u}^{+}(p_0) + \nu\mathbf{u}^{-}(p_0); \lambda > 0, \nu > 0\}$. If S has positive 1-reach, the infinitesimal Voronoi cell at such a point p_0 is equal to the intersection of the cone $Q(\mathbf{u}^{\pm}(p_0))$ and the ball $B(0, R)$. This intersection is a circular slice with radius R and angle $\alpha(p_0)$.

— Unlike the local feature size, which is known to be Lipschitz, the 1-sided lfs is not even necessarily continuous. Indeed, $\mathrm{lfs}^1_S(p)$ can get arbitrarily close to zero as p moves along an edge to one of its extremities p_0, even though $\mathrm{lfs}^1_S(p_0) > 0$.

— According to this definition, the one-sided lfs at a concave corner — *ie.* a corner $p \in S$ whose normal cone only contains the zero vector — is always infinite. However, such corners are discarded by the positive 1-reach assumption, as a consequence of Lemma III.8. Indeed, it suffices to remark that any point in S can be written as a limit of points $p_n \in S$ such that $\dim \mathcal{N}_{p_n} S > 0$; the limit point will then automatically have a non-null normal.

LEMMA III.8. *Let $S \subseteq \mathbb{R}^d$ be a compact set with $\mathrm{reach}_1(S) \geqslant R > 0$, (p_n) be a sequence of points in S converging to a point $p \in S$, and $(\mathbf{n}_n)$ a sequence of unit normals to p_n converging to a unit vector $\mathbf{n}$. Then $\mathbf{n}$ belongs to the normal cone at p.*

Proof of Lemma III.8. For each n, choose a point x_n that projects onto p_n, with $\|p_n - x_n\| \geqslant R$ and such that $(p_n - x_n)$ is collinear to $\mathbf{n}_n$. By compactness of S^R, we can suppose that the sequence x_n converges to some point x. Then, $\|x - p\| = R$, and p is among the (possibly) multiple projections of x. Then for any $r < R$, the point $p + \frac{r}{R}(x - p)$ projects onto p, proving the lemma. □

LEMMA III.9. *If $\gamma :]-\varepsilon, \varepsilon[\to S$ is a curve, differentiable at 0 and such that $p_0 = \gamma(0)$ belongs to $\mathrm{Edg}(S)$, then $\mathbf{e}(p_0)$ is collinear to $\gamma'(0)$.*

Proof. Consider any point $x \in \mathbb{R}^d$ with $p_S(x) = p_0$. By definition of the projection, the open ball $B(x, d(x, p_0))$ does not contain any point in S and in particular points of the curve γ. Hence, it must be tangent to γ at p_0. This proves that $\mathcal{N}_{p_0}S = \{\mathbf{e}(p_0)\}^\perp \subseteq \{\gamma'(0)\}^\perp$, from which one concludes. □

There seems to be no simple way to define a notion of piecewise-smooth surface that allows to give quantitative statements on the combinatorics, total length, geometry *etc.* of $\mathrm{Edg}(S)$. In the remaining of this §, we will suppose that S is a polyhedron with positive 1-reach. Most of these results also apply to more general objects.

In that case, the notion of external dihedral angle defined above coincide with the usual notion for closed polyhedron: if p_0 is on an edge, $\mathbf{u}^\pm(p_0)$ are the projections of the normals to the two faces adjacent to this edge on the plane $\{\mathbf{e}(p_0)\}^\perp$.

LEMMA III.10. *Let $\mathbf{u}, \mathbf{v}$ be two unit vectors in the Euclidean plane, and V be the intersection of the cone $Q(\mathbf{u}, \mathbf{v})$ with the disk $B(0, R)$. The eigenpairs of the covariance matrix $\mathrm{cov}(V, 0)$ are:*

$$\lambda_1 = \frac{R^4}{8}(\alpha - \sin(\alpha)),\ \mathbf{e_1} = \mathbf{u} - \mathbf{v} \quad \textit{and} \quad \lambda_2 = \frac{R^4}{8}(\alpha + \sin(\alpha)),\ \mathbf{e_2} = \mathbf{u} + \mathbf{v}$$

where $\cos(\alpha) = \langle \mathbf{u} | \mathbf{v} \rangle$.

Proof. Consider the covariance in some fixed unit direction $\mathbf{d}$ which is at angle β from $\mathbf{v}$. Then, with $M = \mathrm{cov}(V, 0)$,

$$\begin{aligned} \mathbf{d}^T M \mathbf{d} &= \int_{\theta=0}^{\alpha} \int_{r=0}^{R} r^2 \cos^2(\theta - \beta) r \, dr \, d\theta = \frac{R^4}{4} \int_0^\alpha \frac{\cos(2(\theta - \beta)) + 1}{2} d\theta \\ &= \frac{R^4}{8} \left(\frac{\sin(2(\alpha - \beta)) + \sin(2\beta)}{2} + \alpha \right) \end{aligned} \tag{III.4}$$

By symmetry of the cone, an obvious eigenvalue is the vector $\mathbf{e}_1$ making an angle angle $\frac{1}{2}\alpha$ with $\mathbf{v}$. The second one is orthogonal to $\mathbf{e}_2$ and thus makes an angle of $\frac{1}{2}(\alpha - \pi)$ with $\mathbf{v}$. To compute the eigenvalues associated with these eigenvectors, we plug these two angles into (III.4). □

For any point p in the closed polyhedron S, define $R(p)$ as the smallest (Euclidean) distance between p and any non-adjacent face in S.

PROPOSITION III.11. *Let $S \subseteq \mathbb{R}^3$ be a closed polyhedron with $\mathrm{reach}_1(S) \geqslant R$. One then has the following characterization of points on sharp edges vs points on smooth 2-face:*

(i) For any point p_0 on a 2-face with normal $\mathbf{n}_0$, and $r < R(p)$:

$$\mathcal{V}_{S,S^R}(B(p_0, r)) = \lambda(p_0, r, R)\pi r^2 \frac{R^3}{3} [\mathbf{n}_0 \otimes \mathbf{n}_0]$$

with $8 \geqslant \lambda(p_0, r, R) \geqslant 1$.

(ii) For any point p_0 *on an edge* $e \subseteq \mathrm{Edg}(S)$*, and* $r < R(p_0)$*,*

$$\begin{aligned}\mathcal{V}_{S,R}(B(p_0,r)) = \frac{R^4 r}{4}[&(\alpha(p_0) - \sin(\alpha(p_0)))\, e_1(p_0) \otimes e_1(p_0) \\ &+ (\alpha(p_0) + \sin(\alpha(p_0)))\, e_2(p_0) \otimes e_2(p_0)] \\ &+ \pi r^2 \frac{R^3}{3} \varepsilon(p_0, R, r)\end{aligned}$$

where $\mathbf{e_1}(p_0) = \frac{\mathbf{u}^+(p_0) - \mathbf{u}^-(p_0)}{\|\mathbf{u}^+(p_0) - \mathbf{u}^-(p_0)\|}$, $\mathbf{e_2}(p_0) = \frac{\mathbf{u}^+(p_0) + \mathbf{u}^-(p_0)}{\mathbf{u}^+(p_0) + \mathbf{u}^-(p_0)}$, $\|\varepsilon(p_0, R, r)\|_{\mathrm{op}} \leqslant 1$.

Proof. (i) The intersection of $B(p_0, r)$ with S only contains point from the same facet as p_0 (*ie.* with the same normal). By the 1-reach assumption, for each $p \in B(p_0, r) \cap S$, the length $\ell(p) = b_p - a_p$ of $p_S^{-1}(p)$ is at least R, and at most $2R$. Hence,

$$\begin{aligned}\mathcal{V}_{S,R}(B(p_0,r) \cap S) &= \int_{a_p}^{b_p} \int_{p \in S} t^2 [\mathbf{n}_0 \otimes \mathbf{n}_0] dp dt \\ &= \left(\int_{p \in B(p_0,r) \cap S} \frac{\ell(p)^3}{3} dp \right) [\mathbf{n}_0 \otimes \mathbf{n}_0].\end{aligned}$$

The bound on the scalar factor follows.

(ii) The intersection $S \cap B(p_0, r)$ can be written as $S \cap (e \cup S_1 \cup S_2)$ where e is the edge and S_1 and S_2 are the two adjacent faces (minus their boundaries). The term in $\frac{R^4 r}{4}$ comes from $\mathcal{V}_{S,S^R}(B(p_0, r) \cap e)$ and Lemma III.10. The bound of the error term follows, by the previous computation, and using $\mathcal{H}^2(B(p_0, r) \cap (S_1 \cup S_2)) = \pi r^2$. □

III.2.3 — Robustness of Voronoi covariance measure

This section is devoted to the proof of robustness of the VCM. By robustness we mean that if two compact sets K and K' are very close in the Hausdorff sense, then their VCM are also close. In applications, K will be a sampled surface, and K' will be a point cloud; however the theorem does not make any assumption on the nature of these compact sets.

The distance between two functions $f, g : \mathbb{R}^d \to \mathrm{Sym}(\mathbb{R}^d)$ is given by the norm of uniform convergence, two symmetric matrices being compared by the operator norm:

$$\|f - g\|_\infty := \sup_{p \in \mathbb{R}^d} \|f(p) - g(p)\|_{\mathrm{op}}$$

The Hausdorff-robustness of the convolved VCMs when the convolution kernel $\chi : \mathbb{R}^d \to \mathbb{R}$ is Lipschitz and bounded follows from the projection stability theorem presented in Chapter II. Actually, there is nothing special about convolved VCMs: one could as well define a notion of bounded-Lipschitz distance for tensor-valued measures, and prove convergence of $\mathcal{V}_{K_n,R}$ to $\mathcal{V}_{K,R}$ in the bounded-Lipschitz sense as K_n Hausdorff-converges to K.

THEOREM III.12. *If $\chi : \mathbb{R}^d \to \mathbb{R}$ is a k-Lipschitz function with $\|\chi\|_\infty \leqslant 1$, then for every compact set $K \subseteq \mathbb{R}^d$ and $R > 0$, there is a constant $C'(d, K, R)$ such that*

$$\|\mathcal{V}_{K,R} * \chi - \mathcal{V}_{K',R} * \chi\|_\infty \leqslant C(d, k, K, R) d_H(K, K')^{1/2}$$

as soon as K' is close enough to K,

Proof. We let E be the intersection of K^R and K'^R. By Corollary II.15, we know that the volume $K^R \setminus E$ is in $O(d_H(K, K'))$. Hence, letting $\delta(x) = x - p_K(x)$, we only need to compare

$$\int_E \delta(x) \otimes \delta(x) \chi(p_K(x) - p) dx$$

and the same quantity defined for K', *ie.* we want to bound the operator norm of the difference: $M = \int_E P(x) - P'(x) dx$, where P (resp. P') is defined by $P(x) = \delta(x) \otimes \delta(x) \chi(p_K(x) - p)$. Then,

$$\begin{aligned} P(x) - P'(x) = \chi(p_K(x) - p) \left(\delta(x) \otimes \delta(x) - \delta'(x) \otimes \delta'(x)\right) \\ + \left(\chi(p_K(x) - p) - \chi(p_{K'}(x) - p)\right) \delta'(x) \otimes \delta'(x) \end{aligned}$$

The first term can be written as:

$$\chi(p_K(x) - p) \left[\left((p_K(x) - x) \otimes (p_K(x) - p_{K'}(x)\right) + (p_{K'}(x) - p_K(x)) \otimes (x - p_{K'}(x))\right]$$

Now since χ is bounded by 1, $\|p_K(x) - x\| < R$, and using the triangle inequality, the operator norm of this expression is bounded by $2R \|p_{K'}(x) - p_K(x)\|$. By the k-Lipschitz property of χ, the norm of the second term can be bounded by

$$|\chi(p_K(x) - p) - \chi(p_{K'}(x) - p)| R^2 \leqslant k R^2 \|p_K(x) - p_{K'}(x)\|$$

Hence,

$$\left\|P(x) - P'(x)\right\|_{op} \leqslant (2R + kR^2) \|p_K(x) - p_{K'}(x)\|$$

Integrating and using Theorem II.9 yields the bound

$$\|M\|_{op} \leqslant C(d, K, E)[2R + kR^2] d_H(K, K')^{1/2}$$

□

If S is a piecewise smooth surface, p a point of S and C_n a sequence of point clouds converging to S in the Hausdorff sense, the stability theorem says that $\mathcal{V}_{C_n,R} * \chi_r(p)$ converges to $\mathcal{V}_{S,R} * \chi_r(p)$ w.r.t the operator norm, where χ_r is a Lipschitz approximation of the characteristic function of the ball $B(0, r)$.

Robustness to some outliers. One limitation of the Hausdorff distance in the bound above, is its sensitivity to outliers. Indeed, even under controlled noise, outliers can contaminate the point cloud, and influence the shapes of the Voronoi cells. Nevertheless, as pointed out earlier, intersecting the Voronoi cells with an offset allows us to obtain local information which is

unaffected by of outliers that are sufficiently far away from the point cloud sampling of the surface:

LEMMA III.13. *Let* C *be a point cloud and* O *a set of outliers with* $d(o, C) \geqslant 2R$ *for any* o *in* O. *Then the convolved VCM* $\mathcal{V}_{C,R} * \chi(p)$ *and* $\mathcal{V}_{C',R} * \chi(p)$ *(where* $C' = C \cup O$*) agree on* C *provided that the support of* χ *is contained in a ball of radius smaller than* R.

Proof. For any point $p \in C$, the interection between its Voronoi cell and a ball of radius R remains unchanged after introducing points from O: indeed, the bissecting plane formed between p and any point $p' \in O$ is at least R away from p, since by assumption $\|p - p'\| > 2R$. This implies that for all $p \in C$, $\mathcal{V}_{C,R}(\{p\}) = \mathcal{V}_{C',R}(\{p\})$. The result on convolved VCMs follows by.

$$\mathcal{V}_{C',R} * \chi(p) = \sum_{p_i \in B(p,r)} \operatorname{cov}\left(\operatorname{Vor}(p_i) \cap C'^{R}, p_i\right)$$

and $\|p - p'\| > r$ there will be no outliers in the ball $B(p, r)$, and the desired equality follows. □

Stability of eigenvalues and eigenvectors. As a consequence of Davis-Kahan theorem, one obtains the convergence of the eigenvalues (with multiplicites) and eigenspaces of $\mathcal{V}_{C_n,R} * \chi_r(p)$ to those of $\mathcal{V}_{S,R} * \chi_r(p)$. Below, we give a example of such an analysis. The same analysis can be carried out for sharp features, giving the same speed of convergence once r has been fixed.

Let $p_0 \in S$, and $\chi_r : \mathbb{R}^d \to \mathbb{R}^+$ be the function that decreases linearly from 1 at the origin to 0 on $\partial B(x, r)$, and is null outsize of this ball. It can be proven exactly as in Corollary III.6 that for some constant C_1,

$$\|\mathcal{V}_{S,R} * \chi_r(p_0) - C_1 \mathbf{n}(p_0) \otimes \mathbf{n}(p_0)\|_{op} = O(r)$$

Moreover, for any point cloud C Hausdorff-close enough to S, one has:

$$\|\mathcal{V}_{C,R} * \chi_r(p_0) - \mathcal{V}_{S,R} * \chi_r(p_0)\|_{op} \leqslant O(1 + r^{-1}) d_H(C, S)^{1/2}$$

where r^{-1} is the Lipschitz constant of χ_r. Hence,

$$\|\mathcal{V}_{C,R} * \chi_r(p_0) - C_1 \mathbf{n}(p_0) \otimes \mathbf{n}(p_0)\|_{op} \leqslant O(1 + r^{-1}) d_H(C, S)^{1/2} + O(r)$$

Apply Davis-Kahan theorem to the two matrices inside the operator norm, and to the eigenvector $\mathbf{n}(p_0)$ of the second one. Denoting by $\mathbf{n}(\mathcal{V}_{C,R} * \chi_r(p_0))$ the eigenvector corresponding to the largest eigenvalue of $\mathcal{V}_{C,R} * \chi_r(p_0)$, this gives

$$\|\mathbf{n}(\mathcal{V}_{C,R} * \chi_r(p_0)) - \mathbf{n}(p_0)\| \leqslant O(1 + r^{-1}) d_H(C, S)^{1/2} + O(r)$$

In this equation, one sees that once r has been fixed, it is impossible to go below a $O(r)$ error threshold. With this scale fixed, the speed of convergence is in the square root of the Hausdorff distance.

REMARK. — Very similar arguments prove the convergence and correctness of the normal estimation technique of Alliez *et al.* [ACSTD07], with the same convergence speed.

— This $1/2$-Hölder convergence speed is the same as in the modification of Amenta and Bern's pole-based normal estimation to the noisy case by Dey and Sun [DS05]. It is interesting that this $1/2$ exponent almost always appear in geometric estimation results from point clouds, as soon as the point cloud is allowed to have noise.

III.3 EXPERIMENTAL RESULTS

III.3.1 — Approximate computation by tesselation

In this section, we describe two practical algorithms for computing the VCM of a point cloud. The first method has the advantage of being easy to implement and to be applicable in any ambient dimension. The second algorithm is much faster, and is the one we used for all of our results. We then describe a straightforward way to convolve the VCM.

As remarked earlier, the VCM at a scale R of a point cloud $C = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^3$ is the tensor-valued measure $\mathcal{V}_{C,R}$ concentrated on the point cloud C and such that $\mathcal{V}_{C,R}(\{p_i\})$ is the covariance matrix of the intersection $B_i = B(p_i, R) \cap \mathrm{Vor}(p_i)$, where $\mathrm{Vor}(p_i)$ is the Voronoi cell of p_i.

No simple closed-form expression seems to exist for the covariance matrix of B_i. As remarked in [ABI88], in order to compute the *volume* of the intersection of a Voronoi cell with a ball, one can use the inclusion-exclusion formula to reduce to the case of the intersection of two and three half-planes with the same ball. The same formula can be used to compute the covariance matrix of B_i, reducing it to integrals of the elementary quadratic polynomials $P_{i,j}(x_1, x_2, x_3) = x_i x_j$ over these two type of intersections. However, whereas the integral of the constant function $P(x_1, x_2, x_3) = 1$ admits a (pretty intricate) closed form, there seems to be no such form for the $P_{i,j}$.

Monte-Carlo approximation. The Monte-Carlo algorithm for computing the boundary measures introduced in Chapter II can be easily adapted to compute the VCM of a point cloud (Algorithm 4).

The convergence of the output of this algorithm to a good approximation of the VCM with high probability follows from the same arguments presented. This algorithm has the advantage of being easy to implement, but is too slow in practice for computing the VCM of point clouds with hundreds of thousands of points. In the following, we describe a deterministic approach that can be used to improve the computation speed in low dimensions.

Approximation by tesselation in 3D. We base our second method on the fact that the covariance matrix of a tetrahedron can be computed analytically

Algorithm 4 Monte-Carlo algorithm for VCM

Input: a point cloud C, a scalar R, a number N
Output: an approximation of $\frac{\mathcal{V}_{C,R}}{\mathrm{vol}_d(C^R)} \simeq \frac{1}{N}\sum_{p_i \in C} C_i \delta_{p_i}$
loop {N times}
  [**1.**] Choose a random p_i uniformly in C ;
  [**2.**] Choose a random point X uniformly in $B(p_i, R)$;
  [**3.**] Compute $k = \#(B(X, R) \cap C)$;
  [**4.**] Find the closest point p_j of X in C:
    $C_j \leftarrow C_j + \frac{1}{k}(X - p_j) \otimes (X - p_j)$
end loop

[ACSTD07]. Therefore, using the additivity of the integral, in order to compute the covariance matrix of the intersection of a Voronoi cell with a ball, it is sufficient to approximate it with tetrahedra. For this, we triangulate the boundary of this intersection and build tetrahedra by connecting each of these triangles to the center of the Voronoi cell. This can be done because the intersection is star-shaped with respect to the center of the cell.

We start with an arbitrary triangulation of the boundary of the Voronoi cell $\partial(\mathrm{Vor}(p_i))$. Our goal is subdivide each triangle so that its projection onto the ball $B(p_i, R)$ is a sufficiently good approximation of the corresponding spherical triangle. For this, we process each triangle t in the original triangulation according to three simple rules (see Algorithm 5).

The output of this algorithm yields a tetrahedralization of the intersection of the Voronoi cell with a ball of a given radius, centered at the same point. We then use the closed-form formulas of [ACSTD07] to compute the covariance matrix of this intersection.

Convolution of the VCM. Convolving the VCM of a point cloud using a kernel function χ supported on a ball of radius r can be done by computing for each point $p_i \in C$ the intersection $B(p_i, r) \cap C$, and then summing:

$$\mathcal{V}_{C,R} * \chi(p_i) = \sum_{p_j \in B(p_i, r) \cap C} \chi(p_j - p_i)\mathcal{V}_{C,R}(\{p_j\})$$

The points belonging to $B(p_i, r) \cap C$ are the k nearest neighbor to p_i in C for the suitable value of k, which can be determined by a binary search, using a structure adapted to k-NN queries (such as a kD-tree). As mentioned in §III.5, we mostly use $\chi = \mathbf{1}_{B(0,r)}$.

Implementation. We implemented the two algorithms described above. The tessellation of the Voronoi cells was done using the 3D Delaunay Triangulation package of CGAL [PT06]. Running this algorithm for the Voronoi cell at the origin and 15 random points on the unit sphere, with $R = 1$, yields an average of 120 triangles for tesselating the boundary of $\mathrm{Vor}(p_i) \cap B(p_i, R)$, for a target approximation error of 1%. The running times of this algorithm on more complex point clouds are reported in Table III.1.

Table III.1 – Computation times for VCM of range-scan models (in seconds, on a 3GHz Dual Core CPU)

Model	Delaunay	Tessellation	Total
Blade (195k)	23.73	90.82	114.55
Bimba (502k)	79.04	305.42	384.46
Nicolò (947k)	95.08	465.53	560.61

Algorithm 5 Tessellation of the intersection of a Voronoi cell with a ball

Input: tesselation $\mathcal{T}$ of $\partial\,\mathrm{Vor}(p)$, a scalar R
Output: an approximate tesselation $\mathcal{T}$ of $\partial(\mathrm{Vor}(p) \cap B(p, R))$
for all triangle $t \in \mathcal{T}$ **do**
  if t is completely outside the ball $B(p, R)$ **then**
    1. Recursively subdivide t into a family of smaller triangles $\{t_k\}$, $t_k = (a_k, b_k, c_k)$, until the family of triangles $t'_k = (\pi(a_k), \pi(b_k), \pi(c_k))$ (π being the projection on the ball $B(p, R)$) is a precise enough tessellation of the underlying spherical patch. Then add each t'_k to the final triangulation $\mathcal{T}'$
  else if t is completely inside the ball $B(p, R)$ **then**
    2. Add t to the final triangulation $\mathcal{T}'$.
  else {t crosses the sphere $\partial B(p, R)$}
    Subdivide t by adding points along the circular arc of intersection, and apply **1.** or **2.** to the constructed triangles depending on whether they are inside or outside of the ball.
  end if
end for

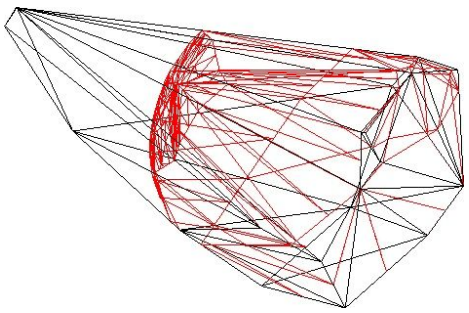


Figure III.6 – The tessellated intersection of a Voronoi cell with a ball.

The convolution step is implemented using the ANN library [AMN+98], which includes a query for finding the set of points of a point cloud contained in a given ball. The time of the convolution step depends on the radius of convolution, but stays within 10 seconds for most models.

Note also that both the tessellation and the convolution steps can be easily parallelized once the Delaunay triangulation and kD-tree have been constructed, since the computations for a given point of the cloud do not involve access to any shared data structures.

III.3.2 — Evaluation on parametric surfaces

In order to understand the relationship between the eigenvectors of the VCM and principal curvature directions, we tested our method on parametric surfaces for which principal curvatures and principal curvature directions can be computed exactly.

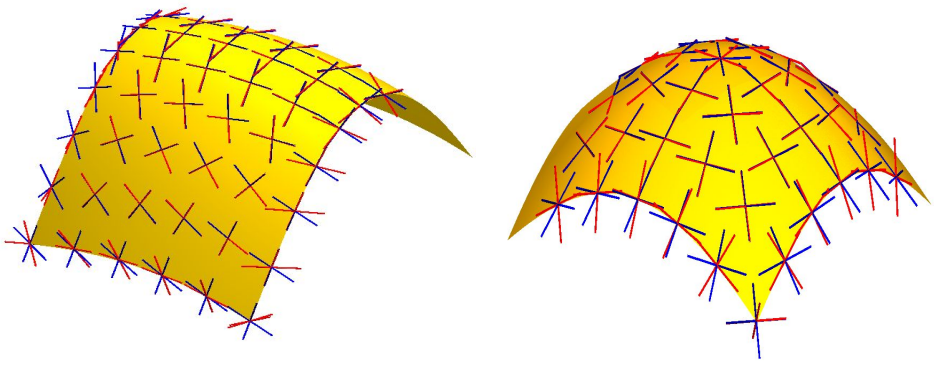
We sampled two functions $z = \sin(3x) + \cos(y)$, where $0 \leqslant x, y \leqslant 1$ and $z = \exp(-x^2) + \exp(-y^2)$, where $-\frac{1}{2} \leqslant x, y \leqslant \frac{1}{2}$. In both examples we used 100,000 points that were chosen uniformly at random within the domain. Figure III.7 shows these two surfaces with the exact and computed principal curvature directions, using $R = 1$ and $r = 0.055$. As expected, away from the boundary the computed and the exact directions match very closely, except possibly at umbilical points (tip of the second surface). Near the boundary of the domain, the principal directions computed using our method follow the edges of the surface, which forms the basis of our feature detection method.

To measure the dependence of our method on the offset and convolution radii, we computed the average deviation (in degrees) of principal directions from exact ones. To minimize the effect of the points close to the boundary, where the directions are not expected to match, we only considered 90 percent of the data-set that is farthest from the boundary of the domain. Figure III.8 shows the average deviation for these two parametric surfaces. As can be seen, the results are quite stable for different choices of the parameters.

III.3.3 — Comparison with polynomial fitting

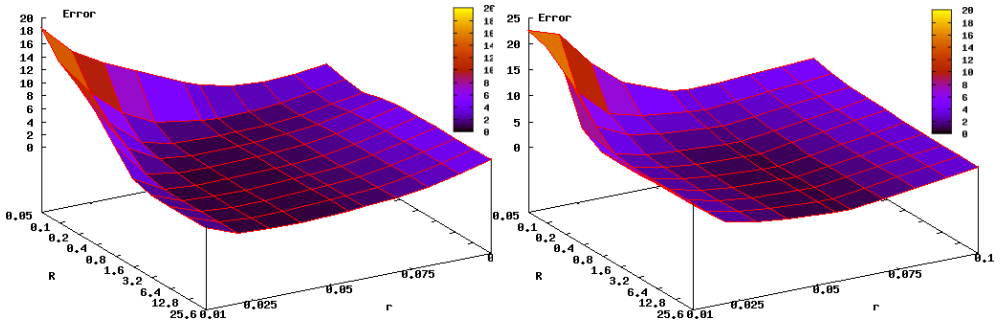
Sampling Conditions. As mentioned in the introduction, the most common method of estimating principal curvatures and principal curvature directions on point clouds is to locally fit a polynomial of a given degree, and to analytically compute principal curvatures of this fitted polynomial [CP05].

Although these methods work well on noiseless and regularly sampled surfaces, polynomial fitting performs poorly if the data has strong noise and sampling bias. Our method, however, is oblivious to the sampling, as a consequence of Hausdorff-robustness proved in Section III.2.3. To illustrate this, we added 50,000 points in a small band along the diagonal (in the parameter space) to the sampling of the surface shown in Figure III.7(a). The results obtained with our method and with the state of the art polynomial jet-fitting algorithm implemented in CGAL [CP05] are reported in Figure III.9. We used second order polynomial fitting with different neighborhood



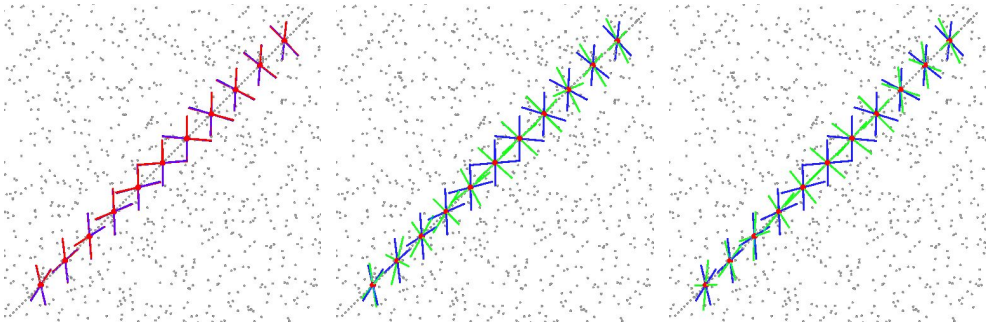
(a) $z = \sin(3x) + \cos(x)$ (b) $z = \exp(-x^2) + \exp(-y^2)$

Figure III.7 – Parametric surfaces with exact (in blue) and computed (in red) principal curvature directions. At the boundary, the computed directions follow the edges of the surface. Note that the tip of the second surface is an umbilical point, and thus the principal curvature directions are not uniquely defined.



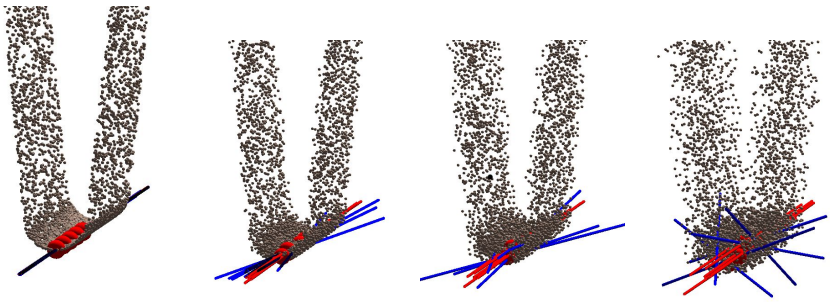
(a) $z = \sin(3x) + \cos(x)$ (b) $z = \exp(-x^2) + \exp(-y^2)$

Figure III.8 – Average deviation (in degrees) of computed principal curvature directions from exact ones for different values of parameters R and r.



(a) Our method (b) Fit with k=100 (c) Fit with k=200

Figure III.9 – Principal curvature directions on a biased dataset. Jet fitting *(b-c)* produces unreliable directions (in green) following oversampled areas.



(a) $r = 0$ (b) $r = 0.01$ (c) $r = 0.04$ (d) $r = 0.08$

Figure III.10 – Principal curvature directions computed with our method (in red) are stable under noise. Directions computed by jet-fitting (in blue), are unreliable, especially when points from separate parts of the shape begin to mix *(d)*.

sizes k, which gave satisfactory results for the original sampling. As can be seen, the extra points do not affect the accuracy of our method. The results obtained with jet-fitting, however, become unreliable and strongly biased in the direction of the oversampling.

Robustness. Other areas that are challenging for polynomial fitting algorithms include parts of the shape with high curvature, and regions where separate parts of the shape come in close contact, thus adversely influencing the quality of the fit. While the first problem can be addressed by fitting higher order polynomials, both of these settings are severely aggravated in the presence of noise. To illustrate this, we sampled a surface made by smoothly joining a small cylinder lying parallel to the z axis, with two planes on either side of the x axis. We used 0.1 as the radius of the cylinder, so the curvature at points along the z axis equals 10. Figure III.10 shows the principal curvature directions obtained on this model with varying levels of noise, using our method and using jet-fitting with 200 nearest neighbors and second order polynomial fitting. Note that when points from different parts of the shape mix, even the robust low order polynomial fitting fails, while our method preserves robustness

III.3.4 — Detection of sharp edges

We evaluated the sharp edge estimation method and its resilience to noise on a unit icosahedron. We sampled 100k points randomly on it, ran the computations with $R = 20$ and various convolution parameters r. We consider a point as a feature if the ratio of the second smallest to the smallest eigenvalue is greater than some threshold parameter. This ratio measures the thickness of the infinitesimal Voronoi cell, as described in Section III.2.2 and on Figure III.4. The results do not seem to be sensitive to this threshold parameter, if it is within the range 10–40; for all the experiments below, we set it to 20.

Resilience to noise. In order to test the resilience of our method to noise, we perturbed the original 100k point cloud on the icosahedron by adding to each point a random vector uniformly chosen in a ball of given radius (the radius of the noise). Figure III.11 shows the estimated feature directions on the icosahedron, for different noise and convolution radii.

We quantify the quality of the approximation using the three following distance measurements between two oriented point clouds $C_1 = (p_i^1, d_i^1)$ and $C_2 = (p_i^2, d_i^2)$:

1. the maximum distance δ_∞ between a point p_i^2 in C_2 and its nearest neighbor in C_1 (this is the half-Hausdorff distance);
2. the average distance δ_{avg} between a point p_i^2 in C_2 and its nearest neighbor in C_1;
3. the average angular deviation (in degrees) α_{avg} between the direction d_i^2 of a point p_i^2 in C_2 and the direction of its nearest neighbor in C_1.

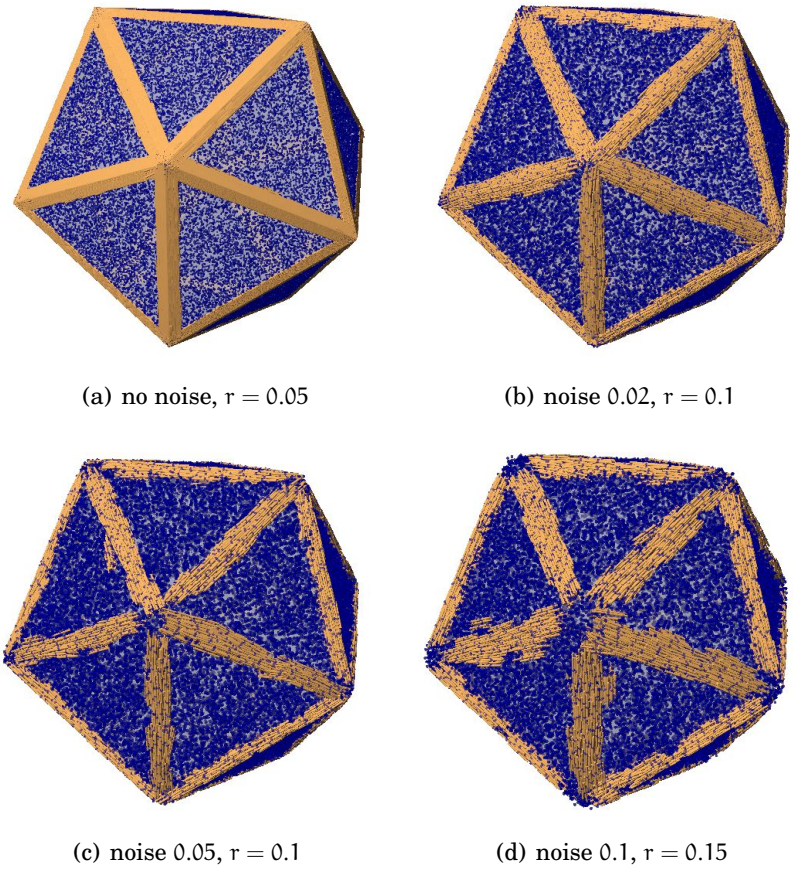
We also consider the symmetric quantities between C_2 and C_1. Table III.2 shows the values of these quantities between the real oriented features C_1 of the icosahedron and the estimated ones C_2, for various noise and convolution radii. The first line of each experiment in the table corresponds to the distances from C_1 to C_2 (as above), while the second line corresponds to the distances from C_2 to C_1.

On the first line of each experiment, δ_∞ measures the presence of isolated outliers, while δ_{avg} measures the spread of the estimated features. One can see that increasing the convolution radius removes isolated outliers, but increases the spread of feature (in fact it seems that $\delta_{avg} \simeq r/2$).

On the second line, δ_∞ and δ_{avg} both evaluate how far every point on an edge of the icosahedron is to an estimated feature. Most of the error here comes from the corners: since we select feature points based on the ratio between the second and third eigenvalues, points nearby corner – at which these two eigenvalues are small – are discarded (see Fig. III.4 and III.11).

Sharp edges with low external angle. In order to understand the effect of sharpness of the edge on the quality of the feature detection, we sampled a surface made of two planar rectangular patches joined by a common edge and whose normals differ by an angle of 2α. As shown in Figure III.12, the feature estimation method described above is able to reliably detect sharp edges whose external dihedral angle is as small as 2°. All the results were produced using the same convolution radius and threshold.

Results on more complex point clouds. In order to further illustrate the robustness of the feature estimation method, we tested it on a 300k point cloud sampled on the standard fandisk model. We then perturbed it by uniform noise of radius 0.01δ, where δ is the radius of the model. The features and feature directions produced by our algorithm are shown in Fig. III.14. While the features are diffused with very strong noise, the feature directions are quite stable and closely follow the edges of the model.



(a) no noise, $r = 0.05$ (b) noise 0.02, $r = 0.1$

(c) noise 0.05, $r = 0.1$ (d) noise 0.1, $r = 0.15$

Figure III.11 – Estimated feature directions on an icosahedron of radius 1, with various convolution radii r and noise values.

Noise	r	δ_∞	δ_{avg}	α_{avg}	δ'_∞	δ'_{avg}	α'_{avg}
0.0	0.05	0.35	0.037	3.25	0.076	0.011	1.42
0.0	0.1	0.118	0.051	0.33	0.124	0.016	1.34
0.02	0.1	0.226	0.049	1.65	0.139	0.020	3.46
0.05	0.1	0.220	0.050	2.82	0.155	0.025	5.45
0.1	0.15	0.271	0.069	3.12	0.178	0.036	7.94

Table III.2 – Distances between the estimated features and real features of an icosahedron, with varying noise radius and convolution radius

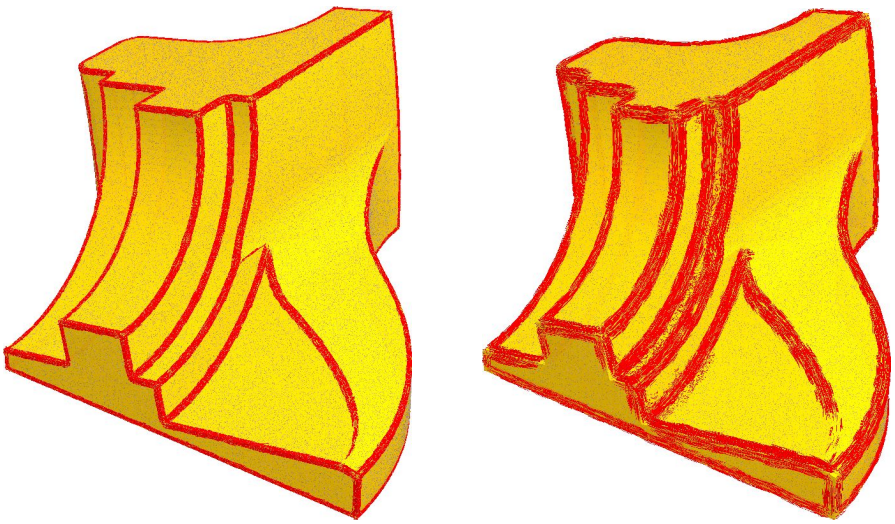
(a) $\alpha = 45°$ (b) $\alpha = 13°$ (c) $\alpha = 1.8°$

Figure III.12 – Estimated feature directions on a folded rectangle, for different values of the external angle α.

On larger range scan point clouds, we were able to decrease both the offset and the convolution radii while keeping the detected features almost noiseless. This enables the algorithm to capture very small and non-sharp features, like the hair of Caesar or the braiding of Bimba, as shown in Fig. III.13 and III.14.

(a) Bimba (b) Bimba

Figure III.13 – Detected features in the Bimba model



(a) Fandisk (b) Fandisk, 0.01 noise

Figure III.14 – Result of the feature detection on the fandisk point cloud with no and 0.01 $\mathrm{diam}(\mathcal{C})$ noise, where $\mathrm{diam}(\mathcal{C})$ is the diameter of the point cloud.

(a) Caesar (b) Nicolo da Uzano

Figure III.15 – Results of the feature detection algorithm on the Julius Caesar *(a)* and Nicolo du Uzano models

Chapter **IV**

GEOMETRIC INFERENCE FOR MEASURES

Abstract

Data often comes in the form of a point cloud sampled from an unknown compact subset of Euclidean space. The general goal of geometric inference is then to recover geometric and topological features (*eg.* Betti numbers, normals) of this subset from the approximating point cloud data. In recent years, it appeared that the study of distance functions allows to address many of these questions successfully. However, one of the main limitations of this framework is that it does not cope well with outliers nor with background noise. In this chapter, we show how to extend the framework of distance functions to overcome this problem. Replacing compact subsets by measures, we introduce a notion of distance function to a probability distribution in $\mathbb{R}^d$. These functions share many properties with classical distance functions, which make them suitable for inference purposes. In particular, by considering appropriate level sets of these distance functions, we show that it is possible to reconstruct offsets of sampled shapes with topological guarantees even in the presence of outliers. Moreover, in settings where empirical measures are considered these functions can be easily evaluated, making them of particular practical interest.

Contents

INTRODUCTION

As we have seen, many geometric inference methods extract information not from the point cloud itself, but from one or several of its offsets, or more generally from its distance function. Stability theoremss then assert that if one takes K and a point cloud C, the information extracted from the distance function to C is close to the one extracted from the distance function to K, provided that the Hausdorff distance between K and C is small. The assumption of Hausdorff closeness is often not met in practice, in particular if C is contaminated with outliers, that is points that are far for the underlying set. Such points badly affect the distance function and all of these stability results break badly, both in theory and practice.

It turns out that most of these inference results rely only on three specific properties of distance functions, two of which are related to the regularity of d_K for a given K:

HAUSDORFF STABILITY: The most important property of the distance function for geometric inference is its Hausdorff stability: $\|d_K - d_{K'}\|_\infty = \sup_{x\in\mathbb{R}^d} |d_K(x) - d_{K'}| \leqslant d_H(K, K')$. This means that if K' is a good Hausdorff approximation of K, then the distance function $d_{K'}$ is close to d_K.

LIPSCHITZ REGULARITY: The distance function is 1-Lipschitz. A consequence of Lipschitz regularity is that the distance function is differentiable almost everywhere (thanks to Rademacher theorem). This means that the set of non-differentiability points of d_K, which is known to coincide with the *medial axis* of K has zero d-volume.

SEMICONCAVITY: The squared distance function d_K^2 is 1-semiconcave. This condition is equivalent to the convexity of the map $x \in \mathbb{R}^d \mapsto \|x\|^2 - d_K^2(x)$. In particular this means, by Alexandrov's theorem [Ale39], that the distance function d_K is not only *almost* $\mathcal{C}^1$, but also twice differentiable almost everywhere. This semiconcavity property is central for the proof of existence of the flow of the distance function in [Pet07], [Lie04, Lemma 5.1]. This flow is one of the main technical tools used in [CCSL09, CCSL08] to prove the stability results. Semiconcavity of the distance function also plays a crucial role in all of the results of Chapters I–III.

A non-negative function φ with the last two properties, and which is *proper* (meaning that $\lim_{\|x\|\mapsto+\infty} |\varphi(x)| = +\infty$) is called *distance-like*. Many Hausdorff-stability results such as those of Chapters II and III can be translated into stability results for uniformly-close distance-like functions.

Dealing with outliers

Unfortunately, offset-based methods do not work well at all in the presence of outliers, since *eg.* the number of connected components will be overestimated if one adds just a single data point far from the original point cloud. The problem here is that while the distance function is only slightly perturbed under Hausdorff noise, adding even a single outlier can change it dramatically.

In order to solve this problem, we have to replace the usual distance function to a set K by another notion of distance function that is robust to the addition of a certain amount of outliers. A way to define what is this *certain amount*, is to change the way point clouds are interpreted: they are no more purely geometric objects, but also carry a notion of *mass*. Formally, we replace compact subsets of $\mathbb{R}^d$ by *finite measures* on the space: a k-manifold will be replaced by the uniform k-dimensional measure on it, a point cloud by a finite sum of Dirac masses, etc. In this setting, the Hausdorff distance is not meaningful any more. The distance between two measures will be measured using the notion of Wasserstein distance, which quantifies the cost of optimally transporting one measure to the other. We refer to §IV.A for more detail. In particular, if S a submanifold of the ambient space, and C is a point cloud uniformly distributed on S, with some noise and a few outliers, then the uniform measure on S and the uniform measure on C are close in the Wasserstein-sense.

Contributions. In order to implement distance-based inference in this setting, we introduce a notion of distance function to a probability measure on the Euclidean space. These distance functions are *distance-like* as defined above; they retain the three properties of the usual distance functions to a compact set that were used in the purely geometric setting. The stability property should of course be understood with respect to Wasserstein distance.

Thanks to this notion of distance function, some of the offset-based inference results mentioned above can be extended to the case where data may be corrupted by outliers. As a proof-of-concept, we generalize the *sampling theory for compact sets* from [CCSL09] to distance-like functions. Using this, we will be able to compare the topology of a compact submanifold of $\mathbb{R}^d$ with the sublevel sets of the distance-to-measure function defined from a finite sample drawn uniformly from S, with some noise and outliers. We also discuss the connections between this distance-based approach and non parametric density estimation as well as mean-shift clustering.

Finally, we study the algorithmic aspects of computing with distance function to measures when the underlying measure is the uniform measure on a finite point cloud. Computing the distance of a given point to such a

measure amounts to a k-nearest neighbor query in the original point cloud, for which plenty of algorithms already exist. However, this does not really help to compute more global properties, such as the topology of the sublevel sets of this distance function. In order to deal with this issue, we approximate the distance function by a power distance to a suitable set of weighted points. We propose two methods for selecting these weighted points. Approximating by a power distance has a big advantage: it very well-known how to work with such a function in a global way, using tools such as weighted Voronoi cells or weighted alpha complexes.

IV.1 DISTANCE FUNCTION TO A PROBABILITY MEASURE

In this section we introduce the notion of distance function to a measure that we consider. As explained in the introduction, there are a few constraints for such a definition to be usable in geometric inference, which we now describe in more detail. Let K be a compact set, and d_K be the distance function to K. Then, one can prove the two following properties:

(i) d_K is 1-***Lipschitz.*** For all x, y in $\mathbb{R}^d$, $|d_K(x) - d_K(y)| \leqslant \|x - y\|$.

(ii) d_K^2 **is** 1-***semiconcave***. This property is equivalent to the concavity of the map $x \in \mathbb{R}^d \mapsto d_K^2(x) - \|x\|^2$.

A consequence of Lipschitz regularity is that the distance function is differentiable almost everywhere; in particular, the *medial axis* of K, defined as the set of non-differentiability points of d_K has zero d-volume. Semiconcavity is a stronger regularity property, as thanks to Alexandrov's theorem it implies that the distance function d_K is not only *almost* $\mathcal{C}^1$, but also twice differentiable almost everywhere. The semiconcavity property plays a central role in the proof of existence of the flow of the gradient of the distance function by [Lie04] (Lemma 5.1), which is the main technical tools used in the topological inference results obtained by [CCSL09]. The semiconcavity of the squared distance function also plays a crucial role in geometric inference results such as [CCSM07] and [MOG09].

This motivates the definition of a *distance-like* function as a non-negative function $\varphi : \mathbb{R}^d \to \mathbb{R}^+$ which is 1-Lipschitz, whose square is 1-semiconcave, and which is *proper* in the sense that $\varphi(x)$ tends to infinity as x does. The following proposition gives a characterization of distance-like functions:

PROPOSITION IV.1. *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a function whose square is 1-semiconcave. There exists a closed subset K of $\mathbb{R}^{d+1}$ such that $\varphi^2(x) = d_K^2(x)$, where a point x in $\mathbb{R}^d$ is identified with the point $(x, 0)$ in $\mathbb{R}^{d+1}$.*

Proof. Let $x \in \mathbb{R}^d$ and v be a subgradient to φ^2 at x, and $v' = v/2$. Define a function ψ_v by $\psi_v(y) = \varphi^2(x) - \|v'\|^2 + \|x - v' - y\|^2$. The 1-semiconcavity of the function φ^2 yields $\psi_v(y) \geqslant \varphi^2(y)$, with equality at $y = x$. Hence, the function φ^2 is the lower envelope of all the functions ψ_v as defined above.

Letting $y = x - v'$, we see that the constant part of ψ_v is positive. Hence, one can define a point z of $\mathbb{R}^{d+1}$, by $(x - v', (\varphi^2(x) - \|v'\|^2)^{1/2})$, such that $\psi_v(x)$ is equal to the squared Euclidean distance between $(x, 0)$ and z in $\mathbb{R}^{d+1}$. Finally, φ^2 is the squared distance to the set $K \subseteq \mathbb{R}^{d+1}$ made of all such points z. □

This proposition proves in particular that a function $\varphi : \mathbb{R}^d \to \mathbb{R}$ whose square is 1-semiconcave and proper is automatically distance-like: the Lipschitz assumption comes with 1-semiconcavity. It also follows from the proof that distance-like functions are simply generalized power distances, with non-positive weights.

IV.1.1 — Definition

The distance function to a compact set K at $x \in \mathbb{R}^d$ is by definition the minimum distance between x and a point of K. Said otherwise, the distance $d_K(x)$ is the minimum radius r such that the ball centered at x of radius r contains at least a point of K. A very natural idea when trying to define the distance function to a probability measure μ on $\mathbb{R}^d$ is to try mimick the definition above. Given a parameter $0 \leqslant m < 1$, define the pseudo-distance $\delta_{\mu,m}$ by

$$\delta_{\mu,m} : x \in \mathbb{R}^d \mapsto \inf\{r > 0\,;\, \mu(\overline{B}(x, r)) > m\}.$$

For instance for $m = 0$, the definition would coincide with the (usual) distance function to the support of the measure μ. For higher values of m, the function $\delta_{\mu,m}$ is 1-Lipschitz, but lacks other features of that a generalization of the usual distance function to a compact should have. For instance, the application that maps a probability measure μ to $\delta_{\mu,m}$ is not continuous in any reasonable sense. Indeed, let δ_x denote the unit Dirac mass at x and $\mu_\varepsilon = (\frac{1}{2} - \varepsilon)\delta_0 + (\frac{1}{2} + \varepsilon)\delta_1$. Then, for $\varepsilon > 0$ one has $\delta_{\mu_\varepsilon,1/2}(t) = |1 - t|$ for $t < 0$ while if $\varepsilon = 0$, one obtains $\delta_{\mu_0,1/2}(t) = |t|$. Said otherwise, the map $\varepsilon \mapsto \delta_{\mu_\varepsilon,1/2}$ is discontinuous at $\varepsilon = 0$.

In order to gain both Wasserstein-stability and regularity, we define the distance function to μ as a L^2 average of the the pseudo-distances $\delta_{\mu,m}$ for a range $[0, m_0]$ of parameters m:

DEFINITION IV.1. Let μ be a (positive) measure on the Euclidean space, and m_0 be a positive mass parameter $m_0 > 0$ smaller than the total mass of μ. We call *distance function to* μ *with parameter* m_0 the function defined by :

$$d^2_{\mu,m_0} : \mathbb{R}^n \to \mathbb{R}^+,\ x \mapsto \frac{1}{m_0}\int_0^{m_0} \delta_{\mu,m}(x)^2 dm$$

As an example, let C be a point cloud with N points in $\mathbb{R}^d$, and μ_C be the uniform measure on it. The pseudo-distance function $\delta_{\mu_C,m}$ evaluated at a point $x \in \mathbb{R}^d$ is by definition equal to the distance between x and its kth nearest neighbor in C, where k is the smallest integer larger than $m|C|$. Hence, the function $m \mapsto \delta_{\mu_C,m}$ is constant on all ranges $(\frac{k}{N}, \frac{k+1}{N}]$. Using this one obtains the following formula for the squared distance d^2_{μ,m_0}, where $m_0 = k_0/|C|$:

$$d^2_{\mu,m_0}(x) = \frac{1}{m_0}\int_0^{m_0} \delta_{\mu,m}(x)^2 = \frac{1}{m_0}\sum_{k=1}^{k_0}\frac{1}{N}\delta_{\mu,k/N}(x)^2$$
$$= \frac{1}{k_0}\sum_{p\in\mathbb{N}_C^{k_0}(x)} \|p-x\|^2$$

where $\mathbb{N}_C^{k_0}(x)$ denote the k_0 nearest neighbors of x in C. In this case, the pointwise evaluation of $d^2_{\mu_C,k/n}(x)$ reduces to a k-nearest neighbor query in C.

IV.1.2 — Equivalent formulation

In this paragraph, we prove that the distance function to a measure d_{μ,m_0} is in fact a real distance to a compact set, but in a infinite-dimensional space. From this fact, we will deduce all of the properties needed for geometric and topological inference.

A measure ν will be called a *submeasure* of another measure μ if for every Borel subset B of $\mathbb{R}^d$, $\nu(B) \leqslant \mu(B)$. This is the same as requiring that $\mu - \nu$ is a measure. The set of all submeasures of a given measure is denoted by $\mathrm{Sub}(\mu)$, while the set of submeasures of μ with a prescribed total mass m_0 is denoted by $\mathrm{Sub}_{m_0}(\mu)$.

PROPOSITION IV.2. *For any measure μ on $\mathbb{R}^d$, the distance function to μ at x is the solution of the following optimal transportation problem:*

$$d_{\mu,m_0}(x) = \min\{m_0^{-1/2} W_2(m_0\delta_x, \nu); \nu \in \mathrm{Sub}_{m_0}(\mu)\} \tag{IV.1}$$

Then, for any measure μ_{x,m_0} that realizes the above minimum one has:

$$d_{\mu,m_0}(x) = \left(\frac{1}{m_0^{1/2}}\int_{\mathbb{R}^d} \|x-h\|^2 \, d\mu_{x,m_0}(h)\right)^{1/2}$$

Said otherwise, the distance d_{μ,m_0} evaluated at a point $x \in \mathbb{R}^d$ is the minimal Wasserstein distance between the Dirac mass $m_0\delta_x$ and the set of submeasures of μ with total mass m_0:

$$d_{\mu,m_0}(x) = \frac{1}{\sqrt{m_0}}\mathrm{dist}_{W_2}(m_0\delta_x, \mathrm{Sub}_{m_0}(\mu)) \tag{IV.2}$$

The set of minimizers in the above expression corresponds to the "orthogonal" projections, or nearest neighbors, of the Dirac mass $m_0\delta_x$ on the set of submeasures $\mathrm{Sub}_{m_0}(\mu)$. As we will see in the proof of the proposition, these are submeasures μ_{x,m_0} of total mass m_0 whose support is contained in the closed ball $\overline{B}(x, \delta_{\mu,m}(x))$, and whose restriction to the *open* ball $B(x, \delta_{\mu,m}(x))$ coincides with μ. Denote these measures by $\mathcal{R}_{\mu,m_0}(x)$.

In order to prove Proposition IV.2, we need a few definitions from probability theory. The *cumulative function* $F_\nu : \mathbb{R}^+ \to \mathbb{R}$ of a measure ν on $\mathbb{R}^+$ is the non-decreasing function defined by $F_\nu(t) = \nu([0,t))$. Its *generalized inverse*, denoted by F_ν^{-1} and defined by $F_\nu^{-1} : m \mapsto \inf\{t \in \mathbb{R}\,;\, F_\nu(t) > m\}$ is left-continuous. Notice that if μ, ν are two measures on $\mathbb{R}^+$, then ν is a submeasure of μ if and only if $F_\nu(t) \leqslant F_\mu(t)$ for all $t > 0$.

Proof. Let first remark that if ν is any measure of total mass m_0, there is only one transport plan between ν and the Dirac mass $m_0\delta_x$, which maps any point of $\mathbb{R}^d$ to x. Hence, the Wasserstein distance between ν and δ_x is given by

$$W_2^2(m_0\delta_x, \nu) = \int_{\mathbb{R}^d} \|h - x\|^2 \, d\nu(h)$$

Let $d_x : \mathbb{R}^d \to \mathbb{R}$ denote the distance function to the point x, and let ν_x be the pushforward of ν by the distance function to x, i.e. for any subset I of $\mathbb{R}$, $\nu_x(I) = \nu(d_x^{-1}(I))$. Using the change-of-variable formula, and the definition of the cumulative function gives us:

$$\int_{\mathbb{R}^d} \|h - x\|^2 \, d\nu(h) = \int_{\mathbb{R}^+} t^2 d\nu_x(t) = \int_0^{m_0} F_{\nu_x}^{-1}(m)^2 dm$$

If ν is a submeasure of μ, then by the remark above, $F_{\nu_x}(t) \leqslant F_{\mu_x}(t)$ for all $t > 0$. From this, one deduces that $F_{\nu_x}^{-1}(m) \geqslant F_{\mu_x}^{-1}(m)$. This gives

$$\begin{aligned} W_2^2(m_0\delta_x, \nu) = \int_{\mathbb{R}^d} \|h - x\|^2 \, d\nu(h) &\geqslant \int_0^{m_0} F_{\mu_x}^{-1}(m)^2 dm \\ &= \int_0^{m_0} \delta_{\mu,m}(x)^2 dm = m_0 d_{\mu,m_0}^2(x) \end{aligned} \tag{IV.3}$$

The second inequality is because $F_{\mu_x}(t) = \mu(B(x,t))$, and thus $F_{\mu_x}^{-1}(m) = \delta_{\mu,m}(x)$. This proves that $d_{\mu,m_0}(x)$ is smaller than the right-hand side of (IV.1).

To conclude the proof, we study the cases of equality in (IV.3). Such a case happens when for almost every $m \leqslant m_0$, $F_{\nu_x}^{-1}(m) = F_{\mu_x}^{-1}(m)$. Since these functions are increasing and left-continuous, equality must in fact hold for every such m. By the definition of the pushforward, this implies that $\nu(\overline{B}(x, \delta_{\mu,m_0}(x))) = m_0$, i.e. all the mass of ν is contained in the closed ball $\overline{B}(x, \delta_{\mu,m_0}(x))$ and $\tilde{\mu}(B(x, \delta_{\mu,m_0}(x))) = \mu(B(x, \delta_{\mu,m_0}(x)))$. Because ν is a submeasure of μ, this can be true if and only iff ν belongs in the set $\mathcal{R}_{\mu,m_0}(x)$ described before the proof.

To finish the proof, we should remark that the set of minimizer $\mathcal{R}_{\mu,m_0}(x)$ always contain a measure μ_{x,m_0}. The only difficulty is when the boundary of the ball carries too much mass. In this case, we uniformly rescale the mass contained in the bounding sphere so that the measure μ_{x,m_0} has total mass m_0. More precisely, we let:

$$\mu_{x,m_0} = \mu|_{B(x,\delta_{\mu,m_0}(x))} + (m_0 - \mu(B(x,\delta_{\mu,m_0}(x)))\frac{\mu|_{\partial B(x,\delta_{\mu,m_0}(x))}}{\mu(\partial B(x,\delta_{\mu,m_0}(x)))}$$

□

IV.1.3 — Stability of the distance function to a measure

The goal of this section is to prove that the notion of distance function to a measure that we defined earlier is stable under change of the measure. This follows rather easily from the characterization of d_{μ,m_0} given by Proposition IV.2.

PROPOSITION IV.3. *Let μ and μ' be two probability measures on $\mathbb{R}^d$. Then,*

$$d_H(\mathrm{Sub}_{m_0}(\mu), \mathrm{Sub}_{m_0}(\mu')) \leqslant W_2(\mu, \mu')$$

Proof. Let ε be the Wasserstein distance of order 2 between μ and μ', and π be a corresponding optimal transport plan, i.e. a transport plan between μ and μ' such that $\int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2 \pi(x,y)dxdy = \varepsilon^2$. Then, given a submeasure ν of μ, one can find a submeasure π' of π that transports ν to a submeasure ν' of μ'. Then,

$$W_2(\nu,\nu')^2 \leqslant \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2 \pi'(x,y)dxdy \leqslant \varepsilon^2$$

This shows that $\mathrm{dist}(\nu, \mathrm{Sub}_{m_0}(\mu')) \leqslant \varepsilon$ for every submeasure $\nu \in \mathrm{Sub}_{m_0}(\mu)$. The same hold by exchanging the roles of μ and μ', thus proving the bound on the Hausdorff distance. □

THEOREM IV.4 ((Distance function stability)). *If μ and μ' are two probability measures on $\mathbb{R}^d$ and $m_0 > 0$, then $\left\|d_{\mu,m_0} - d_{\mu',m_0}\right\|_\infty \leqslant \frac{1}{\sqrt{m_0}} W_2(\mu,\mu')$.*

Proof. The following sequence of equalities and inequalities, that follows from Propositions IV.2 and IV.3, proves the theorem:.

$$\begin{aligned} d_{\mu,m_0}(x) &= \frac{1}{\sqrt{m_0}}\mathrm{dist}_{W_2}(m_0\delta_x, \mathrm{Sub}_{m_0}(\mu)) \\ &\leqslant \frac{1}{\sqrt{m_0}}(d_H(\mathrm{Sub}_{m_0}(\mu), \mathrm{Sub}_{m_0}(\mu')) + \mathrm{dist}_{W_2}(m_0\delta_x, \mathrm{Sub}_{m_0}(\mu'))) \\ &\leqslant \frac{1}{\sqrt{m_0}} W_2(\mu,\mu') + d_{\mu',m_0}(x) \end{aligned}$$

□

IV.1.4 — The distance to a measure is distance-like.

The subdifferential of a function $f : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}$ at a point x, is the set of vectors v of $\mathbb{R}^d$, denoted by $\partial_x f$, such that for all small enough vector h, $f(x+h) \geqslant f(x) + \langle h|v\rangle$. This gives a characterization of convexity: a function $f : \mathbb{R}^d \to \mathbb{R}$ is convex if and only if its subdifferential $\partial_x f$ is non-empty for every point x. If this is the case, then f admits a derivative at a point x if and

only if the subdifferential $\partial_x f$ is a singleton, in which case the gradient $\nabla_x f$ coincides with its unique element.

PROPOSITION IV.5. *The function* $\nu_{\mu,m_0} : x \in \mathbb{R}^d \mapsto \|x\|^2 - d^2_{\mu,m_0}$ *is convex, and its subdifferential at a point* $x \in \mathbb{R}^d$ *is given by*

$$\partial_x \nu_{\mu,m_0} = \left\{ 2x - \frac{2}{m_0} \int_{h\in\mathbb{R}^d} (x-h)\, d\mu_{x,m_0}(h)\, ;\, \tilde{\mu}_{x,m_0} \in \mathcal{R}_{\mu,m_0}(x) \right\}$$

Proof. For any two points x and y of $\mathbb{R}^d$, let μ_{x,m_0} and μ_{y,m_0} be in $R_{\mu,m_0}(x)$ and $R_{\mu,m_0}(y)$ respectively. Thanks to Proposition IV.2 we have the following sequence of equalities and inequalities:

$$\begin{aligned} d^2_{\mu,m_0}(y) &= \frac{1}{m_0} \int_{h\in\mathbb{R}^d} \|y-h\|^2 \, d\mu_{y,m_0}(h) \\ &\leqslant \frac{1}{m_0} \int_{h\in\mathbb{R}^d} \|y-h\|^2 \, d\mu_{x,m_0}(h) \\ &\leqslant \frac{1}{m_0} \int_{h\in\mathbb{R}^d} \|x-h\|^2 + 2\langle x-h|y-x\rangle + \|y-x\|^2 \, d\mu_{x,m_0}(h) \\ &\leqslant d^2_{\mu,m_0}(x) + \|y-x\|^2 + \langle v|y-x\rangle \end{aligned}$$

where v is the vector defined by

$$v = \frac{2}{m_0} \int_{h\in\mathbb{R}^d} [x-h]\, d\mu_{x,m_0}(h).$$

The inequality can be rewritten as:

$$(\|y\|^2 - d^2_{\mu,m_0}(y)) - (\|x\|^2 - d^2_{\mu,m_0}(x)) \geqslant \langle 2x - v|y-x\rangle$$

which shows that the vector $(2x-v)$ belongs to the subdifferential of v at x. By the characterization of convex functions by that we recalled above, one deduces that ν_{μ,m_0} is convex.

We now turn to the proof of the converse inclusion. This proof is slightly more technical, but not really needed for the remaining of the article. First, let

$$\mathcal{D}_{\mu,m_0}(x) := \left\{ 2x - \frac{2}{m_0} \int_{h\in\mathbb{R}^d} (x-h)\, d\mu_{x,m_0}(h)\, ;\, \mu_{x,m_0} \in \mathcal{R}_{\mu,m_0}(x) \right\}.$$

The sets $\mathcal{D}_{\mu,m_0}$ and $\partial_x \nu_{\mu,m_0}$ are both convex, and we have shown that $\mathcal{D}_{\mu,m_0}$ is contained in $\partial_x \nu_{\mu,m_0}$. By Theorem 2.5.1 in [Cla83], the subdifferential $\partial_x \nu_{\mu,m_0}$ can be obtained as the convex hull of the set of limits of gradients $\nabla_{x_n} \nu_{\mu,m_0}$, where (x_n) is any sequence of points converging to x at which ν_{μ,m_0} is differentiable. To sum up, we only need to prove that every such limit also belongs to the set $\mathcal{D}_{\mu,m_0}(x)$. Let (x_n) be a sequence of points at which ν_{μ,m_0} is differentiable, and let μ_n be the unique element in $\mathcal{R}_{\mu,m_0}(x_n)$. Necessarily,

$$\nabla_{x_n} \nu_{\mu,m_0} = 2x_n - 2/m_0 \int_h (x_n - h) d\mu_n(h)$$

where μ_n is in $\mathcal{R}_{\mu,m_0}(x_n)$. Since every μ_n is a submeasure of μ, by compactness one can extract a subsequence of n such that μ_n weakly converges to a measure μ_∞. This measure belongs to $\mathcal{R}_{\mu,m_0}(x)$, and hence the vector

$$D = 2x - 2/m_0 \int_h (x-h) d\mu_\infty(h)$$

is in the set $\mathcal{D}_{\mu,m_0}(x)$. Moreover, the weak convergence of μ_n to μ_∞ implies that the sequence $\nabla_{x_n} v_{\mu,m_0}$ converges to D. This concludes the proof of this inclusion. □

COROLLARY IV.6. *The function d^2_{μ,m_0} is 1-semiconcave. Moreover,*

(i) *d^2_{μ,m_0} is differentiable at a point $x \in \mathbb{R}^d$ if and only if the support of the restriction of μ to the sphere $\partial B(x, \delta_{\mu,m_0}(x))$ contains at most one point;*

(ii) *d^2_{μ,m_0} is differentiable almost everywhere in $\mathbb{R}^d$, with gradient defined by*

$$\nabla_x d^2_{\mu,m_0} = \frac{2}{m_0} \int_{h\in\mathbb{R}^d} [x-h]\, d\mu_{x,m_0}(h)$$

where μ_{x,m_0} is the only measure in $\mathcal{R}_{\mu,m_0}(x)$.

(iii) *the function $x \in \mathbb{R}^d \mapsto d_{\mu,m_0}(x)$ is 1-Lipschitz.*

Proof. For (i), it is enough to remark that $\mathcal{R}_{\mu,m_0}(x)$ is a singleton iff the support of $\mu|_{\partial B(x,\delta_{\mu,m_0}(x))}$ is at most a single point. (ii) This follows from the fact that a convex function is differentiable at almost every point, at which its gradient is the only element of the subdifferential at that point. (iii) The gradient of the distance function d_{μ,m_0} can be written as:

$$\nabla_x d_{\mu,m_0} = \frac{\nabla_x d^2_{\mu,m_0}}{2 d_{\mu,m_0}} = \frac{1}{\sqrt{m_0}} \frac{\int_{h\in\mathbb{R}^d} [x-h]\, d\mu_{x,m_0}(h)}{(\int_{h\in\mathbb{R}^d} \|x-h\|^2\, d\mu_{x,m_0}(h))^{1/2}}$$

Using the Cauchy-Schwartz inequality we find the bound $\|\nabla_x d_{\mu,m_0}\| \leqslant 1$ which proves the statement. □

IV.2 APPLICATIONS TO GEOMETRIC INFERENCE

Reconstruction from point clouds with outliers was the main motivation for introducing the distance function to a measure. In this section, we adapt the reconstruction theorem introduced by [CCSL09] to our setting. The original version of the theorem states that a regular enough compact set K can be faithfully reconstructed from another close enough compact set C. More precisely, for a suitable choice of r, the offsets C^r and K^η have the same homotopy type for any positive η. The regularity assumption on K is expressed as a lower bound on its so-called μ-reach, which is a generalization of the classical notion of reach [Fed59]. In particular, smooth submanifolds, convex sets and polyhedra always have positive μ-reach for suitable μ, hence the reconstruction theorem may be applied to such sets. In these section, we

show that the reconstruction results of [CCSL09] can be easily generalized to compare the sub-level sets of two uniformly-close distance-like functions. It is also possible to adapt most of the topological and geometric inference results of [CCSLT09, CCSL08, CCSM07] in a similar way.

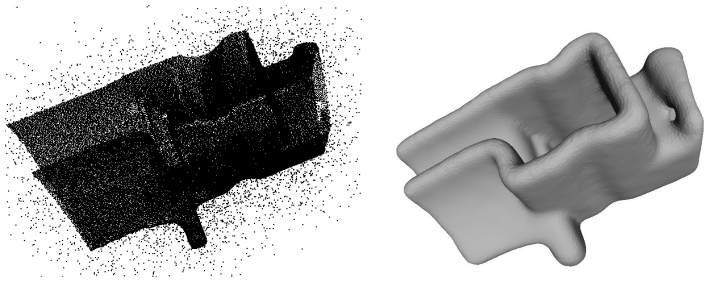


Figure IV.1 – On the left, a point cloud sampled on a mechanical part to which 10% of outliers (uniformly sampled in a box enclosing the model) have been added. On the right, the reconstruction of an isosurface of the distance function d_{μ_C, m_0} to the uniform probability measure on this point cloud.

IV.2.1 — Extending the sampling theory for compact sets

In this paragraph we extend the sampling theory of [CCSL09] for compact sets to distance-like functions. We don't include all of the results of the paper, but only those that are needed to the reconstruction theorem (Th. IV.11). We refer the interested reader to the original paper for more details.

Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a *distance-like* function. The 1-semiconcavity of φ^2 allows to define a notion of gradient vector field $\nabla_x \varphi$ for φ, defined everywhere and satisfying $\|\nabla_x \varphi\| \leqslant 1$. Although not continuous, the vector field $\nabla\varphi$ is sufficiently regular to be integrated in a continuous locally Lipschitz flow [Pet07] $\Phi^t : \mathbb{R}^d \to \mathbb{R}^d$. The flow Φ^t integrates the gradient $\nabla\varphi$ in the sense that for every $x \in \mathbb{R}^d$, the curve $\gamma : t \mapsto \Phi^t(x)$ is right-differentiable, and for every $t > 0$, $\frac{d\gamma}{d\varepsilon}\Big|_{t^-} = \nabla_{\gamma(t)}\varphi$. Moreover, for any integral curve $\gamma : [a, b] \to \mathbb{R}^d$ parametrized by arc-length, one has:

$$\varphi(\gamma(b)) = \varphi(\gamma(a)) + \int_a^b \left\|\nabla_{\gamma(t)}\varphi\right\| dt.$$

DEFINITION IV.2. Let φ be a distance like function. Following the notation for offset of compact sets, we will denote by $\varphi^r = \varphi^{-1}([0, r])$ the r sublevel set of φ.

(i) A point $x \in \mathbb{R}^d$ will be called α-critical (with $\alpha \in [0, 1]$) if the inequality $\varphi^2(x+h) \leqslant \varphi^2(x) + 2\alpha \|h\| \varphi(x) + \|h\|^2$ is true for all $h \in \mathbb{R}^d$. A 0-critical point is simply called a critical point. It follows from the 1-semiconcavity of φ^2 that $\|\nabla_x \varphi\|$ is the infimum of the $\alpha \geqslant 0$ such that x is α-critical.

(ii) The *weak feature size* of φ at r is the minimum $r' > 0$ such that φ doesn't have any critical value between r and $r + r'$. We denote it by $\mathrm{wfs}_\varphi(r)$. For any $0 < \alpha < 1$, the *α-reach* of φ is the maximum r such that $\varphi^{-1}((0, r])$

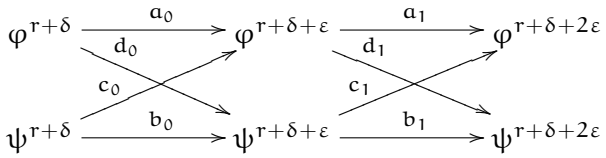
does not contain any α-critical point. Obviously, the α-reach is always a lower bound for the weak-feature size, with $r = 0$.

The proof of the Reconstruction Theorem in [CCSL09] relies on two important observations. The first one is a consequence of a distance-like version of Grove's isotopy lemma [Gro93, Prop. 1.8], which asserts that the topology of the sublevel sets of φ can only change when one passes critical values. As in [CL07, Theorem 3], one deduces that the offsets of two uniformly close distance-like functions with large weak feature size have the same homotopy type:

PROPOSITION IV.7 (Isotopy lemma). *Let φ be a distance-like function and $r_1 < r_2$ be two positive numbers such that φ has no critical points in the subset $\varphi^{-1}([r_1, r_2])$. Then all the sublevel sets $\varphi^{-1}([0, r])$ are isotopic for $r \in [r_1, r_2]$.*

PROPOSITION IV.8. *Let φ and ψ be two distance-like functions, such that $\|\varphi - \psi\|_\infty \leqslant \varepsilon$. Suppose moreover that $\mathrm{wfs}_\varphi(r) > 2\varepsilon$ and $\mathrm{wfs}_\psi(r) > 2\varepsilon$. Then, for every $0 < \eta \leqslant 2\varepsilon$, $\varphi^{r+\eta}$ and $\psi^{r+\eta}$ have the same homotopy type.*

Proof. Let $\delta > 0$ be such that $\mathrm{wfs}_\varphi(r) > 2\varepsilon + \delta$ and $\mathrm{wfs}_\psi(r) > 2\varepsilon + \delta$. Since $\|\varphi - \psi\|_\infty \leqslant \varepsilon$, we have the following commutative diagram where each map is an inclusion.



It follows from the isotopy lemma IV.7 that the inclusions a_0, a_1, b_0 and b_1 are homotopy equivalences. Let s_0, s_1, r_0 and r_1 be homotopic inverses of a_0, a_1, b_0 and b_1 respectively. Now a straightforward computation shows that c_1 is an homotopy equivalence with homotopic inverse $r_1 \circ d_1 \circ s_1$:

$$\begin{aligned} c_1 \circ r_1 \circ d_1 \circ s_1 &\cong c_1 \circ (r_1 \circ b_1) \circ d_0 \circ s_0 \circ s_1 \\ &\cong (c_1 \circ d_0) \circ s_0 \circ s_1 \\ &\cong a_1 \circ a_0 \circ s_0 \circ s_1 \cong \mathrm{id}_{\varphi^{r+\delta+2\varepsilon}} \end{aligned}$$

Similarly, we get $r_1 \circ d_1 \circ s_1 \circ c_1 \cong \mathrm{id}_{\psi^{r+\delta+\varepsilon}}$ proving the proposition IV.8. □

The second key observation made in [CCSL09] is that the critical points of a distance function are stable in certain sense under small Hausdorff perturbations. This result remains true for uniform approximation by distance-like functions:

PROPOSITION IV.9. *Let φ and ψ be two distance-like functions with $\|\varphi - \psi\|_\infty \leqslant \varepsilon$. For any α-critical point x of φ, there exists a α'-critical point x' of ψ with $\|x - x'\| \leqslant 2\sqrt{\varepsilon\varphi(x)}$ and $\alpha' \leqslant \alpha + 2\sqrt{\varepsilon/\varphi(x)}$.*

Proof. The proof is almost verbatim from [CCSL09]. It relies on the following claim: for any α-critical point x of φ and $\rho > 0$, there exists a $\alpha'(\rho)$-critical point x' of ψ with $\alpha'(\rho) \leqslant \alpha + \frac{2\varepsilon}{\rho} + \frac{1}{2}\frac{\rho}{\varphi(x)}$.

Let γ be an integral curve of the flow defined by $\nabla\psi$, starting at x and parametrized by arclength. If γ reaches a critical point of ψ before length ρ, we are done. Assume this is not the case. Then, with $y = \gamma(\rho)$, one has $\psi(y) - \psi(x) = \int_0^\rho \left\|\nabla_{\gamma(t)}\psi\right\| dt$. As a consequence, there exists a point $p(\rho)$ on the integral curve such that $\left\|\nabla_{p(\rho)}\varphi\right\| \leqslant \frac{1}{\rho}(\varphi(y) - \varphi(x))$.

Now, by the assumption on the uniform distance between φ and ψ, $\psi(y) \leqslant \varphi(y) + \varepsilon$ and $\psi(x) \geqslant \varphi(x) - \varepsilon$. Using the fact that x is α-critical, one obtains:

$$\begin{aligned} \varphi(y)^2 &\leqslant \varphi(x)^2 + 2\alpha\|x-y\|\varphi(x) + \|x-y\|^2 \\ \textit{ie.}\quad \varphi(y) &\leqslant \varphi(x)\left(1 + 2\alpha\frac{\|x-y\|}{\varphi(x)} + \frac{\|x-y\|^2}{\varphi(x)^2}\right)^{1/2} \\ &\leqslant \varphi(x) + \alpha\|x-y\| + \frac{1}{2}\frac{\|x-y\|^2}{\varphi(x)} \end{aligned}$$

Putting things together, we get $\left\|\nabla_{p(\rho)}\varphi\right\| \leqslant \alpha + \frac{2\varepsilon}{\rho} + \frac{1}{2}\frac{\rho}{\varphi(x)}$. This proves the claim. The minimum of $\alpha'(\rho)$ is $\alpha + 2\sqrt{\varepsilon/\varphi(x)}$ and is attained for $\rho = 2\sqrt{\varepsilon\varphi(x)}$. This concludes the proof. □

COROLLARY IV.10. *Let φ and ψ be two ε-close distance-like functions, and suppose that* $\mathrm{reach}_\alpha(\varphi) \geqslant R$ *for some $\alpha > 0$. Then, ψ has no critical value in the interval* $\left]4\varepsilon/\alpha^2, R - 3\varepsilon\right[$.

Proof. Suppose the contrary. Then, there exists a critical point x of ψ such that $\psi(x)$ belongs to the range $[4\varepsilon/\alpha^2, R']$. Then, there would exist an α'-critical point y of φ at distance at most D of x. By the previous proposition,

$$\alpha' \leqslant 2\sqrt{\varepsilon/\psi(x)} \leqslant 2\sqrt{\varepsilon/(4\varepsilon/\alpha^2)} = \alpha \text{ and } D \leqslant 2\sqrt{\varepsilon R'}$$

Hence, using the fact that x is a critical point for ψ,

$$\begin{aligned} \varphi(y) \leqslant \psi(y) + \varepsilon &\leqslant \left(\psi^2(x) + \|x-y\|^2\right)^{1/2} + \varepsilon \\ &\leqslant R'\left(1 + D^2/R'^2\right)^{1/2} + \varepsilon \leqslant R' + 3\varepsilon \end{aligned}$$

This last term is less than R if $R' < R - 3\varepsilon$. With these values, one gets the desired contradiction. □

THEOREM IV.11 (Reconstruction). *Let φ, ψ be two ε-close distance-like functions, with* $\mathrm{reach}_\alpha(\varphi) \geqslant R$ *for some positive α. Then, for any $r \in [4\varepsilon/\alpha^2, R - 3\varepsilon]$, and for $0 < \eta < R$, the sublevel sets ψ^r and φ^η are homotopy equivalent, as soon as*

$$\varepsilon \leqslant \frac{R}{5 + 4/\alpha^2}$$

Proof. By the isotopy lemma, all the sublevel sets ψ^r have the same homotopy type, for r in the given range. Let us choose $r = 4\varepsilon/\alpha^2$. We have:

$$\mathrm{wfs}_\varphi(r) \geqslant R - 4\varepsilon/\alpha^2 \text{ and } \mathrm{wfs}_\psi(r) \geqslant R - 3\varepsilon - 4\varepsilon/\alpha^2$$

By Proposition IV.8, the sublevel sets φ^r and ψ^r have the same homotopy type as soon as the uniform distance ε between φ and ψ is smaller than $\frac{1}{2}\,\mathrm{wfs}_\varphi(r)$ and $\frac{1}{2}\,\mathrm{wfs}_\psi(r)$. This is true, provided that $2\varepsilon \leqslant R - \varepsilon(3 + 4/\alpha^2)$. The theorem follows. □

Remark that in the above definition IV.2 the notion of α-reach could be made dependent on a parameter r, i.e. the (r,α)-reach of φ could be defined as the maximum r' such that the set $\varphi^{-1}((r, r + r'])$ does not contain any α-critical value. A reconstruction theorem similar to Theorem IV.11 would still hold under the weaker condition that the (r,α)-reach of φ is positive.

IV.2.2 — Distance to a measure vs. distance to its support

In this paragraph, we compare the distance functions d_{μ,m_0} to a measure μ and the distance function to its support S, and study the convergence properties as the mass parameter m_0 converges to zero. A first obvious remark is that the pseudo-distance δ_{μ,m_0} (and hence the distance d_{μ,m_0}) is always larger than the regular distance function d_S. As a consequence, to obtain a convergence result of d_{μ,m_0} to d_S as m_0 goes to zero, it is necessary to upper bound d_{μ,m_0} by $d_S + o(m_0)$. It turns out that the convergence speed of d_{μ,m_0} to d_S depends on the way the mass of μ contained within any ball $B(p,r)$ centered at a point p of the support decreases with r. Let us define:

(i) We say that a non-decreasing positive function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a *uniform lower bound on the growth of* μ if for every point p in the support of μ and every $\varepsilon > 0$, $\mu(B(p,\varepsilon)) \geqslant f(\varepsilon)$;

(ii) The measure μ has *dimension at most* k if there is a constant $C(\mu)$ such that $f(\varepsilon) = C(\mu)\varepsilon^k$ is a uniform lower bound on the growth of μ, for ε small enough.

LEMMA IV.12. *Let μ be a probability measure and f be a uniform lower bound on the growth of μ. Then $\|d_{\mu,m_0} - d_S\|_\infty < \varepsilon$ as soon as $m_0 < f(\varepsilon)$.*

Proof. Let ε and m_0 be such that $m_0 < f(\varepsilon)$ and let x be a point in $\mathbb{R}^d$, p a projection of x on S, i.e. a point p such that $\|x - p\| = d(p,S)$. By assumption, $\mu(B(x, d_S(x) + \varepsilon)) \geqslant \mu(B(p,\varepsilon)) \geqslant m_0$. Hence, $\delta_{\mu,m_0}(x) \leqslant d_S(x) + \varepsilon$. The function $m \mapsto \delta_{\mu,m}(x)$ being non-decreasing, we get: $m_0 d_S^2(x) \leqslant \int_0^{m_0} \delta_{\mu,m}^2(x)dm \leqslant m_0(d_S(x) + \varepsilon)^2$. Taking the square root of this expression proves the lemma. □

COROLLARY IV.13. *(i) If the support S of μ is compact, then d_S is the uniform limit of d_{μ,m_0} as m_0 converges to 0;*

(ii) If the measure μ has dimension at most $k > 0$, then

$$\|d_{\mu,m_0} - d_S\| \leqslant C(\mu)^{-1/k} m_0^{1/k}$$

Proof. (i) If S is compact, there exists a sequence $x_1, x_2, \cdots$ of points in S such that for any $\varepsilon > 0$, $S \subseteq \cup_{i=1}^n B(x_i, \varepsilon/2)$ for some $n = n(\varepsilon)$. By definition of the

support of a measure, $\eta(\varepsilon) = \min_{i=1\dots n} \mu(B(x_i, \varepsilon/2))$ is positive. Now, for any point $x \in S$, there is a x_i such that $\|x - x_i\| \leqslant \varepsilon/2$. Hence, $B(x_i, \varepsilon/2) \subseteq B(x, \varepsilon)$, which means that $\mu(B(x, \varepsilon)) \geqslant \eta(\varepsilon)$. (ii) Follows straightforwardly from the Lemma. □

For example, the uniform probability measure on a k-dimensional compact submanifold S has dimension at most k. The following proposition gives a more precise convergence speed estimate based on curvature.

PROPOSITION IV.14. *Let S be a smooth k-dimensional submanifold of $\mathbb{R}^d$ whose curvature radii are lower bounded by R, and μ the uniform probability measure on S, then*

$$\|d_S - d_{\mu,m_0}\| \leqslant C(S)^{-1/k} m_0^{1/k}$$

for m_0 small enough and $C(S) = (2/\pi)^k \beta_k / \mathcal{H}^k(S)$ where β_k is the volume of the unit ball in $\mathbb{R}^k$.

Notice in particular that the convergence speed of d_{μ,m_0} to d_S depends only on the *intrinsic* dimension k of the submanifold S, and not on the ambient dimension d. In order to prove this result, we make use of the Günther-Bishop theorem (cf [GHL90, §3.101]).

THEOREM IV.15 (Günther-Bishop). *If the sectional curvatures of a Riemannian manifold M do not exceed δ, then for every $x \in M$, $\mathcal{H}^k(B_M(x, r)) \geqslant \beta_{k,\delta}(r)$ where $\beta_{k,\delta}(r)$ is the volume of a ball of radius r in the simply connected k-dimensional manifold with constant sectional curvature δ, provided that r is smaller than the minimum of the injectivity radius of M and $\pi/\sqrt{\delta}$.*

Proof of Proposition IV.14. Since the intrinsic ball $B_S(x, \varepsilon)$ is always included in the Euclidean ball $B(x, \varepsilon) \cap S$, the mass $\mu(B(x, \varepsilon))$ is always larger than $\mathcal{H}^k(B_S(x, \varepsilon))/\mathcal{H}^k(S)$. Remarking that the sectional curvature of M is upper-bounded by $1/R^2$, Günter-Bishop theorem implies that for any ε smaller than the injectivity radius of S and πR,

$$\mu(B(x, \varepsilon)) \geqslant \frac{\beta_{k,1/R^2}(\varepsilon)}{\mathcal{H}^k(S)}$$

Hence μ has dimension at most k. Moreover, by comparing the volume of an intrinsic ball of the unit sphere and the volume of its orthogonal projection on the tangent space to its center, one has:

$$\beta_{k,1/R^2}(\varepsilon) = R^k \beta_{k,1}(\varepsilon/R) \geqslant R^k [\sin(\varepsilon/R)]^k \beta_k$$

where β_k is the volume of the k-dimensional unit ball. Using $\sin(\alpha) \geqslant \frac{2}{\pi}\alpha$ gives the announced value for $C(S)$. □

IV.2.3 — Shape reconstruction from noisy data

The previous results lead to shape reconstruction theorems from noisy data with outliers. To fit in our framework we consider shapes that are defined as

supports of probability measures. Let μ be a probability measure of dimension at most $k > 0$ with compact support $K \subset \mathbb{R}^d$ and let $d_K : \mathbb{R}^d \to \mathbb{R}_+$ be the (Euclidean) distance function to K. If μ' is another probability measure (*eg.* the empirical measure given by a point cloud sampled according to μ), one has

$$\left\| d_K - d_{\mu', m_0} \right\|_\infty \leqslant \| d_K - d_{\mu, m_0} \|_\infty + \| d_{\mu, m_0} - d_{\mu', m_0} \|_\infty \tag{IV.4}$$

$$\leqslant C(\mu)^{-1/k} m_0^{1/k} + \frac{1}{\sqrt{m_0}} W_2(\mu, \mu') \tag{IV.5}$$

This inequality insuring the closeness of d_{μ', m_0} to the distance function d_K for the sup-norm follows immediately from the stability theorem IV.4 and the corollary IV.13. As expected, the choice of m_0 is a trade-off: small m_0 lead to better approximation of the distance function to the support, while large m_0 make the distance functions to measures more stable. Eq. IV.4 leads to the following corollary of Theorem IV.11:

COROLLARY IV.16. *Let μ be a measure and K its support. Suppose that μ has dimension at most k and that $\mathrm{reach}_\alpha(d_K) \geqslant R$ for some $R > 0$. Let μ' be another measure, and ε be an upper bound on the uniform distance between d_K and d_{μ', m_0}. Then, for any $r \in [4\varepsilon/\alpha^2, R - 3\varepsilon]$, the r-sublevel sets of d_{μ, m_0} and the offsets K^η, for $0 < \eta < R$ are homotopy equivalent, as soon as:*

$$W_2(\mu, \mu') \leqslant \frac{R\sqrt{m_0}}{5 + 4/\alpha^2} - C(\mu)^{-1/k} m_0^{1/k+1/2}$$

Figure IV.1 illustrates the reconstruction Theorem IV.11 on a sampled mechanical part with 10% of outliers. In this case μ' is the normalized sum of the Dirac measures centered on the data points and the (unknown) measure μ is the uniform measure on the mechanical part.

IV.3 RELATION WITH NON-PARAMETRIC DENSITY ESTIMATION

Nearest-neighbor estimators are a family of non-parametric density estimators, that has been extensively used in nonparametric discrimination, pattern recognition and spatial analysis problems. Suppose that C is a point cloud whose points are independently drawn with respect to a probability measure μ on $\mathbb{R}^d$ which admits a densiy f with respect to the Lebesgue measure. Nearest neighbors estimators estimate the density function by:

$$\tilde{f}(x) = \frac{k}{\#C \omega_d(\delta_{C,k}(x))} \tag{IV.6}$$

where $\omega_d(r)$ is the volume of the d-sphere of radius r, and $\delta_{C,k}$ denotes the distance to the kth nearest neighbor in C.

Our claim in this section is that or some of the applications where density estimators are used (such as mean shift, see below), distance to measure

functions could also be used as well. The advantages of the distance function d_{μ,m_0} over nearest-neighbor density estimators are multiple. First of all, they are well defined even when the underlying probability measure does not have a density with respect to the Lebesgue measure, e.g. if it is concentrated on a lower-dimensional subset. Second, the distance to measure d_{μ,m_0} is always uniformly stable with respect to Wasserstein perturbations of the data. Third, because of its 1-semiconcavity, the distance function d_{μ,m_0} is much more regular than the distance δ_{μ,m_0}, as illustrated by Figure IV.2. This has consequences for gradient descent algorithms such as mean shift.

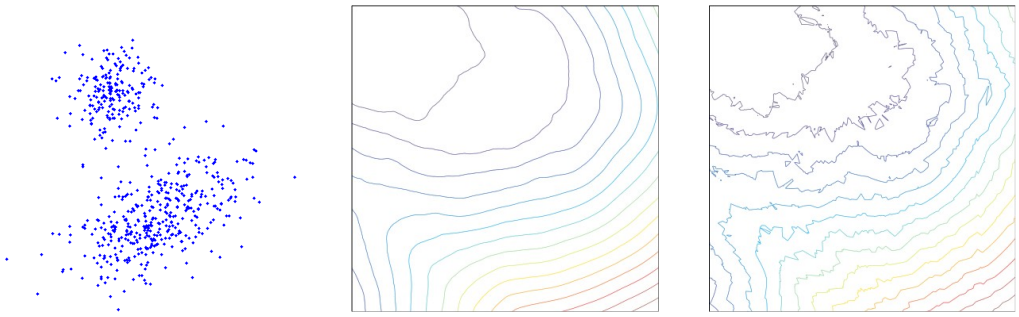


Figure IV.2 – The distance functions to an empirical measure μ associated to a 600 points 2D data set P sampled independently according two gaussians (first figure). The second and third figures represent close-ups on the level sets of the functions $d_{\mu,1/30}$ and $\delta_{\mu,1/30}$ respectively, illustrating the difference of regularity between the two distance functions.

As a consequence of this regularity, it is possible to prove higher order convergence properties of d_{μ_n,m_0} to d_{μ,m_0} as μ_n converges to μ. For example, it can be shown that $\nabla d_{\mu_n,m_0}$ converges to $\nabla d_{\mu,m_0}$ as locally integrable vector fields. Pointwise convergence results can also be obtained at points where $\nabla d_{\mu,m_0}$ is bounded away from 0.

IV.3.1 — Mean-Shift Methods using Distance Functions

Kernel-based mean-shift clustering. Mean-shift clustering [CM02] is a non-parametric clustering method that works on point cloud drawn from an unknown probability measure with density. Specifically, one is given a point cloud $C \subseteq \mathbb{R}^d$ and a radial kernel K. The underlying probability density is estimated by:

$$f(x) = \frac{1}{h^d \#C} \sum_{p \in C} K\left(\frac{p-x}{h}\right)$$

where h is a given bandwidth parameter. Starting from a point x in the space, one iteratively constructs a sequence of points (x_i), descending the gradient

Figure IV.3 – Distance-based mean-shift followed by k-Means clustering on the point cloud made of LUV colors of the pixels of the picture on the right.

of the estimated density:

$$\begin{cases} x_0 = x \\ x_{i+1} = \dfrac{\sum_{p\in C} K\left(\frac{p-x}{h}\right) p}{\sum_{p\in C} K\left(\frac{p-x}{h}\right)} \end{cases} \tag{IV.7}$$

The clustering method works as follows: for each point x_0 in the point cloud, one iterates the sequence x_i until convergence. This defines a mapping from C to the set of critical points of the kernel-based density estimate. A *cluster* of C is simply the set of points of C which correspond to the same critical point under this mapping.

Distance-based mean-shift. We propose a method similar to mean shift, but where the distance function replaces the estimated density. Our iterative scheme is a simple gradient descent for the squared distance function:

$$\begin{cases} x_0 = x \\ x_{i+1} = x_i - \dfrac{1}{2}\nabla_{x_i} d^2_{\mu,m_0} \end{cases} \tag{IV.8}$$

In practice, μ is the uniform probability measure on a point cloud C and $m_0 = k_0/\#C$. In this context, x_{i+1} is simply the isobarycenter of the k_0 nearest neighbor of x_i in C:

$$x_{i+1} = x_i - \frac{1}{k}\sum_{k=1}^{k_0}(x_i - p^k_C(x_i)) = \frac{1}{k}\sum_{k=1}^{k_0} p^k_C(x_i)$$

PROPOSITION IV.17. *Let x be a point in $\mathbb{R}^d$ and $x_t = x - \frac{t}{2}\nabla_{x_i} d^2_{\mu,m_0}$. Then,*

1. $d_{\mu,m_0}(x_t) \leqslant d_{\mu,m_0}(x)$
2. $\langle \nabla_{x_t} d^2_{\mu,m_0} | \nabla_{x_0} d^2_{\mu,m_0} \rangle \geqslant 0$

Proof. 1. This is a simple application of Prop. IV.2.(1).

2. Since d^2_{μ,m_0} is 1-concave,

$$\langle x-y|\nabla_x d^2_{\mu,m_0}(x)-\nabla_x d^2_{\mu,m_0}(y)\rangle \geqslant 2\|x-y\|^2$$

Now, if we set $y = x_t$,

$$\begin{aligned}\langle x-y|\nabla_x d^2_{\mu,m_0}-\nabla_y d^2_{\mu,m_0}\rangle &= \langle x-y|\nabla_x d^2_{\mu,m_0}\rangle - \langle x-y|\nabla_y d^2_{\mu,m_0}\rangle \\ &= \frac{2}{t}\|x-y\|^2 - \frac{t}{2}\langle \nabla_x d^2_{\mu,m_0}|\nabla_y d^2_{\mu,m_0}\rangle\end{aligned}$$

This proves that $\langle \nabla_x d^2_{\mu,m_0}|\nabla_{x_t} d^2_{\mu,m_0}\rangle \geqslant \frac{4}{t}\left(\frac{1}{t}-1\right)\|x-y\|^2$. □

Both properties indicate good convergence properties for our iterative scheme: the first one prevents any infinite loop, while the second shows that trajectories are not too wiggly (more precisely, consecutive edges never make an acute angle). The first property of this proposition had been proved for classical mean-shift, when the kernel K has convex and monotonically decreasing profile (Theorem 1 in [CM02]). On the other hand, the second property is shown to hold for mean-shift when K is Gaussian (Theorem 2 in [CM02]), in which case it is not convex. We are not aware of any choice of kernel such that the resulting mean-shift scheme satisfies these two properties simultaneously.

IV.4 COMPUTING WITH DISTANCE FUNCTIONS

In this §, we are interested with the practical computations using the distance function to the uniform measure on a point cloud P. For simplicity, we will suppose that the parameter m_0 can be written as $k/|P|$. We will call k-distance to P the distance function d_{μ_P,m_0} to the uniform measure on P, with $m_0 = \frac{k}{|P|}$. Equivalently, it is defined by

$$d^2_{P,k}(x) = \frac{1}{k}\min_{c\subseteq P,\#c=k}\|x-c\|^2 = \frac{1}{k}\sum_{p\in NN^k_P(x)}\|x-p\|^2.$$

Here, $NN^k_P(x) \subseteq P$ denote the k nearest neighbors in P of a point $x \in \mathbb{R}^d$. Note that this set is well defined everywhere but on the boundary of the order k Voronoi cells of the point cloud P. If the point x belongs to the boundary of such a cell, then the choice of the nearest neighbors is ambiguous; however the distance function so defined is not.

DEFINITION IV.3. The squared *power distance* to a weighted point $(u, w_u) \in \mathbb{R}^d \times \mathbb{R}$ is defined by $\mathrm{Pow}^2_u(x) = \|x-u\|^2 - w_u$. The squared power distance to a set of weighted points U is defined by $\mathrm{Pow}^2_U(x) = \inf_{u\in U}\mathrm{Pow}^2_u(x)$. In our case, the weights will always be non-positive.

NOTATION. The set of barycenters of k points in P will be denoted by $\mathrm{Bary}^k(P)$. A barycenter of k points in P is called *supporting* if the corresponding order-k Voronoi cell is not empty. The set of supporting k-barycenters is denoted by $\mathrm{Bary}^k_s(P)$.

As we already remarked in §IV.1.1, the (squared) k-distance to a point cloud P is piecewise quadratic on each cell of the kth-order Voronoi diagram. For any subset c of k points in P, define a function $\delta_c : \mathbb{R}^d \to \mathbb{R}$ by $\delta_c^2(x) = \frac{1}{k}\sum_{p\in c}\|x-p\|^2$. Denoting by $\bar{c}$ the barycenter of points in c, we have:

$$\begin{aligned}\delta_c^2(x) &= \frac{1}{k}\sum_{p\in c}\|x-p\|^2 = \frac{1}{k}\sum_{p\in c}\|(x-\bar{c})+(\bar{c}-p)\|^2\\ &= \|x-\bar{c}\|^2 + 2\langle x-\bar{c}|(\bar{c}-\frac{1}{k}\sum_{p\in c}^{k} p)\rangle + \frac{1}{k}\sum_{p\in c}\|\bar{c}-p\|^2\\ &= \|x-\bar{c}\|^2 - w_{\bar{c}}\end{aligned}$$

where the weight $w_{\bar{c}}$ is negative, and given by $-\frac{1}{k}\sum_{p\in c}\|\bar{c}-p\|^2$. From the definition of the k-distance, we are now able to rewrite:

$$d_{P,k}^2(x) = \min_{c\subseteq P,|c|=k}\delta_c^2(x) = \min_{\bar{c}\in \mathrm{Bary}^k(P)}\|x-\bar{c}\|^2 - w_{\bar{c}} \qquad \text{(IV.9)}$$

In other words, the k-distance function to P coincides exactly with the power distance to $\mathrm{Bary}^k(P)$ with the weights defined above. This calculation shows that the order k Voronoi diagram of the original point set is exactly the same as the power diagram of the weighted barycenters. In fact, it is clear that not all of the barycenters actually play a role in the minimum of (IV.9). Equation (IV.9) remains true when replacing the set of barycenters by the set of supporting barycenters.

Approximation of the k-distance. The k-distance to a point cloud P can be interpreted as a power distance to the set of supporting k barycenters. The number of centers involved in this definition is equal to the number of order-k Voronoi cells defined by P. This figure, while considerably smaller than the number of barycenters of any family of k points in the point cloud is still too large for any practical purpose.

However, thanks to the regularity results proved earlier, the geometric properties estimated from a distance function or from a good uniform approximation do not differ much. Likewise, the stability theorem for persistence diagrams [CSEH07] asserts that the persistence diagrams of a function can be reliably estimated from its uniform approximation. These two results mean that it makes sense from both geometric and topological inference viewpoint to try to approximate the k-distance function by a combinatorially simpler function. The challenge is to be able select a small fraction $S \subseteq \mathrm{Bary}^k(P)$ of the supporting barycenters with the property that the power distance to S is a good approximation of the original k-distance function.

Consequences on topology computation. The fact that the k-distance is a power distance makes it possible to compute topological reconstruction of its sublevel sets using weighted alpha-shapes or equivalently Čech complexes. Computing those when the ambient dimension is greater than 3 is impractical.

However, one can use the usual tools defined in computational topology such as Rips complexes or witness complexes in order to achieve approximate computations.

DEFINITION IV.4. The nerve of a family of compact convex sets $\Omega_1, \ldots, \Omega_N \subseteq \mathbb{R}^d$ is an abstract simplicial complex, which contains one vertex for each of the subsets, and contains a k-simplex based on the vertices $x_{i_1}, \ldots, x_{i_k}$ iff the subsets $\Omega_{i_1}, \ldots, \Omega_{i_k}$ have a common point. The nerve theorem asserts that the union of the Ω_N has the same homotopy type as this simplicial complex.

From Eq. IV.9, it follows that the sublevels of the k-distance can be obtained as a finite union of balls:

$$\{x \in \mathbb{R}^d \,;\, d_{P,k}(x) \leqslant \rho\} = \bigcup_{\overline{c} \in \mathrm{Bary}(P)} B\left(\overline{c}, \sqrt{\rho^2 + w_{\overline{c}}}\right) \tag{IV.10}$$

Taking the nerve of this collection of balls gives the notion of k-*distance Čech complex*:

$$\begin{aligned} \check{\mathrm{C}}\mathrm{ech}^k_P(\rho) &= \underset{\overline{c} \in \mathrm{Bary}(P)}{\mathrm{Nerve}} \{B(\overline{c}, \sqrt{\rho^2 + w_{\overline{c}}})\} \\ &\simeq \underset{\overline{c} \in \mathrm{Bary}_s(P)}{\mathrm{Nerve}} \{B(\overline{c}, \sqrt{\rho^2 + w_{\overline{c}}})\}, \end{aligned}$$

where the first definition contains a large redundant subcomplex. By the nerve lemma, both definitions of the k-distance Čech complex are homotopy equivalent to the sublevel sets of $d_{P,k}$. Similarly, if we clip the balls by their power cells $\mathrm{Vor}(\overline{c})$ (recall that these are the order k Voronoi cells of the original point set), we get a k-*distance alpha-shape*, which is a particular case of weighted alpha-shape [Ede92]:

$$\mathrm{AS}^k_P(\rho) = \underset{c \in \overline{c} \in \mathrm{Bary}_w(P)}{\mathrm{Nerve}} \{B\left(\overline{c}, \sqrt{\rho^2 - w_{\overline{c}}}\right) \cap \mathrm{Vor}(\overline{c})\}.$$

Once again, from the Nerve Lemma it follows that $\mathrm{AS}^k_P(\rho)$ is homotopy equivalent to the ρ-sublevel set of the distance function $d_{P,k}$. The definition of the k-distance Čech complex immediately gives rise to the k-distance Vietoris-Rips complex $\mathrm{Rips}^k_P(\rho)$ which is the "Vietoris-Rips closure" (or the clique complex) of the 1-skeleton of $\check{\mathrm{C}}\mathrm{ech}^k_P(\rho)$. Since all the higher dimensional complexes of the Vietoris-Rips complex are defined implicitely from the Rips graph, the complexity (in memory) of the Vietoris-Rips complex is at most $|\mathrm{Bary}_s(P)|^2$. This is a big advantage over Čech or alpha-shape complexes for practical computations. The Rips complex $\mathrm{Rips}^k_P(\rho)$ is generally not homotopy-equivalent to the sublevel set $d^{-1}_{P,k}(\rho)$, but can be used to estimate the homology of the sublevel sets, as in [CO08].

IV.4.1 — Complexity of a distance-like function

NOTATION. Denote by $\mathrm{DL}(\mathbb{R}^d)$ the space of distance-like functions on $\mathbb{R}^d$, *ie.* positive functions whose square is 1-semiconcave. Among these functions,

$\mathrm{DL}^N(\mathbb{R}^d)$ will be the space of functions that can be written as power distances of N points, *ie.* φ is in $\mathrm{DL}^N(\mathbb{R}^d)$ iff there exists points $p_1, \dots, p_N \in \mathbb{R}^d$ and positive weights $w_i, \dots, w_N \in \mathbb{R}^+$ s. t. $\varphi^2(x) = \min_{i\in\{1,\dots,N\}} \|x - p_i\|^2 + w_i$

DEFINITION IV.5. The *complexity at a scale* ε of a distance-like function $\varphi \in \mathrm{DL}(\mathbb{R}^d)$, is the minimum numbers of weighted points needed to define a power distance that is ε-away to it: $\mathrm{Comp}(\varphi, \varepsilon) = \min\left\{N; d(\varphi, \mathrm{DL}^N(\mathbb{R}^d)) \leqslant \varepsilon\right\}$.

REMARKS. — The complexity $\mathrm{Comp}(\varphi, \varepsilon)$ depends continuous on the function φ. More precisely, if φ and ψ are two distance-like functions, with $\|\varphi - \psi\|_\infty \leqslant \eta$, then for any $\varepsilon > \eta$,

$$\mathrm{Comp}(\varphi, \varepsilon - \eta) \leqslant \mathrm{Comp}(\psi, \varepsilon) \leqslant \mathrm{Comp}(\varphi, \varepsilon + \eta)$$

— Suppose that the function φ in the definition above is the distance function d_K to a compact set K. Since the Hausdorff distance between the compact set K and a point cloud C is equal to $\|d_K - d_C\|$, the complexity $\mathrm{Comp}(\varphi, \varepsilon)$ is less than the minimum cardinal of a point cloud C which is ε-Hausdorff close to C. Hence,

$$\mathrm{Comp}(d_K, \varepsilon) \leqslant \mathcal{N}(K, \varepsilon)$$

Relation with convex approximation. There is a nice relation between the complexity of a distance-like function $\varphi : \mathbb{R}^d \to \mathbb{R}$ and the complexity of a convex set in $\mathbb{R}^{d+1}$ that can be defined from it.

DEFINITION IV.6. A weighted point $(u, w_u) \in \mathbb{R}^d \times \mathbb{R}$ can be lifted to $\mathbb{R}^{d+1}$, defining a point $\hat{u} = (u, \|u\|^2 - w_u)$. In our setting, the weights w_u are always negative, and $\hat{u}$ lies above the paraboloid $\{(x, \|x\|^2);\ x \in \mathbb{R}^d\}$. We denote the lifting of a set U of weighted points by $\hat{U}$. The inverse of this lifting map is denoted by $\rho : \mathbb{R}^{d+1} \to \mathbb{R}^d \times \mathbb{R}$.

Denote by $\mathrm{ch}(\hat{U})$ the convex hull of a set of lifted points. The *extended convex hull* of $\hat{U}$, denoted by $\mathrm{ech}(\hat{U})$, is the set of points that are *above* $\mathrm{ch}(\hat{U})$, *ie.* that can be written as $(u, h + \alpha)$ where (u, h) is in the convex hull of $\hat{U}$ and α is any non-negative number. Geometrically, $\mathrm{ech}(\hat{U})$ is the Minkowski sum of the convex hull $\mathrm{ch}(\hat{U})$ with the half line $\{(0, \alpha) \in \mathbb{R}^{d+1};\ \alpha \geqslant 0\}$.

The power distance and the extended convex hull are connected by the following two properties:

PROPOSITION IV.18. *The power distance does not change under extended convex hull of the lifted points. This means that if* U *is a set of weighted points, and* $V = \rho(\mathrm{ech}(\hat{U}))$*, then* $\mathrm{Pow}_U = \mathrm{Pow}_V$.

Proof. If a and b are two weighted points, and $\hat{c}$ is any point on the segment $[\hat{a}, \hat{b}]$ in $\mathbb{R}^{d+1}$, then $\mathrm{Pow}_{\rho(\hat{c})}(x)$ is larger than $\mathrm{Pow}_{\{(a,w_a),(b,w_b)\}}(x)$ for any point $x \in \mathbb{R}^d$. This means that the power distance to the projection of the segment $[\hat{a}, \hat{b}] \subseteq \mathbb{R}^{d+1}$ onto the set of weighted points is equal to the power distance to $\{(a, w_a), (b, w_b)\}$. By induction, and using the definition of the convex hull, one deduces that the power distance Pow_U to the point cloud U coincides with the power distance $\mathrm{Pow}_{\rho(\mathrm{ch}(\hat{U}))}$ to $\rho(\mathrm{ch}(\hat{U}))$.

If (u, w_u) is a weighted point, and $\alpha \geqslant 0$, then $\mathrm{Pow}_{u,w_u+\alpha} \geqslant \mathrm{Pow}_{u,w_u}$. This means that adding the weighted point $(u, w_u + \alpha)$ to U doesn't change anything. This concludes the proof. □

COROLLARY IV.19. *Let U, U' be two sets of weighted point clouds, whose base points are contained in the ball $B(0, R)$, and such that $d_H(\mathrm{ech}(\hat{U}), \mathrm{ech}(\hat{U}')) \leqslant \varepsilon$. Then the power distance they define are uniformly close:*

$$\|\mathrm{Pow}_U - \mathrm{Pow}_{U'}\|_\infty \leqslant \varepsilon + \sqrt{12(2R+1)\varepsilon}$$

Proof. Let (u, w_u) and (v, w_v) be two weighted points. We wish to bound the distance between Pow_u and Pow_v (notice the absence of square):

$$(\|x-u\|^2 - w_u)^{1/2} - (\|x-v\|^2 - w_v)^{1/2} = \frac{\langle 2x - (u+v) | v-u \rangle + w_v - w_u}{(\|x-u\|^2 - w_u)^{1/2} + (\|x-v\|^2 - w_v)^{1/2}}$$

Using the Cauchy-Schwarz inequality, and bounding the denominator from below by $\|x-u\| + \|x-v\|$, we can bound the first term of the sum by $\|u-v\|$. The second term is bounded in the same way, and we obtain:

$$\|\mathrm{Pow}_u - \mathrm{Pow}_v\|_\infty \leqslant \|u-v\| + \left|\sqrt{-w_v} - \sqrt{-w_u}\right|$$

Now, suppose that $\|\hat{u} - \hat{v}\| \leqslant \varepsilon$. Then, $\|u-v\| \leqslant \varepsilon$ and $\left|\|u\|^2 - w_u + \|v\|^2 - w_v\right| \leqslant \varepsilon$. Supposing that u and v are in the ball $B(0, R)$, the second inequality translates as $|w_u - w_v| \leqslant (2R+1)\varepsilon$. Two cases need to be distinguished: if $-w_u \geqslant 2(2R+1)\varepsilon$, then $-w_v \geqslant (2R+1)\varepsilon$, and we have

$$\left|\sqrt{-w_v} - \sqrt{-w_u}\right| = \frac{|w_v - w_u|}{|\sqrt{-w_v} + \sqrt{-w_u}|} \leqslant \frac{(2R+1)\varepsilon}{2\sqrt{(2R+1)\varepsilon}} = 2\sqrt{(2R+1)\varepsilon}$$

In the other case, we can bound $|\sqrt{-w_v} - \sqrt{-w_u}|$ by

$$\sqrt{-w_u} + \sqrt{-w_v} \leqslant \sqrt{2(2R+1)\varepsilon} + \sqrt{3(2R+1)\varepsilon} \leqslant 2\sqrt{3}\sqrt{(2R+1)\varepsilon}$$

□

COROLLARY IV.20. *The complexity $\mathrm{Comp}(\varphi, \varepsilon)$ of a function $\varphi = \mathrm{Pow}_U$ is bounded by the minimum cardinal of a set $V \subseteq \mathbb{R}^{d+1}$ such that the Hausdorff distance between the extended convex hulls $\mathrm{ech}(\hat{U})$ and $\mathrm{ech}(V)$, is at most $O(\varepsilon^2)$.*

IV.4.2 — A random sampling approximation procedure

Following the results of the previous paragraph, we cast the approximation problem for k-distance to a point cloud as a convex approximation problem. The simple Monte-Carlo algorithm for selecting the barycenters that we consider stems from the following observation. In the standard lifting of weighted points onto a paraboloid, the supporting barycenters are the extreme points of the convex hull of all the lifted barycenters. Furthermore, for

any given direction we can find the extreme lifted barycenter by consider the extreme k points of the lifted original points set. In the rest of this section, we study the approximation quality given by the algorithm that draws directions in $\mathbb{R}^{d+1}$ at random and finds the extreme lifted barycenters in those directions.

Let U be the set of weighted barycenters of k points in P, and $\hat{U}$ be their lifts in $\mathbb{R}^{d+1}$. There is an interesting way to interpret the weights that arise in the power distance version of the k-distance: let $C = p_1, \ldots, p_k$ be k points in the original noisy point cloud P, and $(\hat{p}_i)$ be their liftings *onto* the paraboloid, *ie.* $\hat{p}_i = (p_i, \|p_i\|^2)$. The barycenter $\hat{c}$ in $\mathbb{R}^{d+1}$ of these lifted points obviously projects onto the barycenter $\bar{c}$ of the original points. A more interesting fact is that the height (the last coordinate) of $\hat{c}$ is $\|\bar{c}\|^2 - w_{\bar{c}}$, where $w_{\bar{c}}$ is the weight involved in the definition of the k-distance. Geometrically this means that the lifting of the weighted barycenter of the k original points (p_i) coincides with the barycenter of the k lifted points $(\hat{p}_i)$.

Every *extreme point* of the extended convex hull of the lifted weighted barycenters $\mathrm{ech}(\hat{U})$ as defined above can be obtained as the lift of a supporting k-barycenter. Thanks to Corollary IV.19 above, we know that only these points actually play a role in the definition of the k-distance. The same proposition proves that approximating the k-distance in the uniform sense amounts to Hausdorff-approximating the extended convex hull of the lifted weighted barycenters.

Sampling strategy. The main idea behind this sampling strategy is that the importance of an extreme point of the extended convex hull $\mathrm{ech}(\hat{U})$ depends on the sharpness of the convex hull at that point: the bigger the normal cone, the more important the extreme point. A procedure for randomly selecting extreme points with a probability depending on their normal cone is given in Algorithm 6. The analysis of the convergence of this procedure is postponed to the next paragraph.

Given a direction $\mathbf{n}$ in $\mathbb{R}^{d+1}$ whose last coordinate is positive, it is rather easy to find the extreme point of $\mathrm{ech}(\hat{U})$ in that direction. Indeed, any extreme point $\hat{u}$ of $\mathrm{ech}(\hat{U})$ can be written as a barycenter of k points $\hat{p}_1, \ldots, \hat{p}_k$ of the lifting $\hat{P}$ of the original point cloud. In symbols, $\langle \hat{u} | \mathbf{n} \rangle = \frac{1}{k} \langle \hat{p}_1 + \ldots + \hat{p}_k | \mathbf{n} \rangle = \frac{1}{k} \sum_{i=1}^{k} \langle \hat{p}_i | \mathbf{n} \rangle$. The maximum of this scalar product is attained when the points $\hat{p}_1$ to $\hat{p}_k$ are the k most extreme points of the lifted point cloud $\hat{P}$ in the direction $\mathbf{n}$.

IV.4.3 — Normal measure and convex approximation

We analyze the above algorithm by observing that the probability of picking a particular supporting barycenter is equal to the volume of its normal cone. Viewing the algorithm as a way to draw samples from the corresponding normal measure, we relate the quality of convex approximation to the quality of the sample of normal directions.

Algorithm 6 Random Directions Sampling Algorithm

Input: a point cloud $P \subseteq \mathbb{R}^d$, a number N
Construct the lifted point cloud $\hat{P}$ in $\mathbb{R}^{d+1}$.
loop {N times}
  Select a random unit vector $\mathbf{n}$ whose last coordinate is positive
  Select the k extreme points of $\hat{P}$ in the direction given by $\mathbf{n}$.
  Add the weighted barycenter u of these k points to the set of the selected power centers.
end loop

Support function and normal measure. Let us recall a few definitions from convex geometry. If K is a convex subset of $\mathbb{R}^{d+1}$ and x lies on the boundary of K, a *support hyperplane* to x is a hyperplane H such that (i) x is in H, (ii) K is contained entirely in one of the closed half spaces defined by H. The unit normal to H pointing in the opposite direction is called an *outward normal* to K at x. The *normal cone* $\mathcal{N}_x K$ to a point $x \in \partial K$ is the set of outward-pointing unit normals to K at x. The normal cone $\mathcal{N}_B K$ to a subset $B \subseteq K$ is simply the union of the normal cones $\mathcal{N}_x K$, $x \in B$.

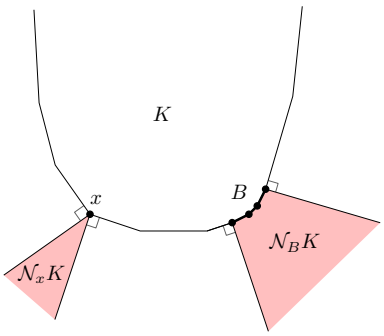


Figure IV.4 – Normal cones of a point x and a subset B of the convex hull K.

The *support function* h_K of the convex set K maps any direction $\mathbf{n}$ in the unit sphere to the distance of the unique support hyperplane of K with normal $\mathbf{n}$ to the origin. It can also be defined by the formula: $h_K : \mathcal{S}^d \to \mathbb{R}, \mathbf{n} \mapsto \max_{x \in K} \langle x | \mathbf{n} \rangle$. It is a well-known fact in convex geometry that the Hausdorff distance between two convex sets is equal to the uniform distance between their support functions: $d_H(K, K') = \|h_K - h_{K'}\|_\infty$.

Using these notions we are able to construct a probability measure ν_K concentrated on the boundary ∂K whose distribution reflects the concentration of normal cones on ∂K. Take ν the uniform probability measure on the unit sphere $\mathcal{S}^{d-1}$, and let ν_K be the probability measure on K defined by $\nu_K(B) = \nu(\mathcal{N}_B K)$. We call this distribution the *normal measure* on K.

EXAMPLE. If K is a polyhedron, ν_K is concentrated on the vertices of K and can thus be written as a sum of Dirac masses. The coefficient of the Dirac mass corresponding to a vertex x is equal to the normalized solid angle of the normal cone to K at x.

We let $\pi_K : \mathbb{S}^d \to \mathbb{R}^{d+1}$ be the function that maps a direction $\mathbf{n}$ to an extreme point $x \in \partial K$ of K with normal $\mathbf{n}$. Equivalently, $\pi_K(\mathbf{n})$ is defined by the property that $\langle \pi_K(\mathbf{n}) | \mathbf{n} \rangle = h_K(\mathbf{n})$. Notice that if K is not strictly convex, several points of K can share a given normal direction $\mathbf{n}$: this means that the function is not uniquely defined everywhere. However, Lemma IV.21 asserts that in all cases the function π_K is uniquely and thus well-defined at almost every direction $\mathbf{n} \in \mathbb{S}^d$. This behavior is similar to the nearest-neighbor function not being defined on the boundary of Voronoi cells but nonetheless being well defined almost everywhere.

LEMMA IV.21. *Let K be a closed convex set contained in the ball $B(0, R) \subseteq \mathbb{R}^{d+1}$. Then,*

(i) h_K is convex, and its subdifferential is $\partial_{\mathbf{n}} h_K = \{x \in K ; \langle x | \mathbf{n} \rangle = h_K(\mathbf{n})\}$;

(ii) h_K is $2R$-Lipschitz.

As a consequence, h_K is differentiable at almost any point $\mathbf{n} \in \mathbb{S}^d$. At these points π_K is well defined and $\nabla_{\mathbf{n}} h_K = \pi_K(\mathbf{n})$.

Proof. The function h_K is convex since it is a supremum of affine functions: $h_K(\mathbf{n}) = \max_{x \in K} \langle x | \mathbf{n} \rangle$. The subdifferential at a point x of a supremum of affine functions is the set of gradients of the affine functions where the maximum is attained, which proves (i).

The proof of (ii) is straightforward: let $\mathbf{n}$ and $\mathbf{m}$ be two vectors in the unit sphere, and x be a point of K where the maximum defining $h_K(\mathbf{n})$ is attained. Then,

$$\begin{aligned} h_K(\mathbf{m}) \geqslant \langle x | \mathbf{m} \rangle = \langle x | \mathbf{m} - \mathbf{n} \rangle + \langle x | \mathbf{n} \rangle &\geqslant - \|x\| \, \|\mathbf{m} - \mathbf{n}\| + h_K(\mathbf{n}) \\ &\geqslant -R \, \|\mathbf{m} - \mathbf{n}\| + h_K(\mathbf{n}) \end{aligned}$$

The Lipschitz regularity follows. To conclude we use the fact that a convex function $f : \mathbb{S}^{d-1} \to \mathbb{R}$ is differentiable at almost every point of $\mathbb{S}^{d-1}$, and that at those points its subdifferential is equal to $\{\nabla_x f\}$. □

Convex approximation and normal measure sampling. We relate the quality of the convex approximation given by the convex hull of the iid samples and the quality of the sampling of the normal measure.

LEMMA IV.22. *Let S be a set of samples on the boundary of K. Denoting by K_S the convex hull of S, and d_S the distance function to S, one has for any $p \geqslant 0$: $\|h_K - h_{K_S}\|_{L^p(\nu)} \leqslant \left(\int_{\mathbb{R}^d} d_S(x)^p \, d\nu_K(x) \right)^{1/p}$.*

Proof. Let us first remark that for any direction $\mathbf{n} \in \mathbb{S}^d$, $h_K(\mathbf{n}) \geqslant h_{K_S}(\mathbf{n}) \geqslant h_K(\mathbf{n}) - d_S(\pi_K(\mathbf{n}))$. The first inequality follows from the inclusion $K_S \subseteq K$. For the second one, let $y \in S$ denote the closest neighbor to $\pi_K(\mathbf{n})$ in S, *ie.* $d_S(y) = \|y - \pi_K(\mathbf{n})\|$. Then,

$$\begin{aligned} h_{K_S}(\mathbf{n}) = \max_{x \in K_S} \langle x | \mathbf{n} \rangle \geqslant \langle y | \mathbf{n} \rangle &= \langle y - \pi_K(\mathbf{n}) | \mathbf{n} \rangle + \langle \pi_K(\mathbf{n}) | \mathbf{n} \rangle \\ &\geqslant h_K(\mathbf{n}) - \|y - \pi_K(\mathbf{n})\| \\ &= h_K(\mathbf{n}) - d_S(\pi_K(\mathbf{n})) \end{aligned}$$

These two inequalities together prove that $\|h_K(\mathbf{n}) - h_{K_S}(\mathbf{n})\| \leqslant d_S(\pi_K(\mathbf{n}))$. By integrating this inequality on the sphere $\mathbb{S}^{d-1}$ one gets:

$$\|h_K - h_{K_S}\|^p_{L^p(\nu)} \leqslant \int_{\mathbb{S}^{d-1}} d_S^p(\pi_K(\mathbf{n}))d\nu(\mathbf{n})$$

Using the fact that ν_K is the pushforward of ν by π_K, and the change-of-variable formula, we have $\int_{\mathbb{S}^{d-1}} d_S^p(\pi_K(\mathbf{n}))d\nu(\mathbf{n}) = \int_{\mathbb{R}^d} d_S^p(x)d\nu_K(x)$. This proves the lemma. □

In the next lemma, we denote by $V_1^k(r)$ the k-volume of a ball of radius r in the k-dimensional unit sphere, and ω_d the volume of the whole unit sphere in $\mathbb{R}^d$.

LEMMA IV.23. *Suppose that* K *is contained in the ball* $B(0, R)$*, and let* $K' \subseteq K$*. Then, for* $d_H(K, K') < \pi$*,*

$$d_H(K, K')^{p+d-1} \frac{1}{2} \frac{\beta_{d-1}}{8^{d-1} R^{d-1} \omega_{d-1}} \leqslant \left\|h_K - h_K'\right\|^p_{L^p(\nu)}$$

Proof. Let ε be the Hausdorff distance between h_K and $h_{K'}$, *ie.* $\varepsilon = \|h_K - h_{K'}\|_\infty$. Since K is contained in K', there is a direction n_0 such that $h_K(n_0) = h_{K'}(\mathbf{n}_0) + \varepsilon$. Both functions h_K and $h_{K'}$ being $2R$-Lipschitz, for $\|\mathbf{n} - \mathbf{n}_0\| \leqslant \frac{\varepsilon}{8R}$, $\|h_K(\mathbf{n}) - h_K(\mathbf{n}_0)\| \leqslant \frac{\varepsilon}{4}$, and likewise for $h_{K'}$.

Hence, $h_K(\mathbf{n})$ is bounded from below by $h_{K'}(\mathbf{n}) + \frac{\varepsilon}{2}$ when $\mathbf{n}$ belongs to the ball $B(\mathbf{n}_0, \varepsilon/8R)$. This proves that $\left\|h_K - h_K'\right\|^p_{L^p(\nu)} \geqslant \frac{\varepsilon^p}{2} \frac{V_1^{d-1}(\varepsilon/8R)}{\omega_{d-1}}$, from which one can deduce the lemma using $V_1^{d-1}(r) \geqslant \beta_{d-1}(r) = \beta_{d-1} r^{d-1}$. □

COROLLARY IV.24. *Given a convex set* K *contained in* $B(0, R) \subseteq \mathbb{R}^d$ *and a finite set of samples* $S \subseteq \partial K$*, the Hausdorff distance between* K *and the convex hull* $\mathrm{ch}(S)$ *can be bounded in term of the Wassertsein distance between the uniform measure* μ_S *on* S *and the distribution of normal cones* ν_K*:*

$$d_H(K, \mathrm{ch}(S)) \leqslant \mathrm{cst}(R) \cdot [W_p(\nu_K, \mu_S)]^{\frac{p}{p+d-1}}$$

REMARK. — *Alas!*, the $1/d$ exponent is not an artifact of the proof. Consider the standard unit d-simplex Δ^d in $\mathbb{R}^{d+1}$. Raise its centroid by ε perpendicularly to the hyperplaine containing Δ^d, and call this point x_ε. Finally, let Δ^d_ε be the convex hull of Δ^d and x_ε. For every d, the Hausdorff distance between Δ^d and Δ^d_ε is exactly ε. However, as a consequence of measure concentration on the sphere, the ν-measure of the normal cone to Δ^d_ε at x_ε decreases exponentially fast to zero.

— On the other hand, using the approach that was used to get the projection stability theorem of Chapter II, one can prove a L^1 stability results in K for the gradient of h_K, *ie.* $\pi_K : \mathbb{S}^{d-1} \to K$. As a consequence, the application that maps a convex set K to the distribution of normal cones ν_K is (locally) $1/2$-Hölder. This result is independent of the ambient dimension.

IV.4.4 — Approximation by witnessed barycenters.

A very natural subset of barycenters that can be chosen to approximate the k-distance is those barycenters that are *witnessed* by the original point cloud P. A barycenter is called witnessed by P if the corresponding k-order Voronoi cell contains a point of P; call $\mathrm{Bary}_w(P)$ the set of witnessed barycenters.

We then define the witnessed k-distance $d^w_{P,k}$ as the power distance obtained from considering witnessed barycenters only. It is clearly unlikely that $d^w_{P,k}$, will be a good approximation of the k-distance regardless of any property of P since generically, the set of witnessed barycenters $\mathrm{Bary}_w(P)$ has the same cardinality as the point cloud P.

Because we consider fewer barycenters, the witnessed k-distance is always greater than the real k-distance. There is a general multiplicative reverse inequality:

LEMMA IV.25. *For any point cloud P and $0 < k < |P|$,*

$$d_{P,k} \leqslant d^w_{P,k} \leqslant (2+\sqrt{2})d_{P,k}$$

Proof. Let $y \in \mathbb{R}^d$ be a point, and $\overline{p}$ the center of (one of) the power cells that contain y. This translatess that $d_{P,k}(y) = d_{\overline{p}}(y)$. In particular, $\|\overline{p} - y\| \leqslant d_{P,k}(y)$ and $\sqrt{-w_{\overline{p}}} \leqslant d_{P,k}(y)$.

Let us find a witnessed barycenter $\overline{q}$ that is close to $\overline{p}$. We know that $\overline{p}$ is the barycenters of k points $x_1, \ldots, x_n$, and that $-w_{\overline{p}} = \frac{1}{k}\sum_{i=1}^{k} \|x_i - \overline{p}\|^2$. Consequently, there should exist an $x \in \{x_i\}$ such that $\|x_i - \overline{p}\| \leqslant \sqrt{-w_{\overline{p}}}$. Let $\overline{q}$ be the barycenter witnessed by x. Then,

$$\begin{aligned} d^w_{P,k}(y) &\leqslant d_{\overline{q}}(y) \\ &\leqslant d_{\overline{q}}(x) + \|x - y\| \\ &\leqslant d_{\overline{p}}(x) + \|x - \overline{p}\| + \|\overline{p} - y\| \end{aligned}$$

Using $d_{\overline{p}}(x) = \left(\|x - \overline{p}\|^2 - w_{\overline{p}}\right)^{1/2} \leqslant \sqrt{2}\sqrt{-w_{\overline{p}}}$ and $\|x - \overline{p}\| \leqslant \sqrt{-w_{\overline{p}}}$, we get

$$\begin{aligned} d^w_{P,k}(y) &\leqslant (1+\sqrt{2})\sqrt{-w_{\overline{p}}} + \|\overline{p} - y\| \\ &\leqslant (2+\sqrt{2})d_{P,k}(y) \end{aligned}$$

□

This Lemma is general and does not exploit any specific property of the input point set P. However, even this coarse estimate can already be used to estimate the Betti numbers of sublevel sets of $d_{P,k}$, using the same approach as in [CO08]. The ith singular homology group of a subset $S \subseteq \mathbb{R}^d$ is denoted by $H_i(S)$; the dimension of this group is the ith Betti number of S, and denoted by $\beta_i(S)$. Also recall that by convention φ^r denotes the r-sublevel of the function φ.

PROPOSITION IV.26. *Suppose that the function $d_{P,k}$ has no critical points in*

the range $[r, \alpha^2 r]$ *(where* $\alpha = 2 + \sqrt{2}$*). Then,*

$$\beta_i(d^r_{P,k}) = \mathrm{rank}(H_i([d^w_{P,k}]^{\alpha r}) \to H_i([d^w_{P,k}]^{\alpha^2 r}))$$

where $H_i([d^w_{P,k}]^{\alpha r}) \to H_i([d^w_{P,k}]^{\alpha^2 r})$ *is the map between homology groups induced by the inclusion between the corresponding sublevel sets.*

Proof. Using Lemma IV.25, one gets the following sequence of inclusions:

$$d^r_{P,k} \subseteq {d^w_{P,k}}^{\alpha r} \subseteq d^{\alpha r}_{P,k} \subseteq {d^w_{P,k}}^{\alpha^2 r} \subseteq d^{\alpha^2 r}_{P,k}$$

Taking the ith singular homology groups of these sublevel sets, we get the sequence:

$$0 \to H_i(d^r_{P,k}) \to H_i([d^w_{P,k}]^{\alpha r}) \to H_i(d^{\alpha r}_{P,k}) \to H_i([d^w_{P,k}]^{\alpha^2 r}) \to H_i(d^{\alpha^2 r}_{P,k}) \to 0$$

If the distance function $d_{P,k}$ doesn't have any critical point between r and $\alpha^2 r$, then the rank of the map $H_i(d^r_{P,k}) \to H_i(d^{\alpha^2 r}_{P,k})$ induced by the inclusion is equal to the kth Betti number of $H_i(d^{r'}_{P,k})$, where r' ranges in $[r, \alpha^2 r]$. In particular, this rank is equal to the dimension of the homology group $H_i(d^{\alpha r}_{P,k})$. Hence, by a standard result of linear algebra, one obtains the result. □

Second bound using the "spread" of a measure. Intuitively, it seems that if P is "concentrated" around a lower dimensional part, the witnessed k-distance function should better approximate the real k-distance than what the bound of Lemma IV.25 gives — at least in a sublevel set $d^{-1}_{P,k}([0, R])$. In this paragraph, we prove such a result. The *spread* at a given scale m_0 is a measure of concentration of a measure μ on $\mathbb{R}^d$. We define it by:

$$\mathrm{spr}_\mu(m_0) = \int_{\mathbb{R}^d} d_{\mu, m_0}(x) d\mu(x).$$

Using the Bishop-Günther theorem, one can prove as in §IV.2.2 that if μ is the uniform measure on a k-dimensional compact submanifold of $\mathbb{R}^d$, then $\mathrm{spr}_\mu(m_0) = O(m_0^{1/k})$. Thus, in this particular case, the spread gives information on the "dimension" of the underlying measure.

This quantity has two other advantages. First and foremost, it can be computed practically when μ is concentrated on a point cloud — the spread is just the sum of k-distance from x to P when x ranges in P. This allows to bound the error made by using the witnessed k-distance instead of the real k-distance, using Proposition IV.28 below. Moreover, thanks to the Wasserstein-stability of the distance to the measure function, the dependence in μ of $\mathrm{spr}^2_\mu(m_0)$ is Lipschitz (the Lipschitz constant is $(1 + m_0^{-1/2})$). This means that if a measure is Wasserstein-close to a measure with low spread, its spread is also low.

The bound of our first version of Proposition IV.28 involved the quantity $\frac{\mathrm{spr}_\mu(m_0)}{m_0}$, a quantity doesn't converge to zero as m_0 does. Moreover, there was a gap between the error measured in practice and the bound we had. In

order to get a tighter bound, we introduce the notion of *local spread* of μ. Its definition depend on a parameter $\alpha > 1$ (this is explained by Lemma IV.27). Recall that $\mu_{x,m}$ denotes the restriction of μ to the ball $B(x, \delta_{\mu,m})$ where the mass on the sphere has been rescaled properly so that $\mu_{x,m}(\overline{B}(x, \delta_{\mu,m})) = m$. Then, setting $m = (1 - \alpha^{-2})m_0$, one let

$$\text{loc. spr}_\mu(m_0, \alpha) := \sup_{x \in \mathbb{R}^d} \left[\frac{1}{m} \int_{\mathbb{R}^d} d_{\mu,m_0}(y) d\mu_{m,x} \right]$$

The reason for the choice of m in the definition of the local spread is the following lemma:

LEMMA IV.27. *Let $\alpha > 0$ and $m = (1 - \alpha^{-2})m_0$. Then, $\delta_{\mu,m}(x) \leqslant \alpha d_{\mu,m_0}(x)$, or equivalently $\mu(B(x, \alpha d_{\mu,m_0}(x))) \geqslant m$.*

Proof. For any point x in $\mathbb{R}^d$, the Chebyshev inequality applied to $y \mapsto \|x - y\|^2$ gives the following bound on $m' = \mu(B(x, \alpha d_{\mu,m_0}(x)))$:

$$\begin{aligned} m_0 d^2_{\mu,m_0}(x) &= \int_{\mathbb{R}^d} \|x - y\|^2 \, d\mu_{x,m_0} \\ &\geqslant \int_{\mathbb{R}^d \setminus B(x, \alpha d_{\mu,m_0}(x))} \alpha^2 d^2_{\mu,m_0}(x) d\mu_{x,m_0} \\ &= (m_0 - m')\alpha^2 d^2_{\mu,m_0}(x) \end{aligned}$$

This means that m' is at least $m_0(1-\alpha^{-2}) = m$, and proves the statement. □

PROPOSITION IV.28. *Let P be a point cloud, μ the uniform measure on P, and $m_0 = k/|P|$. Then, for any $x \in \mathbb{R}^d$ and $\alpha > 1$,*

$$d_{P,k}(x) \leqslant d^w_{P,k}(x) \leqslant \alpha d_{P,k}(x) + \text{loc. spr}_{m_0,\alpha}(\mu) \tag{IV.11}$$

Moreover, on $S := d^{-1}_{P,k}([0, R[)$, one has:

$$\|d_{P,k} - d^w_{P,k}\|_{\infty,S} \leqslant (\alpha - 1)R + \text{loc. spr}_{m_0,\alpha}(\mu) \tag{IV.12}$$

Proof. The first part of this proof works for any probability measure μ on $\mathbb{R}^d$. Observing that the spread upper bounds the integral of d_{μ,m_0} w.r.t the measure $\mu_{x,m}$, we get:

$$\begin{aligned} \text{loc. spr}_\mu(m_0, \alpha) &\geqslant \frac{1}{m} \int_{\mathbb{R}^d} d_{\mu,m_0}(y) d\mu_{x,m}(y) \\ &\geqslant \min \{ d_{\mu,m_0}(y); y \in \text{supp}\, \mu_{x,m} \} \end{aligned}$$

Take the y realizing the above minimum above. It satisfies $\|x - y\| \leqslant \alpha d_{\mu,m_0}(x)$ and $d_{\mu,m_0}(y) \leqslant \text{loc. spr}_\mu(m_0, \alpha)$.

If μ is the uniform measure on P and $m_0 = k/|C|$, since y is in the support of μ (*ie.* P), the k-distance and witnessed k-distance agree at y:

$$d^w_{P,k}(x) \leqslant d^w_{P,k}(y) + \|x - y\| = d_{P,k}(y) + \|x - y\|$$

The inequality follows using the given bounds on $\|x - y\|$ and $d_{P,k}(y)$. □

REMARK. If μ is the uniform measure on a k-dimensional submanifold S of $\mathbb{R}^d$, $\text{loc.spr}_\mu(m_0, \alpha) = O\left(m_0^{1/k}\right)$ and converges to zero as m_0 does — the dependence of this bound in α is being hidden by the uniform convergence of d_{μ,m_0} to zero on S. This makes it reasonable that the bound of Proposition IV.28 is better than the one given in Lemma IV.25 when the uniform measure on the point cloud P is Wasserstein-close to the uniform measure on a k-dimensional submanifold (and the lower the dimension, the better the approximation).

Proposition IV.28 is still probably not the definite answer to the question of the approximation of the k-distance by the witnessed k-distance. This topic still needs further investigation, both theoretical and experimental.

IV.A MEASURES AND WASSERSTEIN DISTANCES

Measure theory A *measure* μ on the space $\mathbb{R}^d$ is a mass distribution. Mathematically, it is defined as a function that maps every (Borel) subset B of $\mathbb{R}^d$ to a non-negative number $\mu(B)$, which is *countably additive* in the sense that whenever (B_i) is a countable family of disjoint Borel subsets of $\mathbb{R}^d$, $\mu(\cup_{i\in\mathbb{N}} B_i) = \sum_i \mu(B_i)$. The *total mass* of a measure is $\mu(\mathbb{R}^d)$. A measure with finite total mass is called *finite*, while a measure with total mass one is a *probability measure*. The *support* of a measure μ is the smallest closed set K on which the mass of μ is concentrated, i.e. $\mu(\mathbb{R}^d \setminus K) = 0$.

Given a set of N points C, the *uniform measure* on C, which we denote by μ_C, can be defined by $\mu_C(B) = \frac{1}{N}|B \cap C|$. More intuitively, it is the sum of N Dirac masses of weight $1/N$, centered at each point of C. When the points in C are chosen randomly and independently according to an underlying, unknown measure, the measure μ_C is called an *empirical measure*. Formally, we are given a family of independent identically distributed random variables $X_1, \dots X_N$ who are distributed according to a common measure μ. The uniform probability measure carried by the point cloud $C_N = \{X_1, \dots, X_n\}$ is known as the *empirical measure*, and simply denoted by μ_N. The uniform law of large numbers asserts that, as N goes to infinity, the empirical measure converges to the underlying measure with probability one — in a sense that will be explained in the next paragraph.

The approach we will describe in this article applies to any measure on Euclidean space. However, to fix ideas, let us describe a family of measures with geometric content that we have in mind when thinking of the underlying measure. One starts from the probability measure μ_M on a compact k-dimensional manifold $M \subseteq \mathbb{R}^d$ given by the rescaled volume form on M, possibly with a non-uniform density. Such measures can be combined, yielding a measure supported on a union of submanifolds of $\mathbb{R}^d$ with various intrinsic dimensions: $\nu = \sum_{i=1}^{\ell} \lambda_i \mu_{M_i}$. Finally, as a simple model of noise, this measure can be convolved with a Gaussian distribution: $\mu = \nu * \mathcal{N}(0, \sigma)$. This is the same as assuming that each sample that is drawn according to ν is known up to an independant Gaussian error term.

The empirical measure defined by the measure μ we just described could then be obtained by repeatedly *(i)* choosing a random integer $i \in \{0, \dots, \ell\}$, *(ii)* picking a random sample X_n uniformly distributed in M_i, *(iii)* adding a random Gaussian vector of variance σ^2 to X_n.

Wasserstein distances. The definition of Wasserstein W_p ($p \geqslant 1$) distance between probability measures rely on the notion of transport plan between measures. It is related to the theory of *optimal transportation* (see e.g. [Vil03]). The Wasserstein distance W_1 is also known as the earth-mover distance, and has been used in vision by [PWR89] and in image retrieval by [RTG00] and others.

A *transport plan* between two probability measures μ and ν on $\mathbb{R}^d$ is a probability measure π on $\mathbb{R}^d \times \mathbb{R}^d$ such that for every $A, B \subseteq \mathbb{R}^d$ $\pi(A \times \mathbb{R}^d) =$

$\mu(A)$ and $\pi(\mathbb{R}^d \times B) = \nu(B)$. Intuitively $\pi(A \times B)$ corresponds to the amount of mass of μ contained in A that will be transported to B by the transport plan. Given $p \geqslant 1$, the p-cost of such a transport plan π is given by

$$\mathcal{C}_p(\pi) = \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^p \, d\pi(x, y) \right)^{1/p}$$

This cost is finite if both measures μ and ν both have finite p-*moments*, i.e. $\int_{\mathbb{R}^d} \|x\|^p \, d\mu(x) < +\infty$ and $\int_{\mathbb{R}^d} \|x\|^p \, d\nu(x) < +\infty$. The set of probability measures on $\mathbb{R}^d$ with finite p-moment includes all probability measures with compact support, such as empirical measures.

DEFINITION IV.7. The *Wasserstein distance* of order p between two probability measures μ and ν on $\mathbb{R}^d$ with finite p-moment is the minimum p-cost $\mathcal{C}_p(\pi)$ of a transport plan π between μ and ν. It is denoted by $W_p(\mu, \nu)$.

As a first example, consider a reference point cloud C with N points, and define a noisy version C' by replacing n points in C by *outliers*, i.e. points o such that $d_C(o) \geqslant R$. The Wasserstein distance between the uniform measures μ_C and μ is at most $\frac{n}{N}(R + \operatorname{diam}(C))$. This can be seen by considering the cost of the transport plan between C' and C that moves the outliers back to their original position, and keeps the other points fixed. On the other hand, the Hausdorff distance between C and C' is at least R. Hence, if the number of outliers is small, i.e. $n \ll N$, the Wasserstein distance is much smaller than the Hausdorff distance.

As mentioned earlier, the question of the convergence of the empirical measure μ_N to the underlying measure μ is fundamendal in the measure-based inference approach we propose. It has been a subject of study in probability and statistics for a long time. If μ is concentrated on a compact set, then μ_N converges almost surely to μ in the W_p distance. More quantitative convergence statement under different assumptions can be given, as in [BGV07].

If $\chi : \mathbb{R}^d \to \mathbb{R}^+$ defines a probability distribution with finite p-moment $\sigma^p := \int_{\mathbb{R}^d} \|x\|^p \chi(x) dx$, the Wasserstein distance of order p between any probability measure μ and the convolved measure $\mu * \chi$ can be bounded by: $W_p(\mu, \mu * \chi) \leqslant \sigma$. If one considers again the example given in the end of §IV.A of an empirical measure μ_N whose samples are drawn according to a "geometric" measure ν convolved with a Gaussian distribution $\mathcal{N}(0, \sigma)$, the combination of the two previous facts gives:

$$\lim_{N \to +\infty} W_2(\mu_N, \mu) \leqslant \sigma \text{ with probability one}$$

Similar bounds are also possible with convolution kernels that are not translation invariant, such as the ones defining the noise model used in [NSW08]. This being said, we would like to stress that the stability results we obtain for the distance functions introduced below do not depend on any noise model; they just depend on the Wasserstein distance between the two probability measures being small.

CONCLUSION

In this thesis, we made some contributions to geometric inference. Instead of recalling them in detail, let us just state a few messages that are worth (in our mind) remembering. In a very general way, we observed that the semi-concavity of the distance function to a compact set can be used to better understand the local properties of the medial axis, and to prove quantitivative stability results for its first order derivatives (such as the projection function). The second message that this thesis tries to convey is that some differential quantities such as curvature or normal cones are better represented by measures rather than in a point-wise way. Trying to approximate them point-wise is at the same time less general, because it forbids sharp features, and leads to more cumbersome stability results under noisy sampling. Moreover, taking this measure-theoretic approach allows to use Monte-Carlo methods for estimation of normal cones, which are very easy to implement (for rapid prototyping) and scale well to higher ambient dimension.

Chapter IV can almost be considered as a second part of the thesis. There are essentially two contributions: first the claim that many distance-based geometry inference results can be applied to compare the geometry of two distance-like functions (*ie.* positive, 1-semiconcave and proprer function) that are uniformly close to each other. Then the construction of a natural notion of distance to a measure that is distance-like and stable under Wasserstein approximation. Putting these two together allow to devise (or rather, adapt) geometric inference methods that are robust to outliers. It is also noticeable that global computations involving the distance to the uniform measure on a point cloud raise interesting questions related to convex approximation, and put us back in the realms of "classical" computational geometry.

There are numerous questions and possible extensions that are left open by this work. Let us mention a few of them, that are of interest to us. The question of reach estimation presented at the end of Chapter II is still completely unsolved. Forgetting the specific approach we suggest, it raises

an interesting question: since claiming that the reach of C is at least R at scale ε means that there is a compact set with reach at least R at distance ε of C. Answering this question requires to better understand how to construct a compact set reach at least R, and more precisely how to construct such a compact set that is Hausdorff-close to C. This "theoretical reconstruction" problem seems interesting: it could take inspiration from existing practical reconstruction methods — especially those that reconstruct an approximating manifold as the level set of a function that minimizes an energy that balances smoothness and attach to the data (called variational methods) — while hopefully shedding a new light on them.

In Chapter II and III, the approximate location of features is extracted from measures through convolution and thresholding. This is theoretically sound and works well in practice; however, it would be nice to find a better way to exploit these measures especially if the goal is to extract an embedded sharp feature graph. A first embedding of such a graph could be provided by using sublevel sets of distance-to-measure to the boundary measure. This raises computability issues. One could then apply techniques similar to active contours in computer vision in order to improve the embedding of the edges, both in precision and smoothness.

More generally, working with sum of Dirac masses instead of point clouds has barely been studied if at all. A problem that would probably benefit from this approach is shape matching: given two point clouds C and D, find (one of) the rigid transformation φ that minimizes the Hausdorff distance between $\varphi(C)$ and D. When C is already close to D, the problem is known as registration, and the standard algorithm known as "iterated closest points" (ICP) is a simple gradient descent. In practice, it often gets stuck at local minima, especially at the later stages of the descent. Using local geometric information such as the one provided by boundary measures could help find good matchings. The matching problem between two measures μ and ν could be stated as: find the rigid transformation that minimizes energy $E(\varphi) = W_2^2(\varphi_\# \mu, \nu)$. This energy seems to have less critical points, and more generally be better behaved than ICP energy. Moreover, the gradient descent could benefit from a hierarchical approach, starting with rough approximations of the measures μ_C and μ_D and refine them as the descent goes on. This might speed the convergence of the descent, and hopefully find better local minima of the energy. Understanding this problem and implementing a solution requires to have quantitative stability results for W_2-optimal transportation as well as being able to compute optimal transportation plans very efficiently.

The last set of questions is related to the computational issues in distance-to-measure simplification presented in §IV.4. We have a bound on the approximation error made by replacing the real k-distance to a point cloud C by the witnessed k-distance, using the "spread" of the uniform measure on C. This bound seems far from optimal in practice. Improving it requires to get a better understanding of what it means for a measure to be concentrated, in a quantitative way, and how this impacts the complexity of the k-distance. A more practical question arise when the size of the input data

is huge — 10M points for laser scans is now very common —, in which case even the witnessed k-distance migh not be a compact enough approximation of the k-distance. In that case, one could try some heuristics to select fewer barycenters among the witnessed barycenters, *eg.* in a greedy way: select the barycenter that improves the approximation the most (on a given domain). Another possibility is to start from a set of barycenters that yield a decent approximation of the k-distance, and then try to optimize their location and weight in order to improve the approximation.

NOTATIONS

BIBLIOGRAPHY

[AAC92] G. Alberti, L. Ambrosio, and P. Cannarsa, *On the singularities of convex functions*, Manuscripta Math. **76** (1992), no. 3, 421–435. 25

[AB99] N. Amenta and M. Bern, *Surface reconstruction by Voronoi filtering*, Discrete and Computational Geometry **22** (1999), no. 4, 481–504. 18, 42, 78

[ABE07] D. Attali, J.D. Boissonnat, and H. Edelsbrunner, *Stability and computation of medial axes-a state-of-the-art report*, Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration. (2007). 18

[ABI88] D. Avis, B. K. Bhattacharya, and H. Imai, *Computing the volume of the union of spheres*, The Visual Computer **3** (1988), no. 6, 323–328. 91

[ACSTD07] P. Alliez, D. Cohen-Steiner, Y. Tong, and M. Desbrun, *Voronoi-based variational reconstruction of unoriented point sets*, Proceedings of the Eurographics Symposium on Geometry Processing, 2007, pp. 39–48. 73, 78, 81, 91, 92

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, 2000. 56, 57

[AK04] N. Amenta and Y.J. Kil, *Defining point-set surfaces*, ACM Transactions on Graphics **23** (2004), no. 3, 264–270. 71

[Alb94] G. Alberti, *On the structure of singular sets of convex functions*, Calculus of Variations and Partial Differential Equations **2** (1994), no. 1, 17–27. 25

[Ale39] A. D. Alexandrov, *Almost everywhere existence of second differentials of convex functions and convex surfaces connected with it*, Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser. **6** (1939), 3–35. 102

[AMCO08] D. Aiger, N. J. Mitra, and D. Cohen-Or, *4-points congruent sets for robust pairwise surface registration*, ACM Transactions on Graphics **27** (2008), no. 3. 70

[AMN+98] S. Arya, D.M. Mount, N.S. Netanyahu, R. Silverman, and A.Y. Wu, *An optimal algorithm for approximate nearest neighbor searching fixed dimensions*, Journal of the ACM (JACM) **45** (1998), no. 6, 891–923. 94

[Ban82] V. Bangert, *Sets with positive reach*, Archiv der Mathematik **38** (1982), no. 1, 54–57. 43

[BC94] J. Berkmann and T. Caelli, *Computation of surface geometry and segmentation using covariance techniques*, IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (1994), no. 11, 1114–1116. 71

[BGV07] F. Bolley, A. Guillin, and C. Villani, *Quantitative concentration inequalities for empirical measures on non-compact spaces*, Probability Theory and Related Fields **137** (2007), no. 3, 541–593. 133

[CCG07] P. Cannarsa, P. Cardaliaguet, and E. Giorgieri, *Hölder regularity of the normal distance with an application to a PDE model for growing sandpiles*, Transactions of the American Mathematical Society **359** (2007), no. 6, 2741. 28

[CCL03] F. Cazals, F. Chazal, and T. Lewiner, *Molecular shape analysis based upon the Morse-Smale complex and the Connolly function*, ACM Symposium on Computational Geometry (2003), 351–360. 73, 76

[CCSL08] F. Chazal, D. Cohen-Steiner, and A. Lieutier, *Normal cone approximation and offset shape isotopy*, Computational Geometry: Theory and Applications (2008). 4, 11, 47, 80, 102, 111

[CCSL09] ———, *A sampling theory for compact sets in Euclidean space*, Discrete and Computational Geometry **41** (2009), no. 3, 461–479. 4, 11, 26, 27, 47, 86, 102, 103, 104, 110, 111, 112

[CCSLT09] F. Chazal, D. Cohen-Steiner, A. Lieutier, and B. Thibert, *Stability of curvature measures*, Computer Graphics Forum **28** (2009), no. 5, 1485–1496. 73, 111

[CCSM07] F. Chazal, D. Cohen-Steiner, and Q. Mérigot, *Stability of boundary measures*, Preprint (2007). 54, 104, 111

[CGOS09] F. Chazal, L.J. Guibas, S.Y. Oudot, and P. Skraba, *Analysis of scalar fields over point cloud data*, Proceedings of the ACM/SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2009, pp. 1021–1030. 18

[CL05] F. Chazal and A. Lieutier, *The λ-medial axis*, Graphical Models **67** (2005), no. 4, 304–331. 21

[CL07] _____, *Stability and computation of topological invariants of solids in $\mathbb{R}^n$*, Discrete and Computational Geometry **37** (2007), no. 4, 601–617. 112

[Cla83] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley New York, 1983. 80, 109

[CM02] D. Comaniciu and P. Meer, *Mean shift: a robust approach toward feature space analysis*, IEEE Trans. Pattern Analysis and Machine Intelligence **24** (2002), no. 5, 603–619. 117, 119

[CO08] F. Chazal and S.Y. Oudot, *Towards persistence-based reconstruction in Euclidean spaces*, ACM Symposium on Computational Geometry (2008), 232–241. 4, 12, 121, 128

[Con86] ML Connolly, *Measurement of protein surface shape by solid angles*, Journal of Molecular graphics **4** (1986), no. 1, 3–6. 73, 76

[CP05] F. Cazals and M. Pouget, *Estimating differential quantities using polynomial fitting of osculating jets*, Computer Aided Geometric Design (2005), 121–146. 72, 94

[CR09] M. Castelpietra and L. Rifford, *Regularity properties of the distance functions to conjugate and cut loci for viscosity solutions of Hamilton-Jacobi equations and applications in Riemannian geometry*, ESAIM: Control, Optimisation and Calculus of Variations (2009). 28

[CRT04] U. Clarenz, M. Rumpf, and A. Telea, *Robust feature detection and local classification for surfaces based on moment analysis*, IEEE Transactions on Visualization and Computer Graphics **10** (2004), no. 5, 1000–9999. 76

[CS04] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Birkhauser, 2004. 25, 43, 80

[CSEH07] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, *Stability of persistence diagrams*, Discrete and Computational Geometry **37** (2007), 103–120. 4, 12, 120

[CSM03] D. Cohen-Steiner and J.M. Morvan, *Restricted Delaunay triangulations and normal cycle*, ACM Symposium on Computational Geometry (2003), 312–321. 73

[DGG03] T.K. Dey, J. Giesen, and S. Goswami, *Shape segmentation and matching with flow discretization*, Lecture Notes in Computer Science **2748** (2003), 25–36. 18

[DS05] T. K. Dey and J. Sun., *Normal and feature estimations from noisy point clouds*, Tech. Report OSU-CISRC-7/50-TR50, Ohio State University, 2005. 78, 91

[Ede92] H. Edelsbrunner, *Weighted alpha shapes*, Report (1992). 121

[Ede95] ———, *The union of balls and its dual shape*, Discrete and Computational Geometry **13** (1995), 415–440. 4, 12

[EH08] H. Edelsbrunner and J. Harer, *Persistent homology -— a survey*, Surveys on discrete and computational geometry, vol. 453, Amer Mathematical Society, 2008, p. 257. 4, 12

[Erd46] P. Erdös, *On the Hausdorff dimension of some sets in Euclidean space*, Bulletin of the American Mathematical Society **52** (1946). 24

[FCOS05] S. Fleishman, D. Cohen-Or, and C. T. Silva, *Robust moving least-squares fitting with sharp features*, ACM Transactions on Graphics **24** (2005), no. 3, 544–552. 72

[Fed59] H. Federer, *Curvature measures*, Transactions of the American Mathematical Society **93** (1959), no. 3, 418–491. 42, 46, 62, 63, 110

[Fed69] ———, *Geometric Measure Theory*, Springer New York, 1969. 22, 62, 63, 64

[Fu85] J.H.G. Fu, *Tubular neighborhoods in Euclidean spaces*, Duke Math. J **52** (1985), no. 4, 1025–1046. 25

[GHL90] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian Geometry*, Springer-Verlag, 1990. 115

[Gro93] K. Grove, *Critical point theory for distance functions*, Proc. of Symposia in Pure Mathematics, vol. 54, 1993. 112

[GTE+06] J. P. Gois, E. Tejada, T. Etiene, L. G. Nonato, A. Castelo, and T. Ertl, *Curvature-driven modeling and rendering of point-based surfaces*, Braz. Symp. Comp. Graph. Imag. Proc, 2006, pp. 27–36. 70

[HdRT+92] H. Hugues, T. de Rose, D. Tom, et al., *Surface reconstruction from unorganized points*, Computer Graphics **26** (1992), no. 2, 71–78. 77

[HHL04] M. Heveling, D. Hug, and G. Last, *Does polynomial parallel volume imply convexity?*, Mathematische Annalen **328** (2004), no. 3, 469–479. 40, 64, 65

[HLW04] D. Hug, G. Last, and W. Weil, *A local Steiner–type formula for general closed sets and applications*, Math. Z. **246** (2004), no. 1, 237–272. 64

[HM02] J. Huang and C. H. Menq, *Combinatorial manifold mesh reconstruction and optimization from unorganized points with arbitrary topology*, Computer Aided Geometric Design **34** (2002), no. 2, 149–165. 71

[HT03] D. Hulin and M. Troyanov, *Mean curvature and asymptotic volume of small balls*, American Mathematical Monthly (2003), 947–950. 76

[IOHS08] J. Daniels II, T. Ochotta, L.K. Ha, and C. T. Silva, *Spline-based feature curves from point-sampled geometry*, The Visual Computer **24** (2008), no. 6, 449–462. 72

[IT01] J. Itoh and M. Tanaka, *The Lipschitz continuity of the distance function to the cut locus*, Transactions of the American Mathematical Society (2001), 21–40. 28

[Jun10] H.W.E. Jung, *Über den kleinsten Kreis, der eine ebene Figur einschließt*, J. Angew. Math. (1910), 310–313. 31

[JWS08] P. Jenke, M. Wand, and W. Straßer, *Patch-graph reconstruction for piecewise smooth surfaces*, Proceedings Vision, Modeling and Visualization, 2008. 70, 72

[Kle80] N. Kleinjohann, *Convexity and the unique footpoint property in Riemannian geometry*, Archiv der Mathematik **35** (1980), no. 1, 574–582. 43

[LCOL07] Y. Lipman, D. Cohen-Or, and D. Levin, *Data-dependent mls for faithful surface approximation*, Proceedings of the Eurographics Symposium on Geometry Processing, 2007, pp. 59–67. 72

[Lie04] A. Lieutier, *Any open bounded subset of $\mathbb{R}^n$ has the same homotopy type as its medial axis*, Computer Aided Geometric Design **36** (2004), no. 11, 1029–1046. 18, 25, 26, 102, 104

[LN05] Y. Li and L. Nirenberg, *The distance function to the boundary, Finsler geometry, and the singular set of viscosity solutions of some Hamilton-Jacobi equations*, Communications on Pure and Applied Mathematics **58** (2005), 0085–0146. 28

[LP05] C. Lange and K. Polthier, *Anisotropic smoothing of point sets*, Computer Aided Geometric Design **22** (2005), no. 7, 680–692. 70, 71

[Mat95] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press Cambridge, 1995. 22, 62

[MD02] J.-L. Maltret and M. Daniel, *Discrete curvatures and applications : a survey.*, Tech. report, Laboratoire des Sciences de l'Information et des Systemes, 2002. 71

[MOG09] Q. Mérigot, M. Ovsjanikov, and L. Guibas, *Robust Voronoi-based Curvature and Feature Estimation*, Proc. SIAM/ACM Joint Conference on Geometric and Physical Modeling, 2009, pp. 1–12. 70, 104

[Mor88] F. Morgan, *Geometric Measure Theory: A Beginner's Guide*, Academic Press, 1988. 22, 55

[MSR07] E. Magid, O. Soldea, and E. Rivlin, *A comparison of Gaussian and mean curvature estimation methods on triangular meshes of range image data*, Computer Vision and Visual Understanding **107** (2007), no. 3, 139–159. 71

[NSW06] P. Niyogi, S. Smale, and S. Weinberger, *Finding the homology of submanifolds with confidence from random samples*, Discrete and Computational Geometry (2006). 4, 11

[NSW08] ———, *A topological view of unsupervised learning from noisy data*, Preprint (2008). 133

[OGG09] C. Oztireli, G. Guennebaud, and M. Gross, *Feature preserving point set surfaces based on non-linear kernel regression*, Eurographics, 2009. 72

[OP85] I.Y. Oleksiv and NI Pesin, *Finiteness of Hausdorff measure of level sets of bounded subsets of Euclidean space*, Mathematical Notes **37** (1985), no. 3, 237–242. 29

[Pet07] A. Petrunin, *Semiconcave functions in Alexandrov's geometry*, Surveys in differential geometry. Vol. XI, Int. Press, Somerville, MA, 2007, pp. 137–201. MR2408266 25, 26, 102, 111

[PKG03] M. Pauly, R. Keiser, and M. Gross, *Multi-scale feature extraction on point-sampled surfaces*, Computer Graphics Forum, vol. 22, Blackwell Synergy, 2003, pp. 281–289. 70

[PMW+08] M. Pauly, N. J. Mitra, J. Wallner, H. Pottmann, and L. J. Guibas, *Discovering structural regularity in 3d geometry*, ACM Transactions on Graphics **27** (2008), no. 3. 70

[PT06] S. Pion and M. Teillaud, *3d triangulations*, CGAL-3.2 User and Reference Manual (CGAL Editorial Board, ed.), 2006. 92

[PWHY09] H. Pottmann, J. Wallner, Q. Huang, and Y.L. Yang, *Integral invariants for robust geometry processing*, Computer Aided Geometric Design **26** (2009), 37–60. 76

[PWR89] S. Peleg, M. Werman, and H. Rom, *A unified approach to the change of resolution: space and gray-level*, IEEE Trans. Pattern Analysis and Machine Intelligence **11** (1989), no. 7, 739–742. 45, 132

[Riv01] A. Rivière, *Dimension de Hausdorff de la Nervure*, Geometriae Dedicata **85** (2001), no. 1, 217–235. 20, 24

[RTG00] Y. Rubner, C. Tomasi, and L.J. Guibas, *The earth mover's distance as a metric for image retrieval*, International Journal of Computer Vision **40** (2000), no. 2, 99–121. 45, 132

[SHB07] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, Thomson-Engineering, 2007. 18

[SS90] G.W. Stewart and J. Sun, *Matrix Perturbation Theory*, Computer Science and Scientific Computing, 1990. 75

[Ste63] S. B. Stečkin, *Approximation properties of sets in normed linear spaces*, Rev. Roumaine Math. Pures Appl **8** (1963), no. 1, 5–18. 20

[Tan05] X. Tang, *A sampling framework for accurate curvature estimation in discrete surfaces*, IEEE Transactions on Visualization and Computer Graphics **11** (2005), no. 5, 573–583. 71

[TT05] W.-S. Tong and C.-K. Tang, *Robust estimation of adaptive tensors of curvature by tensor voting*, IEEE Trans. Pattern Analysis and Machine Intelligence **27** (2005), no. 3, 434–449. 71

[Vil03] C. Villani, *Topics in Optimal Transportation*, American Mathematical Society, 2003. 45, 132

[Wey39] H. Weyl, *On the volume of tubes*, American Journal of Mathematics **61** (1939), no. 2, 461–472. 41

[YQ07] P. Yang and X. Qian, *Direct computing of surface curvatures for point-set surfaces*, Proceedings of the IEEE/Eurographics Symposium on Point-based Graphics, 2007. 71

[Zam04] T. Zamfirescu, *On the cut locus in Alexandrov spaces and applications to convex surfaces*, Pacific J. Math. **217** (2004), 375–386. 20, 21

Détection de structure géométrique dans les nuages de points

Quentin MÉRIGOT

Résumé

Cette thèse s'inscrit dans la problématique générale de l'inférence géométrique. Étant donné un objet qu'on ne connaît qu'à travers un échantillon fini, à partir de quelle qualité d'échantillonage peut-on estimer de manière fiable certaines de ses propriétés géométriques ou topologique?

L'estimation de la topologie est maintenant un domaine assez mûr. La plupart des méthodes existantes sont fondées sur la notion de fonction distance. Nous utilisons cette approche pour estimer certaines notions de courbure dues à Federer, définies pour une classe assez générale d'objets non lisses. Nous introduisons une version approchée de ces courbures dont nous étudions la stabilité ainsi que calcul pratique dans le cas discret. Une version anisotrope de ces mesures de courbure permet en pratique d'estimer le lieu et la direction des arêtes vives d'une surface lisse par morceaux échantillonnée par un nuage de point. En chemin nous sommes amenés à étudier certaines propriétés de régularité de la fonction distance, comme le volume de l'axe médian.

Un défaut des méthodes qui utilisent la fonction distance est leur extrême sensibilité aux points aberrants. Pour résoudre ce problème, nous sortons du cadre purement géométrique en remplaçant les compacts par des mesures de probabilité. Nous introduisons une notion de fonction distance à une mesure, robuste aux perturbations Wasserstein (et donc aux points aberrants) et qui partage certaines propriétés de régularité et de stabilité avec la fonction distance usuelle. Grâce à ces propriétés, il est possible d'étendre de nombreux théorèmes d'inférence géométrique à ce cadre.

Abstract

This thesis deals with the general question of geometric inference. Given an object that is only known through finite sampling, what conditions are required on the sampling in order to be able to estimate correctly some of its topological or geometric properties ?

Topological estimation is by now quite well understood. Most existing approaches rely on the notion of distance function. We use the distance function in order to estimate a notion of curvature due to Federer, that is defined for a rather general class of non-smooth objects. We study the stability of an approximate version of these measures when the unknown object is replaced by a discrete approximation; we also deal with the practical computation of these measures in the discrete setting. An anisotropic notion of these curvature measures can be used to robustly estimate the locus and the direction of sharp edges of a piecewise smooth surface from a point-cloud sampling. Theoretical results required to study some regularity properties of the distance function, such as the volume of the medial axis.

A drawback of distance-based methods is their extreme sensibility to outliers. In order to overcome this problem we propose to leave the purely geometric setting, and replace compact sets with measures (as in Lebesgue theory). We introduce a notion of distance function to a measure, which is robust to Wasserstein perturbations – hence in particular to the addition of outliers. This distance function shares any regularity and stability properties with the usual distance function; this allows to extend many existing geometric inference theorems to this new setting.