Algorithmic Foundations of Geometric Modeling in Higher Dimensions (GeHDi)

Jean-Daniel Boissonnat

January 19, 2012

Abstract

During the past decade, exceptional progress was made with Geometric Modeling and 3D data processing. The new field of Geometry Processing emerged, theoretical foundations together with very efficient practical solutions were obtained for ubiquitous/emblematic problems like surface meshing and surface reconstruction. Extending these techniques to higher dimensions would have tremendous applications in science and engineering but is currently extremely limited and asks for new algorithmic breakthrough. This project aims at settling the algorithmic foundations of Geometric Modeling in higher dimensions and to propose a well-principled ground-breaking software platform allowing technological advances for varied applications in science and engineering.

TODO. Expand the abstract. Add references. Bag of features. Challenges : prciser les questions ouvertes.

Questions. Geometric Modeling sounds outdated. List of WP.

1 The principal investigator

1.1 Curriculum Vitae

I was born in Nice, May 18, 1953. I am married and have 3 children. I am a french citizen.

Education

• I graduated from the Ecole Supérieure d'Electricité (Supélec) in 1976. Supélec is among the top "Grandes Ecoles" in France and the reference in the field of electric energy and information sciences. Supélec is on a par with the best departments of electrical and computer engineering of the top American or European universities. • I obtained my PhD Thesis in Control Theory from the University of Rennes. • I obtained the Habilitation diploma (the highest grade at french universities) in Computer Science from the University of Nice in 1992.

Professional academic experience

• I have been a researcher at INRIA since 1980, first in Rocquencourt and since 1986 in Sophia Antipolis. • I am currently a *Research Director* at INRIA Sophia-Antipolis (France) where I lead the Geometrica project-team. I was promoted in 2009 at the highest rank (Class Exceptional) for my contributions to research, formation and dissemination. • I also had eminent positions within INRIA, most notably the *VP for Science* at INRIA Sophia-Antipolis (500 employees, 30 research groups) (2001-2005). I have also been the *chairman of the Evaluation Commitee* of the institute, one of the most important position at INRIA with a key role in the definition and the implementation of the scientific policy of the institute. This position gave me a comprehensive view of the research performed in the eight research centers of INRIA.

Academic awards and honors

I received two highly prestigious prizes, the IBM award in Computer Science in 1987 and the Grand prize EADS in Information Sciences in 2006 (awarded by the French Academy of Science). I was nominated to the Roberval prize for the french version of my book Algorithmic Geometry coauthored by M. Yvinec. I have also been nominated Chevalier de l'Ordre National du Mérite in 2006.

Publications, patents

I am the author of over 150 technical publications including 1 text book, 4 edited books, 57 journal articles, 89 refereed international conference articles, 12 book chapters. My h-number is 47 with 7786 citations according to Google Scholar.

My publications cover *several fields* : computational geometry, geometric modeling, algorithmic robotics, medical imaging and to a lower extent structural biology and control theory.

My main *contributions* are on mesh generation, surface reconstruction, motion planning, robust geometric computing, randomized algorithms, Voronoi diagrams, Delaunay triangulations, manifold learning.

I am the author of *4 patents* : 2 on mesh generation (Assignee: Institut Francais du Pétrole (IFP)), robotic surgery (Assignee: Intuitive Surgical Inc.), virtual endoscopy (Assignee: Siemens Coroporate Research)).

Software

I am the author of two software that have been commercialized at large scale by major companies, one by Siemens (Flying Through, installed on Siemens scanners) and one by Dassault Systems (integrated in Catia V5 (Shape Editor)).

My group Geometrica is one of the leader teams of the CGAL Open Source Project. The CGAL library is now regarded as the gold standard in Computational Geometry with a huge impact, both in academia and in industry (e.g., the triangulation package of CGAL, developed within Geometrica, is now integrated in the heart of Matlab)

Scientific leadership

My first group Prisme has been the birth place of Computational Geometry in France and played a prominent role in shaping the field. I promoted research on Delaunay triangulations, randomized algorithms, exact computing, and, most importantly, launched 15 years ago the development of the *CGAL library* in collaboration with partners in Europe [27].

In 2003, I founded the Geometrica project-team in replacement of Prisme. The main objective was to develop Nonlinear Computational Geometry. Geometrica's research in this area has been flourishing with seminal contributions on mesh generation and surface reconstruction [8]. Part of the research agenda was devoted to the development of CGAL components that are now used worldwide in academia and in industry for various applications in geometric modeling, medical imaging and geology.

In 2006, I created an antenna of Geometrica in Saclay (Paris's area) to work on the emerging field of *geometric inference*. Together with my colleagues, we made influential contributions to the analysis of distance functions, persistent topology, manifold learning and the development of a theory of geometric sampling.

Supervision of Ph.D. students, postdocs and young researchers

I have supervised 24 Ph.D. students. All of them are enjoying successful careers in academia or industry. I am currently advising two Ph.D. students. One of my former students, Andreas Fabri, founded in 2003 GeometryFactory, a startup company that commercializes CGAL.

Five members of my research group successfully defended their Habilitation. Three created their own research teams at INRIA : J-P. Merlet (Robotics), F. Cazals (Structural Biology), S. Lazard (Computational Geometry). Another member of the group, P. Alliez (who received a Jr. ERC grant), is in the process of creating his own group on Geometry Processing.

Funding ID

I have been the site leader of 8 European projects and the project leader of the IST Project ECG ("Effective Computational Geometry") (2001-2004).

I have been the principal investigator of 8 collaborations with french industry.

I am a site leader of the ICT Fet-Open project Computational Geometric Learning (CGL) which is closely related to this project (http://cglearning.eu/).

1.2 10-year track record

Top 10 publications as senior researcher

Citations are according to Google Scholar. For journal articles, I added the citations of the conference version of the article. Discrete and Computational Geometry and the Symposium on Computational Geometry are regarded as the most prestigious journal and conference in Computational Geometry.

1. J-D. Boissonnat, F. Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. Comput. Geom. Theory Appl. Vol. 22 (2002) 185-203. (421 citations)

2. J-D. Boissonnat, F. Cazals. Natural neighbour coordinates of points on a surface. Comput. Geom. Theory Appl., Vol. 19, No 2-3, July 2001. (74 citations)

3. D. Attali, J-D. Boissonnat, A. Lieutier. Complexity of the Delaunay Triangulation of Points on Surfaces : the Smooth Case. 20th ACM Symposium on Computational Geometry, 2003. (74 citations)

4. D. Attali, J-D. Boissonnat. A Linear Bound on the Complexity of the Delaunay Triangulation of Points on Polyhedral Surfaces. Discrete and Comp. Geometry 31: 369–384 (2004). (61 citations)

5. J-D. Boissonnat, S. Oudot. Provably good sampling and meshing of surfaces. Graphical Models, 67 (2005) 405-451. (198 citations. Graphical Models Top-Cited Article 2005-2010)

6. D. Attali, J-D. Boissonnat, H. Edelsbrunner. Stability and computation of medial axes : a state-of-the-art report. In *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*, T. Moeller, B. Hamann and B. Russell Ed., Springer, series Mathematics and Visualization, 2007. (93 citations)

7. J-D. Boissonnat, D. Cohen-Steiner, G. Vegter. Isotopic implicit surface meshing. Discrete and Computational Geometry, 39: 138-157, 2008. (55 citations)

8. J-D. Boissonnat, C. Wormser and M. Yvinec. Locally uniform anisotropic meshing. 24th ACM Symposium on Computational Geometry, SoCG'08. (16 citations)

9. J-D. Boissonnat, L. Guibas, S. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. Discrete and Comp. Geom. Vol 42, No 1, 2009. (46 citations)

10. J-D. Boissonnat, F. Nielsen, R. Nock. On Bregman Voronoi diagrams. Discrete and Comp. Geom. (2), 2010. (76 citations)

Edited Books and Proceedings

• Algorithmic Foundations of Robotics V, Springer 2004. Coeditors: J. Burdick, K. Goldberg, S. Hutchinson. • Effective Computational Geometry for Curves and Surfaces, Springer, 2006. Coeditor: M. Teillaud. I coauthored two chapters of this book. • Curves and Surfaces. Coeditors: P. Chenin, A. Cohen, C. Gout, T. Lyche, M-L. Mazure and L. Schumaker, Springer Verlag LNCS Vol. 6920, 2012. • Geometric Computing, special issue of the International Journal of Computational Geometry and Applications, Vol. 11, No. 1, 2001. • Computational Geometry, Theory and Applications, Vol. 35 No. 1-2, August 2006. Special issue on the 20th Symposium on Computational Geometry. • Discrete and Computational Geometry, Vol. 36, No 4, December 2006. Special issue on the 20th Symposium on Computational Geometry.

Granted patents

• Methods and apparatus for planning robotic surgery. United States Patent Application 20030109780.

Assignee INRIA and Intuitive Surgical Inc. (2002). Coauthors: E. Coste-Manière, L. Adhami, A. Carpentier, G. Guthart. • Method and apparatus for fast automatic centerline extraction for virtual endoscopy. United States Patent Application 20050033114. Siemens Corporate Research (2004). Coauthor B. Geiger.

Keynote presentations (since 2004)

• International Symposium on Voronoi Diagrams, Tokyo (2004). • Workshop "The World a Jigsaw: Tessellations in the Sciences", Leiden (2006). • French Academy of Science (2 talks, 2006). • Franco Preparata's schriftfest, Brown university (2007). • Seventh conference on "Mathematical Methods for Curves and Surfaces", Toensberg, Norway, 2008. • Colloquium on Emerging Trends in Visual Computing (ETVC, Ecole Polytechnique, 2008). • 22th Sibgrapi, Rio de Janeiro (2009). • ATMCS 2012 (Algebra and Topology; Methods, Computation, and Science), Edinburgh (2012).

Membership to editorial board of international journals

I am on the editorial board of 5 international scientific journals, including two among the most prestigious journals in Computer Science, the *Journal of the ACM* and *Algorithmica*, and the first venue in my field *Discrete and Computational Geometry.*

Organization of international conferences

• I co-chaired in 2004 the program committee of the Symposium on Computational Geometry (SCG), the top conference of the field. • I chaired the Workshop on Algorithmic Foundations of Robotics (WAFR) in 2002. • I have been a member of the steering committee of SCG (1999-2001) and of the scientific committees of the International Conference on Curves and Surfaces (2010) and of the eighth conference on "Mathematical Methods for Curves and Surfaces" (2012). • I have been on the program committee of the following international conferences : Symposium on Geometry Processing (each year since its creation in 2003), STACS 2001 (Symposium on Theoretical Aspects of Computer Science), ESA 2003 (European Symposium on Algorithms), SCG'04 (Symposium on Computational Geometry) SMI'05 (Shape Modelling International), SMP'05 (ACM Symposium on Solid and Physical Modeling), Curves and Surfaces 2006, GMP 2012 (Geometric Modeling and Processing).

International prizes/awards/academy memberships I received the Grand prize EADS in Information Sciences in 2006 (awarded by the French Academy of Science). I received the Graphical Models Top-Cited Article for the period 2005-2010 for my paper with S. Oudot (ref 5 above).

Scientific councils and international visiting committees (2001-)

• Scientific Council of the Ecole Normale Supérieure de Lyon (2000-2003) • Member of the AERES Board (French Evaluation Agency for Research and Higher Education) • Member of working groups 1 (Modèles et calcul) and 2 (Logiciels et systmes informatiques) of Allistène (Alliance des sciences et technologies du numérique) • Chair of the Visiting Committee of LIAMA (Beijing, 2010) • Member of the Visiting Committee of the Computer Science department of ULB (Free University of Bruxels, 2011) • Member of the Visiting Committee of the Geometric Modeling and Scientific Visualization (GMSV) Center of the King Abdullah University of Science and Technology (KAUST, Saudi Arabia, 2012)

2 Extended synopsis of the project

The need for higher-dimensional geometric modeling. Geometric modeling refers to the construction and the manipulation of computerized representations of shapes. Since shape is the first and foremost attribute of our familiar 3D world, geometric modeling has been and still is a core discipline with a vast array of applications ranging from multimedia to engineering through computer-aided medicine. It has been on the research agenda of several communities like computer-aided design, computer graphics, computer vision, computational geometry, geometry processing and numerical analysis.

Until now, most of the work on geometric modeling has been restricted to 3D shapes. However, the need to represent higher-dimensional spaces and shapes is ubiquitous in science. Physicists are used to combine space and time into a single 4-dimensional *space-time* continuum. In particle physics, the *phase space* consists of all possible values of position and momentum variables and is 6-dimensional. *Configuration spaces* of mechanical systems, *conformational spaces* of macromolecules are other examples of common high-dimensional geometric objects. A maybe less intuitive place where high-dimensional geometry can be found is in *data*. Natural and artificial systems like biological or sensor networks are often described by a large number of real parameters, whereas a collection of text documents can be represented as a set of term frequency vectors in Euclidean space; similar interpretations can be given for image and video data [25]. Data analysis can then be turned into a geometric problem by encoding those heteregeneous data as clouds of points in high dimensional spaces equiped with some appropriate distance function. Usually those points are not distributed in the whole embedding space but, due to the very nature of the system that produced those data, lie close to some subset of much smaller intrinsic dimension. Hence, the data conveys some geometric structure whose extraction and modeling are key to understanding the underlying system.

Since data are produced at an unprecedented rate in all sciences, Geometric modeling in higher dimensions has become a core task in science and engineering. Still, it is an emerging and extremely challenging discipline whose full potential has not been harnessed and which has not yet penetrated applied fields.

The curses of higher-dimensional geometry. This situation has several causes. High-dimensional data are much more difficult to process than 3D data. The dimensionality severely restricts our intuition and ability to visualize the data. Moreover, the complexity of the data structures and the algorithms used to approximate shapes rapidly grows as the dimensionality increases, which makes them intractable in high dimensions. This curse of *dimensionality* is exemplyfied by the size of one of the simplest representations of a point set, namely its convex hull, whose complexity depends exponentially on the dimension.

In addition, the data often suffer from significant *defects*, including sparsity, noise, and outliers which make them much more difficult to process than 3D data routinely provided by scanning devices. This is particularly so in the case of biological data, such as high throughput data from microarray or other sources. Moreover, the structure and occurrence of geometric features in the data may depend on the *scale* at which it is considered, thus requiring the inference process to be multiscale.

The emergence of a new research area. Computational geometry and topology have been very successful at providing solid foundations for 3D geometric modeling [8]. The concepts of ε-samples, restricted Delaunay triangulation, anisotropic meshes emerged together with efficient and provably correct algorithms for emblematic problems like mesh generation and surface reconstruction. The attempt to extend these results to higher dimensions led to the development of beautiful pieces of theory with deep roots in various areas of mathematics like Riemannian geometry, geometric measure theory, differencial and algebraic topology. Let us mention the invention of a sampling theory of geometric objects [11], the new concept of topological persistence [18] or the design of new simplicial complexes with good complexity and approximation properties [10, 5]. Although of a fundamental nature, these advances attracted interest in several fields like data analysis, computer vision or sensor networks.

The computational bottleneck. However, the current theory has not demonstrated its scalability to real problems and, up to now, it has only been applied to rather simple cases and in low dimensions. This is due to the lack of satisfactory algorithmic solutions in high-dimensional spaces. Breaking the computational bottleneck is now the main issue and settling the *algorithmic foundations* of higher dimensional geometric modeling is a grand challenge of great theoretical and practical significance.

The tenet of this proposal is that, to take up the challenge, we need a global approach involving tight and long-standing interactions between mathematical developments, algorithmic design and advanced programming. We believe that this is key to obtaining methods with built in robustness, scalability and guarantees, and therefore impact in the long run. Such an approach has been successfully carried out in low dimensions with the development of the Computational Geometric Algorithms Library CGAL [27]. We want to build upon this success and to give to higher dimensional Geometric Modeling an effective theory and a reference software platform.

We strongly believe that this ambitious objective is realistic and can be reached. To pave the way towards this goal, we have identified four main scientific challenges that will also correspond to the main worktasks of the project.

Scientific challenge 1 : Going beyond affine models and Euclidean spaces. In the last decades, a set of new geometric methods, known as manifold learning, have been developed with the intent of parametrizing nonlinear shapes embedded in high-dimensional spaces. Although widely used, those methods assume very restrictive hypotheses on the geometry of the manifolds sampled by the datapoints to ensure correctness. A different route, inspired by what has been done in 3-dimensions, consists in approximating highly nonlinear shapes by *simplicial complexes*, the analogue of triangulations in higher dimensions. Recent research has exhibited sampling conditions under which topological or geometric properties, or even a full approximation of the sampled shape can be recovered [24, 11, 5]. Still the sampling conditions are quite stringent, the simplicial complexes are huge objects and their construction may be problematic.

A fundamental issue is the choice of a metric which determines the type and quality of an approximation. The simple Euclidean distance in the ambient space, while easy to deal with, is often not the right choice. As already mentionned, most of the time, when working in high dimensional spaces, we are in fact interested by objets of much smaller intrinsic dimension. The intrinsic geometry on the objects provides the right framework. Computational intrinsic geometry has not been seriously tackled yet and a basic question like the existence of Delaunay triangulations on Riemannian manifolds is still open. This question is of utmost importance for anisotropic mesh generation and optimal approximations. Another important situation is when the data are not given as a point cloud in some Euclidean space, but rather as a matrix of pairwise distances (i.e. a discrete metric space). Although such data may not be sampled from geometric subsets of Riemannian manifolds, it may still carry some interesting topological structures that need to be understood.

In information theory, signal and image processing, other pseudo-distances such as Kullback-Leibler, Itakura-Saito or Bregman divergences are prefered. These divergences are usually not true distances (they may not be symmetric nor satisfy the triangular inequality) and it is necessary to revisit geometric data structures and algorithms in this context [7].

Scientific challenge 2 : Bypassing the curse of dimensionality. The complexities of many geometric algorithms and data structures grow exponentially with increasing dimension. This behavior is commonly called the curse of dimensionality after Bellman. An immediate consequence is that it is no longer possible to partition a high dimensional space. This rules out most, even if not all, geometric algorithms developped in low dimensions. Hence, extending Computational Geometry in high dimensions cannot be done in a straightforward manner and one has to take advantage of additional structure of the problem or to restrict attention to approximation. This motivated a number of new algorithmic paradigms such as locality-sensitive hashing, smoothed analysis, intrinsic algorithms. In this project, we will address the curse of dimensionality by focusing on the inherent structure in the data which in some sense needs to be sparse or of low intrinsic dimension as well as by putting the emphasis on output-sensitive algorithms and average-case analysis [25]. An important feature is that, even if we go to approximate solutions, we do not want to sacrifice guarantees. This is mandatory since the behaviour of algorithms in high dimensions is much less intuitive and easy to predict than in small dimensions.

Scientific challenge 3 : Searching for stable models. When dealing with approximation and samples, one needs stability results to ensure that the quantities that are computed are good approximations of the real ones. This is especially true in higher-dimensions where data are usually corrupted by various types of noise. A number of groundbreaking new approaches appeared recently. *Topological persistence* was recently introduced as a powerful tool for the study of the topological invariants of sampled spaces [18].

To deal with *non-local noise* and outliers, another new paradigm for point cloud data analysis has emerged recently. Point clouds are no longer treated as mere compact sets but rather as empirical measures. A notion of distance to such measures has been defined and shown to be stable with respect to perturbations of the measure [12]. A big challenge is to find efficient algorithms in arbitrary dimensions to compute or approximate the topological structure of the sublevel-sets of the distance to a measure. Such algorithms would naturally find applications in topological inference in the presence of significant noise and outliers, but also in other less obvious contexts such as stable clustering.

Multiscale reconstruction is another novel approach [6]. Taking advantage of the ideas of persistence, the approach consists in building a one-parameter family of simplicial complexes approximating the input at various scales. Determining the topology and shape of the original object reduces to finding the stable sequences in the one-parameter family of complexes. Despite its nice features, multiscale reconstruction, in its current form, can only be applied to low-dimensional data sets in practice.

Scientific challenge 4 : Building up the reference platform for high-dimensional geometric modeling. Software development is a central issue in this project and we will devote a substantial part of the effort to build up a software platform that will provide easy access to efficient and reliable high-dimensional geometric algorithms in the form of a C++ library. We will follow the example of the Computational Geometry Algorithms Library (CGAL), one of the major success stories of computational geometry, by now the gold standard for low dimensional geometric computing [27]. Our project aims at extending this success story, combining efficiency with correctness guarantees to high-dimensional geometric computation.

The development of a reference platform for high-dimensional geometric modeling is dictated by three main motivations. *First*, the sofware platform will allow to experiment at a large scale, which is mandatory to design the right models and data structures. It will boost theoretical research in geometric modeling, computational geometry in high dimensions, and computational topology, leading towards a virtuous circle between theory and experimental research. This has proven to be of utmost importance when developing the CGAL library and will be even truer in high dimensional geometry.

Second, maintaining such a platform will help further effort and consolidation in the long run. Having a library with interoperable modules will allow to incrementally add more and more sophisticated tools based on solid foundations. This is consistent with our long-term vision and our conviction that it is only through such a long standing effort that true impact, both theoretical and applied, can be gained.

Third, the platform will serve as a unique tool to communicate with the computational geometry community and with researchers from other fields. In return, we will get feedback from practionners that will help shaping the platform.

Risks and feasibility of the project. Simultaneously pursuing basic research at the best international level and developing industrial strength software is not without risks. My personnal record as well as the record of the members of Geometrica involved in this project are strong indications of our ability to take up the challenge with good chances of success. Geometrica has been at the cutting edge of research in geometric data structures and algorithms, mesh generation, shape approximation, geometric inference and computational topology. Geometrica is also one of the leader teams of the CGAL project. We were at the source of successfull developments in CGAL like interval arithmetics, triangulations (now integrated in the heart of Matlab) and meshing packages. We also took a prominent part in the animation of the CGAL project and community, and in the creation of the spinoff GeometryFactory. It can be argued that my research group Geometrica is the best team worldwide to take up this dual challenge and to make this project a success.

I will devote 70% of my time to this project, and I will dedicate all my expertise and efforts to conduct and supervise the research work. To this end, I will receive the precious help of 2 permanent researchers of the Geometrica team : Frédéric Chazal who is a leading researcher in geometric inference and computational topology and Mariette Yvinec who is an expert in geometric computing and a member of the CGAL Editorial Board. They will devote 20% of their time to this project to co-supervise with me the research and implementation work of the students, postdocs and engineers to be engaged in this project. Other members of Geometrica, not financially supported by this project, will also collaborate to the project.

We will pursue our collaborations with the best groups in Europe and in the US, most notably Stanford university (Pr. Guibas) and Ohio State university (Pr. Dey). We will also collaborate with researchers in Brazil and China.

New horizons and opportunities. If successful, the project will put higher-dimensional geometric modeling on new theoretical and algorithmic ground. It will also provide an open platform with no equivalent in USA or Asia. We forsee the platform to become a catalyst for research in high-dimensional geometry inside and outside of the project. The platform will implement the most effective techniques in a reliable and scalable way, opening the way to groundbreaking technological advances for applications as varied as numerical simulation and visualization, machine learning, computer vision, data analysis, robotics or molecular biology. We will keep close contacts with research groups working in those domains and leap on opportunities arising from our new results and tools. In return, we will get feedback from practitionners which will help shaping the theoretical models and the software platform. Based on our experience with CGAL, we will undertake a vigorous action towards code diffusion in applied domains.

An opportunity for Europe. Research in computational geometry and topology is very active in Europe. The academic European community has been supported through several research projects by the European Commission leading to the successful development of CGAL over the years. The ICT Fet-Open project Computational Geometric Learning (CG-Learning) is closely related to this project. The focus and the timetable are different though. The proposed project wants to take over the results of CG-Learning and to go beyong prototype developments. This new project will further strenghten the leadership of Europe in Geometric Computing.

On the industrial side, Dassault Systèmes is a world leader in Geometric Modeling and we have been collaborating with them for more than 10 years (commercialization of software, co-advised Ph.D. student, joint publications). We intend to pursue this fruitful collaboration and to keep our collaborators at Dassault-Systèmes informed of our advances. CGAL is commercialized by the start-up company GeometryFactory. The small size of the company and the fact that most of its members are alumni of the Geometrica team makes contacts very easy and allows fast transfer of results.

It is to be noted that a start-up company has just been created in California. Its motto "using shapes to find patterns in data" is close to ours (http://www.ayasdi.com/).

References

[1] P. Alliez, J.-D. Boissonnat, and M. Yvinec. Cgalmesh: An open platform for mesh generation. http://www-sop.inria.fr/geometrica/software/cgalmesh/.

[2] P. Alliez, L. Rineau, S. Tayeb, J. Tournois, and M. Yvinec. 3D mesh generation. In *CGAL User and Reference Manual*. CGAL Editorial Board, 3.8 edition, 2011.

[3] J.-D. Boissonnat, D. Cohen-Steiner, B. Mourrain, G. Rote, and G. Vegter. Meshing of surfaces. In J.-D. Boissonnat and M. Teillaud, editors, *Effective Computational Geometry for Curves and Surfaces*, pages 181–229. Springer-Verlag, Mathematics and Visualization, 2006.

[4] J.-D. Boissonnat, O. Devillers, and S. Hornus. Incremental construction of the Delaunay graph in medium dimension. In *Proc. 25th Annual Symposium on Computational Geometry*, pages 208–216, 2009.

[5] J.-D. Boissonnat and A. Ghosh. Manifold reconstruction using tangential Delaunay complexes. In *Proc. 26th Annual Symposium on Computational Geometry*, 2010.

[6] J.-D. Boissonnat, L. J. Guibas, and S. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. *Discrete and Computational Geometry*, 42(1):37–70, 2009.

[7] J.-D. Boissonnat, F. Nielsen, and R. Nock. Bregman Voronoi diagrams. *Discrete and Computational Geometry*, 44(2), 2010.

[8] J.-D. Boissonnat and M. Teillaud, editors. *Effective Computational Geometry for Curves and Surfaces*. Springer-Verlag, 2006.

[9] J.-D. Boissonnat and M. Yvinec. *Algorithmic geometry*. Cambridge University Press, 1998.

[10] G. Carlsson and V. de Silva. Topological estimation using witness complexes. In *Symposium on Point-Based Graphics*, 2004.

[11] F. Chazal, D. Cohen-Steiner, and A. Lieutier. A sampling theory for compact sets in euclidean space. *Discrete Comput. Geom.*, 41(3):461–479, 2009.

[12] F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Journal on Foundations of Computational Mathematics*, 11(6):733–751, 2011.

[13] F. Chazal and S. Y. Oudot. Towards Persistence-Based Reconstruction in Euclidean Spaces. In *Proc. ACM Symp. on Computational Geometry*, pages 232–241, 2008.

[14] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.

[15] V. de Silva. A weak characterisation of the Delaunay triangulation. *Geometriae Dedicata*, 135:39–64, 2008.

[16] T. Dey. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis*. Cambridge University Press, 2007.

[17] H. Edelsbrunner. The union of balls and its dual shape. *Discrete Comput. Geom.*, 13:415–440, 1995.

[18] H. Edelsbrunner and J. Harer. Persistent homology — a survey. In J. P. J. E. Goodman and R. Pollack, editors, *Surveys on Discrete and Computational Geometry. Twenty Years Later*, Contemporary Mathematics 453, pages 257–282. Amer. Math. Soc., 2008.

[19] H. Edelsbrunner, D. G. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29:551–559, 1983.

[20] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, 28:511–533, 2002.

[21] R. Ghrist. Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc., 45(1)*, pages 61–75, 2008.

[22] L. J. Guibas and S. Oudot. Reconstruction using witness complexes. *Discrete and Computational Geometry*, 40(3):325–356, 2008.

[23] S. Har-Peled. *Geometric Approximation Algorithms*. Amer. Math. Soc., 2011.

[24] P. Niyogi, S. Smale, and S. Weinberger. Finding the Homology of Submanifolds with High Confidence from Random Samples. *Discrete and Computational Geometry*, 39(1):419–441, 2008.

[25] H. S. Seung and D. D. Lee. The manifold ways of perception. *Science*, 290:2268–2269, 2000.

[26] D. Spielmann and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.

[27] Cgal. Computational Geometry Algorithms Library. http://www.cgal.org.

3 Research proposal (15p.)

3.1 State-of-the-art and objectives

The research and development pipeline in Computational Geometry. Computational geometry emerged as a discipline in the seventies and has met with considerable success in providing foundations to solve basic geometric problems including data structures, convex hulls, triangulations, Voronoi diagrams, geometric arrangements and geometric optimisation [9]. This initial development of the discipline has been followed, in the mid-nineties, by a vigorous effort to make computational geometry more effective. Situated in-between basic theoretical research and the development of robust software is the emerging study of effective methods (algorithms and data structures) of geometric computing, namely methods that are not only theoretically proved but also work well in practice. Several EC projects (CGAL, GALIA) established an outstanding research momentum and gave a leading role to Europe in this context. They led to successful techniques and tools, most notably the CGAL library [27], a unique tool that provides a well-organised, robust and efficient software environment for developing geometric applications. CGAL is considered as one of the main achievements of the field and is by now the standard in Geometric Computing, with a large diffusion worldwide and varied applications in both academia and industry. CGAL has no equivalent counterpart in the world.

3D Geometric Modeling. Approximating complex shapes through meshing is a fundamental problem on the agenda of several communities like Numerical Analysis, Computer Graphics, Geometry Processing and Computer-Aided Design. Emblematic problems are mesh generation that aims at sampling and meshing a given domain, and surface reconstruction that constructs an approximation of a surface which is only known through a set of points. Although these problems have received considerable attention in the past, it is only during the last 10 years that the Computational Geometry community established solid theoretical foundations to the problem, most notably in the emerging new area of Computational Topology. This approach has shown to be very successful and led to recent breakthroughs in *mesh generation* [3] and *surface reconstruction* [16]. The Geometrica group took a leading role in this research and contributed major theoretical advances as well as practical developments in the form of fast, safe and quality-guaranteed CGAL components for mesh generation and shape reconstruction [2]. Those components are now used worldwide in academia and in industry for various applications in Geometric Modeling, Medical Imaging and Geology.

Noisy data. When dealing with approximation and samples, one needs stability results to ensure that the quantities that are computed, geometric or topological invariants, are good approximations of the real ones. *Topological persistence* was recently introduced as a powerful tool for the study of the topological invariants of sampled spaces [18, 21]. Given a point cloud in Euclidean space, the approach consists in building a simplicial complex whose elements are filtered by some user-defined function. This filter basically gives an order of insertion of the simplices in the complex. The persistence algorithm, first introduced by Edelsbrunner, Letscher and Zomorodian [20], is able to track down the topological invariants of the filtered complex as the latter is being built. As proved by Cohen-Steiner et al. [14], under reasonable conditions on the input point cloud, and modulo a right choice of filter, the most persistent invariants in the filtration correspond to invariants of the space underlying the data. Thus, the information extracted by the persistence algorithm is global, as opposed to the locality relationships used by the dimensionality reduction techniques. In this respect, topological persistence appears as a complementary tool to dimensionality reduction. In particular, it enables to determine whether the input data is sampled from a manifold with trivial topology, a mandatory condition for dimensionality reduction to work properly. Note however that it does not tell how and where to cut the data to remove unwanted topological features.

Multiscale reconstruction is a novel approach [6]. Taking advantage of the ideas of persistence, the approach consists in building a one-parameter family of simplicial complexes approximating the input at various scales. Differently from above, the family may not necessarily form a filtration, but it has other nice properties. In particular, for a sufficiently dense input data set, the family contains a long sequence of complexes that approximate the underlying space provably well, both in a topological and in a geometric sense. In fact, there can be several such sequences, each one corresponding to a plausible reconstruction at a certain scale. Thus, determining the topology and shape of the original object reduces to finding the stable sequences in the one-parameter family of complexes. However, multiscale reconstruction, at least in its current form, still has a complexity that scales up exponentially with the dimension of the ambient space. Hence, it can only be applied to low-dimensional data sets in practice.

High dimensional spaces. Dimensionality reduction is certainly one of the most popular approaches to high-dimensional data analysis. It consists in projecting the data points down to a linear subspace, whose dimension supposedly coincides with the intrinsic dimension of the data. This approach is elegant in that it helps detect the intrinsic parameters of the data, and by doing so it also reduces the complexity of the problem. Dimensionality reduction techniques fall into two classes: linear methods, e.g. principal component analysis (PCA) or multi-dimensional scaling (MDS), and non-linear methods, e.g. isomap or locally-linear embedding (LLE). The second class of algorithms is more powerful in that it computes more general (in fact, non-linear) projections. On the whole, dimensionality reduction works well on data sets sampled from manifolds with low curvature and trivial topology. Although the condition on the curvature is mainly a sampling issue, the condition on the topology is mandatory for the projection onto a linear subspace to make sense.

Many of the results of Computational Geometry have been extended to arbitrary dimension. In particular, worst-case optimal algorithms are known for computing convex hulls, Voronoi diagrams and Delaunay triangulations in any dimension. Hence, in principle, the methods developed for 3D applications should be extendable to higher dimensions. However, the size of these structures depends exponentially on the dimension of the embedding space, which makes them only useful in moderate dimensions [4]. Many ideas have been suggested to bypass this curse of dimensionality. A first approach looks for more realistic combinatorial analyses such as smoothed analysis that bounds the expected complexity under some small random perturbation of the data. This led to the celebrated analysis of Linear Programming of Spielmann and Teng [26]. Another route is to trade exact for approximate algorithms [23]. An important example is the search for approximate nearest neighbours. Another example is the theory of core sets that was shown to provide good approximate solution to some optimization problems like computing the smallest enclosing ball, or computing an optimal separating hyperplane (SVM). These tools are both extremely useful but limited to basic operations and have not been yet applied in the context of geometric modeling. A third approach assumes that the intrinsic dimension of the object of interest has a much lower intrinsic dimension than the dimension of the embedding space. This is the usual assumption in Manifold Learning. It is then possible to resort to techniques derived from the 3D case and to approximate complex shapes by simplicial complexes (the analogue of triangulations in higher-dimesnional spaces). Various types of simplicial complexes have been proposed such as the Czech and the RIPS complexes, and the more recent Delaunay-like complexes such as the α-complex [19, 17], the witness complex [15, 10] and the Delaunay tangential complex [5]. They differ by their combinatorial and algorithmic complexities, and their power to approximate a shape. Under appropriate sampling conditions, we have shown that one can reconstruct a provably correct approximation of a smooth k-dimensional manifold M embedded in $\mathbb{R}^d$ in a time that depends only linearly on d [5]. Researchers have also turned their focus to the somewhat easier problem of inferring topological invariants of the shape without explicitly reconstructing it with the hope that more lightweight data structures and weaker sampling conditions would be appropriate for this simpler task [24, 13].

3.2 Methodology

Our overall goal is to settle the foundations for Geometric Modeling in higher dimensions by developing

- Sounded approaches providing guarantees even in the presence of noise or outliers,
- Algorithms that are of both theoretical and practical interest, that are amenable to theoretical analysis and fully validated experimentally,
- A software platform that is generic, robust and efficient.

The proposal is structured into the following four workpackages:

- WP 1: Computational geometry in non euclidean spaces.
- WP 2: Dimension-sensitive algorithms and data structures.
- WP 3: Geometric modelling.
- WP 4: Platform for geometric modelling in higher dimensions.

WP 1: Computational geometry in non euclidean spaces

Delaunay triangulation and Voronoi diagrams in non-euclidean spaces, e.g. riemannian manifolds, Bregman and statistical spaces, intrinsic algorithms (discrete metric spaces). Approximate nearest-neighbour search and other basic algorithms.

WP 2: Dimension-sensitive algorithms and data structures

Data structures with an intrinsic dimension sensitivity : small simplicial complexes, compact representation of simplicial complexes. Random simplicial complexes. More general complexes (e.g.cubical)? Efficient construction of simplicial complexes from data points. Major progress in mesh generation and surface reconstruction has been obtained through the elaboration of the theory of Delaunay triangulations. Although not usable in high dimensions, the search for tractable Delaunay-like simplicial complexes is an avenue for research.

WP 3 : WP 3: Geometric modelling

Triangulating Riemannian manifolds, stratified shapes, mesh generation, reconstruction. Parameterization of data/nonlinear shapes. Geometric inference. Feature extraction. Persistent homology. Stability with respect to perturbation of the data. Topology preserving approximation/simplification. Clustering. Applications in Data Analysis, Computer Vision, Numerical Simulation, Robotics, Molecular Biology.

WP 4 : Platform for geometric modelling in higher dimensions

Development of a C++ platform that implements the best algorithms developed during the project or by external collaborators. Set up datasets for controlled experiments in support of WP 1-3, and more generally, research on Geometric Modeling in higher dimensions. Diffuse the software and promote its use in fields like Machine Learning, Data Analysis, Numerical Simulation, Visualization, Structural Biology and others.

A small but significant part of the project will be dedicated to gathering datasets. These data sets will be made publicly available. They will include both synthetic datasets and datasets coming from the application domains listed above. This will be done in collaboration with expert groups with whom we are in close contact.

4 References

References

[1] P. Alliez, J.-D. Boissonnat, and M. Yvinec. Cgalmesh: An open platform for mesh generation. http://www-sop.inria.fr/geometrica/software/cgalmesh/.

[2] P. Alliez, L. Rineau, S. Tayeb, J. Tournois, and M. Yvinec. 3D mesh generation. In *CGAL User and Reference Manual*. CGAL Editorial Board, 3.8 edition, 2011.

[3] J.-D. Boissonnat, D. Cohen-Steiner, B. Mourrain, G. Rote, and G. Vegter. Meshing of surfaces. In J.-D. Boissonnat and M. Teillaud, editors, *Effective Computational Geometry for Curves and Surfaces*, pages 181–229. Springer-Verlag, Mathematics and Visualization, 2006.

[4] J.-D. Boissonnat, O. Devillers, and S. Hornus. Incremental construction of the Delaunay graph in medium dimension. In *Proc. 25th Annual Symposium on Computational Geometry*, pages 208–216, 2009.

[5] J.-D. Boissonnat and A. Ghosh. Manifold reconstruction using tangential Delaunay complexes. In *Proc. 26th Annual Symposium on Computational Geometry*, 2010.

[6] J.-D. Boissonnat, L. J. Guibas, and S. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. *Discrete and Computational Geometry*, 42(1):37–70, 2009.

[7] J.-D. Boissonnat, F. Nielsen, and R. Nock. Bregman Voronoi diagrams. *Discrete and Computational Geometry*, 44(2), 2010.

[8] J.-D. Boissonnat and M. Teillaud, editors. *Effective Computational Geometry for Curves and Surfaces*. Springer-Verlag, 2006.

[9] J.-D. Boissonnat and M. Yvinec. *Algorithmic geometry*. Cambridge University Press, 1998.

[10] G. Carlsson and V. de Silva. Topological estimation using witness complexes. In *Symposium on Point-Based Graphics*, 2004.

[11] F. Chazal, D. Cohen-Steiner, and A. Lieutier. A sampling theory for compact sets in euclidean space. *Discrete Comput. Geom.*, 41(3):461–479, 2009.

[12] F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Journal on Foundations of Computational Mathematics*, 11(6):733–751, 2011.

[13] F. Chazal and S. Y. Oudot. Towards Persistence-Based Reconstruction in Euclidean Spaces. In *Proc. ACM Symp. on Computational Geometry*, pages 232–241, 2008.

[14] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.

[15] V. de Silva. A weak characterisation of the Delaunay triangulation. *Geometriae Dedicata*, 135:39–64, 2008.

[16] T. Dey. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis.* Cambridge University Press, 2007.

[17] H. Edelsbrunner. The union of balls and its dual shape. *Discrete Comput. Geom.*, 13:415–440, 1995.

[18] H. Edelsbrunner and J. Harer. Persistent homology — a survey. In J. P. J. E. Goodman and R. Pollack, editors, *Surveys on Discrete and Computational Geometry. Twenty Years Later*, Contemporary Mathematics 453, pages 257–282. Amer. Math. Soc., 2008.

[19] H. Edelsbrunner, D. G. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29:551–559, 1983.

[20] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, 28:511–533, 2002.

[21] R. Ghrist. Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc., 45(1)*, pages 61–75, 2008.

[22] L. J. Guibas and S. Oudot. Reconstruction using witness complexes. *Discrete and Computational Geometry*, 40(3):325–356, 2008.

[23] S. Har-Peled. *Geometric Approximation Algorithms.* Amer. Math. Soc., 2011.

[24] P. Niyogi, S. Smale, and S. Weinberger. Finding the Homology of Submanifolds with High Confidence from Random Samples. *Discrete and Computational Geometry*, 39(1):419–441, 2008.

[25] H. S. Seung and D. D. Lee. The manifold ways of perception. *Science*, 290:2268–2269, 2000.

[26] D. Spielmann and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.

[27] Cgal. Computational Geometry Algorithms Library. http://www.cgal.org.

5 Ressources