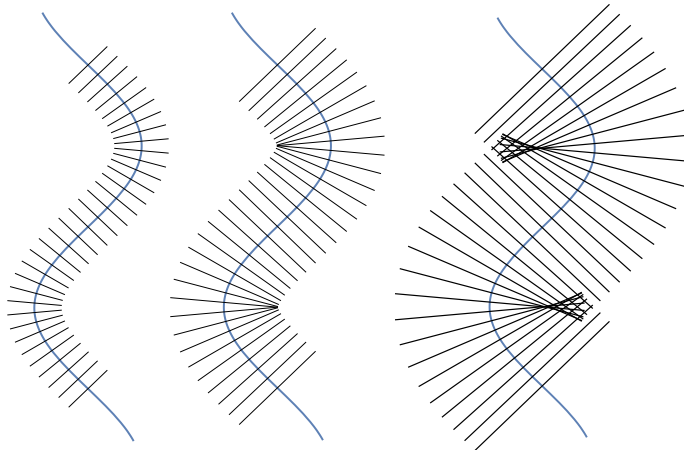


1 Definitions and pictures

How big a neighborhood of a manifold can be embedded along with the manifold?

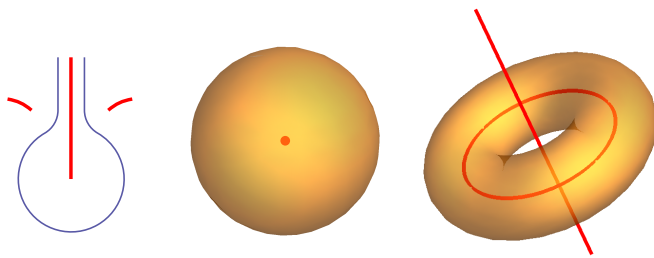
Defn 1. The *conditioning number* of M is

$$\tau := \sup_{\substack{\text{embeddings} \\ N^\epsilon M}} \epsilon.$$


embedding of $N^\epsilon M$, $\epsilon < \tau$ embedding of $N^\tau M$ non-embedding of $N^\epsilon M$, $\epsilon > \tau$

This is also known as the *reach*, the *critical radius*, or the *injectivity radius*. It is determined by some point(s) in the ambient space where their normal spaces intersect.

Defn 2. The *medial axis* of M is the closure of the set of all $x \in \mathbf{R}^n$ for which there exist $p \neq q \in M$ such that

$$\inf_{y \in M} |x - y| = |x - p| = |x - q|.$$


The medial axis is affected by local properties (how “curved” the manifold is) and by global properties (how “close to itself” the manifold is).

2 Curves in \mathbf{R}^n

Let C be a smooth curve embedded in \mathbf{R}^n . For local properties, take $p, q \in C$ close to each other and intersect their normal planes $N_p C$ and $N_q C$. Measuring the shortest distance from p to q along these planes will describe τ “locally.” “Globally,” we only need to consider pairs of points (p, q) that are in each other’s normal planes. Fix p and set

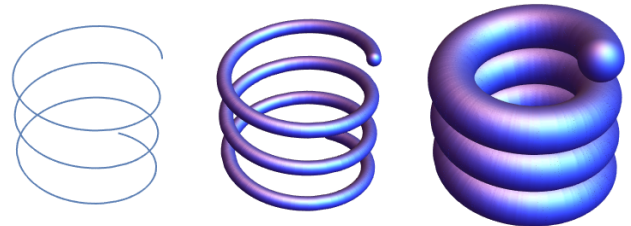
$$\tau_p^\ell = \frac{1}{2} \lim_{q \rightarrow p} \left[\inf_{x \in N_p C \cap N_q C} \{d(p, x) + d(q, x)\} \right],$$

$$\tau_p^g = \frac{1}{2} \inf_{\substack{q \in N_p C \\ p \in N_q C}} d(p, q)$$

as the *local* and *global* conditioning numbers at p , respectively. Then the conditioning number of the curve is

$$\tau = \min \left\{ \inf_{p \in C} \tau_p^\ell, \inf_{p \in C} \tau_p^g \right\}.$$

In special cases, we may take advantage of the regularity or symmetry of the curve. For example, consider the helix:



With radius r and period $2\pi c$, its equations are

$$f_1(x, y, z) = x - r \cos(z/c),$$

$$f_2(x, y, z) = y - r \sin(z/c).$$

The local conditioning number is $\tau^\ell = (r^2 + c^2)/r$, which simplifies to $\tau^\ell = r$ when $c = 0$, when the helix has collapsed to a circle. The global conditioning number is also found easily, by observing that the helix rises at an angle of $\arctan(c/r)$. Hence

$$\tau_{\text{helix}} = \min \left\{ \frac{r^2 + c^2}{r}, \frac{\pi cr}{\sqrt{r^2 + c^2}} \right\}.$$

3 Settings

- Ideal:** We know the (equations of the) manifold.
- Semi-ideal:** We may sample as many points as we want.
- Realistic:** We are given a finite point sample.

So far we have considered the ideal case. Now we turn to situations when all is not known about the manifold.

4 Finite point sampling

Sample points on M according to a random variable X . It may not be truly random (*sampling bias*) and may have support strictly larger than M (*noise*).

Defn 3. The *probability density function* of X is the function $f : M \rightarrow \mathbf{R}$ satisfying
 (a) $f(p) \geq 0$ for all $p \in M$, and
 (b) $\int_N f(p) dp = P(X \in N)$ for any $N \subseteq M$.

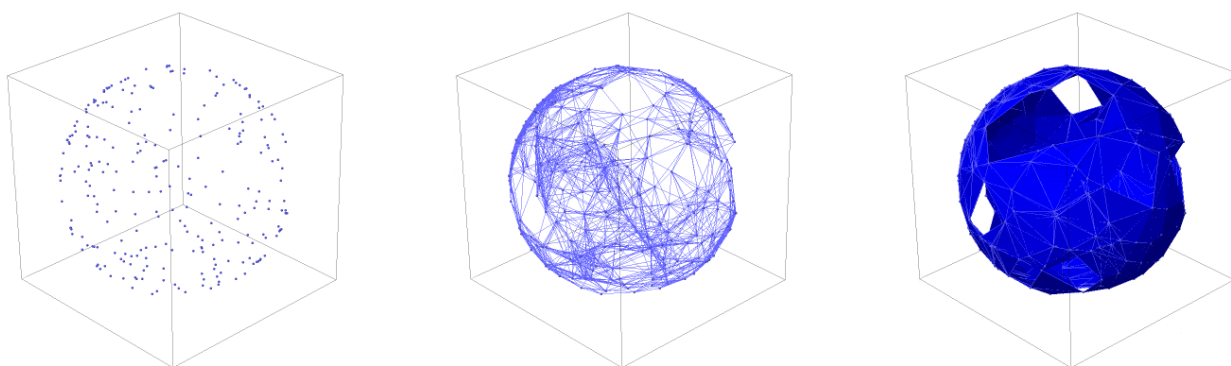
A lower bound on the number of sampled points necessary to find the homology of M is given in [2].

Thm 1. There is an algorithm that will find the homology of a d -manifold M , with probability $1 - \delta$, from a random point sample of at least

$$\log\left(\frac{\delta}{\ell}\right) \log\left(1 - \frac{\pi^{d/2}(\delta/4)^d}{\Gamma(d/2 + 1)}\right)$$

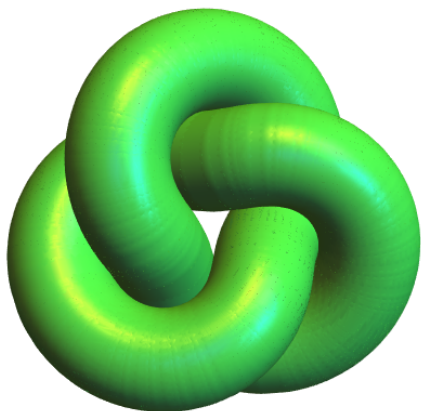
points on M , for ℓ a function in $1/\tau^d$.

For example, consider the unit 2-sphere, with $\tau = 1$. For $\delta = 0.5$, the algorithm then needs at least 174 points; for $\delta = 0.1$, at least 6919 points. Below we have 250.



5 Applications

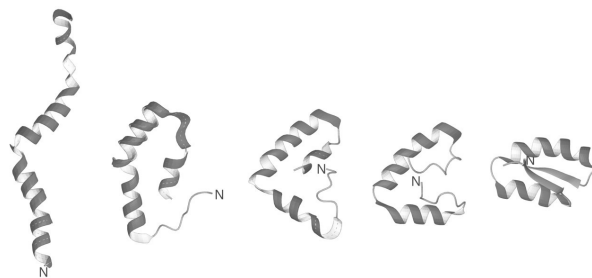
Knot theory: Given a knot K of fixed length, which embedding of K into \mathbf{R}^3 will have the largest τ ?



Such an embedding, described in [1], gives the *ideal shape* of K , representing the “tightest” a knot can be pulled.

Mechanics: A wire of diameter d and length ℓ , when coiled into a spring of radius r , can be compressed down to a height of no less than $2\pi r\ell/d$.

Chemistry: A protein is a chain of amino acids and *folding* describes how its physical states change. Some states, “tightest” in some sense, make the organism function properly, while others cause problems.



Given an potential embedding of a protein, knowing τ indicates if it is not a valid embedding.

6 References

- [1] Oscar Gonzalez and John H. Maddocks. “Global curvature, thickness, and the ideal shapes of knots”. In: *Proc. Natl. Acad. Sci. USA* 96.9 (1999), 4769–4773 (electronic). ISSN: 1091-6490.
- [2] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. “Finding the homology of submanifolds with high confidence from random samples”. In: *Discrete Comput. Geom.* 39.1-3 (2008), pp. 419–441. ISSN: 0179-5376.