Lecture notes for Math 635: quantitative aspects of Morse and Floer theory
USC Spring 2022

Kyler Siegel

Contents

LECTURE 1

Introduction

Let M^n be a closed smooth manifold of dimension n. There are many equivalent models for its (co)homology:

- singular homology $H_*^{\mathrm{sing}}(M;\mathbb{Z})$
- simplicial homology $H_*^{\mathrm{simp}}(M;\mathbb{Z})$
- de Rham cohomology $H^*_{\mathrm{dR}}(M;\mathbb{R})$
- Morse homology $H_*^{\mathrm{Morse}}(M;\mathbb{Z})$,

etc. A basic fact from algebraic topology is that singular and simplicial homology are isomorphic, although the latter requires a choice of triangulation. Similarly, a basic fact from smooth manifold theory is that de Rham cohomology (i.e. closed differential forms modulo exact ones) is isomorphic to singular or simplicial cohomology over the real numbers.

The last model, Morse homology, is the one most relevant for this course. We begin by recalling the basics of Morse theory.

1.1. First look at Morse theory

DEFINITION 1.1. *A smooth function $f : M \to \mathbb{R}$ is* **Morse** *if all of its critical points are nondegenerate. Here $p \in M$ is a* **critical point** *if $df|_p = 0$ (i.e. $\partial_{x_1} f(p) = \cdots = \partial_{x_n} f(p) = 0$ in local coordinates $x_1, \ldots, x_n$ near p) and such a critical point is* **nonsingular** *if the Hessian $d^2 f|_p$ is nonsingular (i.e. the determinant of the symmetric $n \times n$ matrix $(\partial_{x_i}\partial_{x_j} f(p))$ is nonzero).*

LEMMA 1.2 (Morse lemma). *If $f : M \to \mathbb{R}$ is a Morse function and $p \in M$ is critical point, we can find local coordiantes $x_1, \ldots, x_n$ near p such that*

$$f(x_1, \ldots, x_n) = f(p) - x_1^2 - \cdots - x_s^2 + x_{s+1}^2 + \cdots + x_n^2.$$

Here s is called the **(Morse) index** *of f at p, denoted by $\mathrm{ind}_f(p)$.*

Remark 1.3. If $p \in M$ is a **regular point** (i.e. not a critical point), then we can find local coordinates near p such that $f(x_1, \ldots, x_n) = x_1$. This is a consequence of the inverse function theorem. ◊

Example 1.4. We have $\mathrm{ind}_f(p) = 0$ if and only if p is a local minimum, $\mathrm{ind}_f(p) = n$ if and only if p is a local maximum, and otherwise p is a saddle point. ◊

The basic philosophy of Morse theory is to study the relationship between the topology of M and properties of a Morse function $f : M \to \mathbb{R}$. Consider the sublevel sets

$$S_{\leq a} := f^{-1}((-\infty, a]).$$

We have:

- $S_{\leq a} = \varnothing$ for $a < \min f$
- $S_{\leq a} = M$ for $a > \max f$
- $S_{\leq a} \subset S_{\leq a'}$ for $a \leq a'$.

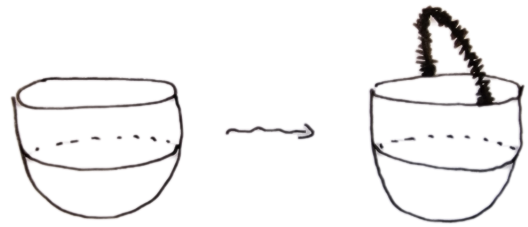

FIGURE 1.1. Attaching a 2-dimensional 1-handle.

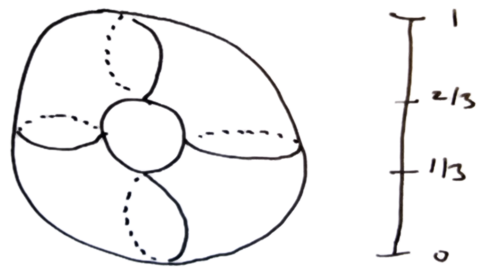

FIGURE 1.2. The height function $\mathbb{T}^2 \to \mathbb{R}$.

NOTATION 1.5. *Let* $\mathrm{crit}_i(f)$ *denote the set of index i critical points of f, and put* $\mathrm{crit}(f) := \cup_{i=0}^n \mathrm{crit}_i(f)$. *Note that* $f(\mathrm{crit}(f)) \subset \mathbb{R}$ *is the set of critical values of f.*

The following could be called the "fundamental theorem of Morse theory":

THEOREM 1.6. *For $a \leq a'$, $S_{\leq a'}$ deformation retracts onto $S_{\leq a}$ if $[a, a'] \cap f(\mathrm{crit}(f)) = \varnothing$. On the other hand, if (for simplicity) there is a single critical point p in $f^{-1}([a, a'))$ with $a < f(p) < a'$, then $S_{\leq a}$ is homotopy equivalent to $S_a \cup H$, where H is an n-dimensional handle of index* $\mathrm{ind}_f(p)$.

Recall that a n-dimensional handle of index k is of the form $\mathbb{D}^k \times \mathbb{D}^{n-k}$, glued to a manifold with boundary along $\partial\mathbb{D}^k \times \mathbb{D}^{n-k}$. This can be viewed as a "thickened" version of attaching a k-cell to a CW complex. See Figure 1.1 for a cartoon.

Example 1.7. Figure 1.2 depicts a Morse function $\mathbb{T}^2 \to \mathbb{R}$ given by the height. Figure 1.3 shows how the sublevel sets vary with a and the corresponding handle attachments as promised by Theorem 1.6 $\Diamond$

Remark 1.8. A function $f : \mathbb{T}^2 \to \mathbb{R}$ must have a global minimum and a global maximum since $\mathbb{T}^2$ is compact. From the above theorem, it's easy to see that it must also have at least one saddle point, i.e. there's no way to get $\mathbb{T}^2$ by a sequence of 0-handle and 2-handle attachments. $\Diamond$

1.2. Morse homology basics and consequences

Let $f : M^n \to \mathbb{R}$ be a Morse function. Let μ be a Riemannian metric on M which satisfies the Morse–Smale condition (roughly this means that μ is "generic"). Put

$$C_i^{\mathrm{Morse}}(f; \mathbb{K}) := \mathbb{K}\langle \mathrm{crit}_i(f) \rangle$$

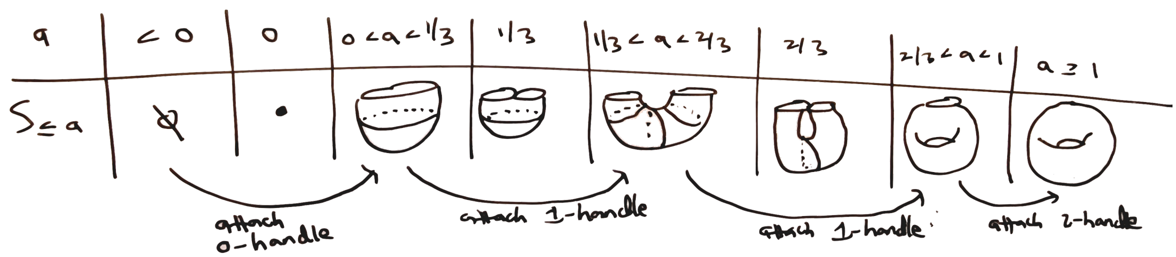

FIGURE 1.3. The sublevel sets $S_{\leq a}$ of the height function $\mathbb{T}^2 \to \mathbb{R}$ as a increases.

denote the module over some chosen coefficient ring $\mathbb{K}$ which is freely generated by the index i critical points of f. We have boundary operators:

$$C_n^{\mathrm{Morse}}(f;\mathbb{K}) \xrightarrow{\partial_n} C_{n-1}^{\mathrm{Morse}}(f;\mathbb{K}) \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_1} C_0^{\mathrm{Morse}}(f;\mathbb{K}).$$

The precise definition of ∂_k is somewhat technical, but roughly it counts "gradient flow lines" between critical points with Morse index differing by 1. Namely, given $p_+ \in C_k^{\mathrm{Morse}}(f;\mathbb{K})$ and $p_- \in C_{k-1}^{\mathrm{Morse}}(f;\mathbb{K})$, we consider maps $u : \mathbb{R} \to M$ satisfying

$$\begin{cases} \lim\limits_{s \to \pm\infty} u(s) = p_{\pm} \\ (\partial_s u)(s) = -(\nabla_\mu f)(u(s)) \text{ for all } s \in \mathbb{R}. \end{cases}$$

Then we have $\partial_k(p_-) = \mathrm{sign}(u)p_+ + \cdots$, with contributions from all other gradient flow lines with negative asymptotic p_-. Here the gradient flow lines are counted modulo translation in the s variable, and with a certain sign, $\epsilon(u) \in \{1, -1\}$ (which we can ignore if say $\mathbb{K} = \mathbb{Z}/2$). We then put

$$H_k^{\mathrm{Morse}}(f;\mathbb{K}) := \ker(\partial_k)/\mathrm{im}(\partial_{k+1}).$$

THEOREM 1.9. *We have* $H_k^{\mathrm{Morse}}(f;\mathbb{K}) \cong H_k^{\mathrm{sing}}(M;\mathbb{K})$.

As a first consequence, observe that we must have

$$|\mathrm{crit}_i(f)| \geq \mathrm{rank}\, H_i(M;\mathbb{Z}).$$

Example 1.10. Any Morse function $f : \mathbb{T}^2 \to \mathbb{R}$ must have a least *two* saddle points, since $H_2(\mathbb{T}^2;\mathbb{Z}) \cong \mathbb{Z}^2$ has rank two. $\Diamond$

A second consequence is:

$$\sum_{i=0}^{n} (-1)^i |\mathrm{crit}_i(f)| = \chi(M),$$

where $\chi(M)$ denotes the Euler characteristic of M. Indeed, a basic fact about chain complexes is that the alternating sum of ranks is unchanged after passing to homology.

Example 1.11. Figure 1.4 depicts a Morse function $S^2 \to \mathbb{R}$ with four critical points. Namely, we have $\mathrm{crit}_0 = \{p_0\}$, $\mathrm{crit}_1 = \{p_1\}$, $\mathrm{crit}_2 = \{p_2, p_3\}$. $\Diamond$

Example 1.12. There exists a Morse function $f : \mathbb{CP}^n \to \mathbb{R}$ all of whose critical points have even Morse index. Hence the Morse complex is

$$C_{2n}^{\mathrm{Morse}}(f;\mathbb{K}) \to 0 \to C_{2n-2}^{\mathrm{Morse}}(f;\mathbb{K}) \to \cdots \to 0 \to C_0^{\mathrm{Morse}}(f;\mathbb{K}),$$

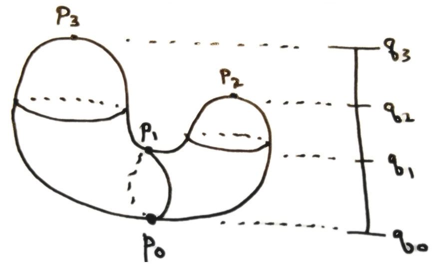

FIGURE 1.4. A Morse function $S^2 \to \mathbb{R}$ with four critical points.

and therefore we have $H_i^{\text{Morse}}(f;\mathbb{K}) = C_i^{\text{Morse}}(f;\mathbb{K})$ for $i = 0, \ldots, n$. In particular,

$$H_i(\mathbb{CP}^n;\mathbb{Z}) \cong \mathbb{Z}^{|\text{crit}_i(f)|}.$$

$\Diamond$

Example 1.13. We have $p \in \text{crit}_i(f)$ if and only if $p \in \text{crit}_{n-i}(-f)$. By dualizing the Morse complex for f, we get a complex

$$C_n^\vee(f) \xleftarrow{\partial_n^\vee} C_{n-1}^\vee(f) \xleftarrow{\partial_{n-1}^\vee} \leftarrow \cdots \xleftarrow{\partial_1^\vee} C_0^\vee(f)$$

which computes the Morse *co*homology of f, and this is precisely identified with the Morse complex of $-f$:

$$C_0(-f) \leftarrow C_1(-f) \leftarrow \cdots \leftarrow C_n(-f).$$

We then have $H^i_{\text{Morse}}(-f) \cong H_{n-i}^{\text{Morse}}(f)$, which is Poincaré duality. $\Diamond$

In Example 1.11, we have $\partial_1(p_1) = p_0 - p_0 = 0$, $\partial_2(p_2) = p_1$, and $\partial_2(p_3) = p_1$. Note that there are e.g gradient flow trajectories from p_2 to p_0, but these do not appear in the appear complex since these critical points index differing by two rather than one. We see that p_1 is a boundary, and hence does not contribute to the homology of S^2. A natural question is whether the critical point p_1, and particularly the gradient flow trajectory from p_2 to p_1, is "visible" in some sense. It turns out that this flow line does not contribute to the Morse homology, but it does contribute to the *filtered* Morse homology. In fact, we will see the filtered Morse homology imposes restrictions on the "geometry" of the function $f : S^2 \to \mathbb{R}$.

LECTURE 2

Quantitative Morse homology

2.1. Toy problem

Let M^n be a closed manifold. For a function $f : M \to \mathbb{R}$, let $||f|| := \max_{x \in M} |f(x)|$ denote its "uniform norm". In the following, let $h : S^2 \to \mathbb{R}$ denote the Morse function with four critical points depicted in Figure 1.4. Let $\mathcal{F}$ denote the set of Morse functions on S^2 with exactly two critical points (note that these necessarily have index 0 and 2 respectively).

PROBLEM 2.1. *What is $\inf_{f \in \mathcal{F}} ||h - f||$? In other words, how well can the Morse function h with four critical points be approximated by a Morse function with only two critical points?*

We give a (partial) answer to the above toy problem with the following:

PROPOSITION 2.2. *We have $\inf_{f \in \mathcal{F}} ||h - f|| \geq \frac{1}{2}(q_2 - q_1)$.*

IDEA OF PROOF. We proceed as follows (with the various ingredients to be introduced shortly):

(1) to f we associate a *filtered chain complex* $C_*^{\mathrm{Morse}}(f)$
(2) to this filtered chain complex $C_*^{\mathrm{Morse}}(f)$ we associate a *persistence module* $V(f)$
(3) to the persistence module $V(f)$ we associate a *barcode* $\mathcal{B}(V(f))$
(4) barcode $\mathcal{B}(V(f))$ has a *boundary depth* $\beta_1(\mathcal{B}(V(f))) \in \mathbb{R}_{\geq 0}$.

Roughly, a persistence module is a collection of $\mathbb{K}$ modules V_t indexed by the real numbers, along with maps $V_s \to V_t$ for all $s < t$ which are coherent under compositions. A barcode is roughly a multiset of interals - see Figure 2.1 for a pictorial representation. The boundary depth of a barcode is by definition the length of the longest finite bar, or zero in the event that there are no finite bars.

In turns out that all of these objects admit natural metrics:

(1) for functions $f, g : M \to \mathbb{R}$ we use the uniform distance $||f - g||$
(2) for two persistence modules V, W, we have the *interleaving distance* $d_{\mathrm{int}}(V, W)$
(3) for two barcodes $\mathcal{B}, \mathcal{B}'$, we have the *bottleneck distance* $d_{\mathrm{bot}}(\mathcal{B}, \mathcal{B}')$.

Moreover, we have:

(1) the association $f \mapsto V(f)$ is a 1-Lipschitz map from the set of Morse functions equipped with the uniform distance to the set of persistence modules equipped with the interleaving distance
(2) the association $V \mapsto \mathcal{B}(V)$ is an isometry from the set of persistence modules equipped with the interleaving distance to the set of barcodes equipped with the bottleneck distance
(3) the boundary depth is a 2-Lipshitz map from the set of barcodes equipped with the bottleneck distace to $\mathbb{R}$ (with the Euclidean metric).

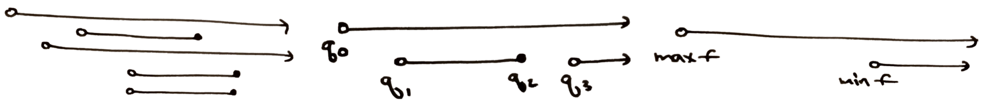

FIGURE 2.1. Left: an example of a barcode, i.e. a multiset of intervals. Note that some are finite and some are infinite, and some intervals may be repeated. Center: the barcode of the Morse function $h : S^2 \to$ depicted in Figure 1.4. Right: the barcode of a Morse function $f : S^2 \to \mathbb{R}$ with exactly two critical points.

Recall that a map $F : X \to Y$ between metric space (X, μ_X) and (Y, μ_Y) is *K-Lipschitz* if $\mu_Y(F(x_1), F(x_2)) \leq K\mu_X(x_1, x_2)$ for all $x_1, x_2 \in X$. Moreover, we claim that the barcodes of h and $f \in \mathcal{F}$ are as depicted in Figure 2.1. As we will see, the infinite bars correspond to homology classes, while the finite bars record features of a more quantitative nature.

Taking this all for granted for the moment, we can now finish off the proof. Namely, for $f \in \mathcal{F}$ we have:

$$\begin{aligned}||f - h|| \geq d_{\text{int}}(V(f), V(h)) = d_{\text{bot}}(\mathcal{B}(V(f)), \mathcal{B}(V(g))) &\geq \tfrac{1}{2}|\beta_1(\mathcal{B}(V(f))) - \beta_1(\mathcal{B}(V(h)))| \\ &= \tfrac{1}{2}|0 - (q_2 - q_1)| \\ &= \tfrac{1}{2}|q_2 - q_1|.\end{aligned}$$

□

2.2. Filtered chain complexes

Let us now introduce some of these objects in the above proof more carefully. We begin with:

DEFINITION 2.3. *A* **filtered chain complex** *is a chain complex (C, ∂) and a collection of subcomplexes $C^{<r} \subset C$, $r \in \mathbb{R}$, such that $C^{<r} \subset C^{<r'}$ for $r < r'$ and $\bigcup_{r \in \mathbb{R}} C^{<r} = C$.*

Here $C^{<r}$ being a subcomplex of C means that from the differential ∂ preserves $C^{<r}$, given an induced differential $C^{<r} \to C^{<r}$ (which we often still denote by ∂). Note that the chain complex in C need not be graded, although it will often be the case that our chain complex has a natural $\mathbb{Z}$ grading (as in singular chains) or at least a $\mathbb{Z}/2$ grading. We typically write C_* instead of C if we wish to emphasize the grading.

Remark 2.4. In the future we will often want to impose some additional technical assumptions on our filtered chain complexes. For example, it is often useful to assume that $C^{<r} = \{0\}$ for r sufficiently small. ◊

Example 2.5. Given a continuous function $f : M \to \mathbb{R}$, put

$$C_k^{\text{sing},<r} := C_k^{\text{sing}}(S_{<a}),$$

with $S_{<a} := f^{-1}((-\infty, a))$. In other words, C_k^{sing} consists of linear combinations of continuous maps from the k-simplex Δ^k to M whose images lie "below" the level set $\{f = a\}$. This makes C_*^{sing} into the structure of a filtered chain complex. ◊

2.3. Filtered Morse theory

If $f : M \to \mathbb{R}$ is a Morse function, then the Morse complex $C_*^{\text{Morse}}(f)$ is also naturally a filtered chain complex. In order to explain this, let us expose a few more details about the Morse complex. Recall that we put $C_k^{\mathrm{Morse}}(f) = \mathbb{K}\langle \mathrm{crit}_k(f)\rangle$. The differential $\partial_k : C_k^{\mathrm{Morse}}(f) \to C_{k-1}^{\mathrm{Morse}}(f)$ is a $\mathbb{K}$-linear map whose value on a basis element $p_- \in \mathrm{crit}_k(f)$ takes the form

$$\partial_k(p_-) = \sum_{p_+ \in \mathrm{crit}_{k-1}(f)} \#\left(\mathcal{M}(p_-;p_+)/\mathbb{R}\right) p_+.$$

Here $\mathcal{M}(p_-;p_+)$ denotes the moduli space of gradient descent flow lines from p_- to p_+, i.e. the space of maps $u : \mathbb{R} \to M$ such that $\partial_s u = -\nabla_\mu f \circ u$ and $\lim_{s\to\pm\infty} u(s) = p_\pm$.

Recall that μ is a chosen Riemannian metric on M satisfies the Morse–Smale condition, defined as follows. Since $-\nabla_\mu f$ is a vector field on a closed manifold M, it has a time-t flow for all $t \in \mathbb{R}$, which we denote by $\phi_t \in \mathrm{Diff}(M)$.

Definition 2.6. *Given a critical point $p \in \mathrm{crit}(f)$, the* **stable manifold** *is*

$$W^s(p) := \{x \in M \mid \lim_{t\to\infty} \phi_t(x) = p\}.$$

Similarly, the **unstable manifold** *is*

$$W^u(p) := \{x \in M \mid \lim_{t\to-\infty} \phi_t(x) = p\}.$$

Roughly, $W^s(p)$ is the set of all points which "descend" to p, and $W^u(p)$ is the set of all points which 'ascend" to p. It turns out that $W^u(p)$ is diffeomorphic to an open disk of dimension $\mathrm{ind}_f(p)$, and $W^s(p)$ is diffeomorphic to an open disk of codimension $\mathrm{ind}_f(p)$, i.e. dimension $n - \mathrm{ind}_f(p)$.

Definition 2.7. *The pair (f,μ) is* **Morse–Smale** *if $W^s(p)$ and $W^u(q)$ intersect transversely for all $p,q \in \mathrm{crit}(f)$.*

Recall that two submanifolds $A, B \subset M$ are said to intersect transversely if for every $x \in A \cap B$ we have $T_xA + T_xB = T_xM$. In particular, this holds vacuously if $A \cap B = \varnothing$.

A consequence of the above definition is that for any $p_-, p_+ \in \mathrm{crit}(f)$ with $\mathrm{ind}(p_-) > \mathrm{ind}(p_+)$, $\mathcal{M}(p_-;p_+)$ is a smooth manifold of dimension $\mathrm{ind}(p_-) - \mathrm{ind}(p_+)$. Moreover, there is a free $\mathbb{R}$-action on $\mathcal{M}(p_-;p_+)$ given simply by translating. Namely, for $r \in \mathbb{R}$ we put $r \cdot u := u_r$, where $u_r(s) = u(s+r)$. Then the quotient manifold $\mathcal{M}(p_-;p_+)/\mathbb{R}$ is also a smooth manifold of dimension $\mathrm{ind}(p_-) - \mathrm{ind}(p_+) - 1$ (not necessarily closed or compact!).

It turns out that we can orient each of the moduli spaces $\mathcal{M}(p_-;p_+)/\mathbb{R}$, at least if M is oriented. In brief, these orientations are naturally induced by picking (arbitrarily) orientations on all of the stable manifolds of critical points of f, which in turn induce orientations on all of the unstable manifolds. In the case $\mathrm{ind}(p_-) = \mathrm{ind}(p_+) + 1$, $\mathcal{M}(p_-;p_+)/\mathbb{R}$ becomes an oriented 0-manifold, which means a collection of points, each with an associated sign $\pm$ attached to it. We then denote by $\#\mathcal{M}(p_-;p_+)$ the signed count of such points. Note that this count is only legitimate if the number of such points is finite. This turns out to be the case by a compactness theorem.

Now observe that $\#\mathcal{M}(p_-;p_+)/\mathbb{R}$ can only be nonzero if $f(p_-) \geq f(p_+)$. Indeed, f decreases along gradient descent trajectories. Therefore if we put

$$C_k^{\mathrm{Morse},<r}(f) := \mathbb{K}\langle p \in \mathrm{crit}_k(f) \mid f(p) < r\rangle,$$

then the Morse differential ∂_k maps $C_k^{\mathrm{Morse},<r}(f)$ to $C_{k-1}^{\mathrm{Morse},<r}(f)$. This makes $C_*^{\mathrm{Morse}}(f)$ into a filtered chain complex.

Remark 2.8. Various nontrivial analytic facts go into the proof that $\partial_{\mathrm{Morse}}$ squares to zero (and similarly for various other structural properties of Morse homology). In brief, the proof proceeds by analyzing the moduli spaces $\mathcal{M}(p_-;p_+)/\mathbb{R}$ whenever $\mathrm{ind}(p_-) = \mathrm{ind}(p_+) + 2$. These are 1-dimensional manifolds which are not typically compact, but they admit natural compactifications in terms of "broken flow lines" going from p_- to an intermediate point q with $\mathrm{ind}(q) = \mathrm{ind}(p_-) - 1$, and then proceeding from q to p_+. Proving this requires a compactness theorem, which states that we really do get something compact after adding in these broken flow lines, and also gluing theorem, which states that this compactification really does have the structure of a manifold with boundary. Combining this with the fact that the (signed) count of boundary points of an (oriented) 1-dimensional manifold with boundary is zero translates into the algebraic relation $(\partial_{\mathrm{Morse}})^2 = 0$ (check this!). $\Diamond$

For two different Morse functions $f, g : M \to \mathbb{R}$, we know that the corresponding Morse homologies $H_*^{\mathrm{Morse}}(f)$ and $H_*^{\mathrm{Morse}}(g)$ are isomorphic, since these are both isomorphic to the singular homology of M. However, the filtered chain complexes $C_*^{\mathrm{Morse}}(f)$ and $C_*^{\mathrm{Morse}}(g)$ are *not* isomorphic. There are natural chain homotopy equivalences

$$C_*^{\mathrm{Morse}}(f) \underset{\psi}{\overset{\phi}{\rightleftarrows}} C_*^{\mathrm{Morse}}(g) \ ,$$

meaning that ϕ and ψ are chain maps and the compositions $\phi \circ \psi$ and $\psi \circ \phi$ are homotopic to the identity, but these do not preserve filtrations. More precisely, we have

$$\phi(C_*^{\mathrm{Morse},<r}(f)) \subset C_*^{\mathrm{Morse},<r+\delta}(g) \quad \text{and} \quad \psi(C_*^{\mathrm{Morse},<r}(g)) \subset C_*^{\mathrm{Morse},<r+\delta}(f),$$

where $\delta = ||f - g||$.

LECTURE 3

Persistence modules and barcodes

We begin with:

Definition 3.1. *A* **persistence module** *(over a field $\mathbb{F}$) is a family of finite dimensional vector spaces $\{V_t\}_{t\in\mathbb{R}}$, along with linear maps $\pi_{s,t} : V_s \to V_t$ for all $s < t$ such that we have $\pi_{t,r} \circ \pi_{s,t} = \pi_{s,r}$ for all $s < t < r$.*

Example 3.2. Given a filtered chain complex $\{C_*^{<r}\}_{r\in\mathbb{R}}$, we get a persistence module by putting $V_t := H(C_*^{<r})$, and letting $\pi_{s,t} : H(C_*^{<s}) \to H(C_*^{<t})$ be the map induced by the inclusion $C_*^{<s} \subset C_*^{<t}$. Note that $\pi_{s,t}$ is not necessarily injective. $\Diamond$

Example 3.3. Given a Morse function $f : M \to \mathbb{R}$, we get a persistence module $V(f)$ with $V(f)_t := H_*^{\mathrm{Morse}}(S_{<t})$, where $S_{<t} := \{x \in M \mid f(x) < t\}$. We also get persistence submodules for each $k = 0, \dots, \dim M$ by looking only at degree k homology classes. Note that we could also replace Morse homology with e.g. singular homology. $\Diamond$

Definition 3.4. *Let us call a persistence module V* **finite type** *if*

- *for all but finitely many $t \in \mathbb{R}$, there is an open neighborhood U of t such that $\pi_{r,s}$ is an isomorphism for all $r, s \in U$ with $r < s$*
- *for some t_0, we have $V_t = \{0\}$ for all $t \le t_0$.*

Definition 3.5. *A persistence module V is* **lower semicontinuous** *if for any $t \in \mathbb{R}$ there exists ϵ such that $\pi_{s,t}$ is an isomorphism for any $t - \epsilon < s < t$.*

The following example plays the role of a basic building block in the theory of persistence modules:

Example 3.6. For $a < b \le \infty$, we have a persistence module $\mathbb{F}(a, b]$, called an "interval module", where

$$\mathbb{F}(a,b]_t = \begin{cases} \mathbb{F} & t \in (a,b] \\ 0 & \text{otherwise,} \end{cases}$$

and we put

$$\pi_{s,t} = \begin{cases} \mathbb{1} & s,t \in (a,b],\ s < t \\ 0 & \text{otherwise.} \end{cases}$$

$\Diamond$

Remark 3.7. Observe that the interval module $\mathbb{F}(a, b]$ is finite type and lower semicontinuous. However, if we were define e.g. $\mathbb{F}[a, b)$ analogously, the lower semocontinuity property would no longer hold. $\Diamond$

Definition 3.8. *A* **morphism** *$F : V \to W$ between persistence modules V, W consists of a family of linear maps $F_t : V_t \to W_t$ for $t \in \mathbb{R}$ such that for each $s, t \in \mathbb{R}$ the following* *diagram commutes:*

$$\begin{tikzcd} V_s \arrow[r] \arrow[d, "F_s"'] & V_t \arrow[d, "F_t"] \\ W_s \arrow[r] & W_t. \end{tikzcd},$$

where the horizontal arrows are the structural maps $\pi_{s,t}$ for the persistence modules V and W.

Example 3.9. There is a natural morphism F of persistence modules $\mathbb{F}(1,2] \to \mathbb{F}(0,2]$ where $F_t = \mathbb{1}$ for $t \in (1,2]$, and $F_t = 0$ otherwise. On the other hand, any morphism from $\mathbb{F}(0,1] \to \mathbb{F}(0,2]$ is necessarily trivial (check this). $\Diamond$

In order to define the interleaving distaince d_{int}, we introduce a little bit more formalism.

Definition 3.10. *Given a persistence module V and $\delta \in \mathbb{R}$, let $V[\delta]$ denote its* **δ-shift***, the persistence module with $V[\delta]_t = V_{t+\delta}$ and structural maps $\pi[\delta]_{s,t} = \pi_{s+\delta,t+\delta}$.*

Similarly, given a morphism $F : V \to W$, we have the **δ-shifted morphism** *$F[\delta] : V[\delta] \to W[\delta]$ defined by $F[\delta]_t = F_{t+\delta}$.*

Note that the structural maps $\pi_{t,t+\delta}$ always define a morphism from V to its shift $V[\delta]$:

Definition 3.11. *Given a persistence module V and $\delta \in \mathbb{R}$, the* **shift morphism** *$\mathrm{Sh}^\delta_V : V \to [\delta]$ is defined by $(\mathrm{Sh}^\delta_V)_t = \pi_{t,t+\delta}$.*

Definition 3.12. *Given $\delta > 0$, two persistence modules V, W are* **δ-interleaved** *if there exist persistence morphisms $F : V \to W[\delta]$ and $G : W \to V[\delta]$ such that the following diagrams commute:*

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[rr, bend right, "\mathrm{Sh}^{2\delta}_V"'] & W[\delta] \arrow[r, "G[\delta]"] & V[2\delta] \end{tikzcd} \qquad \begin{tikzcd} W \arrow[r, "G"] \arrow[rr, bend right, "\mathrm{Sh}^{2\delta}_W"'] & V[\delta] \arrow[r, "F[\delta]"] & W[2\delta] \end{tikzcd}.$$

There is a natural way of composing two morphisms $F : V \to W$ and $G : W \to Q$, and there is also a notion of identity morphism $\mathbb{1}_V$ from a persistence module V to itself. We will say that two persistence modules V and W are *isomorphic* if there exist morphisms $F : V \to W$ and $G : W \to V$ such that $G \circ F = \mathbb{1}_V$ and $F \circ G = \mathbb{1}_W$. One can view a δ-interleaving as an "isomorphism up to an error of δ".

Definition 3.13. *Given two persistence modules V, W, their* **interleaving distance** *is defined by*

$$d_{\text{int}}(V, W) := \inf\{\delta > 0 \mid V, W \textit{ are } \delta\textit{-interleaved}\}.$$

Exercise 3.14. *The interleaving distance is a* **pseudometric***, i.e. it is symmetric and satisfies the triangle inequality. However, a priori $d_{\text{int}}(V, W)$ could be infinity.*

Exercise 3.15. *Given a finite type persistence module V, there exists $t_1 \in \mathbb{R}$ such that $\pi_{s,t}$ is an isomorphism for all $s, t \geq t_1$. Let V_∞ denote V_t for $t \geq t_1$ (or, more precisely, V_∞ is the direct limit of our direct system).*

For finite type persistence modules V, W, show that $d_{\text{int}}(V, W) < \infty$ if and only if $\dim V_\infty = \dim W_\infty$.

In the future, it will sometimes be convenient to use the language of multisets:

Definition 3.16. *A* **multiset** *is a set S together with a function $m : S \to \mathbb{Z}_{\geq 1}$. We view a multiset as a set except that each element can be repeated a finite number of times, and we view $m(x)$ as the "multiplicity" of the element $x \in S$.*

Equivalently, we can view a multiset as a set of pairs $(x, m_x) \in S \times \mathbb{Z}_{\geq 1}$*, where each* x *appears only once.*

THEOREM 3.17 (Normal form theorem for persistence modules). *Let* V *be a persistence module over a field* $\mathbb{F}$ *which is finite type and lower semicontinuous. Then there is a unique finite multiset of left open right closed intervals* $\{(I_i, m_i)\}_{i=1}^N$ *such that*

$$V \cong \mathbb{F}(I_1)^{\oplus m_1} \oplus \cdots \oplus \mathbb{F}(I_N)^{\oplus m_N}.$$

Here each I_i *is of the form* $(a_i, b_i]$ *with* $a_i < b_i \leq \infty$ *and* $m_i \in \mathbb{Z}_{\geq 1}$.

Let $\mathcal{B}(V) := \{(I_i, m_i)\}_{i=1}^N$ denote the multiset as in Theorem 3.17. We call this the *barcode* associated to the persistence module V.

We next define the bottleneck distance between barcodes. Given a barcode $\mathcal{B}$ and $\delta > 0$, let $\mathcal{B}_\delta$ denote the result after throwing away all bars of length $\leq \delta$. For an interval $I = (a, b]$, put $I^{-\delta} := (a - \delta, b + \delta]$.

DEFINITION 3.18. *Two barcodes* $\mathcal{B}, \mathcal{B}'$ *are* δ**-matched** *if there is a bijection between the bars of* $\mathcal{B}_{2\delta}$ *and* $\mathcal{B}'_{2\delta}$ *such that the endpoints of corresponding intervals lie at distance* $< \delta$ *from each other.*

The last condition is equivalent to having $I \subset J^{-\delta}$ and $J \subset I^{-\delta}$ whenever I and J are paired under the bijection.

DEFINITION 3.19. *The* **bottleneck distance** *between two barcodes* $\mathcal{B}, \mathcal{B}'$ *is*

$$d_{\mathrm{bot}}(\mathcal{B}, \mathcal{B}') := \inf\{\delta \mid \mathcal{B}, \mathcal{B}' \text{ are } \delta\text{-matched}\}.$$

EXERCISE 3.20. *The bottleneck distance* d_{bot} *satisfies the axioms of a metric (at least for barcodes at a finite distance from a given one), i.e. it is symmetric, satisfies the triangle inequality, and is nondegenerate.*

By Theorem 3.17, there is a bijective correspondence between persistence modules and barcodes. In fact, we have:

THEOREM 3.21. *The association* $V \mapsto \mathcal{B}(V)$ *is an isometry from the set of persistence modules equipped with the interleaving distance to the set of barcodes equipped with the bottleneck distance.*

In particular, d_{int} is nondegenerate.