VI.3 An Application to Curves

In this section, we use the stability of persistence to prove an inequality that connects the length and total curvature of two curves. We begin by recasting the statement of stability in terms of continuous functions instead of filtrations.

Sublevel sets. Let $\mathbb{X}$ be a topological space and $g : \mathbb{X} \to \mathbb{R}$ a continuous function. Given a threshold $a \in \mathbb{R}$, the *sublevel set* consists of all points $x \in \mathbb{X}$ with function value less than or equal to a, $\mathbb{X}_a = g^{-1}(\infty, a]$. Similar to the complexes in a filtration, the sublevel sets are nested and give rise to sequences of homology groups connected by maps induced by inclusion, one for each dimension. Writing $f_p^{a,b} : \mathsf{H}_p(\mathbb{X}_a) \to \mathsf{H}_p(\mathbb{X}_b)$ for the map from the homology group of the sublevel set for a to that for b, we call its image a *persistent homology group*. The corresponding *persistent Betti number* is $\beta_p^{a,b} = \operatorname{rank} \operatorname{im} f_p^{a,b}$. Furthermore, $a \in \mathbb{R}$ is a *homological critical value* if there is no $\varepsilon > 0$ for which $f_p^{a-\varepsilon,a+\varepsilon}$ is an isomorphism for each dimension p. We assume that g is *tame*, by which we mean that it has only finitely many homological critical values and every sublevel set has only finite rank homology groups. Let $a_1 < a_2 < \ldots < a_n$ be the homological critical values and $b_0 < b_1 < \ldots < b_n$ interleaved values with $b_{i-1} < a_i < b_i$ for $1 \leq i \leq n$. The 1-parameter family of p-th homology groups can therefore be replaced by the finite sequence

$$0 = \mathsf{H}_p^{b_{-1}} \to \mathsf{H}_p^{b_0} \to \mathsf{H}_p^{b_1} \to \ldots \to \mathsf{H}_p^{b_n} \to \mathsf{H}_p^{b_{n+1}} = 0,$$

where $\mathsf{H}_p^b = \mathsf{H}_p(\mathbb{X}_b)$ and the groups at the two ends are added for convenience. Finally, we add $a_0 = -\infty$ and $a_{n+1} = \infty$ to the list of critical values. For $0 \leq i < j \leq n+1$, the *multiplicity* of the pair (a_i, a_j) is now defined as

$$\mu_p^{i,j} \quad = \quad (\beta_p^{b_i,b_{j-1}} - \beta_p^{b_i,b_j}) - (\beta_p^{b_{i-1},b_{j-1}} - \beta_p^{b_{i-1},b_j}).$$

To get the *dimension p persistence diagram* of g we draw each point (a_i, a_j) with multiplicity $\mu_p^{i,j}$. In contrast to the case of filtrations in which contiguous complexes differ by only one simplex, the multiplicities are no longer restricted to 0 and 1 and there can be points at infinity. In particular for a bounded function g the homology classes that get born but do not die correspond to points in the diagrams with ∞ as their second coordinate. With these definitions, we have the following stability result which we state without proof.

STABILITY THEOREM FOR FUNCTIONS. Let $\mathbb{X}$ be a triangulable topological space, $g, g_0 : \mathbb{X} \to \mathbb{R}$ two tame functions, and p any dimension. Then the bottleneck distance between their diagrams satisfies $d_B(\mathrm{Dgm}_p(g), \mathrm{Dgm}_p(g_0)) \leq \|g - g_0\|_\infty$.

Closed curves. We consider a closed curve $\gamma : \mathbb{S}^1 \to \mathbb{R}^2$, with or without self-intersections. Assuming γ is smooth, we have derivatives of all orders. The *speed at a point* $\gamma(s)$ is the length of the velocity vector, $\|\dot{\gamma}(s)\|$. We can use it to compute the length as the integral over the curve,

$$L(\gamma) \quad = \quad \int_{s \in \mathbb{S}^1} \|\dot{\gamma}(s)\| \, \mathrm{d}s.$$

It is convenient to assume a constant speed parametrization, that is, $\varrho = \|\dot{\gamma}(s)\| = \frac{1}{2\pi} L(\gamma)$ for all $s \in \mathbb{S}^1$. With this assumption, the *curvature at a point* $\gamma(s)$ is the norm of the second derivative divided by the square of the speed, $\kappa(s) = \|\ddot{\gamma}(s)\|/\varrho^2$. One over the curvature is the radius of the circle that best approximates the shape of the curve at the point $\gamma(s)$. To interpret this formula geometrically, we follow the velocity vector as we trace out the curve. Since its length is constant it sweeps out a circle of radius ϱ, as illustrated in Figure VI.10. The curvature is the speed at which the unit tangent vector

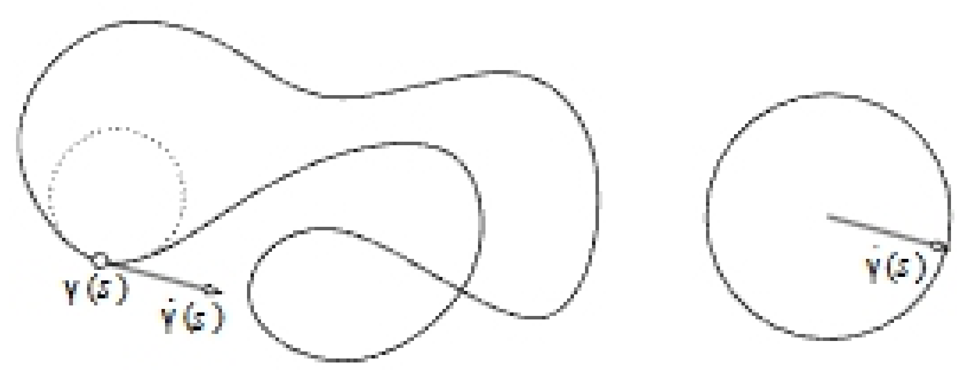

Figure VI.10: A curve with constant speed parametrization and its velocity vector sweeping out a circle with radius equal to the speed.

sweeps out the unit circle as we move the point with unit speed along the curve. This explains why we divide by the speed twice, first to compensate for the length of the velocity vector and second to compensate for the actual speed. The *total curvature* is the distance traveled by the unit tangent vector,

$$K(\gamma) \quad = \quad \int_{s \in \mathbb{S}^1} \varrho \kappa(s) \, \mathrm{d}s.$$

As an example consider the constant speed parametrization of the circle with radius r, $\gamma(s) = rs$. Writing a point of the unit circle in terms of its angle we get

$$s = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}, \quad \gamma(s) = \begin{bmatrix} r\cos\varphi \\ r\sin\varphi \end{bmatrix}, \quad \dot{\gamma}(s) = \begin{bmatrix} -r\sin\varphi \\ r\cos\varphi \end{bmatrix}.$$

The constant speed is therefore $\varrho = r$ and the length is $L(\gamma) = \int r \, \mathrm{d}s = 2\pi r$. The second derivative is $\ddot{\gamma}(s) = -\gamma(s)$ and the curvature is $\kappa(s) = \frac{1}{r}$ which is independent of the point on the circle. The total curvature is $K(\gamma) = \int \frac{r}{r}\,\mathrm{d}s = 2\pi$, which is independent of the radius. Indeed, the unit tangent vector travels once around the unit circle, no matter how small or how big the parametrized circle is.

Integral geometry. The length and total curvature of a curve can also be expressed in terms of integrals of elementary quantities. We begin with the length. Take a unit length line segment in the plane. The lines that cross the line segment at an angle φ form a strip of width $\sin\varphi$. Integrating over all angles gives $\int_{\varphi=0}^{\pi} \sin\varphi\,\mathrm{d}\varphi = [-\cos\varphi]_0^{\pi} = 2$. In words, the integral of the number of intersections over all lines in the plane is twice the length of the line segment. Since we can approximate the curve by a polygon whose total length approaches that of the curve, the same holds for the curve. To express this result, we write each line as the preimage of a linear function. Given a direction $u \in \mathbb{S}^1$ let $h_u : \mathbb{R}^2 \to \mathbb{R}$ be defined $h_u(x) = \langle u, x\rangle$. The line with normal direction u and offset z is $h_u^{-1}(z)$. The intersections between γ and this line corresponds to the preimage of the composition, $h^{-1}(z)$, where $h = h_u \circ \gamma$. The length of the curve is therefore as given by the Cauchy-Crofton formula,

$$L(\gamma) = \frac{1}{4}\int_{u\in\mathbb{S}^1}\int_{z\in\mathbb{R}} \mathrm{card}\,(h^{-1}(z))\,\mathrm{d}z\,\mathrm{d}u.$$

To get an alternative interpretation of the total curvature, we again consider a direction $u \in \mathbb{S}^1$ and the height function in that direction, $h_u : \mathbb{S}^1 \to \mathbb{R}$ defined by $h_u(s) = \langle u, \gamma(s)\rangle$. For generic directions u, this height function has a finite number of minima and maxima, as illustrated in Figure VI.11. Recall that the

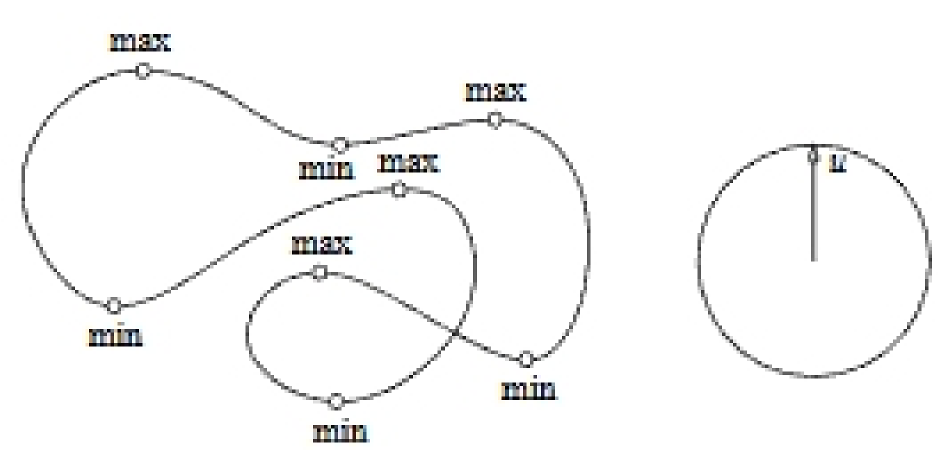

Figure VI.11: The vertical height function defined on the curve has four local minima which alternate with the four local maxima along the curve.

total curvature is the length traveled by the unit tangent vector. Equivalently, it is the length traveled by the outward unit normal vector. The number of maxima of h_u is the number of times the unit normal passes over $u \in \mathbb{S}^1$ and