A Point Calculus for Interlevel Set Homology*

Paul Bendich[†], Sergio Cabello[‡] and Herbert Edelsbrunner[§]

Abstract

The theory of persistent homology opens up the possibility to reason about topological features of a space or a function quantitatively and in combinatorial terms. We refer to this new angle at a classical subject within algebraic topology as a *point calculus*, which we present for the family of interlevel sets of a real-valued function. Our account of the subject is expository, devoid of proofs, and written for non-experts in algebraic topology.

Keywords. Continuous functions, interlevel sets, homology, vector spaces, persistence diagrams, exact sequences.

1 Introduction and Background

We write this paper to describe a recent development within algebraic topology that makes is possible to reason combinatorially about algebraic concepts. The intention is not to divorce the topology from the algebra, but rather to find a more direct, combinatorial route at understanding the algebra. In principle, this opens a road to topology in terms of homology without going through the demanding rigor these subjects usually require. Such a short-cut has of course its dangers, but we feel the risk is worth taking.

From geometry to topology to algebra. Geometry studies properties of a space that are invariant under transformations preserving a measure, such as volume, area, or distance. It is an ancient subject which underlies much of mathematics and is also widely used in non-mathematical disciplines. Indeed, geometric thinking is an important pillar of human comprehension of this world. A first step toward a more abstract, topological understanding of space was taken by Leonhard

Euler in the 18th century, when he noticed that the number of vertices, edges, and faces of every (3-dimensional) convex polytope satisfies

$$\#\text{vertices} - \#\text{edges} + \#\text{faces} \quad = \quad 2; \qquad (1)$$

see [13]. Today, this is considered the first invariant in algebraic topology. More than a century later, Enrico Betti introduced what are now known as *Betti numbers*, one per dimension, which quantify the connectivity of a space in each dimension [4]. These numbers are still of central importance in modern topology today. The big breakthrough toward the establishment of topology as a field within mathematics came with the work of Henri Poincaré around the year 1900; see [20]. In a series of papers, he laid the foundations by realizing, among other things, that the alternating sum of Betti numbers is an algebraic invariant of a space; that is: the alternating sum does not change when we deform the space, provided we refrain from cutting and gluing. Today, the alternating sum of Betti numbers is known as the *Euler characteristic* of the space, and together with the easy realization that it equals the alternating sum of faces, we get a generalization of (1) known as the *Euler-Poincaré formula*.

In topology, two spaces are considered to be the same when there is a *homeomorphism* between them, that is, a continuous bijective mapping whose inverse is also continuous. This is where topology takes its departure from geometry: a small circle is the same as a big circle or indeed as (the boundary of) a square. To establish sameness, it suffices to exhibit a homeomorphism, while to establish the opposite requires a proof that no homeomorphism exists. Commonly, this is done by constructing an algebraic invariant, such as the alternating sum of Betti numbers, that differentiates between the two spaces and thus implies the impossibility of a homeomorphism. Under the influence of Emmy Noether, the initial emphasis on numbers was changed in favor of a more complete, algebraic characterization with Abelian groups, the *homology groups* whose ranks are the Betti numbers. Crucially, her group formulation allows the systematic transfer of algebraic information from one space to another via continuous maps, a concept known as *functoriality*, which lies at the heart of our approach to homology, as we will explain shortly.

†IST Austria (Institute of Science and Technology Austria), Klosterneuburg, Austria.

‡IMFM and FMF, University of Ljubljana, Ljubljana, Slovenia. Research was partially done while visiting IST Austria.

§IST Austria (Institute of Science and Technology Austria), Klosterneuburg, Austria, Departments of Computer Science and of Mathematics, Duke University, Durham, North Carolina, and Geomagic, Research Triangle Park, North Carolina.

From homology to persistence to applications. For the next step, we take broad inspiration from a mathematical theory named after Marston Morse, see eg. the classic book on the subject by John Milnor [17]. Using tools from differential geometry, this theory gives a vast generalization of the early observation by Arthur Cayley [8] and James Clerk Maxwell [16] that for a smooth surface generically placed in 3-dimensional space, the number of peaks minus the number of passes plus the number of pits equals the alternating sum of Betti numbers. In other words, we can recover the topological information of the surface by considering the sequence of contour lines or, alternatively, the *critical points* at which all partial derivatives of the height function vanish.

Combining the Morse theoretic outlook with the algebraic theory of homology, we are now only two small steps away from the notion of persistent homology. One step is the focus on the algebraic expression of the change that happens to a contour line when it sweeps over a critical point, the other is the realization that critical points generically come in pairs, and that there is valuable information to be gained from these pairs. Indeed, the measurement of the difference between the values of the paired critical points forms a bridge from topology back to geometry, an ingredient that brought about a watershed when topological theories suddenly became accessible to applications in the sciences and engineering.

To apply persistent homology, we need both a space and a real-valued function. In contrast to Morse Theory, the function is no longer subordinate to the space and moves into the center of the investigation. Different application areas distinguish themselves in the spaces of interest and the functions they employ. In *data analysis*, the data is often put inside a finite-dimensional Euclidean space, and a common function is the Euclidean distance to the nearest data point. The analysis then proceeds through the sequence of spaces formed by progressively thickening the data set; see eg. [7, 14]. In *visualization*, the space is typically our ordinary 2- or 3-dimensional Euclidean space, and the function is a measurement of interest, such as the heat distribution in a combustion chamber or the proportionality factor inside mixing fluids [19]. In *shape analysis*, the space can be the shape itself, while the function is chosen to emphasize interesting features; see eg. [15]. In *structural molecular biology*, the space may be the surface of a protein and the function may be a measurement of local protrusions and cavities [1, 9]. Our running example is an instance of perhaps the simplest setup, that of a *height function*. Here, we embed a surface in Euclidean space and map each point to its last Cartesian coordinate.

Outline. In Section 2, we explain how to draw a diagram of dots that captures persistent homology. We also provide the necessary background in topology. In Section 3, we de-scribe how we can read the homology of interlevel sets directly from the persistence diagram of a function. In Section 4, we explain how the diagram can be used to reason about the interaction between different interlevel sets using functoriality. We conclude in Section 5 with a discussion of this paper's contributions.

2 Drawing

In this section, we describe how we make a record of the topological properties of a real-valued function. The details will be important, and we will use a running example to illustrate all steps. As we will see in the subsequent sections, careful book-keeping will give handsome returns.

(Absolute) homology. We now give an intuitive description of the theory of homology. A formal definition requires a good deal of algebra, and for that we refer the reader to a standard textbook, such as [18]. Here, we will confine ourselves to homology groups for coefficients from the binary field, which consists of elements 0 and 1 and addition modulo 2. Our results generalize to other fields, such as the real or the rational numbers, but not to non-fields, such as the integers. Given a space $\mathbb{X}$, we have a homology group $H_p(\mathbb{X})$ for each integer dimension p, but we will restrict our attention to $p = 0, 1, 2$.

Briefly, a p-dimensional homology class is an equivalence class of p-cycles, as we now discuss. A 0-cycle is a collection of points from $\mathbb{X}$, a 1-cycle is a collection of closed curves in $\mathbb{X}$ (possibly with intersections and self-intersections), and a 2-cycle is a collection of closed surfaces in $\mathbb{X}$ (again possibly with intersections and self-intersections). Given any two p-cycles α and γ, we can formally add them to form the p-cycle $\alpha + \gamma$, which we can think of as the closure of their symmetric difference. We say that two p-cycles are *homologous* within $\mathbb{X}$ if their sum forms the boundary of a $(p + 1)$-dimensional subspace of $\mathbb{X}$; for example, two points are homologous if we can draw a path between them entirely within $\mathbb{X}$. A p-cycle α is said to *bound* if it is itself the boundary of a $(p + 1)$-dimensional subspace of $\mathbb{X}$; in this case, we think of α as being homologous to zero and refer to it as a *trivial p-cycle*. A p-dimensional *homology class* is a collection of mutually homologous p-cycles. For intuition, one might imagine a 0-dimensional class as a connected component of $\mathbb{X}$, a 1-dimensional class as a loop going around a tunnel, and 2-dimensional class as a closed surface enclosing a void inside $\mathbb{X}$. We can add p-dimensional homology classes to produce other ones, and thus the set of p-dimensional homology classes forms a vector space $H_p(\mathbb{X})$, called the *p-th homology group* of the space $\mathbb{X}$. We denote the rank of this group by $\beta_p(\mathbb{X})$ and call it the *p-th Betti number* of $\mathbb{X}$. Often we wish to refer to homology without sticking to a particular

dimension p. For this reason, we sometimes use the direct-sum notation $\mathsf{H}(\mathbb{X}) = \bigoplus_p \mathsf{H}_p(\mathbb{X})$, although we stress this is only used to think of all homology groups at once, and we never add two cycles of different dimension to one another.
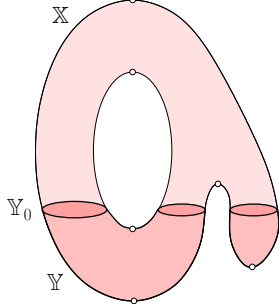

Figure 1: A (hollow) torus $\mathbb{X}$, a closed subspace $\mathbb{Y}$ of $\mathbb{X}$, and the boundary $\mathbb{Y}_0$ of $\mathbb{Y}$. The way we have drawn it, $\mathbb{Y}_0$ is a level set of the height function on the torus, and $\mathbb{Y}$ is a sublevel set of that function.

For example, suppose that $\mathbb{X}$ is the torus and $\mathbb{Y} \subseteq \mathbb{X}$ is the shaded region as drawn in Figure 1. We will compute the Betti numbers of both spaces. The torus is connected, while $\mathbb{Y}$ has two connected components; thus $\beta_0(\mathbb{X}) = 1$ and $\beta_0(\mathbb{Y}) = 2$. Within the torus, we can find at least two non-bounding 1-cycles: the horizontal loop α formed by the leftmost circle at the top of the shaded region, and the vertical loop γ which goes entirely around the hole. With some thought, one can see that every other non-bounding 1-cycle must be homologous to exactly one of α, γ, or $\alpha + \gamma$. In other words, $\beta_1(\mathbb{X}) = 2$. On the other hand, $\beta_1(\mathbb{Y}) = 1$, since α still forms a non-bounding 1-cycle in $\mathbb{Y}$, but γ does not. Note that the 1-cycle at the top of the rightmost component of $\mathbb{Y}$ is trivial, since it is the boundary of the component itself. Finally, the torus is hollow and therefore surrounds a void, so $\beta_2(\mathbb{X}) = 1$, while $\beta_2(\mathbb{Y}) = 0$.

Relative homology. We also give an intuitive description of the relative homology groups $\mathsf{H}_p(\mathbb{Y}, \mathbb{Y}_0)$, associated to any nested pair of spaces $(\mathbb{Y}, \mathbb{Y}_0)$ such that $\mathbb{Y}_0$ is a closed subspace of $\mathbb{Y}$. In brief, one adjusts the above definitions of cycle, boundary, and homology, making them relative to $\mathbb{Y}_0$. More precisely, we define the boundary *relative to* $\mathbb{Y}_0$ of any subspace γ to be the portion of the boundary of γ outside $\mathbb{Y}_0$. A *relative p-cycle* is a p-dimensional subspace α of $\mathbb{Y}$ whose boundary relative to $\mathbb{Y}_0$ is empty; of course, this includes the possibility that the boundary of α is actually empty. Adding relative p-cycles exactly as above, we say that two relative p-cycles α and γ are *homologous* if there exists a $(p+1)$-dimensional subspace of $\mathbb{Y}$ whose boundary relative to $\mathbb{Y}_0$ is $\alpha + \gamma$. Furthermore, a relative p-cycle α *bounds* if there exists such a subspace whose boundary relative to $\mathbb{Y}_0$ is exactly α. For example, a point $y \in \mathbb{Y}$ is a relative 0-cycle, and

it bounds iff there is a path within $\mathbb{Y}$ that connects y to some point in $\mathbb{Y}_0$. A p-dimensional *relative homology class* is a set of mutually homologous relative p-cycles. The collection of such classes, along with the obvious additive structure, forms the vector space $\mathsf{H}_p(\mathbb{Y}, \mathbb{Y}_0)$, which we call the *$p$-th relative homology group* of the pair $(\mathbb{Y}, \mathbb{Y}_0)$, and its rank, $\beta_p(\mathbb{Y}, \mathbb{Y}_0)$ is the *p-th relative Betti number* of the pair.

For example, suppose again that $\mathbb{Y}$ is the shaded region drawn in Figure 1 and that $\mathbb{Y}_0$ consists of the three circles that make up its upper boundary. Every point in $\mathbb{Y}$ can be connected within $\mathbb{Y}$ to some point in $\mathbb{Y}_0$. Hence all relative 0-cycles are trivial and $\beta_0(\mathbb{Y}, \mathbb{Y}_0) = 0$. We can also compute $\beta_1(\mathbb{Y}, \mathbb{Y}_0) = 1$; as representative for the only non-trivial 1-dimensional relative homology class, we could take a path connecting the two leftmost circles in $\mathbb{Y}_0$. As for dimension 2, we find $\beta_2(\mathbb{Y}, \mathbb{Y}_0) = 2$, with two independent relative 2-cycles formed by the two connected components of $\mathbb{Y}$. The symmetry $\beta_p(\mathbb{Y}) = \beta_{2-p}(\mathbb{Y}, \mathbb{Y}_0)$ is not accidental but rather an example of Lefschetz Duality, which implies that $\beta_p(\mathbb{Y}) = \beta_{d-p}(\mathbb{Y}, \partial\mathbb{Y})$ whenever $\mathbb{Y}$ is a d-manifold with boundary $\partial\mathbb{Y}$; see eg. [18, Chapter 8]. As a special case, we have Poincaré Duality, which states that $\beta_p(\mathbb{X}) = \beta_{d-p}(\mathbb{X})$ if $\mathbb{X}$ is a d-manifold without boundary; we see an example of this when $\mathbb{X}$ is the torus above.

Finally, suppose that we have another pair of spaces $(\mathbb{X}, \mathbb{X}_0)$ such that $\mathbb{Y} \subseteq \mathbb{X}$, $\mathbb{Y}_0 \subseteq \mathbb{X}_0$, and $\mathbb{X} - \mathbb{X}_0 = \mathbb{Y} - \mathbb{Y}_0$. For example, $\mathbb{X}$ could be the torus in Figure 1 and $\mathbb{X}_0$ could be everything at the level of $\mathbb{Y}_0$ and above. Then the Principle of Excision tells us that replacing the pair $(\mathbb{Y}, \mathbb{Y}_0)$ with $(\mathbb{X}, \mathbb{X}_0)$ has no effect on relative homology, that is, $\mathsf{H}_p(\mathbb{X}, \mathbb{X}_0)$ and $\mathsf{H}_p(\mathbb{Y}, \mathbb{Y}_0)$ are isomorphic for all p; see eg. [18, Chapter 3].

Ordinary persistence. The reader may have noticed that homology classes do not come with a notion of size; for example, if we were to make the hole in the torus from Figure 1 smaller, the Betti numbers would be unchanged. To provide a richer picture, we employ persistent homology which, briefly, takes a compact (closed and bounded) topological space $\mathbb{X}$ along with a real-valued function f and returns the size, as measured by f, of each homology class in $\mathbb{X}$, as well as the size of transient homological features created by the function itself. For a precise algebraic definition, see for example [10, Chapter VII]. We imagine persistent homology as a two-stage filtering process. In the first stage, we filter $\mathbb{X}$ via the *sublevel sets* $\mathbb{X}_r = f^{-1}(-\infty, r]$ of f, where r can be any real number. As r runs from negative to positive infinity, the sublevel sets include into one another and get bigger, eventually forming the space $\mathbb{X}$ itself. During this process, homology classes appear and disappear; persistent homology tracks and quantifies this evolution, as we now explain via an example.
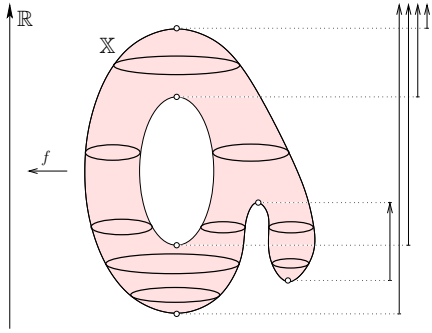
Figure 2: Height function on the nosy torus. We observe five homology classes during the filtration, drawn as vertical bars on the right. The four that go to infinity correspond to the essential homology classes of the torus.

We consider the function $f : \mathbb{X} \to \mathbb{R}$ depicted in Figure 2, where $\mathbb{X}$ is the same torus from before and f measures height in the vertical direction. A value r is critical if $f^{-1}(r)$ contains a point where the tangent plane is horizontal, and regular otherwise. It should be intuitively clear that the only homological changes happen when we pass the six critical values of f; indeed, this is one of the fundamental principles of Morse Theory; see eg. [17]. Passing the minimum such value changes $\mathbb{X}_r$ from the empty set to something homeomorphic to a disk; we say that a component is *born* at this critical value. Another component is born upon passing the second minimum, and this component remains distinct until it merges with the main component at the fourth critical value. At this time, the component that was born later *dies* and we pair the second and fourth critical values. We represent the component as a directed line between the two relevant critical values, as shown in Figure 2. Two 1-dimensional classes are born at the third and fifth critical values, and a 2-dimensional class is born at the maximum. We call these latter three classes, along with the original component, *essential classes*, as they represent the actual homology of $\mathbb{X}$. The other component is an *inessential class* caused by the function itself. Despite their name, inessential classes may actually represent very interesting features of the space; here the inessential component picks up the protrusion on the right side of the torus.

Extended persistence. It should now be clear what we mean by the size of an inessential class, as measured by the function f, namely the difference between the critical values that gave birth and death. The essential classes, on the other hand, have yet to be measured effectively, since they have a birth but no death value. To correct this, we consider, in the second filtering stage, pairs of spaces $(\mathbb{X}, \mathbb{X}^r)$, where $\mathbb{X}^r = f^{-1}[r, \infty)$ is a *superlevel set*, and we watch

the changes in relative homology that occur as r decreases from positive to negative infinity. At the end of this process, the superlevel set equals $\mathbb{X}$ itself, so there can be no remaining live homology classes; in particular, all of the essential classes will obtain a death value. For example, the essential component of our torus dies at the maximum, since this is the first point where it becomes a relative boundary. We thus pair the global minimum with the global maximum, and represent the essential component as a directed curve which starts at the birth value, moves up to infinity, and then ends at the death value; see Figure 3, which also displays the birth and death values for the other three essential classes.
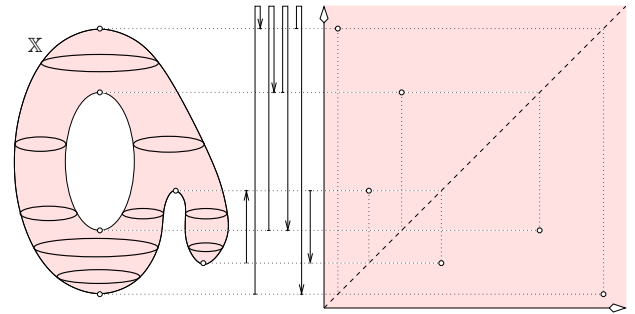

Figure 3: From left to right: the height function of the nosy torus, the 'curvy' barcode in which the extended classes go to infinity and then come back, and the traditional persistence diagram in which every interval is drawn as a point.

Every class which is born at some point of the two-stage process will eventually die, and is thus associated with a pair of critical values. These pairs fall into three types. *Ordinary* pairs have birth and death during the first stage, *extended* pairs have birth during the first stage and death during the second, while *relative* pairs come entirely within the second stage. For example, the 2-dimensional relative class represented by the downward-directed line segment in Figure 3 gives rise to a pair of this latter type. Note that death values can be lower than birth values; this is always true for relative pairs, and sometimes true for extended pairs. Whatever the case, we define the *persistence* of a class to be the absolute difference between its birth and death values.

Diagram. We represent this homological information in compact form via a multi-set of dots in the 2-dimensional plane, containing one dot (r_i, r_j) for each pair of critical values r_i, r_j, such that a class is born at r_i and dies at r_j. We denote this multi-set by $\mathrm{Dgm}(f)$, and call it the *persistence diagram* of the function f. By $\mathrm{Dgm}_p(f)$ we mean the sub-diagram of $\mathrm{Dgm}(f)$ corresponding to p-dimensional homology classes. Within each $\mathrm{Dgm}_p(f)$ are contained the *ordinary*, *extended*, and *relative subdiagrams*, $\mathrm{Ord}_p(f)$,

4

$\mathrm{Ext}_p(f)$, and $\mathrm{Rel}_p(f)$, defined in the obvious way. For technical reasons apparent below, it is convenient to include in $\mathrm{Ord}_p(f)$ and $\mathrm{Rel}_p(f)$ infinitely many copies of every dot along the major diagonal. Note that dots in the ordinary diagrams are above the major diagonal, dots in the extended diagrams can be on either side, and dots in the relative digram are below. For each type of dot, its vertical distance from the diagonal equals the persistence of the associated class. The persistence diagram for our height function f is displayed in Figure 3. The reader will notice an obvious symmetry across the major diagonal; this is a consequence of the dualities mentioned above, and it will happen whenever the domain $\mathbb{X}$ is a manifold without boundary.
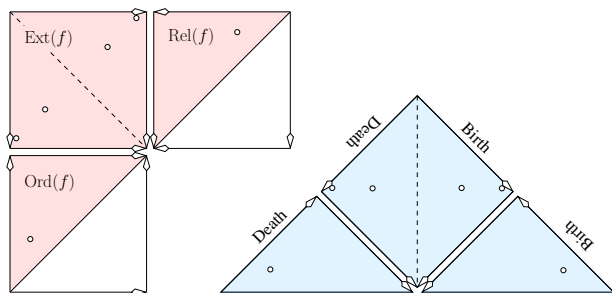

Figure 4: The three overlaid subdiagrams in Figure 3 are unfolded by flipping pages: keeping $\mathrm{Ord}(f)$ fixed, $\mathrm{Ext}(f)$ flips up, followed by $\mathrm{Rel}(f)$ which flips up and then to the right. Finally, we clip the ordinary and relative subdiagrams along the diagonal and rotate the entire design by 45 degrees so it rests on its long side. The arrows of the diagram go from negative to positive infinity.

We will see in Section 3 that in some contexts the three subdiagrams play very different roles. To avoid confusion, we often redraw $\mathrm{Dgm}(f)$ as two triangles and a diamond, with coordinate systems as shown in Figure 4. Note that the horizontal line along the bottom of the new drawing corresponds to the major diagonal in both the ordinary and relative subdiagrams, while the vertical dashed line across the middle diamond is the major diagonal for the extended subdiagram. The middle diamond contains one dot per essential class, and thus in our example there are four dots. In cases in which the space $\mathbb{X}$ is contractible, for example when we have an image considered as a real-valued function on the unit-cube, this diamond will contain exactly one dot.

Stability. The persistence diagram $\mathrm{Dgm}(f)$ is a stable representation of the homological information carried by the function f, in the sense that replacing f by a noisy version g will result in a diagram which is not too different. There are a variety of ways to make this statement more precise. Most simply, we consider a bijection between $\mathrm{Dgm}(f)$ and $\mathrm{Dgm}(g)$, recalling that the ordinary and relative subdiagrams contain infinitely many copies of each major diagonal dot,

and we find the length of the longest edge in the matching. We then minimize this length over all possible matchings, and call this minimum the *bottleneck distance* between the two diagrams, denoted as $W_\infty(\mathrm{Dgm}(f), \mathrm{Dgm}(g))$. Since the distance between dots in different subdiagrams is infinity, it suffices to consider matchings within each subdiagram. One can then show, under mild conditions on f and g, that $W_\infty(\mathrm{Dgm}(f), \mathrm{Dgm}(g))$ is bounded from above by the L_∞-distance between the two functions; see [10, Chapter XIII]. There is also a result which bounds the Wasserstein distance between the diagrams, although this requires stronger assumptions on both the space and the two functions.
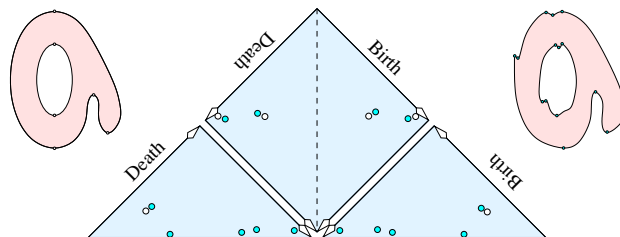

Figure 5: On the right, we see the distorted nosy torus that mimics the effect of a slight perturbation on the height function. In the middle, we see the persistence diagrams of the two height functions superimposed. The best bijection matches each of the six old dots with a nearby new dot, and it matches the remaining eight new dots with points on the diagonal.

For example, suppose we embed the torus $\mathbb{X}$ into $\mathbb{R}^3$ as shown in Figure 5 on the right. Letting $g : \mathbb{X} \to \mathbb{R}$ measure height in the vertical direction, we note that g is indeed a noisy version of our original height function f. The persistence diagrams $\mathrm{Dgm}(g)$ and $\mathrm{Dgm}(f)$ are overlaid in the middle of the same figure. Note that the ordinary and relative subdiagrams of g contain more dots than those of f, but these extra dots are all close to the major diagonal. The two extended subdiagrams will of course have the same number of dots, since the actual homology of $\mathbb{X}$ is unaffected by our choice of real-valued function.

3   Reading

In this section, we discuss how one can mine a great deal of homological information from the persistence diagram of a real-valued function f on a topological space. By its definition, the diagram reflects a filtration by sublevel sets, so it is not surprising that one can read their homology groups from $\mathrm{Dgm}(f)$. But the diagram also displays the homology of every *level set* $f^{-1}(a)$, as well as the absolute and relative homology of every *interlevel set* $f^{-1}[a, b]$. Level sets and sublevel sets are of course just special cases of interlevel sets, but we explain the results separately for greater clarity.

For simplicity, we will assume that the extremes a, b of the intervals are regular values of f.

Abstract vector space. A slight algebraic digression is needed in order to even make the statements in this section. First, we recall that every absolute or relative homology group is a vector space, and as such has a basis, or a set of vectors that uniquely generate all other vectors in the vector space. In [11], the authors show that the notions of birth and death allow one to choose bases for the groups $\mathsf{H}(\mathbb{X}_r)$ in a compatible way. This means that for $r < s$, we obtain the basis of $\mathsf{H}(\mathbb{X}_s)$ from a subset of the basis of $\mathsf{H}(\mathbb{X}_r)$ by adding new basis vectors, but without any need for recombining already chosen basis vectors. In a nutshell, birth corresponds to adding a basis vector and death corresponds to deleting one. Furthermore, these basis vectors are in one-to-one correspondence with the dots in $\mathrm{Dgm}(f)$.

More precisely, one can think of the dots themselves as defining a basis $\mathcal{B}$ for an abstract vector space V. One then defines a special set $\mathcal{V}$ of linear subspaces of V, namely those spanned by all subsets of vectors from $\mathcal{B}$, and one then proves that each $\mathsf{H}(\mathbb{X}_r)$ is isomorphic to a particular element of $\mathcal{V}$; this isomorphism, which we discuss below, defines the chosen basis. This result was significantly extended in [3], following work in [5] and [6]: it turns out that the dots in $\mathrm{Dgm}(f)$ also correspond, with some dimensional modification, to bases for a much wider class of absolute and relative homology groups. We explain this in several stages, although we omit proofs.

Sublevel sets. Given a real number b, we stress that the homology of the sublevel set $\mathbb{X}_b$ can be quite different from that of $\mathbb{X}$: there can be essential classes in $\mathbb{X}$ that are born after b, and there can be inessential classes which at b have yet to die. Nonetheless, a basis for the homology group $\mathsf{H}(\mathbb{X}_b)$ can be read off $\mathrm{Dgm}(f)$. For each regular value b and each dimension p, we define the following sub-multi-set of $\mathrm{Dgm}_p(f)$:

$$
\begin{aligned}
\mathcal{W}_p(-\infty, b] \;=\; & \{(x, y) \in \mathrm{Ord}_p(f) \mid x < b < y\} \\
\sqcup\; & \{(x, y) \in \mathrm{Ext}_p(f) \mid x < b\}.
\end{aligned}
$$

In other words, $\mathcal{W}_p(-\infty, b]$ contains dots that lie within the rectangular shaded region on the left side of Figure 6. Note that dots in the $\mathrm{Ord}_p(f)$ portion of this region correspond to p-dimensional classes that are inessential but still alive at b, while dots in the $\mathrm{Ext}_p(f)$ portion give essential p-dimensional classes that have formed before b. Hence the following result should not be too surprising.

CLOSED SUBLEVEL SET LEMMA. For each regular value b and each dimension p, there exists a natural isomorphism that takes $\mathsf{H}_p(\mathbb{X}_b)$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{W}_p(-\infty, b]$.
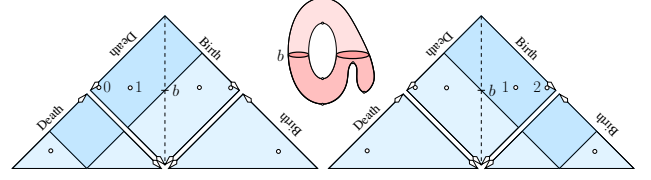

Figure 6: Left: the rectangular region of the closed sublevel set that cuts the hole in the middle. Right: the rectangular region of the corresponding open sublevel set.

For this result and several others to follow, it is illustrative to consider how the sets $\mathcal{W}_p$ change as the extreme values of the intervals change continuously. For example, if we let b go to ∞, or indeed just set it higher than the maximum value of f, then the $\mathrm{Ord}_p(f)$ portion of $\mathcal{W}_p(-\infty, b]$ disappears and we are left with $\mathcal{W}_p(-\infty, +\infty) = \mathrm{Ext}_p(f)$. That is, the essential classes make up the actual homology of $\mathbb{X}$. The second stage of the filtration concerns pairs of spaces $(\mathbb{X}, \mathbb{X}^b)$. We now explain how to read the relative homology of such pairs from the diagram. For each regular value b and each dimension p, we define:

$$
\begin{aligned}
\mathcal{W}_p(-\infty, b) \;=\; & \{(x, y) \in \mathrm{Ext}_p(f) \mid y < b\} \\
\sqcup\; & \{(x, y) \in \mathrm{Rel}_p(f) \mid y < b < x\}.
\end{aligned}
$$

Here we have also introduced the notation $(-\infty, b)$ to stand for the pair of spaces $((-\infty, b], \{b\})$, whose preimage under f has, by excision, the same relative homology as the pair $(\mathbb{X}, \mathbb{X}^b)$. Again, $\mathcal{W}_p(-\infty, b)$ is the sub-multi-set of dots within a particular rectangular region of $\mathrm{Dgm}_p(f)$, this time the shaded one on the right side of Figure 6.

OPEN SUBLEVEL SET LEMMA. For each regular value b and each dimension p, there exists a natural isomorphism that takes $\mathsf{H}_p(\mathbb{X}, \mathbb{X}^b)$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{W}_p(-\infty, b)$.

Level sets. The diagram also encodes the homology of all level sets $f^{-1}(a)$. The situation is almost identical to that of sublevel sets, in that a basis for $\mathsf{H}(f^{-1}(a))$ is in one-to-one correspondence with the dots in certain rectangular regions; the only difference is that we now have two rectangles and we have to shift the dimension inside one. We state the formula first, and then give some intuition behind the dimension shift; for a rigorous explanation, see [3]. For each a and each dimension p, we distinguish the dots in $\mathrm{Dgm}(f)$ within two rectangular regions,

$$
\begin{aligned}
\lambda_p(a) \;=\; & \{(x, y) \in \mathrm{Ord}_p(f) \mid x < a < y\} \\
\sqcup\; & \{(x, y) \in \mathrm{Ext}_p(f) \mid x < a < y\}, \\
\varrho_p(a) \;=\; & \{(x, y) \in \mathrm{Ext}_p(f) \mid y < a < x\} \\
\sqcup\; & \{(x, y) \in \mathrm{Rel}_p(f) \mid y < a < x\},
\end{aligned}
$$

and we define $\mathcal{W}_p(a) = \lambda_p(a) \sqcup \varrho_{p+1}(a)$. Motivated by their graphical appearance, we call the two rectangles a *pair of wings*. Finally, we have:

LEVEL SET LEMMA. For each regular value a and each dimension p, there exists a natural isomorphism that takes $\mathsf{H}_p(f^{-1}(a))$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{W}_p(a)$.
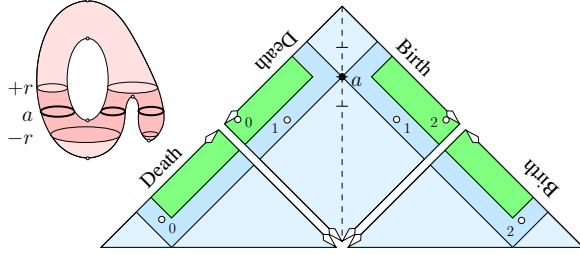

Figure 7: The pair of wings defined by a contains six dots corresponding to the three components and the three 1-cycles of the level set. If we allow a perturbation of the height function of strength up to r, then we capture only two dots.

This result is illustrated in Figure 7. On the left side, we see a level set $f^{-1}(a)$ which consists of three disjoint circles, and on the right we focus on the pair of wings. From the lemma, we know that a basis for $\mathsf{H}_p(f^{-1}(a))$ can be derived in some way from the p-dimensional dots in the left wing and the $(p+1)$-dimensional dots in the right wing.

To see this explicitly, note that $\mathsf{H}_0(f^{-1}(a))$ is rank three. The most obvious choice of basis would be the three 0-dimensional homology classes α, γ, δ, representing, respectively, the three connected components of the level set as read from left to right in the picture. Note that the left wing, $\lambda_0(a)$, contains two dots: the ordinary one gives us the right-most component δ, while the extended dot corresponds either to α or γ, depending on an arbitrary choice made during the persistence computation. On the other hand, the right wing, $\varrho_1(a)$, contains one dot: an extended one, and this corresponds to the essential 1-cycle represented by the vertical circle around the hole in the torus. To transform this cycle into a basis element for $\mathsf{H}_0(f^{-1}(a))$, we take its cross-section at a, which results in a pair of points, one from each of the two left components. In other words, this dot corresponds to the basis element $\alpha + \gamma$ in $\mathsf{H}_0(f^{-1}(a))$. For the expert, we note that this cross-section is a substitute for the connecting homomorphism in the Mayer-Vietoris sequence. A similar explanation gives the transformation between the three dots in $\mathcal{W}_1(a)$ and a basis for the rank three group $\mathsf{H}_1(f^{-1}(a))$.

Robustness. Note that the diagram in Figure 7 also contains a pair of smaller wings, which do not meet and answer the following question: for each homology class in

$\mathsf{H}(f^{-1}(a))$, how strong a perturbation h of f is required so that the level set $h^{-1}(a)$ no longer supports this class? This required strength, measured in the L_∞-distance, is a real number r, and we call the collection of all such numbers, over all homology classes in $f^{-1}(a)$, the *robustness* of the level set; we refer the reader to [12] for a fully rigorous definition. To read the robustness from the diagram, we shrink $\lambda_p(a)$ and $\varrho_p(a)$ as indicated in Figure 7. If a dot leaves the wings after a shrinking by distance r, then we know that there exists a perturbation of strength r that destroys this class. In the example drawn, we note that four dots have already left the pair of wings, two representing components and the other two giving 1-cycles. This reflects the fact that we can move the lower saddle above a and the end of the nose below a, while perturbing f by at most r. Equivalently, we can deform the three circles of the current level set until they merge through addition into a single circle, and we can do this without leaving the interlevel set defined by $[a - r, a + r]$.

Interlevel sets. By making a slight adjustment in the placement of our pair of wings in the diagram, we can also capture the absolute homology of all interlevel sets $f^{-1}[a, b]$, as well as the relative homology of the pairs of spaces $(f^{-1}[a, b], f^{-1}(a) \cup f^{-1}(b))$; we use the notation $f^{-1}(a, b)$ as shorthand for the latter pair. To do this, we set

$$
\begin{aligned}
\lambda_p[a, b] \;=\;& \{(x, y) \in \mathrm{Ord}_p(f) \mid x < b < y\} \\
\sqcup\;& \{(x, y) \in \mathrm{Ext}_p(f) \mid x < b, a < y\}, \\
\varrho_p[a, b] \;=\;& \{(x, y) \in \mathrm{Ext}_p(f) \mid b < x, y < a\} \\
\sqcup\;& \{(x, y) \in \mathrm{Rel}_p(f) \mid y < a < x\}, \\
\lambda_p(a, b) \;=\;& \{(x, y) \in \mathrm{Ord}_p(f) \mid x < a < y\} \\
\sqcup\;& \{(x, y) \in \mathrm{Ext}_p(f) \mid x < a, b < y\}, \\
\varrho_p(a, b) \;=\;& \{(x, y) \in \mathrm{Ext}_p(f) \mid a < x, y < b\} \\
\sqcup\;& \{(x, y) \in \mathrm{Rel}_p(f) \mid y < b < x\},
\end{aligned}
$$

and we define

$$
\mathcal{W}_p(I) \;=\; \begin{cases} \lambda_p[a, b] \sqcup \varrho_{p+1}[a, b] & \text{if } I = [a, b], \\ \lambda_{p-1}(a, b) \sqcup \varrho_p(a, b) & \text{if } I = (a, b). \end{cases}
$$

INTERLEVEL SET LEMMA. For each open or closed interval I and each dimension p, there exists a natural isomorphism that takes $\mathsf{H}_p(f^{-1}(I))$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{W}_p(I)$.

In the closed case, the two wings meet to the right of the vertical axis, so that the left wing ends up larger than the right wing; in the open case, the situation is reversed. For each pair, the dots in the smaller wing must undergo a dimension shift in order to give the required homology basis. For example, if we take (a, b) as shown on the right side
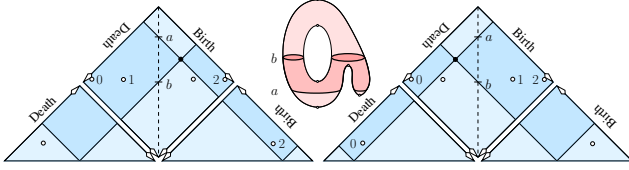
Figure 8: The wings in the left diagram represent the absolute homology of the closed interlevel set, while the wings in the right diagram represent the relative homology of the open interlevel set.

of Figure 8, we have a rank-three relative homology group $H_1(f^{-1}(a,b))$, and there are three dots in $\mathcal{W}_1(a,b)$. The rightmost one of these corresponds to the vertical 1-cycle in the torus, which becomes a relative 1-cycle in the open interlevel set. The connection between the two 0-dimensional dots in the left wing and the other two 1-cycles in the open interlevel set is a bit more complicated. Each dot corresponds to a component within the sublevel set defined by a, and these components then get suspended to the top and bottom boundaries of the interlevel set to form relative 1-cycles. Again for the expert, a rigorous explanation of the dimension shift comes via the Mayer-Vietoris sequence.

4 Reasoning

In this section, we discuss the long exact sequence of a pair, which is an algebraic relationship between the relative homology of a pair of spaces $\mathbb{Y} \subseteq \mathbb{X}$, and the absolute homology of the individual spaces $\mathbb{Y}$ and $\mathbb{X}$. Second, we describe the Mayer-Vietoris exact sequence, which is an algebraic expression of a divide-and-conquer technique that connects the homology of a perhaps complicated space $\mathbb{X} = \mathbb{A} \cup \mathbb{C}$ to the homology of the hopefully simpler subspaces $\mathbb{A}$, $\mathbb{C}$, and $\mathbb{A} \cap \mathbb{C}$. We begin with a detailed discussion of maps between homology groups.

Functoriality. As we have seen, homology assigns a discrete set of algebraic invariants, namely the homology groups along with their Betti numbers, to a topological space. Crucially, this algebraic information can be transferred from one space to another via a continuous map. More precisely, such a map $i : \mathbb{Y} \to \mathbb{X}$ induces linear transformations $i_p : H_p(\mathbb{Y}) \to H_p(\mathbb{X})$, one for each homological dimension p, in such a way that homeomorphisms between spaces induce isomorphisms between groups. This entire process is an example of *functoriality*, one of the key tools in algebra. The induced maps are easiest to understand when i is simply the inclusion of a subspace $\mathbb{Y}$ into a larger space $\mathbb{X}$; for example, let $\mathbb{Y}$ and $\mathbb{X}$ be as in Figure 1. In these situations, the linear map i_p takes a homology class $\alpha \in H_p(\mathbb{Y})$ with cycle representative A, and maps it to the homology class

$i_p(\alpha)$ consisting of all cycles in $\mathbb{X}$ that are homologous to A. For example, the homology class given by the sum of the two components in our $\mathbb{Y}$ gets mapped to zero, since it is a boundary within $\mathbb{X}$. On the other hand, nothing in $H(\mathbb{Y})$ gets mapped to the 2-dimensional homology class of $\mathbb{X}$.

For any pair of spaces $\mathbb{Y} \subseteq \mathbb{X}$, there is also a series of linear transformations $j_p : H_p(\mathbb{X}) \to H_p(\mathbb{X}, \mathbb{Y})$ between the absolute homology of the larger space and the relative homology of the pair. Briefly, j_p takes each homology class $\alpha \in H_p(\mathbb{X})$ with cycle representative A and maps it to the relative homology class $j_p(\alpha) \in H_p(\mathbb{X}, \mathbb{Y})$ consisting of all relative cycles that, relative to $\mathbb{Y}$, are homologous to A. If $(\mathbb{X}, \mathbb{Y})$ is as in Figure 1, then j_0 sends the component of $\mathbb{X}$ to zero, while j_1 maps the 1-dimensional homology class in $\mathbb{X}$ represented by the vertical circle to the relative homology class represented by an arc connecting the two leftmost boundary circles of $\mathbb{Y}$.

Linear transformations and exact sequences. To state our results formally, we need to first recall some terminology from linear algebra, remembering that our homology groups are just vector spaces. Let $j : U \to W$ be a linear transformation between vector spaces. Its *kernel* consists of all elements in the domain that map to zero in the range: $\ker j = \{u \in U \mid j(u) = 0\}$. The *image* of j consists of all elements in the range that have a preimage in the domain: $\operatorname{im} j = \{j(u) \mid u \in U\}$. We note that the kernel is a linear subspace of U, while the image is a linear subspace of W. We also have the *cokernel* of j which we obtain by taking the quotient of the range and the image: $\operatorname{cok} j = W/\operatorname{im} j$. Going back to our example illustrated in Figure 1, the sum of components in $\mathbb{Y}$ is in the kernel of i_0, while the 2-cycle in $\mathbb{X}$ is in the cokernel of i_2. Consider now a sequence of three vector spaces connected by linear transformations: $i : U \to V$ and $j : V \to W$. This sequence is said to be *exact* at V if $\operatorname{im} i = \ker j$. A *long exact sequence* is a doubly infinite sequence of vector spaces and linear maps that is exact at every node. In the examples considered in this paper, all but finitely many vectors spaces in the sequence will be zero.

Finally, we define the *direct sum* $V \oplus V'$ of two vector spaces V, V' to be the set of coordinate pairs (v, v'), with $v \in V$ and $v' \in V'$. Vector operations are defined componentwise. To illustrate this concept in action, we consider again a linear transformation $j : U \to W$. Note that the preimage of each $w \in W$ is of the form $j^{-1}(w) = u + \ker j$. The preimages of $w, w' \in W$ are therefore either disjoint or the same. It follows that the domain is isomorphic to the direct sum of the kernel and the image. We state this more formally, together with a similar decomposition of the range:

$$
\begin{aligned}
U &\cong \ker j \oplus \operatorname{im} j, \\
W &\cong \operatorname{im} j \oplus \operatorname{cok} j.
\end{aligned}
$$

Indeed, every set in the cokernel is of the form $w + \operatorname{im} j$, and the sets are again either disjoint or the same.

Inclusions. Using the diagram, we now describe the images, kernels, and cokernels of the homology maps induced by the inclusion $i : f^{-1}[c, d] \to f^{-1}[a, b]$ whenever $a \le c \le d \le b$ are regular values of f. Fix a homological dimension p. The Interlevel Set Lemma distinguishes bases for the homology groups $H_p(f^{-1}[c, d])$ and $H_p(f^{-1}[a, b])$ which are in one-to-one correspondence with the dots in the regions $\mathcal{W}_p[c, d]$ and $\mathcal{W}_p[a, b]$. It turns out that basic set-theoretic manipulations of these regions suffice to describe the image, kernel, and cokernel of the induced maps i_p; we give intuitive justifications of each result here, leaving the rigor to [3]. Since a class in the larger interlevel set is in $\operatorname{im} i_p$ iff it has a cycle representative in the smaller interlevel set, the following result should not be surprising.

IMAGE LEMMA. For each p, there exists a natural isomorphism that takes $\operatorname{im} i_p$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{R}(\operatorname{im} i_p) = \mathcal{W}_p[a, b] \cap \mathcal{W}_p[c, d]$.

Since $H_p(f^{-1}[a, b])$ is isomorphic to the direct sum of the image and the cokernel of i_p, we also obtain:

COKERNEL LEMMA. For each p, there exists a natural isomorphism that takes $\operatorname{cok} i_p$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{R}(\operatorname{cok} i_p) = \mathcal{W}_p[a, b] - \mathcal{W}_p[c, d]$.

Finally, the kernel of i_p consists of all homology classes in the smaller set that disappear under i_p, and thus cannot be homologous to anything in the larger set.

KERNEL LEMMA. For each p, there exists a natural isomorphism that takes $\ker i_p$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{R}(\ker i_p) = \mathcal{W}_p[c, d] - \mathcal{W}_p[a, b]$.

These claims are illustrated in Figure 9 for the special case $c = d = b$; that is, we include the top level into the interlevel set. In this case, we get five rectangles, two each for the image and cokernel, and one for the kernel. As mentioned above, the domain of i_p is isomorphic to the direct sum of the kernel and the image, and the range of i_p is isomorphic to the direct sum of the image and the cokernel. Accordingly,

$$\begin{aligned} \mathcal{W}_p[b, b] &= \mathcal{R}(\ker i_p) \cup \mathcal{R}(\operatorname{im} i_p), \\ \mathcal{W}_p[a, b] &= \mathcal{R}(\operatorname{im} i_p) \cup \mathcal{R}(\operatorname{cok} i_p); \end{aligned}$$

compare with Figure 9.

Long exact sequence of a pair. Suppose that $\mathbb{Y} \subseteq \mathbb{X}$ is a pair of topological spaces. Then the absolute homology groups $H(\mathbb{X})$ and $H(\mathbb{Y})$ fit together with the relative homology group $H(\mathbb{X}, \mathbb{Y})$ into the following long exact sequence:

$$\overset{k_{p+1}}{\to} H_p(\mathbb{Y}) \overset{i_p}{\to} H_p(\mathbb{X}) \overset{j_p}{\to} H_p(\mathbb{X}, \mathbb{Y}) \overset{k_p}{\to} H_{p-1}(\mathbb{Y}) \overset{i_{p-1}}{\to};$$
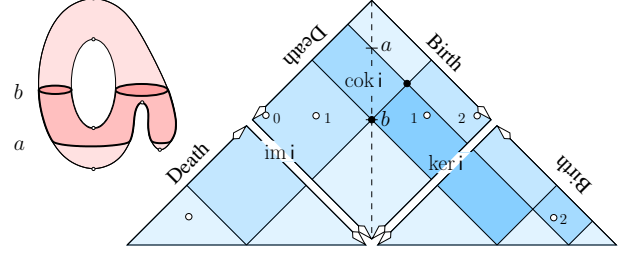

Figure 9: Left: the nosy torus with a shaded closed interlevel set. Right: the persistence diagram in which we show the pair of wings for the interlevel set as well as for the level set defined by the upper endpoint of the interval. Subregions of the wings correspond to the kernel, image, and cokernel of the homology map induced by the inclusion of the level in the interlevel set.

see eg. [18]. Here i_p is the p-dimensional homology map induced by the inclusion, and j_p is the map between absolute and relative homology discussed above. The map k_p is often called the *connecting homomorphism* and is a bit more difficult to understand. We will give a partial explanation of it below, but first we note that it is, at least abstractly, fully characterized by the following three easy consequences of exactness:

$$\begin{aligned} \operatorname{cok} k_{p+1} &\cong \operatorname{im} i_p &\cong \ker j_p, & \text{(2)} \\ \operatorname{cok} i_p &\cong \operatorname{im} j_p &\cong \ker k_p, & \text{(3)} \\ \operatorname{cok} j_p &\cong \operatorname{im} k_p &\cong \ker i_{p-1}. & \text{(4)} \end{aligned}$$

There is a good deal of algebra to unpack here. To illustrate, we again let f be the height function defined on our torus $\mathbb{X}$, and we take $[a, b]$ as indicated on the left side of Figure 9. We explain this exact sequence in terms of the level set $f^{-1}(b)$, the interlevel set $f^{-1}[a, b]$, and the pair $(f^{-1}[a, b], f^{-1}(b)) = f^{-1}[a, b)$, where we have made an obvious adjustment to previous notation. First, we already understand the linear transformation

$$i_p : H_p(f^{-1}(b)) \to H_p(f^{-1}[a, b]).$$

Indeed, the Image, Kernel, and Cokernel Lemmas tell us how to read bases for $\operatorname{im} i_p$, $\ker i_p$, and $\operatorname{cok} i_p$, for each dimension p, from the diagram: they correspond to the dots in the regions indicated on the right side of Figure 9. We encourage the reader to verify that the dots in the rectangular regions give the correct ranks in all cases. Using (2) to (4), we immediately have bases for $\operatorname{im} j_p$, $\ker j_p$, and $\operatorname{cok} j_p$, again for all dimensions p, where

$$j_p : H_p(f^{-1}[a, b]) \to H_p(f^{-1}[a, b))$$

maps absolute to relative homology. In particular, we can use that $H_p(f^{-1}[a, b))$ is isomorphic to the direct sum of the image and the cokernel to derive a further diagram-reading rule:

9

HALF-OPEN INTERLEVEL SET LEMMA. For each half-open interval $[a, b)$ and each dimension p, there exists a natural isomorphism that takes $H_p(f^{-1}[a, b))$ to the vector space in $\mathcal{V}$ whose basis corresponds to $\mathcal{W}_p[a, b) = \mathcal{R}(\operatorname{im} j_p) \cup \mathcal{R}(\operatorname{cok} j_p)$.

In Figure 9, we see this union formed by $\mathcal{R}(\operatorname{cok} i_p)$ and $\mathcal{R}(\ker i_{p-1})$. Counting the dots in these regions and recalling the dimension shift in the right wing, we note that the Betti numbers of $f^{-1}[a, b)$ in dimensions 0, 1, and 2 are indeed 0, 2 and 0, as they should be. Incidentally, we could use the exact sequence of a pair to deduce how to read the homology of an open interval set. The result is of course the formula we claimed for $I = (a, b)$ in the Interlevel Set Lemma. Another application of (2) to (4) gives the regions corresponding to the image, kernel, and cokernel of the connecting homomorphism,

$$k_p : H_p(f^{-1}[a, b)) \to H_{p-1}(f^{-1}(b)).$$

For example, focusing our attention on dimension $p = 1$, we note that $\operatorname{im} k_1 \cong \ker i_0$ and $\mathcal{R}(\ker i_0)$ contains a single dot. In the relative homology of the pair $f^{-1}[a, b)$, this dot corresponds to the relative 1-dimensional homology class represented by an arc A connecting the two portions of $f^{-1}(b)$. This class is mapped, via k_1, to the sum of the two components in $H_0(f^{-1}(b))$. In other words, k_1 takes the relative class represented by A and maps it to the absolute class represented by the boundary of A. This is indeed the general rule for the boundary map in the exact sequence of a pair, although we omit further details.

Mayer-Vietoris exact sequence. As promised, we now explain the Mayer-Vietoris divide-and-conquer technique via the diagram. First, suppose $\mathbb{X} = \mathbb{A} \cup \mathbb{C}$ is a decomposition of a topological space into two closed subspaces, and set $\mathbb{B} = \mathbb{A} \cap \mathbb{C}$. We let $i^{\mathbb{A}}$ and $i^{\mathbb{C}}$ denote the homology maps induced by the inclusions of $\mathbb{B}$ into $\mathbb{A}$ and into $\mathbb{C}$, and we denote by $j^{\mathbb{A}}$ and $j^{\mathbb{C}}$ the homology maps induced by the inclusions of $\mathbb{A}$ and of $\mathbb{C}$ into $\mathbb{X}$. Then the *Mayer-Vietoris sequence*:

$$\overset{k_{p+1}}{\to} H_p(\mathbb{B}) \overset{i_p}{\to} H_p(\mathbb{A}) \oplus H_p(\mathbb{C}) \overset{j_p}{\to} H_p(\mathbb{X}) \overset{k_p}{\to} H_{p-1}(\mathbb{B}) \overset{i_{p-1}}{\to},$$

is exact, where i_p is defined by $i_p(\alpha) = (i_p^{\mathbb{A}}(\alpha), i_p^{\mathbb{C}}(\alpha))$, while j_p is given by the formula $j_p(\alpha, \gamma) = j_p^{\mathbb{A}}(\alpha) + j_p^{\mathbb{C}}(\gamma)$. The map k_p is again a connecting homomorphism, and we will illustrate it, as well as the other two maps, via the following example.

As always, we let f be the height function on our torus $\mathbb{X}$, and we choose $a < b < c$ as indicated in Figure 10. We will discuss the Mayer-Vietoris sequence in terms of the decomposition $f^{-1}[a, c] = f^{-1}[a, b] \cup f^{-1}[b, c]$. To use the notation above, we set $\mathbb{A} = f^{-1}[a, b]$ and $\mathbb{C} = f^{-1}[b, c]$, noting
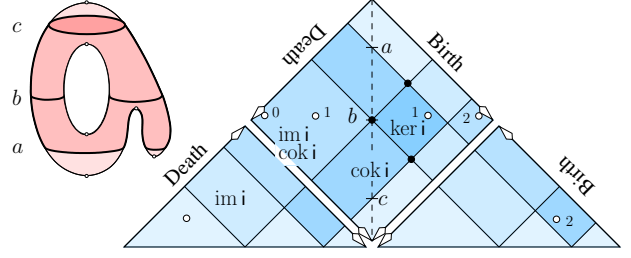

Figure 10: Left: the decomposition of $f^{-1}[a, c]$ into the interlevel sets defined by $[a, b]$ and $[b, c]$. Right: the wings of the two interlevel sets as well as the level set at which the interlevel sets intersect. The wings are decomposed into rectangular regions that correspond to the kernels, images, and cokernels of the various maps. Two of the rectangles belong to both the image and the cokernel of the homology map induced by the inclusion of $\mathbb{B}$ in $\mathbb{A}$ and $\mathbb{C}$.

that $\mathbb{A}$ and $\mathbb{C}$ intersect in the level set $\mathbb{B} = f^{-1}(b)$. Now the maps $i^{\mathbb{A}}$ and $i^{\mathbb{C}}$ are fully characterized by the lemmas above. To understand the map i_p from $H_p(f^{-1}(b))$ to the direct sum of $H_p(f^{-1}[a, b])$ and $H_p(f^{-1}[b, c])$, we exploit the following formulas:

$$\begin{aligned} \mathcal{R}(\operatorname{im} i_p) &= \mathcal{R}(\operatorname{im} i_p^{\mathbb{A}}) \cup \mathcal{R}(\operatorname{im} i_p^{\mathbb{C}}), \\ \mathcal{R}(\ker i_p) &= \mathcal{R}(\ker i_p^{\mathbb{A}}) \cap \mathcal{R}(\ker i_p^{\mathbb{C}}), \\ \mathcal{R}(\operatorname{cok} i_p) &= \mathcal{R}(\operatorname{cok} i_p^{\mathbb{A}}) \cup \mathcal{R}(\operatorname{cok} i_p^{\mathbb{C}}) \\ &\cup \ [\mathcal{R}(\operatorname{im} i_p^{\mathbb{A}}) \cap \mathcal{R}(\operatorname{im} i_p^{\mathbb{C}})]. \end{aligned}$$

Only the third formula needs an explanation. If the regions of $\operatorname{im} i_p^{\mathbb{A}}$ and $\operatorname{im} i_p^{\mathbb{C}}$ share ℓ dots, then they contribute 2ℓ dimensions to the direct sum. Only the ℓ-dimensional diagonal of this linear subspace belongs to the image of i_p. The remaining ℓ dimensions belong to the cokernel of i_p. Indeed, the dots in $\mathcal{R}(\operatorname{im} i_p^{\mathbb{A}}) \cap \mathcal{R}(\operatorname{im} i_p^{\mathbb{C}})$ do double duty to cover the extra dimensions of the direct sum construction. As in the case of the exact sequence of a pair, we can exploit the isomorphism formulas (2) to (4) to understand the images, kernels, and cokernels of j_p and k_p. For instance, from (4) we get

$$\mathcal{R}(\operatorname{im} k_1) = \mathcal{R}(\ker i_0) = \mathcal{R}(\ker i_0^{\mathbb{A}}) \cap \mathcal{R}(\ker i_0^{\mathbb{C}}).$$

This region contains a single dot contributed by the vertical circle going around the hole of the torus. This class is in the kernel of both $i_0^{\mathbb{A}}$ and $i_0^{\mathbb{C}}$; indeed, both $\mathbb{A}$ and $\mathbb{C}$ are connected, so $x + y$ forms the boundary of an arcs A in $\mathbb{A}$ and another arc C in $\mathbb{C}$. On the other hand, notice that A and C then have a common boundary, and so they glue together to form the 1-dimensional homology class $\gamma \in H_1(\mathbb{A} \cup \mathbb{C})$ represented by $A + C$. The map k_1 is defined in such a way that $k_1(\gamma) = \alpha$, and hence $\alpha \in \operatorname{im} k_1$, as our formulas require. Intuitively, γ is mapped by k_1 into the 0-cycle created by taking a cross-section at b.

5 Discussion

The contribution of this paper is the exposition of a new, combinatorial angle at a classic subject within algebraic topology: the characterization of topological features of a space or a function through homology groups. This new angle is facilitated by the relatively recent addition of persistence to the theory of homology groups. In particular, the persistence diagram of a function forms a comprehensive book-keeping tool on which we build our point calculus. We believe that topological information is useful also to non-topologists, even to non-mathematicians, so we have written this paper with an eye on minimizing the formalism and adding intuitive explanations where the formalism seemed necessary or useful.

In the interest of brevity, we have not described the algorithms needed to compute persistence diagrams. Suffice to say that they exist and are fast; see [10] for a general introduction and [2] for algorithms appropriate for image data.

References

[1] P. K. AGARWAL, H. EDELSBRUNNER, J. HARER AND Y. WANG. Extreme elevation on a 2-manifold. *Discrete Comput. Geom.* **36** (2006), 553–572.

[2] P. BENDICH, H. EDELSBRUNNER AND M. KERBER. Computing persistence and robustness for images. *IEEE Trans. Visual. Comput. Graphics* **16** (2010), 1251-1260.

[3] P. BENDICH, H. EDELSBRUNNER, D. MOROZOV AND A. PATEL. Homology and robustness of level and interlevel sets. Manuscript, IST Austria, 2010.

[4] E. BETTI. Sopra gli spazi di un numero qualunque di dimensioni. *Ann. Mat. Pura Appl.* **2/4** (1871), 140–158.

[5] G. CARLSSON AND V. DE SILVA. Zigzag persistence. *Found. Comput. Math.* **10** (2010), 367–405.

[6] G. CARLSSON AND V. DE SILVA AND D. MOROZOV. Zigzag persistent homology and real-valued functions. *In* "Proc. 25th Ann. Sympos. Comput. Geom., 2009", 247–256.

[7] G. CARLSSON, T. ISHKHANOV, V. DE SILVA AND A. ZOMORODIAN. On the local behavior of spaces of natural images. *Internat. J. Comput. Vision* **76** (2008), 1–12.

[8] A. CAYLEY. On contour and slope lines. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **18** (1859), 264–268.

[9] M. L. CONNOLLY. Shape complementarity at the hemo-globin albl subunit interface. *Biopolymers* **25** (1986), 1229–1247.

[10] H. EDELSBRUNNER AND J. L. HARER. *Computational Topology. An Introduction.* Amer. Math. Soc., Providence, Rhode Island, 2010.

[11] H. EDELSBRUNNER, D. LETSCHER AND A. ZOMORODIAN. Topological persistence and simplification. *Discrete Comput. Geom.* **28** (2002), 511–533.

[12] H. EDELSBRUNNER, D. MOROZOV AND A. PATEL. Quantifying transversality by measuring the robustness of intersections. Manuscript, Dept. Comput. Sci., Duke Univ., Durham, North Carolina, 2009.

[13] L. EULER. Demonstratio nonnullarum insignium proprietatum, quibus solida hedris planis inclusa sunt praedita. *Novi Commentarii academiae scientiarum Petropolitanae* **4** (1758), 140–160.

[14] R. GHRIST. Barcodes: the persistent topology of data. *Bulletin Amer. Math. Soc.* **45** (2008), 61–75.

[15] M. HILAGA, Y. SHINAGAWA, T. KOHMURA AND T. L. KUNII. Topology matching for fully automatic similarity estimation of 3D shapes. *In* "Proc. SIGGRAPH, 2001", 203–212.

[16] J. C. MAXWELL. On hills and dales. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **40** (1870), 421–427.

[17] J. MILNOR, *Morse Theory*, Princeton Univ. Press, Princeton, New Jersey, 1963.

[18] J. R. MUNKRES. *Elements of Algebraic Topology.* Perseus, Cambridge, Massachusetts, 1984.

[19] V. PASCUCCI. Topology diagram of scalar fields in scientific visualization. In *Topological Data Structures for Surfaces*, ed. S. Rana, Wiley, Chicester, England, 2004.

[20] H. POINCARÉ. *Papers on Topology: Analysis Situs and Its Five Supplements.* Translated by John Stillwell, Amer. Math. Soc., Providence, Rhode Island, 2010.